TRAGEDIES of Regional Conflicts & PROMISES of Peacebuilding
Responding to Disruptors and Enablers of MENA Development Pathway
ERF 30th Annual Conference

www.erf.org.eg

Agglomeration Economies and Wage Differences:

Evidence from Turkey

Serife Genc Ileri
and Muhammad Moiz

ECONOMIC RESEARCH FORUM 30 years
منتدى البحوث الاقتصادية

2024

Agglomeration Economies and Wage Differences: Evidence from Turkey

Cities are growing across the globe, increasing the regional concentration of people and firms. This proximity between workers and firms can have many benefits such as labor pooling, knowledge exchange, availability of resources, and development of infrastructure. The benefits arising from this concentration are known as agglomeration economies. In this study, we investigate agglomeration economies and their effects on wages using a micro dataset at the city and district levels within Turkey. The purpose of this research is to add to the literature of agglomeration economies in the Turkish context, to compare those results with other developing countries, and to provide policymakers with empirical evidence to base their policies regarding industry and labor markets.

Keywords: Agglomeration economies, Wage differences, Turkey, agglomeration, panel data.

Agglomeration Economies and Wage Differences: Evidence from Turkey

The concentration of humans has increased greatly over the past few decades, mass migration from rural to urban areas continues to rise. Recent data shows that 55% of the world's population lived in urban areas in 2018 and was projected to increase to 68% by 2050 (United Nations, 2018). Urbanization has also given rise to mega cities.[1]. There are more than 30 mega cities in the world today and they continue to expand. People generally migrate to cities for better opportunities which can range from education, work, to improved facilities and resources. As a result, the concentration of people and economic activity goes hand in hand. This concentration can provide great benefits to firms in terms of availability of skilled workers, proximity to suppliers, proximity to customers, exchange of information and ideas, and others. Employees in the region also benefit through greater opportunities, bargaining power, lower switching costs, and learning from others.

The concentration of firms in cities, whether large or small, influences productivity and wages within the region. However, the concentration of the type of industries, or the skill-level of the workers can affect the benefits of agglomeration. Agglomeration economies can differ with respect to region and time, and may show more distinct effects in one region but not in another. While the worker and firm characteristics are an important factor in the determination of total factor productivity and wages, agglomeration externalities also play a significant role. Furthermore, the goals of the firms and workers are different, while a firm aims to maximize profits, a worker maximizes its utility. Workers may also be quick to relocate, reacting to market conditions. Therefore, it becomes essential that agglomeration effects be studied and empirically sought out.

For policy makers, it is important to realize the impact of agglomeration economies as they are responsible for deciding whether to focus on larger or smaller cities. It would be equally important to understand the productivity or wage differences arising from an increase in density, specialization, and diversity among other factors. How does the number of workers per square kilometer affect the wage premiums in the region? Does a higher level of specialization (concentration of employment in a few industries) or diversity (employment spread over a number of industries) have greater impact on productivity? Moreover, is competition good or bad for productivity or wage growth? These are all important questions that can be posed by policy makers. A better understanding of agglomeration economies can help form economic policies that provide answers to these questions and more. From a policy perspective, it becomes essential to know which factors besides the firm and worker characteristics account for the change in productivity and wages. Once learned, economic policies can be formed that take into account different agglomeration externalities and how they apply to firms and workers within a specific region. The importance of agglomeration can only be understood after having empirical evidence of different factors that determine productivity at a regional level. It may also be more beneficial to study at micro-level as differences may arise among micro regions whether these be districts or metropolitan areas.

