Marianna Sávai (ed) 2024: Green and Digital Transitions: Global Insights into Sustainable Solutions, Szeged, 10.14232/gtk.gdtgiss.2024.13

Explaining income inequality by the relationship between social network fragmentation and social segregation indicators

Byambasuren Dorjnyambuu

This paper investigates the way social networks and social indicators of segregation interact and their relationship with income disparity for 426 towns and cities in Hungary. Three social indicators of segregation are used to capture different characteristics of social segregation in towns: (i) ethnic fragmentation, (ii) religious fragmentation, and (iii) education inequality. Using open-access data from Tóth et al. (2021), non-spatial and spatial two-stage least square models are estimated for income inequality at the town level. The study finds that these social segregation indicators positively correlate with income inequality through social network fragmentation. Also, the spatial model shows that income inequality has a strong spatial relationship across towns.

Keywords: Inequality; social network; fragmentation; social segregation; spatial two-stage model
JEL classifications: J31, D31, C36

1. Introduction

Rising income inequality remains one of today's significant social and economic difficulties. Widening income inequality is harmful to socioeconomic well-being through many channels. It is the cause of many social issues, such as poverty, crime, unemployment, health problems, lower life expectancy, and lower levels of education (Stiglitz, 2012). A large amount of literature has been written to study the reasons for income inequality, with most studies focusing more on evidence from the United States and the OECD than other nations (Cavanaugh–Breau, 2018). The factors explaining the income disparities can be categorized into six broad explanations: (i) structural macroeconomic sectoral changes, (ii) globalization and technology change, (iii) labor market and other relevant institutions, (iv) demographic and microstructural changes, (v) politics and political processes, and (vi) tax/transfer schemes (Förster–Tóth, 2015). However, many objections still exist over why the expansion in income disparity occurred. The interaction of structural elements such as social networks and geography that influence inequality is less known (Tóth et al., 2021).

Network effects arise when a person's likelihood of adopting a habit rises with the number of other people in their social network who have already done so. Individuals' choices are influenced by their network peers in several domains, such as network externalities, social learning and peer assistance, and normative influence (DiMaggio–Garip, 2012). Studies on social networks highlight how social relationships enable individuals to take advantage of economic opportunities (Granovetter, 1985). However, inequality is aggravated when the impacts of individual differences are magnified via social networks, particularly when economic status influences the formation of social relationships (DiMaggio−Garip, 2012). Social networks are ingrained in geography, which has profound implications for inequality. For instance, much of an individual's economic potential and access to possibilities through education is determined by where they live (Chetty et al., 2014). Even within relatively small geographical units such as cities and towns, there is a concomitant difference in outcomes between neighborhoods (Glaeser et al., 2009).

There has been very little research into the regional aspects of income disparity. This is due, in part, to a lack of suitably disaggregated data linking geography to income at an individual or household level. Data from wider geographical areas cannot indicate heterogeneity between small geographical units or locations with easy access to a prosperous labor market and more isolated ones. The main motivation of this paper is the open-access data and findings of Tóth et al. (2021) on the joint relationship between social network structure, urban topology, and inequality in Hungarian towns. Tóth et al. (2021) investigated how social networks and urban topology interact and how they affect inequality. They employed three urban topology indicators to capture different characteristics of social segregation in towns: (i) the average distance from the center, (ii) the extent of spatial concentration of amenities in towns, and (iii) the degree to which physical barriers divide residential areas. They found that social network fragmentation is substantially higher in towns where residential communities are divided by physical barriers like rivers and railroads. Towns with relatively remote neighborhoods from the center of town and spatially concentrated amenities are also more socially divided. Finally, they concluded that the spatial characteristics of a location could exacerbate economic inequities through social network fragmentation using a two-stage model. If social network fragmentation is related to income inequality, then the question arises whether other social segregation metrics can explain network relations with inequality. In this regard, this paper aims to examine how social segregation measures connect to inequality via their relationships to social network fragmentation. I use three social measures to capture segregation characteristics: ethnic fragmentation, religious fragmentation, and education disparities. This paper relies on work by Tóth et al. (2021) in data and approach but changes the channels explaining social network fragmentation and extends the estimation method using a spatial two-stage model. The data includes social network, population, and socioeconomic information for 426 towns in Hungary from 2011 to 2016.

This study offers two contributions to the existing literature. First, it stands out for its focus on inequality at the town level, departing from the more common country- or regional-level analyses. Second, exploring inequality through the lenses of spatial and social network fragmentation and conventional indicators adds depth to current research in this area.

The paper is organized as follows. Section 2 examines the relevant literature. The data are introduced in Section 3, which describes spatial characteristics of income inequality across towns, social network fragmentation, and control and instrumental variables. The key findings are then reported in Section 4, and the conclusion is provided in Section 5.

2. Literature review

Income inequality trends and drivers: Medgyesi and Tóth (2021) documented that Hungary had the lowest inequality group during the pre-transition period of 1980-1984, with Gini values below 0.25. However, in the early 1990s, Eastern European (EE) countries underwent a transitional recession, resulting in a significant fall in GDP and increased income inequality (Flemming–Micklewright, 2000). Kattuman and Redmond's (2001) study on income inequality in Hungary from 1987 to 1996 revealed that inequality remained stable until 1991, with factors like falling real incomes and changes in taxes and state transfers playing a role. However, after 1991, income inequality significantly increased due to growing earnings disparity, diverse sources of household income, and emerging disparities in state transfers. Furthermore, throughout the transition period, some sources provide cross-country evidence on trends and determinants of inequality in EE countries (Milanovic, 1999; Flemming–Micklewright, 2000; Aristei–Perugini, 2015).

Authors primarily concluded that the key contributors to rising income inequality in EE countries throughout the transition period were widening differences in labor income distribution, the increasing importance of capital income, and the deterioration of the redistributive impact of welfare state programs. Milanovic (1999) is the first of the studies to give a cross-country perspective on the determinants of income inequality during the transition period in six countries, including Hungary. His findings revealed that the growing wage distribution discrepancy, exacerbated by the reduction in employment, was the fundamental cause of rising overall inequality. Pensions, somewhat unexpectedly, also contributed to increased inequality in EE countries. Meanwhile, both inadequately funded and poorly targeted social transfers had minimal impact on mitigating inequality across countries (Milanovic, 1999). Profits and capital income became more important as the private sector emerged and state-owned firms were privatized, increasing economic disparity in EE countries. This is because capital income is more unequally distributed than labor income (Medgyesi–Tóth, 2022).

Nonetheless, variations in policies and results were evident among EE nations throughout the transition period. Hungary's encounter with inequality during this transition period deviated slightly from that of other countries, marked by a relatively smaller uptick in inequality. This disparity can be attributed to a blend of factors, including Hungary's early initiation of economic liberalization, influential dynamics driving inequality, and diverse social policies that offered some safeguards for individuals in the lower echelons of the labor market (Tóth, 2008). As a result, before becoming a member of the EU, income inequality in Hungary was roughly in line with the EU-15 average (Medgyesi–Tóth, 2021). However, during the global financial crisis, income inequality in Hungary saw a significant and alarming increase of 13.5%. This spike in inequality can largely be attributed to the reduced progressiveness of Hungary's tax system, which had a substantial impact. At the same time, the decline in the full-time employment rate also played a contributing but comparatively smaller role (Brzezinski, 2018).

Aside from the previously identified transitional and structural issues, the next strand of literature has taken a partial approach to exploring the sources of income inequality by keeping a cross-country perspective in EE countries. These sources include economic development (Tsaurai, 2020), globalization and technological progress (Esposito–Stehrer, 2009; Josifidis et al., 2021; OECD, 2011), labor market institutions (OECD, 2011; Szczepaniak–Szulc-Obłoza, 2020), education and human capital accumulation (Omoeva et al., 2018; OECD, 2011), demographic shifts (Dolls et al. 2019), migration (Docquier et al., 2019), and firm characteristics (Magda et al., 2021).