Literature Review

[1] The UN defines megacity as a city with more than 10 million inhabitants

Agglomeration economies have been a topic of study for more than a century. Marshall (1890) is considered as one of the earliest works in the field with three main categories of agglomeration including labor market interactions, firm linkages, and knowledge spillovers. The micro-foundations of agglomeration economies can be better described by three sources of agglomeration as stated by Duranton and Puga (2004), these include sharing, matching, and learning.[2]. Sharing is based on the concept of indivisible goods, where everyone would benefit from sharing a good that can be resourced jointly Sharing can be achieved by producing non-tradeable goods in a single city (Abdel-Rahman, 1990) or by several sectors operating within a city leading to economies of scale (Abdel-Rahman, 1994). Diversity can play an important role in agglomeration, creating a positive effect on productivity and wages. On the other hand, specialization can also have its own benefits, when studied at a micro level, specialized workers can provide higher productivity as they gain more experience in their work and become more efficient. Similarly, companies can also be specialized in their production processes. Labor pooling is another factor that can be considered under sharing, that is the sharing of the labor force within the region. The firms can benefit from a larger labor pool by having the ability to lay off workers at a time of a negative market shock, whereas employees prefer denser markets to minimize their risks of possible unemployment that may result from shocks. In this way, both the firms and employees prefer to agglomerate in certain regions.

A related concept is urbanization economies which refer to externalities that arise due to the characteristics of the location such as density (Cohen et al., 2008). These economies are more pronounced in regions that are closest to economic activity. The proximity of firms plays an important role in determining the effects of agglomeration in urbanization, where closeness can lead to cost benefits that arise from various industries established in the same region. Therefore, distance between firms may lead to a decline in agglomeration effects (Audretsch & Feldman, 2004; van Soest et al., 2006). These cost benefits may be a result of firms using various services that are common among them such as legal services, logistics, marketing, etc. Hence, the concentration of the city (region) in terms of employment is an important factor when discussing urbanization economies. In other words, one would observe the concentration of different industries (employment) within a region rather than being concerned with the specialization of the region.

There are mixed findings when the role of technology is assessed in terms of geographical concentration of industries. In the case of UK, Devereux et al. (2004) found geographical firm density in low-tech firms, that is low-tech firms tend to be in geographically dense regions. In addition, they also find higher survival rates in agglomerated industries, giving another perspective . Whereas Sedgley and Elmslie (2004) find high tech firms to be located in economically dense regions where they can benefit from the concentration of industries and labor pooling . Strange et al. (2006) find that innovation within firms is a determinant of agglomeration, that is innovative firms will tend to agglomerate more. This gives evidence for the argument of knowledge spillovers that are a result of agglomeration. Innovative firms may prefer to be in proximity of other innovative firms to benefit from labor pooling, sharing of ideas, and learning. Some studies find localization effects to be more prevalent while other studies highlight the benefits of urbanization.

Localization Economies, on the other hand, refer to externalities arising from the characteristics of the industry in a specific location such as specialization. (Combes & Gobillon, 2015) Desmet and Fafchamps

[2] (Duranton and Puga, 2004 provide an overview of the micro foundations of agglomeration economies)

(2005) find that localization economies are of more importance than urbanization economies when assessing the employment growth in non-service sectors.

), these are

The geographical location of firms depends on these factors among others when making decisions on where they will locate whether if they are in an industrially specialized region where they seek benefits of localization economies or a more diverse region where they can take advantage of urbanization economies.

Despite the abundance of research agglomeration economics being vastly in developed countries, there are also studies analyzing the same question in the context of developing countries. Pan et al. (2016) uses Chinese data and finds that city size plays an important role in the wage premium of workers in China

Wage premium also differs with the set of skills where high-skilled workers maintain a higher premium than their low-skilled counterparts. Differences in wage premiums may also be heterogenous across regions, where inland cities show negative effects of city size on wages in the long-run as compared to coastal cities having a positive long-run effect (Chen at al., 2021). These differences may be attributed to the development of firms in the different regions, where firms in the later stages of the development will experience a positive effect of city size on wages in the long run. Development of the city can also be a factor in these heterogenous wage premiums across regions, improved infrastructure should provide more positive outcomes. The results may also differ with respect to industries, Barufi (2015) finds that high and low-tech firms gain more advantages from urban scale along with the service industry in Brazil. In the case of Egypt, service firms are found to benefit more from diversification while manufacturing firms benefit more from specialization (Badr et al., 2019). Evidence of agglomeration economies is also found in Colombia where Garcia (2019) finds a positive effect of employment density on the wage premium. The differences are more profound in the informal sector than the formal sector where the informal sector has a higher effect on productivity than within the formal sector. In Indonesia, urban density is found to have a positive agglomeration effect on individual wages in addition to the significant positive effects of specialization (Ridhwan, 2021).

The literature provides mixed evidence on the impacts of diversity on wages in in developing countries. While some lend support to a positive effect of diversity effect(Barufi, 2015; Torres-Gutiérrez et al., 2019) others display a negative diversity effect (Ridhwan, 2021; Prasertsoong & Puttanapong, 2022). Different methodologies utilized to measure diversity or other agglomeration variables can be one reason for the difference in the results. This is evident in the study by Barufi (2015) where a number of different methods are used to measure agglomeration economies to observe how the results may vary. In the case of diversity measured by either the Herfindhal Index or Shannon’s Entropy, Barufi (2015) finds a positive diversity effect in Brazil however, when using the share of 5 biggest sectors as an indicator of diversity, a negative effect is of diversity is observed.

In terms of Turkey, there are a limited number of studies and most of them focus only on the manufacturing sector. Alkay and Hewings (2012) studied the manufacturing sector in Istanbul using three different geographical areas namely the European side, the Asian side, and Istanbul Metropolitan Area level. They find that urbanization economies can be explained through density, market area potential, and labor market potential. Their results also show that labor market pooling is in effect within their study area and raw materials play an important role in agglomeration at the geographical level, however, they do not find any evidence of knowledge spillover. Since they studied the manufacturing sector in Istanbul, the results cannot to generalized to a wider geographical area of Turkey due to the status of Istanbul as an economic hub. Istanbul is the largest city and has the most economic activity of any city in Turkey by a large margin. Though, these results may explain how agglomeration operates within large cities where there is high population and high firm density within proximity.

Kent (2018) uses a similar strategy to Alkay and Hewings (2010), utilizing the Ellison and Glaeser (1997) index to study the concentration of manufacturing sector in Turkey. While this study captures a wider region (whole of Turkey), it is limited to the NUTS-2 level creating a geographical restriction. The study finds that there is higher concentration in low-tech and medium-tech firms. Their findings concur with literature on demonstrating that low-tech firms tend to agglomerate more, similar has been found by Devereux et al. (2004) for the case of the UK.

Özgüzel (2023) conducts a more comprehensive and recent study which analyzes the effects of agglomeration in Turkey at the NUTS-III level using an administrative dataset from 2008 to 2013 The study has utilized a dataset provided by the Social Security Administration of Turkey where the data is aggregated at NUTS-III level and respective sector. Like other studies, Özgüzel (2023) finds a positive effect of density on labor productivity whereas no effects of sorting are found.

Different from the existing studies using Turkish data, the current study utilizes a rich employer-employee matched data setto investigate the agglomeration determinants in Türkiye. We also contribute to the existing literature through the exploration of with the use of firm-level data. and also by furthering literature on developing or emerging economies in the case of agglomeration and wage premiums.

Data

Administrative data on employees and firms was used in this study through the Enterprise Information System (EIS) provided by the Ministry of Industry and Technology (STB). EIS is a comprehensive database of enterprise, plant, and employee data which covers all registered firms and employees that are registered with the Social Security Institution (SGK). Furthermore, there is detailed financial data (balance sheet) for firms across Turkey. It also has trade (import and export) data for all countries and Turkey, as well as data on firm support programs initiated by different institutions such as Science Ministry (TUBITAK), Small and Medium Enterprises Development Organization (KOSGEB), and others. EIS provides data from 2006 onwards and has been available to researchers and other institutions since 2016.