As of 2020, inequality at the national level is moving closer to the EU average, but inequality within the country is becoming more heterogeneous. Hungary's Gini index of equalized disposable income was slightly lower than the EU average, but it has increased by 3.9 percentage points over the last decade. The top 20% of Hungarian households earn four times more than the bottom 20% of households in the income distribution. These findings are supported by examining changes in quintile distributions from 2010 to 2020. Income transfers in Hungary demonstrated a regressive tendency, with the lower quintiles either losing their share of income or seeing no improvement, while the top quintile had the most significant improvements (Table 1).

Table 1. The main statistics in income inequality in Hungary, 2020

Country	Gini index of equalized disposable[1] income[(a)]		Income quantile share ratio S80/S20 for disposable income[(b)]		Share of national equivalized income by[(c)]									
					First quintile		Second quintile		Third quintile		Fourth quintile		Fifth quintile	
	2020	Change 2010-2020	2020	Change 2010-2019	2020	Change 2010-2020	2020	Change 2010-2020	2020	Change 2010-2020	2020	Change 2010-2020	2020	Change 2010-2020
EU-27	30.2	0	4.9	0.1	8.0	0	13.4	0	17.7	0.2	22.8	0.1	38.1	-0.3
Hungary	28.0	3.9	4.2	0.8	8.8	-1.2	13.8	0	17.8	-0.6	22.9	0.2	36.6	2.4

Sources: Eurostat (2022a); Eurostat (2022b); Eurostat (2022c)
Note: Online codes: (a) ilc_di12, (b) ilc_di11, (c)ilc_di01

Regarding spatial income inequality, Pénzes et al. (2014) investigated the trajectory of spatial income disparities in Hungarian regions (NUTS-2) from 1988 to 2012. The study revealed that the increase in income inequalities was notably more pronounced in underdeveloped areas. Conversely, developed regions experienced a more modest rise in income inequalities, demonstrating their ability to attract new investments and restructure their economies. Vida (2022) examined regional income

[1] Equalized disposable income is determined by taking a household's overall income, which includes earnings, after tax contributions, and other deductions, and then dividing it by an equivalency scale.

disparities in Hungary from 2010 to 2019, focusing on their correlation with geographic trends in "realized competitiveness." The findings reveal significant spatial variations in regional performance and residents' income levels during this period, influenced by economic, spatial, and social factors. While some areas saw reduced spatial disparities, income gaps among different regions widened. Importantly, these changes also manifested in the spatial patterns of regional competitiveness within districts. Zoltán (2022) underscored the importance of geographic proximity in shaping income inequalities and dynamics among local settlements in Hungary, highlighting the formation and evolution of income clubs as a significant aspect of this phenomenon.

Network effects on inequality. Researchers who have examined network effects in various fields often find that these effects tend to exacerbate social inequality. For example, in the context of health, Pampel et al. (2010) suggested that individuals with higher socioeconomic status tend to adopt healthier behaviors and form connections with others of similar status, strengthening their social networks and ultimately contributing to improved health outcomes and widening health disparities. In education, Gamoran (2011) highlighted that school tracking, a form of induced homophily, typically amplifies disparities in academic achievement, thus exacerbating educational inequalities.

According to DiMaggio and Garip (2012), network effects can potentially worsen disparities among different groups in adopting beneficial practices. Their financial or cultural resources should positively influence an individual's likelihood of adopting a useful practice. Financial resources enhance a person's capability to engage in the practice due to improved affordability, while cultural resources, typically assessed through years of formal education, play a role by elevating awareness of new practices, enhancing comprehension of intricate innovations, or facilitating more efficient utilization of these practices. Moreover, individuals' social networks should comprise people who resemble them regarding traits indicative of their likelihood to adopt the new practice. This phenomenon is known as homophily, where individuals tend to establish connections with those who share similar socioeconomic and demographic characteristics. Homophily is observed across various contexts, including adult friendship networks, children's friendship networks, and even marital choices (Bianconi et al., 2014; Rivera et al., 2010). Homophily can arise due to structural factors or individual decisions, but in either case, it can create conditions that lead to increased inequality when network effects are at play (McPherson et al., 2001). This ubiquity of network effects on the micro-level can lead to social segregation on a larger scale: groups with various socioeconomic statuses are segregated in social networks (Stadtfeld, 2018). If access to resources or information passes across the network, this macro-scale network structure might lead to divergent economic potentials between groups (Tóth et al., 2021).

Social network and geography. Social networks are closely intertwined with geography, and this relationship has significant implications for inequality (Tóth et al., 2021). For instance, where an individual lives, often indicated by their home location, can strongly predict a substantial portion of their future economic prospects. Namely, individuals living in certain neighborhoods or areas may have better access to educational opportunities, job markets, and other economic resources that can positively impact their economic outcomes (Chetty et al., 2014). Interestingly, economic disparities are not solely limited to comparisons between regions or countries. Even within relatively small geographical units like cities and towns, significant divergence in economic outcomes can be observed. Glaeser et al. (2009) reviewed the economic causes of income inequality in metropolitan areas and found that differences in both skill distribution and returns to skill play essential roles in explaining income inequality variations across metropolitan areas. This suggests that disparities in economic well-being can be highly localized and not just a macro-level phenomenon.

Furthermore, people build social relationships with others who are physically close to them. This "local bias" indicates that people are more inclined to form ties with neighbors, coworkers, or community members in the same geographical area. Local relationships can impact many aspects of life, including job opportunities, social support, and resource access (Sampson, 2008). Geography primarily shapes economic outcomes by influencing the composition of social networks. When social interactions are constrained to a particular geographical region, it can limit personal and collective advancement. This is because access to a wide range of resources, information, and opportunities provided by socially distant connections is vital for progress. When social networks are spatially bounded, people may miss out on valuable connections and resources that could enhance their economic well-being (Bailey et al., 2018; Eriksson–Lengyel, 2019).

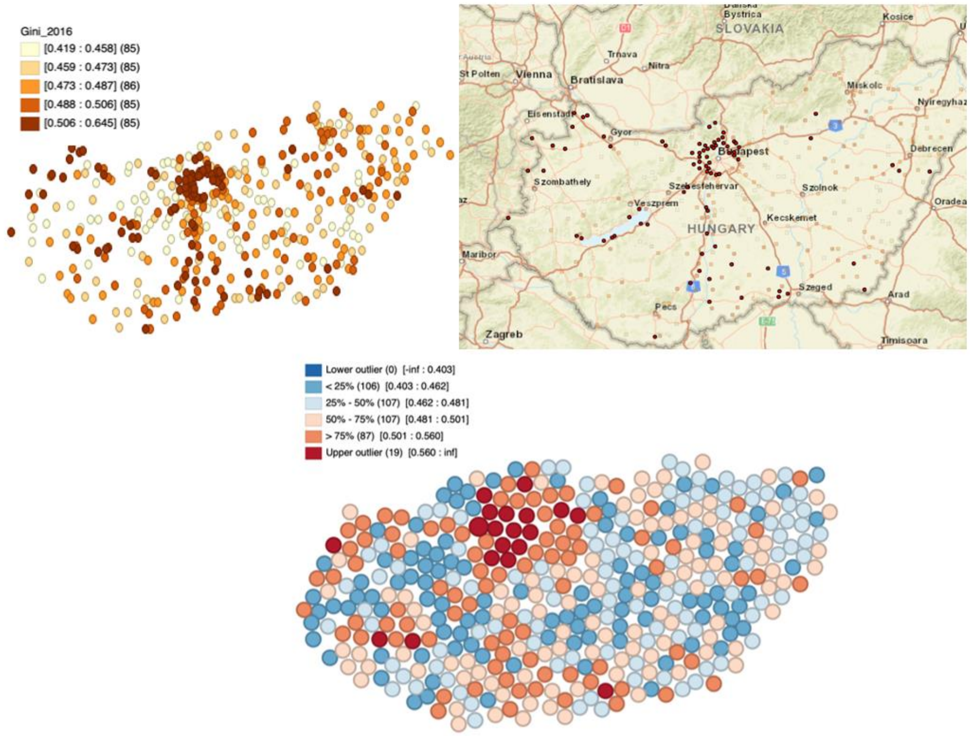
In essence, geography plays a pivotal role in shaping economic inequality by influencing the structure of social networks. Social ties formed within specific geographic boundaries can either enhance or limit individuals' and communities' access to opportunities, resources, and economic success.