EIS providesWe match the individual employee data with the firm-level administrative data to create the dataset for our analysis. Matching the firm data with the employee data, we are able to incorporate employee locational data on the city and district levels. Data available over the time frame since 2006 provides us with all registered employees with SGK, making the dataset high in observations. In our dataset which has a cross-section of 2009 and panels from 2014-2019, there are more than 7 million observations in 2009 and more than 11 million annual observations each year from 2014 onwards. Firm data is also vast where more than 10 million observations are recorded over the study period.

Both district-level and city-level datasets are used in our analysis. time and resource constraints. City-level dataset consists of 81 Turkish cities over 6-year period and the district-level dataset consists of 957 districts across Turkey over the same period. The employee data provided includes the employee's age, gender, daily wages, and occupation. The firm data available includes firm size, number of employees, location (city and district), wages paid to employees, 4-digit NACE classification, establishment year, whether if the firm is operational, the technology level (only among manufacturing firms), and detailed balance sheet data. We do not use any of the balance sheet data for this study.

Variables

Wage – The monthly and daily wage data is available for all employees by quarter. We use the daily wages because monthly wages can vary depending on the number of working days. Using the daily wage level would provide more accurate results and has been used in literature. The year end daily wage data is used therefore only the individuals employed in the last quarter of each year are a part of the dataset.

Age – The age of employees is an indicator of their work experience and therefore their wages. Age and age squared are widely used in labor economics literature.

Gender – This captures the wage differences among males and females.

Occupation – Education is used in much of the literature to measure the wage differences, however, education data is not available through the EIS database. We use data on occupations as a proxy for education which is available after 2013. Besides serving as a proxy for education and this variable also gives us an idea of how wages differ with respect to the occupation of an individual. There are 9 occupational categories used in this dataset.

Table 1: ISCO Occupations matched with ISCED-97 Levels of Education

Occupation	Skill Level	ISCED-97 Levels of Education
Managers	4	Second stage of tertiary education (leading to an advanced research qualification)
		First stage of tertiary education, 1st degree (medium duration)
	3	First stage of tertiary education (short or medium duration)
Professionals	4	Second stage of tertiary education (leading to an advanced research qualification)
		First stage of tertiary education, 1st degree (medium duration)
Technicians and Associate Professionals	3	First stage of tertiary education (short or medium duration)
Clerical Support Workers	2	Post-secondary, non-tertiary education
		Upper secondary level of education
		Lower secondary level of education
Services and Sales Workers		
Skilled Agricultural, Forestry and Fishery Workers		
Craft and Related Trades Workers		
Plant and Machines Operators, and Assemblers		
Elementary Occupations	1	Primary level of education

Source: International Standard Classification of Occupations (ISCO-08 Vol. 1, ILO, 2012)

ILO has also matched some of these occupations with their education levels. Although, it does not mean that those education levels are the only consideration that may go into these occupations and the skills required for each individual position may differ according to the responsibilities. For example, a manager in the manufacturing sector may require a different skillset than a manager in the service industry, however some level of education will be attained by both of them. The ISCED-97 levels of education are defined for occupations where it is used for measurement of the skill within these occupations. This identification of education levels with their respective occupations provides us with an idea of how we can treat them in our assessment in the absence of an employee's education.

Firm Size – The firm size is defined by the number of employees and the amount of annual revenue generated by the firm. The size of the firm can have an effect on the wages where it is expected that bigger firms will provide higher wages. This can be a result of higher profits and productivity in larger firms leading to better incentives for employees.

Table 2: Firm size by number of employees and Net Sales

Firm Size	# of Employees	Net Sales or Balance Statement (TL)
Micro	0-9	≤3,000,000
Small	10-49	≤25,000,000
Medium	50-249	≤125,000,000
>SME	250 or more	>125,000,000

Fixed effects are applied through the use of year, region, firm size, and sector dummies. The sector dummy is based on NACE Rev. 2 (Eurostat, 2008) categorization of industries, where 19 sectors are identified and used as a part of this study.