3. Data and variables

This paper uses open-source geographic data (Open Street Map data) and town-level aggregate data applied in the recent paper by Tóth et al. (2021), which includes population and socioeconomic information for 474 towns in Hungary. These towns and cities comprise about 60% of Hungary's population. The capital city of Budapest was excluded from the data due to its considerable disparities with other cities and towns. Because 48 cities were not matched when the data was matched to the map, this paper used 423 cities. Tóth et al. (2021) estimated the Gini index at the town level in 2011 and 2016 to quantify income inequality based on total income distributions across HNSO income categories. This paper uses the town Gini index 2016 ($G_{(i,2016)}$) as a dependent variable.

According to Eurostat, Hungary's income disparity has been lower than the EU-27 average for the past decade (the 10-year average is 0.28), but increases in regional disparities within the country have moderated it. Specifically, Hungary has come closer to EU averages while moving further away from other marginalized parts of its territory. The income inequalities in 426 towns and cities range from 0.42 (Füzesgyarmat) to 0.65 (Telki). The towns neighboring Budapest, those along Lake Balaton, certain towns near the western and southern borders, and those along highway M6 have the highest income inequalities among Hungarian towns (Figure 1.). It is worth noting that inequality is low in towns in the east and north of the country.

Figure 1. Town Gini Index 2016 (Quintile and Cartogram map)



Source: own elaboration based on the dataset of Tóth et al. (2021)

Figure 2a depicts box plots of the towns' Gini index in 2011 and 2016 with descriptive statistics. In terms of inequality developments, there was a strong correlation between inequality in 2011 and 2016 (Figure 2b.). The overall level of inequality in most towns increased somewhat, from an average Gini index of 0.476 in 2011 to an average of 0.485 in 2016. Between 2011 and 2016, inequality increased in around 69% of total towns, remained unchanged in 12%, and declined in 19% of towns; the highest increases and decreases in town Gini were reported in Zalakomár (15%) and Balatonkenese (10%), respectively. High-income disparities have also been observed in high-income and high-density towns, as indicated by the town Gini coefficient's positive correlation with income per capita and population density (Figure 2b.).

Figure 2a. Box plot for the town Gini index for 2011 and 2016

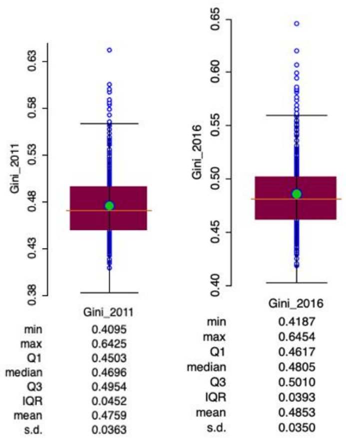
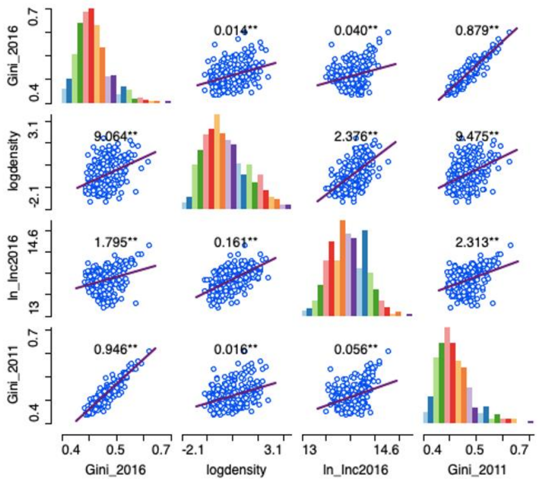


Figure 2b. Correlation and distribution of the variables



Source: own elaboration based on the dataset of Tóth et al. (2021)
Note: (Gini index 2016; ln (population density), ln (income per capita 2016), and Gini index 2011)

Regarding spatial statistics, Moran's I statistic was 0.220 in 2016, indicating that income inequality has a weak spatial autocorrelation. Focusing on the towns where inequality is positively and strongly linked with its spatially lagged counterparts (the first or High-High quadrant of Moran scatter plot), spatial autocorrelation has increased (Moran's I is 0.287) than the overall statistic of 0.220, but unselected observations indicate no spatial autocorrelation (a value of -0.052). This result would imply the presence of spatial heterogeneity in the strength of the spatial autocorrelation, as the subset chosen exhibits a significantly different degree of dependency than its complement or the whole dataset (Figure 3.).

The local indices for spatial analysis (LISA) show that 113 of 426 towns have a significant local spatial association. Based on the position of the value and its spatial lag in the Moran scatter plot, the LISA's cluster map illustrates the significant places with an indicator of the type of spatial relationship. Among the 113 towns, 38 are in high-high clusters, 44 are in low-low clusters, 17 are in low-high clusters, and 14 are in high-low clusters (Figure 4.). Looking at the relationship between the Morgan scatter plot and the cluster map, are 81 towns selected in the High-High quadrant of the Moran scatter plot, of which 38 are significant on the cluster map (see Annex-1).

Figure 3. Moran scatter plot in the town Gini index 2016 (focusing on a subset, including 81 towns)*

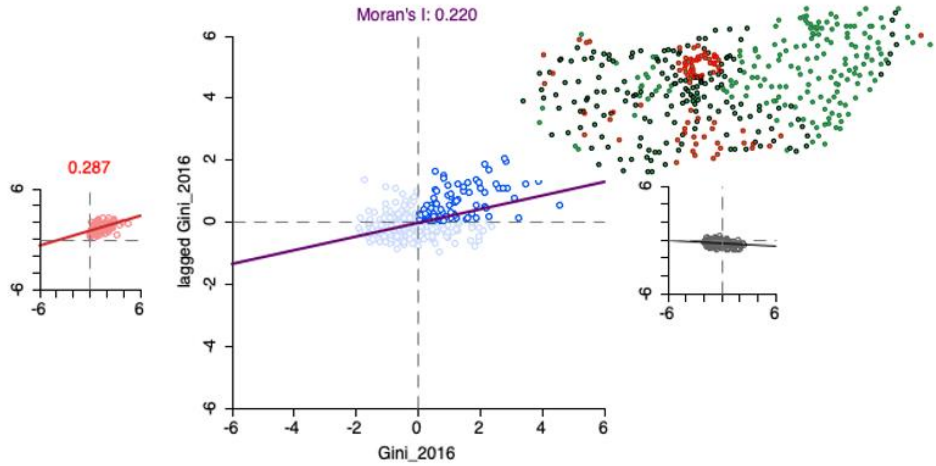
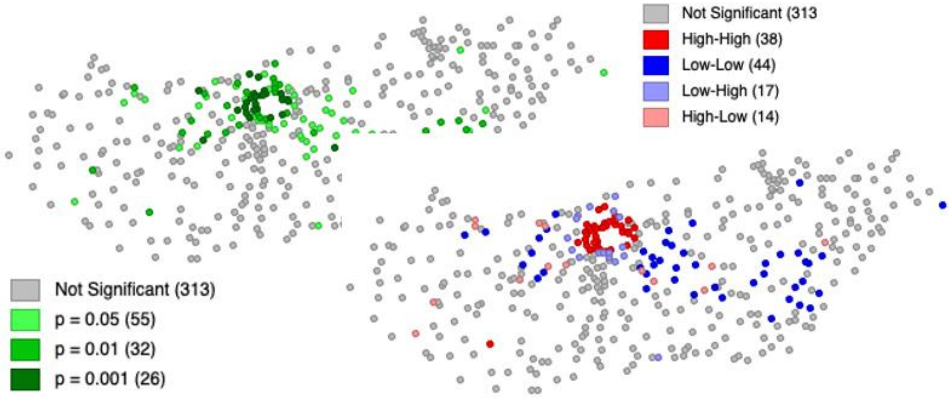


Figure 4. LISA Significance and cluster maps



Source: own elaboration based on the dataset of Tóth et al. (2021)
Note: *The one on the left (in red) is for the selected observations, while the one on the right is for the complement, often known as unselected observations (in black)

Another important variable in this paper is social network fragmentation within a town. Tóth et al. (2021) used data from a Hungarian online social network, iWiW, to describe social network structure within towns. The iWiW (International Who Is Who) network was founded in 2002 and quickly became Hungary's most popular online social network. At its peak in 2010, it was one of the most popular national websites, reaching about 40% of the country's population (Tóth et al., 2021). The site was permanently shut down in 2014 due to heavy competition from Facebook. It is now being utilized as a large-scale dataset, including location (self-reported), birthday, gender, date of registration and last login, the ID of friends, and ID of inviters, to research the social interactions of the Hungarian people. Regarding population representativity, the iWiW was accessible to those aged 14 and above, potentially reaching 8.2 million people in Hungary. By early 2013, roughly 33% of Hungarians aged 14 and older were members of this network. When compared to nationally representative internet usage surveys conducted in 2013, approximately half of the adult online population was iWiW users. The age distribution of iWiW users closely aligns with the estimated number of internet users in Hungary up to age 60 but decreases significantly after age 70. The network is well-represented among Hungary's economically active population (Supplementary information of Tóth et al. 2021). Tóth et al. (2021) substituted online social ties for in-person social relationships. While this method oversimplifies the complexities of social interactions, they argue that it is still the most accurate data source available. It is important to acknowledge that data is imperfect and comes with limitations, such as not having insights into the specifics of these social connections, including their nature, strength, or how often people communicate. Nevertheless, they believe that there is no systematic bias in the data that would undermine the credibility of the analysis.

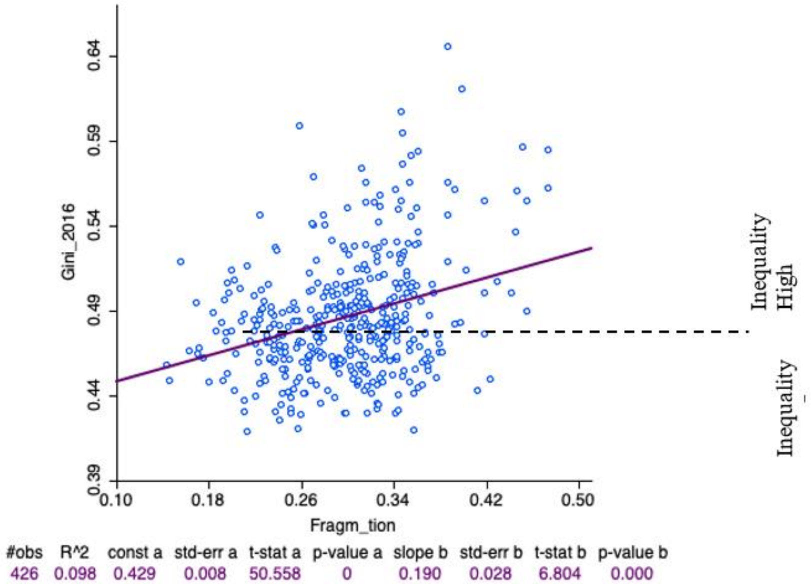
It is important to note potential geographical biases in representativeness, as innovations like iWiW were notably influenced by age, education, and location dynamics. Initially, it gained popularity among young, educated urban demographics, and while later adoption included elderly individuals from rural areas, it never matched the earlier rates. The overall iWiW user rate spans from 23% in small villages to 42% in major cities, with smaller settlements being over-represented by the elderly and those with lower educational levels. Nonetheless, the relatively limited number of outlier settlements in this pattern strengthens the reliability of iWiW data.

When examining social network fragmentation within a town, Tóth et al. (2021) only look at ties between iWiW users that live in the same town through the Louvain algorithm, which detects communities. This approach divides the people in the town *i*'s network into groups by optimizing a modularity metric Q_i, which compares the density of edges within groups to the density across groups. To eliminate the dependence on the size and density of the network, they scaled it by the maximum value of Q_i. Then, the fragmentation is

$$Frag_i = \frac{Q_i}{Q_i^{max}} \tag{1}$$

The fragmentation values in the data were created for the end of 2011. There is a positive correlation (ρ=0.3) between the town Gini 2016 and social network fragmentation (Figure 5.).

Figure 5. The correlation between town Gini 2016 and social network fragmentation



Source: own elaboration based on the dataset of Tóth et al. (2021)
Note: The fitted line indicates a linear regression: $G_{i,2016} = 0.429 + 0.19F_i$; and the dashed horizontal line represents the town Gini 2016 mean (0.485).

Lengyel and Jakobi (2016) found that the economic development of towns influences the adoption of online social networks, including iWiW, with wealthier towns having a higher share of iWiW users. As a result, low-income people may be underrepresented on iWiW, thus skewing the statistical association between social network fragmentation and income disparity. Tóth et al. (2021), on the other hand, noted that those who do not use iWiW may be socially segregated from those who do, implying that social network fragmentation in poorer towns may be even more pronounced than observed in iWiW, underestimating the correlation between fragmentation and income inequality.

The table below describes all the variables used in this study and sample average values. These variables span from 2011 to 2016 based on data availability. However, given the relatively consistent annual changes in most variables, it is believed that having data spanning five years does not have a significant effect.

Table 2. List of variables and sample average

Name of variables	Notation	Type	Definition	Sample average (n=426)
Dependent variable				
Inequality	G_i	Numeric	Gini index calculated at the town level on equivalized household income 2016	0.485
Independent variables				
Distance to border	$Dist_i$	Numeric	The distance in kilometers from the nearest border	60.673
Town size	$Size_i$	Numeric	The town's total area, in km2	7,896.7
Population density	$Dens_i$	Numeric	The population is divided by the size of the residential area.	2.031
Unemployment ratio	$Unemp_i$	Numeric	Unemployed people as a percentage of the total labor force	0.063
Employment in the manufacturing sector	Emp_i	Numeric	Number of persons employed in the manufacturing sector	1,557.69
Business tax	Tax_i	Numeric	Taxes on corporate income. It is used to measure the level of economic efficiency of a town.	560,469.8
Foreign investment	$FInv_i$	Numeric	Revenue capital owned by foreign firms in 2011, measured in 1,000 Hungarian Forint	9,367,646
Fragmentation	$Frag_i$	Numeric	Social network fragmentation (see Equation 1)	0.297
Age	Age_i	Numeric	The ratio of residents older than 60 years	0.245
High school	$Hschool_i$	Numeric	The ratio of residents with high school degrees or above	0.301
Income per capita	Inc_i		The income per capita in 2016 at the town level, measured in Hungarian Forint	999,005.1
Instruments				
Ethnic fragmentation	$Ethnic_i$	Numeric	*The entropy of ethnic distribution* is estimated using ethnic group size distribution. The indicator is high if the town's ethnic groups are of similar size.	0.126
Religious fragmentation	Rel_i	Numeric	*The entropy of religious distribution* is calculated using population distribution across confession groups. The indicator is high if the town's religious groups are of similar sizes.	0.677
Education inequalities	$Educ_i$	Numeric	*Coefficient of variance* in the 6th-grade math exam.[2] The indicator is high if there are significant discrepancies between primary schools in the town's commuter zone.	0.068
User rate	UR_i	Numeric	The fraction of the population of a town on iWiW	0.344