The current research uses a 6-year dataset covering the 2014-2019 period becausethe employee occupational data is not available before 2014We also present results for a cross-section from 2009 and 2021 to analyze and compare the pre and post pandemic periods.

We follow an empirical strategy proposed by Combes et al. (2008) and later used by Groot et al. (2014) in their study of micro-data from Netherlands. We conduct a two-stage analysis to check for the effects of various agglomeration factors on wage premium. The first step uses a Mincerian wage regression to analyze the wage differences. These wage differences can be a result of various employee characteristics such as age, education level, gender, and occupation. In addition to the employee characteristics, wage differences may also occur due to regional characteristics such as infrastructure, resources, and agglomeration effects.

$$\log w_{i,t} = \tau_{i,t} + \delta_o + \rho_r + \beta_j + \mu_t + \varepsilon_{i,t}$$

$\tau_{i,t}$ is a vector of employee characteristics, these include age, age squared, and gender. Various dummy variables are used, δ_o represents the occupation dummy, ρ_r represents the regional dummy, β_j represents the sector dummy, and μ_t represents the time dummy.

Specialization

$$Spec_{ir} = \frac{E_{jr}}{E_r}$$

Share of employment in a specific industry compared to the total employment in the region. The higher the specialization index, the more concentrated an area is in terms of industry. This indicator shows the concentration of an industry within a region. A higher specialization could mean benefits for the industry because a concentration of companies can result in placement of suppliers in the same region. In addition, more potential employees may be attracted to the region.