Source: own construction based on the dataset of Tóth et al. (2021) and TEIR database

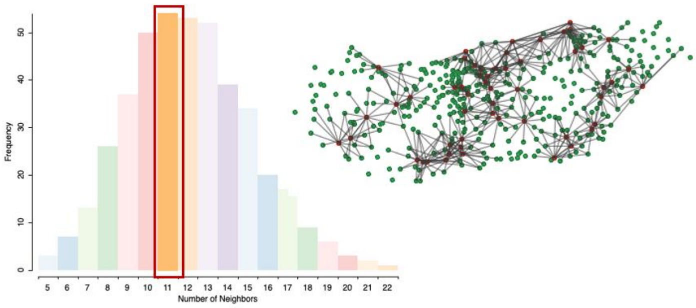
[2] Data were gathered through the national 6th grade mathematical competence test in 2011.

4. Results

The estimation strategy of this paper is twofold. First, I use the non-spatial simple regression models to estimate the relationship between social network fragmentation and income inequality. Second, I apply a spatial two-stage least square (2SLS) regression model to estimate how social segregation measures are associated with income inequality through their relationship to social network fragmentation.

The spatial weights need first be established before analyzing the spatial dependency. Spatial weights are important in constructing spatial autocorrelation statistics because they enable the generation of spatially explicit variables, such as spatially lagged variables and spatially smoothed rates. Technically, the weights are a square matrix that describes the neighbor structure between the observations. This study uses a contiguity weight because the towns with high inequality tend to agglomerate in certain areas (see Figure 1.). Namely, the queen continuity with the second order is applied since it is a little broader, defining neighbors as geographical units with a shared edge or the same vertex.

Figure 6. Queen weights characteristics: Connectivity histogram, towns with 11 neighbors



Source: own elaboration based on the dataset of Tóth et al. (2021)

Non-spatial model:

$$G_i = \beta_0 + \beta_1 Frag_i + \beta_2 \log(Dens_i) + \beta_3 Hschool_i + \beta_4 Age_i + \beta_5 Unemp_i + \beta_6 \log(Emp_i) + \beta_7 \log(FInv_i) + \beta_8 Tax_i + \beta_9 \log(Dist_i) + \beta_{10} \log(Size_i) + \varepsilon \quad (2)$$

The result of OLS regression indicates the ten predictors explained 37.7% of the variance in the town Gini index. Social network fragmentation, population density, foreign investment, business taxes, and town size are positively and statistically significantly related to income inequality in towns. Social network fragmentation, in particular, has the greatest influence on income disparity; one unit increase in fragmentation corresponds to a 0.164 unit increase in income inequality. Employment in the manufacturing sector, on the other hand, has a negative and significant impact on income disparity in towns (Table 4.). There is no multicollinearity (mean VIF=2.09) between independent variables, and heteroskedasticity-robust standard errors are applied to avoid the presence of heteroskedasticity. However, the normality test confirms that the distribution of the residuals is significantly different from normal (see test results from Annex-2).

Furthermore, Ramsey's regression specification error test (RESET) utilizing powers of the fitted values of the Gini index reveals that the model includes omitted variables, assuming an endogeneity problem in equation (2) (Table 4). To reveal the endogenous variable in the model, I manually performed the Durbin-Wu-Hausman (DWH) test for the potential endogenous regressor, social network fragmentation ($Frag_i$). The test result (F (1,388) =4.37, p=0.037) leads to a rejection of the null hypothesis that social network fragmentation ($Frag_i$) is exogenous (see the test performance from Table A 1 in the Annex-2). Also, diagnostics Moran I of the residuals (I= 7.336, p < 0.001) indicates that the OLS residuals are spatially autocorrelated, meaning that the spatial relationship is identified in the error terms (see Diagnostics for spatial dependence from Table A 2 in the Annex-2). Therefore, I use the spatial two-stage least square regression model with spatial lag.

The spatial 2SLS estimation model:

$$G_i = \beta_0 + \rho W G_i + \beta_1 \widehat{Frag}_i + \beta_2 \log(Dens_i) + \beta_3 Hschool_i + \beta_4 Age_i + \beta_5 Unemp_i + \beta_6 \log(Emp_i) + \beta_7 \log(FInv_i) + \beta_8 Tax_i + \beta_9 \log(Dist_i) + \beta_{10} \log(Size_i) + \varepsilon_i \tag{3}$$

IVs for *Frag*:

$$\widehat{Frag}_i = \alpha_0 + \underbrace{\alpha_1 Etnic_i + \alpha_2 Rel_i + \alpha_3 Educ_i}_{Social\ segregation\ measures} + \alpha_4 UR_i + v_i \tag{4}$$

WG_i in equation (3) is the spatial lagged term of the dependent variable and $Etnic_i, Rel_i, and\ Educ_i$ in equation (4) are the social measures of segregation. The spatial model generally expresses how the dependent variable *G* directly affects its immediate neighbors. In other words, widening income inequality in one town will directly impact towns nearby.

Before proceeding to model estimation, let us first determine whether these variables are suitable instruments for social network fragmentation ($Frag_i$). Key assumptions are that the instrumental variables (IVs) are correlated with social network fragmentation ($Frag_i$) but uncorrelated with residuals of equation (3). An F-test of the first stage regression confirms that social segregation indicators are strong instruments of social network fragmentation (Table 3.). Stock and Yogo (2005) proposed the rule concerning the size of F statistics (with more than one instrument) from the first stage regression to detect weak instruments. The IVs are strong if the first-stage F statistic is more than 10. Thus, the relevant F statistic is 18.97, higher than 10, showing that I do not need to be concerned about weak instruments.

Table 3. Estimation of the first stage

	Fragmentation
Ethnic fragmentation	-0.017
	(0.031)
Religious fragmentation	0.05*
	(0.031)
Education	-0.256**
	(0.112)
User rate	0.229***
	(0.03)
Constant	0.204***
	(0.023)
Observations	426
Adj R-squared	0.145
F statistic (4, 421)	18.97***
Standard errors are in parentheses	
**** p<0.01, ** p<0.05, * p<0.1*	

Source: own computation based on the dataset of Tóth et al. (2021)

The estimation results of non-spatial and spatial 2SLS are presented in Table 3. The social network fragmentation, instrumented by social segregation indicators, is positively and significantly associated with income inequality. Its magnitude has increased dramatically in both 2SLS models compared to the OLS result. This finding provides strong evidence for a link between social network fragmentation and inequality. It also implies that social segregation measures significantly predict social network outcomes associated with inequality.

Regarding control variables, only business tax is turned into a non-significant variable compared to the OLS result. Other relevant controls indicate that highly populated towns, larger towns, and towns with significant foreign investment have higher levels of inequality. In contrast, towns with higher employment in the manufacturing sector have lower levels of inequality. As a result, indicators favoring agglomeration and income tend to increase inequality, whereas ones favoring productivity tend to decrease inequality. This result can be explained as an “inverted-U” Kuznets curve indicating the relationship between a country’s income per capita and interpersonal income inequality. It suggested that inequality would be minor within low-income countries, rise as development progressed, and narrow as growth's advantages spread.

Table 4. Estimation results of non-spatial and spatial models

	OLS	**2SLS**	**Spatial 2SLS**
	Gini 2016	Gini 2016	Gini 2016
WGini 2016			0.233**
			(0.116)
Estimated Fragmentation[(a)]	0.164***	0.594**	0.466***
	(0.029)	(0.241)	(0.162)
Log (Pop. density)	0.048***	0.045***	0.044***
	(0.006)	(0.007)	(0.006)
High school	0.010	0.017	0.014
	(0.012)	(0.016)	(0.014)
Age	0.005	-0.025	-0.019
	(0.031)	(0.037)	(0.033)
Unemployment ratio	-0.017	-0.014	-0.026
	(0.047)	(0.061)	(0.056)
Log (Employment in manufacturing)	-0.055***	-0.063***	-0.058***
	(0.006)	(0.008)	(0.007)
Log (Foreign investment)	0.006**	0.011**	0.009**
	(0.001)	(0.005)	(0.004)
Log (Business tax)	0.007***	0.002	0.003
	(0.002)	(0.004)	(0.003)
Log (Distance to the border)	0.001	0.001	0.002
	(0.002)	(0.002)	(0.002)
Log (Town size)	0.041***	0.048***	0.045***
	0.005	(0.007)	(0.006)
Constant	0.360***	0.299***	0.199***
	(0.024)	(0.045)	(0.077)
Observations	426	426	426
Adj R-squared/Pseudo R-squared/Spatial Pseudo R-squared	0.377	0.274	0.316
F stat/ First stage F-test	26.73***	18.97***	-
RESET	12.31***	-	-
DWH test	-	0.014 (p=0.906)	-
Anselin-Kelejian Test	-	12.288***	2.136 (p=0.144)
*Robust standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1*			

Source: own construction based on the dataset of Tóth et al. (2021)

Note: [(a)]-The level of fragmentation is used in the OLS regression

5. Conclusion

Using Hungarian town-level data from the paper by Tóth et al. (2021), this study examines how social networks and social segregation measures interact and their relationship with income disparity. Hungary's income inequality has approached the

EU average nationally, but the disparity in cities and towns has widened. At the town level, income disparities range from 0.418 to 0.645. The towns surrounding Budapest, those along Lake Balaton, and certain towns along the country's western and southern borders have higher income inequality. In contrast, places in the country's east and north have lower inequality. In this paper, I have found that social network fragmentation, as instrumented by social segregation indicators, significantly impacts income inequality at the town level. When there is ethnic fragmentation, religious fragmentation, and educational inequality in a town, social networks tend to be more fragmented. Therefore, like the urban indicators used by Tóth et al. (2021), these social segregation measures can also be significant predictors of social network outcomes associated with inequality. This study supports the findings of previous studies on the role of ethnicity and identity (e.g. Fundamental Rights Agency, 2014; Omoeva et al., 2018). Furthermore, a positive significant spatial lagged term suggests that inequality in neighboring towns has a strong spillover effect.

Acknowledgments

First, I am grateful to Tóth et al. (2021) for making their aggregated city-level data available. I want to acknowledge Dr. Mónika Galambosné Tiszberger, assistant professor, and Dr. Richard Farkas, assistant professor from the Faculty of Business and Economics at the University of Pécs. Their guidance and advice are valuable for writing this paper. I also express my gratitude to an anonymous reviewer for their insightful feedback. Finally, I am grateful to the Stipendium Hungaricum Scholarship for giving me all these opportunities.

Annex-1: Morgan scatter plots

Figure A-7. High-High Morgan scatter plot locations (81 observations selected)

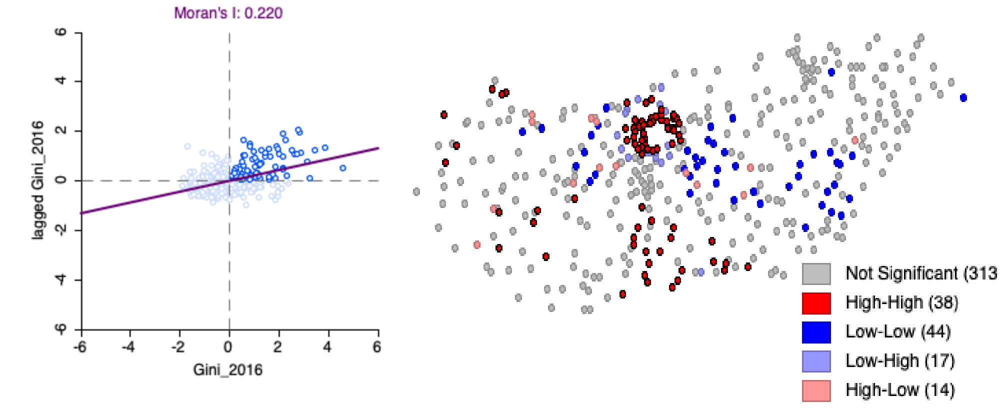


Figure A-8. High-Low Morgan scatter plot locations (99 observations selected)

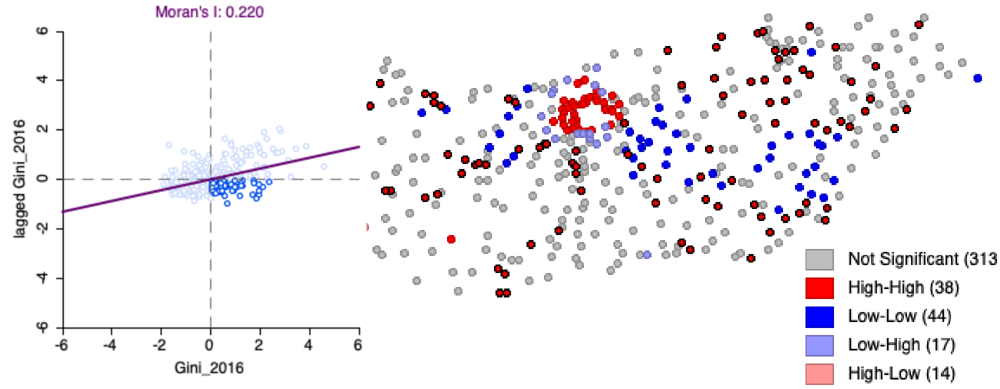


Figure A-9. Low-Low Morgan scatter plot locations (178 observations selected)