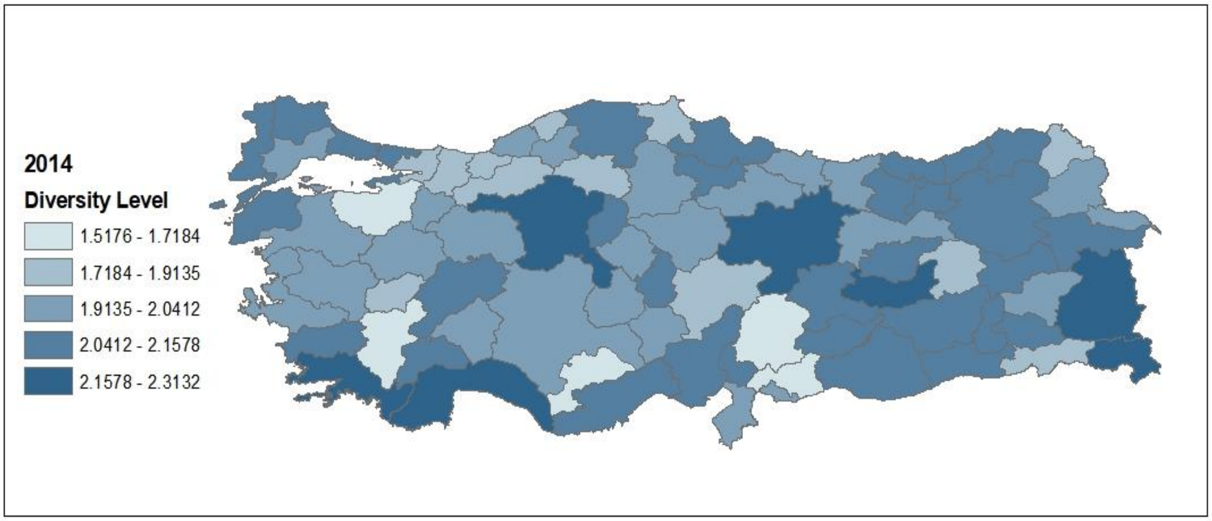
Diversity

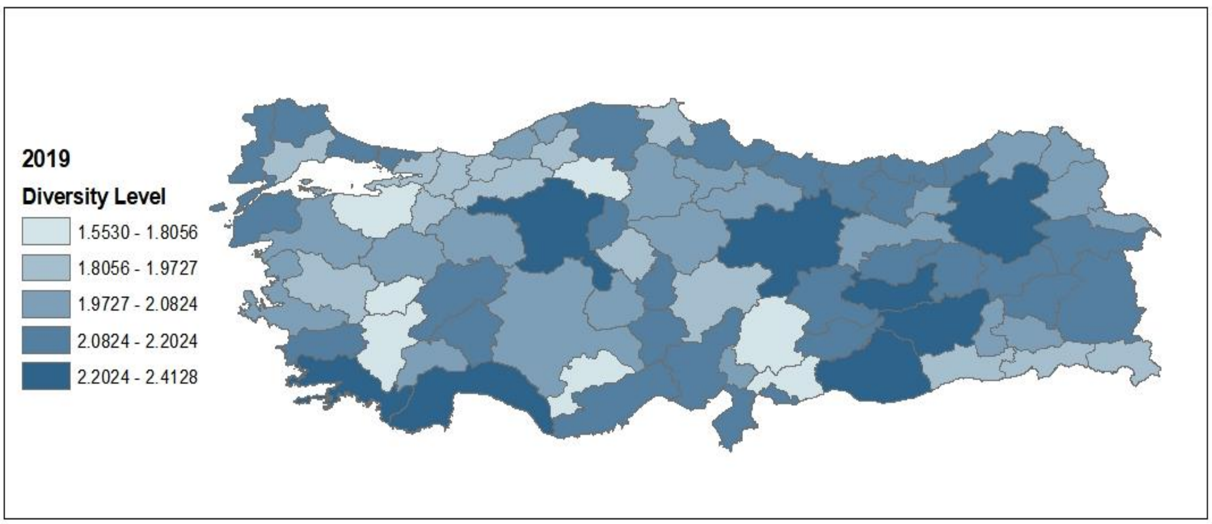
$$Div_r = -\sum\left(\frac{E_{jr}}{E_r}\ x\ ln\frac{E_{jr}}{E_r}\right)$$

Shannon's entropy is used to capture the diversity of a region. High value means the region is diversified whereas a low value indicates that the region has a high concentration of a few industries. Diversity indicates the number of different industries that are operating within a region. This indicator can help us understand whether the region is more diversified or more specialized or neither. Diversity could benefit the region through greater spillover results and the availability of a diverse workforce to cater to the needs of the different industries.

The highest levels of diversity are observed in southern two cities, in the far east, and central Turkey. Cities close to Istanbul are more specialized and host a smaller number of industries. In Central Turkey, there is also lower diversity, where this trend continues as we move further east with the exceptions of a few cities such as Ankara, Sivas, Elazig, Hakkari, and Van. In 2019, more diversification of industries is observed in the Southeastern cities of Turkey, namely Şanlıurfa, Diyarbakir, and Elazig. The overall diversity across Turkey has also increased over this period as seen in the increase of diversity levels.

Graphic 1: Comparison of Diversity from 2014 and 2019





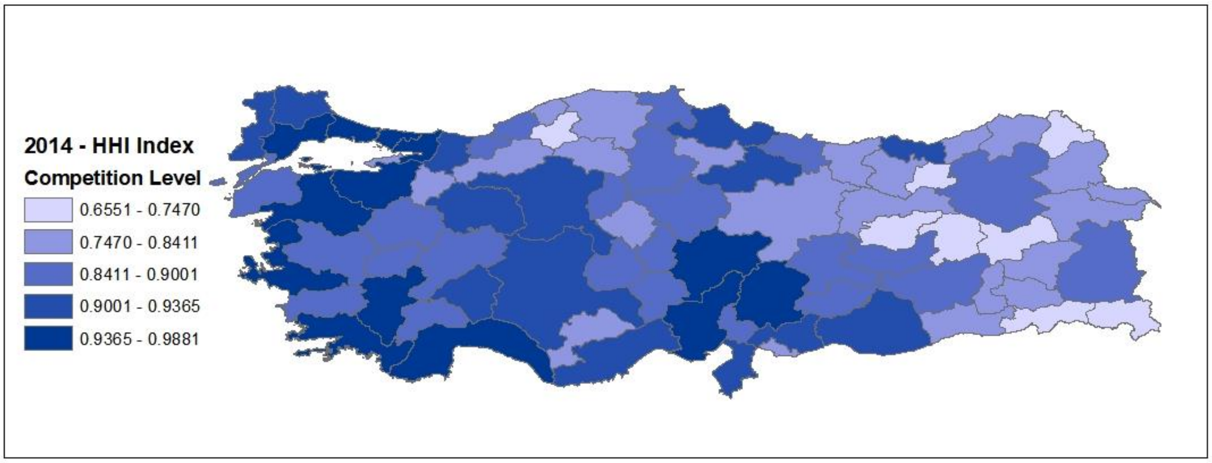
Competition

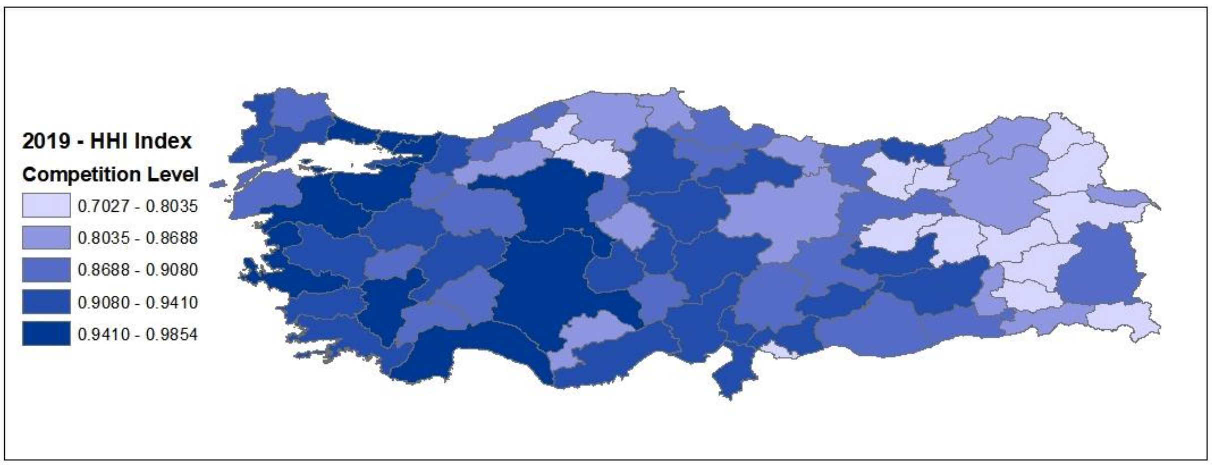
$$Comp_{ir} = 1 - \sum_{f} (\frac{E_{fjr}}{E_{jr}})^2$$

The Hirschman-Herfindahl Index is used to measure the competition among firms within a region. This is another important agglomeration indicator as it shows whether the regional employment is concentrated among a few firms or spread out. This can be interpreted as a low value indicating lower competition (regional employment concentrated among a small number of firms) and a value closer to one indicating intense competition (regional employment spread out over a large number of firms).

There is high competition among firms in Western Turkey and it decreases as one moves across the region towards the East, though there are a few exceptions as seen in graphic 2. The competition is becoming fiercer as one moves across time, which may also be linked to the increase in the number of firms that are established across Turkey. The overall competition has more or less remained similar, however the concentration in the Western provinces has increased while the industries in Eastern cities are becoming less competitive.