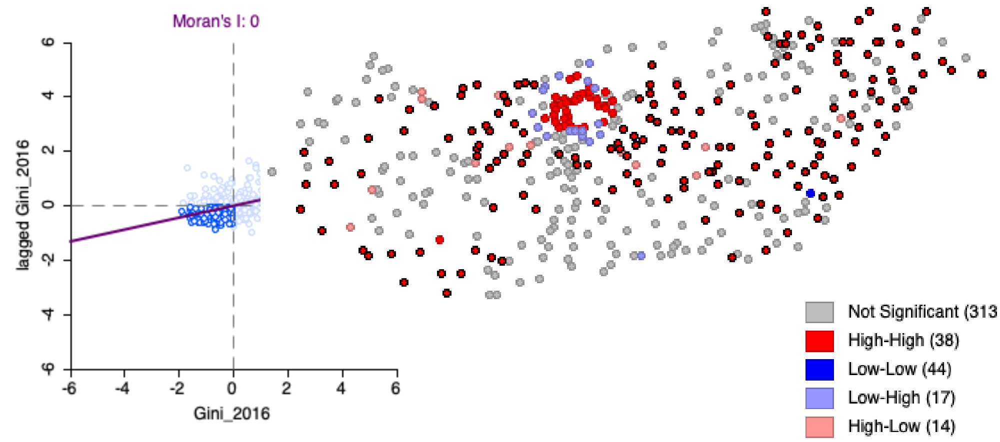
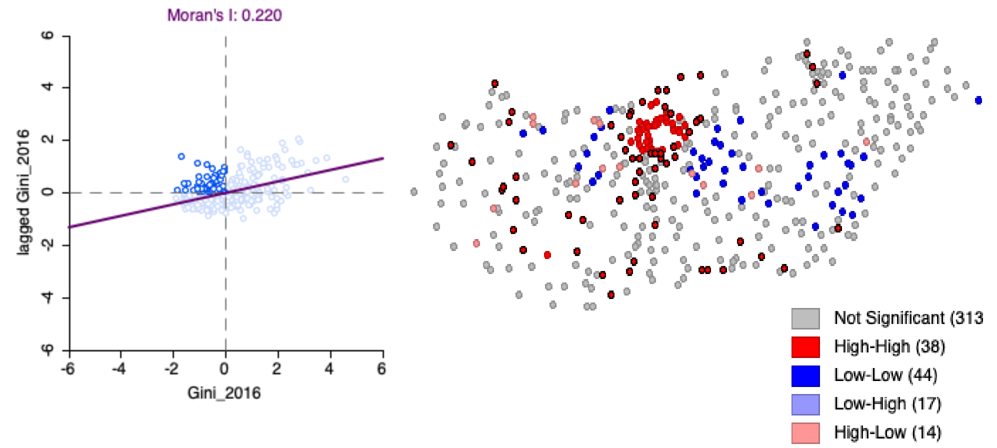


Figure A-10. Low-High Morgan scatter plot locations (68 observations selected)



Source: own elaboration based on the dataset of Tóth et al. (2021)

Annex-2: Test results

Non-spatial model (OLS regression): Test results

1. Heteroskedasticity tests

H0: Constant variance

Test	df	Value	Prob
Breusch-Pagan test	10	97.770	<0.001
Koenker-Bassett test	10	53.699	<0.001

The Breusch-Pegan test result rejects the null hypothesis of homoskedasticity in the OLS regression. Thus, I used a robust heteroskedasticity standard error because significant heteroskedasticity generates biased standard errors and invalidates the resulting hypothesis tests.

2. Test on the normality of errors

H0: Residuals are normally distributed

Test	df	Value	Prob
Jarque-Bera	2	66.159	<0.001

The Jarque-Bera test results show that the distribution of the residuals is significantly different from normal.

3. Durbin–Wu–Hausman (DWH) test of endogeneity (implemented manually):

In the 1st stage, the potential endogenous variable, *Fragmentation*, was estimated on other explanatory variables in equation (2) and instrumental variables. Equation (2) with an additional variable, *Res* (the error from the first stage equation for *Fragmentation*), is estimated in the second stage.

$$G_i = \beta_1 Frag_i + \beta_2 X_i + \gamma Res_i + \varepsilon_i$$

Under the null hypothesis, *Fragmentation* is exogenous. If *Res* could be observed, the exogeneity test would be the test of $H_0: \gamma = 0$. The estimation results in Table A 1 show that the coefficient of fragmentation's residual is different from zero, implying *Fragmentation* is endogenous.

Table A-1. Estimation results of the DWH test

Variables	Stage 1	Stage 2
	Fragmentation	Gini 2016
Fragmentation		.491***
		(.165)
Log (Pop.density)	.016*	.053***
	(.009)	(.005)
High school	-.016	.01
	(.022)	(.012)
Age	.055	-.007
	(.046)	(.03)
Unemployment ratio	-.035	-.036
	(.09)	(.042)
Log (Employment in manufacturing)	.011	-.068***
	(.009)	(.006)
Log (Foreign investment)	-.013**	.01***
	(.005)	(.003)
Log (Business tax)	.011***	.004
	(.004)	(.003)
Log (Distance to the border)	-.002	.004*
	(.004)	(.002)
Log (Town size)	-.008	.051***
	(.009)	(.006)
Ethnic fragmentation	-.043	
	(.032)	
Religious fragmentation	-.031	
	(.033)	
Education	-.238**	
	(.108)	
User rate	-.002	
	(.038)	
Fragmentation residual		-.345**
		(.165)
Constant	.178***	.304***
	(.053)	(.031)
Observations	400	400
R-squared	.331	.478
Robust standard errors are in parentheses		
**** p<0.01, ** p<0.05, * p<0.1*		

H0: Fragmentation Residual =0
F(1,388)=4.37
Prob =0.0371

Source: own elaboration based on the dataset of Tóth et al. (2021)

4. Diagnostics for spatial dependence

Table A-2. Diagnostics for spatial dependence test for the OLS regression model

Test	MI/DF	Value	Prob
Moran's I (error)	0.137	7.336	<0.001
Lagrange Multiplier (lag)	1	35.947	<0.001
Robust LM (lag)	1	1.394	0.2378
Lagrange Multiplier (error)	1	46.49	<0.001
Robust LM (error)	1	11.936	0.0006
Lagrange Multiplier (SARMA)	2	47.883	<0.001

Source: own elaboration based on the dataset of Tóth et al. (2021)

References

Aristei, D. – Perugini, C. (2015): The drivers of personal income inequality in transition, and the role of reform approaches. In: Perugini, C. – Pompei, F. (eds.): *Inequalities during and after Transition in Central and Eastern Europe*. London: Springer, 41-58.

Bailey, M. – Cao, R. – Kuchler, T. – Stroebel, J. – Wong, A. (2018): Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32, 259-280. https://doi.org/10.1257/jep.32.3.259

Bianconi, G. – Darst, R. K. – Iacovacci, J. – Fortunato, S. (2014): Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E*, 90, 042806. https://doi.org/10.1103/PhysRevE.90.042806

Brzezinski, M. (2018): Income inequality and the Great Recession in Central and Eastern Europe. *Economic Systems*, 42, 219-247. https://doi.org/10.1016/j.ecosys.2017.07.003

Cameron, A. C. – Trivedi, P. K. (2010): *Microeconometrics using Stata*. Texas, Stata Press.

Cavanaugh, A. – Breau, S. (2018): Locating geographies of inequality: publication trends across OECD countries. *Regional Studies*, 52, 1225-1236. https://doi.org/10.1080/00343404.2017.1371292

Chetty, R. – Hendren, N. – Kline, P. – Saez, E. (2014): Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129, 1553-1623. https://doi.org/10.1093/qje/qju022

DiMaggio, P. – Garip, F. (2012): Network effects and social inequality. *Annual Review of Sociology*, 38, 93-118. https://doi.org/10.1146/annurev.soc.012809.102545

Docquier, F. – Kone, Z. L. – Mattoo, A. – Ozden, C. (2019): Labor market effects of demographic shifts and migration in OECD countries. *European Economic Review*, 113, 297-324. https://doi.org/10.1016/j.euroecorev.2018.11.007

Dolls, M. – Doorley, K. – Paulus, A. – Schneider, H. – Sommer, E. (2019): Demographic change and the European income distribution. *Journal of Economic Inequality*, 17, 337-357. https://doi.org/10.1007/s10888-019-09411-z

Eriksson, R. H. – Lengyel, B. (2019): Coworker networks and agglomeration externalities. *Economic Geography*, 95, 65-89. https://doi.org/10.1080/00130095.2018.1498741

Esposito, P. – Stehrer, R. (2009): The sector bias of skill-biased technical change and the rising skill premium in transition economies. *Empirica*, 36, 351-364. https://doi.org/10.1007/s10663-008-9097-9

Eurostat (2022a): Distribution of income by quantiles - EU-SILC and ECHP surveys. https://data.europa.eu/data/datasets/v6gke8fztkdfoqkblozwa?locale=en

Eurostat (2022b): Gini coefficient of equivalised disposable income - EU-SILC survey. https://data.europa.eu/data/datasets/dvrrgg5nu7galdtl3xsyq?locale=en

Eurostat (2022c): Income quintile share ratio S80/S20 for disposable income by sex and age group - EU-SILC survey. https://eige.europa.eu/gender-statistics/dgs/indicator/eustrat_epsr_eoalm_ium__ilc_di11

Flemming, J. S. – Micklewright, J. (2000): Chapter 14 Income distribution, economic systems and transition. *Handbook of Income Distribution*, 1, 843-918. https://doi.org/10.1016/S1574-0056(00)80017-2

Förster, M. F. – Tóth, I. G. (2015): Chapter 19, Cross-country evidence of the multiple causes of inequality: Changes in the OECD Area. In: Atkinson, A. B. – Bourguignon, F. (eds.): *Handbook of income distribution*. North-Holland: Elsevier, 1729-1843. https://doi.org/10.1016/B978-0-444-59429-7.00020-0

Fundamental Rights Agency (2014): *Poverty and employment: The situation of Roma in 11 EU Member States*. https://fra.europa.eu/sites/default/files/fra_uploads/fra-2014-roma-survey-dif-employment-1_en.pdf

Gamoran, A. (2011): Designing instruction and grouping students to enhance the learning of all: New Hope or false promise? In: Hallinan, M. T. (ed.): *Frontiers in sociology of education*. Dordrecht: Springer, 111-126. https://doi.org/10.1007/978-94-007-1576-9_6

Glaeser, E. L. – Resseger, M. – Tobio, K. (2009): Inequality in cities. *Journal of Regional Science*, 49, 617-646. https://doi.org/10.1111/j.1467-9787.2009.00627.x

Granovetter, M. (1985): Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481-510.

Josifidis, K. – Supic, N. – Bodor, S. (2021): Distributional effects of foreign versus domestic investment: Evidence from post-communist EU member states. *Panoeconomicus*. https://doi.org/10.2298/PAN2102187J

Kattuman, P. – Redmond, G. (2001): Income inequality in early transition: The case of Hungary 1987-1996. *Journal of Comparative Economy*, 29, 40-65. https://doi.org/10.1006/jcec.2000.1701

Lengyel, B. – Jakobi, Á. (2016): Online social networks, location, and the dual effect of distance from the centre. *Tijdschriftvoor Economische en Sociale Geografie*, 107, 298–315. https://doi.org/10.1111/tesg.12150

Magda, I. – Gromadzki, J. – Moriconi, S. (2021): Firms and wage inequality in Central and Eastern Europe. *Journal of Comparative Economy*, 49, 499-552. https://doi.org/10.1016/j.jce.2020.08.002

McPherson, M. – Smith-Lovin, L. – Cook, J. M. (2001): Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415-444. https://doi.org/10.1146/annurev.soc.27.1.415

Medgyesi, M. – Tóth, I. G. (2021): Income, wealth, employment, and beyond: Central and eastern Europe. In: Fischer, G. – Strauss, R. (eds.): *Europe's income, wealth, consumption, and inequality*. New York: Oxford University Press, 177-217.

Medgyesi, M. – Tóth, I. G. (2022): Inequality and welfare. In: Mátyás, L. (ed.): *Emerging European economies after the pandemic: Stuck in the middle income trap?* Cham: Springer International, 575-610. https://doi.org/10.1007/978-3-030-93963-2_12

Milanovic, B. (1999): Explaining the increase in inequality during transition. *Economics of Transition*, 7, 299-341. https://doi.org/10.1111/1468-0351.00016

OECD (2011): Divided We Stand: Why Inequality Keeps Rising. Paris:OECD Publishing. https://doi.org/10.1787/9789264119536-en

Omoeva, C. – Moussa, W. – Gale, C. (2018): The economic costs of educational inequality in developing countries. In: Ben David-Hadar, I. (ed.): *Education finance, equality, and education equity economy*. Cham: Springer, 187-217. https://doi.org/10.1007/978-3-319-90388-0_10

Pampel, F. C. – Krueger, P. M. – Denney, J. T. (2010): Socioeconomic disparities in health behaviors. *Annual Review of Sociology*, 36, 349-370. https://doi.org/10.1146/annurev.soc.012809.102529

Pénzes, J. – Bujdosó, Z. – Dávid, L. – Radics, Z. – Kozma, G. (2014): Differing development paths of spatial income inequalities after the political transition-by the example of Hungary and its regions. *Экономика региона*, (1), 73-84. https://doi.org/10.17059/2014-1-6

Rivera, M. T. – Soderstrom, S. B. – Uzzi, B. (2010): Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. *Annual Review of Sociology*, 36, 91-115. https://doi.org/10.1146/annurev.soc.34.040507.134743

Sampson, R. J. (2008): Moving to inequality: Neighborhood effects and experiments meet social structure. American *Journal of Sociology*, 114, 189-231. https://doi.org/10.1086/589843

Stadtfeld, C. (2018): The micro-macro link in social networks. In: Scott R. A. – Buchmann M. – Kosslyn S. (eds): *Emerging Trends in the Social and Behavioral Sciences*, Hoboken: John Wiley & Sons. 1-15.

Stiglitz, J. E. (2012): *The price of inequality: How today's divided society endangers our future*. New York: W.W. Norton Co.

Stock, J. – Yogo, M. (2005): Asymptotic distributions of instrumental variables statistics with many instruments. In: Andrews, D. W. K. (ed.): *Identification and inference for econometric models*. New York: Cambridge University Press, 109-120.

Szczepaniak, M. – Szulc-Obłoza, A. (2020): Labour market institutions and income inequalities in the Visegrad group countries. Comparative Economic Research.

Central and Eastern Europe, 23, 75-90. https://doi.org/10.18778/1508-2008.23.21

Tóth, G. – Wachs, J. – Di Clemente, R. – Jakobi, Á. – Ságvári, B. – Kertész, J. – Lengyel, B. (2021): Inequality is rising where social network segregation interacts with urban topology. *Nature Communications*, 12. https://doi.org/10.1038/s41467-021-21465-0

Tóth, I. G. (2008): Economic transition and income distribution in Hungary. 1987-2001, In: Fanelli, J. M. – Squire, L. (eds.): *Economic reform in developing countries: Reach, range, reason*. Cheltenham, Edward Elgar Publishing, 3-34. https://doi.org/10.4337/9781781007655.00009

Tsaurai, K. (2020): An empirical study of the determinants of income inequality in transitional economies using panel data analysis methods. *Academy of Accounting and Financial Studies Journal*, 24, 1-15.

Vida, Gy. (2022): Spatial characteristics of some aspects of regional 'realised competitiveness' in Hungary between 2010 and 2019. *Területi Statisztika*, 62, 538-569. https://doi.org/10.15196/TS620503

Wooldridge, J. M. (2019): Introductory econometrics: A modern approach. South-Western: Cengage Learning.

Zoltán, E. (2022): Characteristics of settlement income inequality in Hungary by geographical proximity, 2012-2019. *Területi Statisztika*, 62, 625-662. https://doi.org/10.15196/TS620601

Internet sources

GeoDa workbook (https://geodacenter.github.io/documentation.html)

GeoSpace (https://geodacenter.github.io/GeoDaSpace/)

National Regional Development and Spatial Planning Information System (TEIR). (https://www.teir.hu)

Open Street Map data (https://download.geofabrik.de)

Town-level data by Tóth et al. (2021): (https://zenodo.org/record/4448183#.YAXjOOhKg2w)