Graphic 2: Comparison of Competition between 2014 and 2019 across Turkey





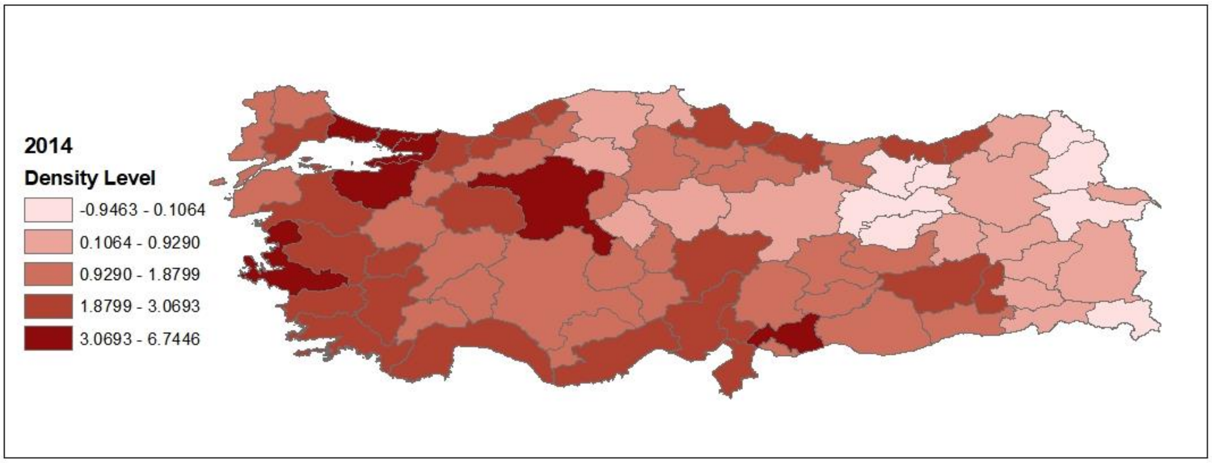
Density

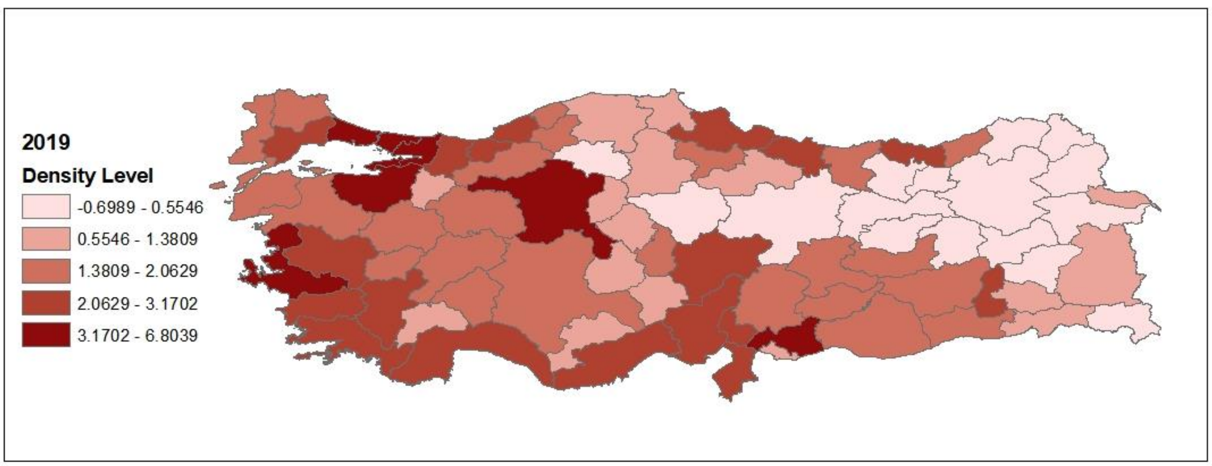
$$Den_r = \ln(\frac{E_r}{Area_r})$$

Employment density is used to capture the urbanization effects. It should be noted that we use employment density and not population density because we want to check for the concentration of employment in different regions. A high employment density means there is higher urbanization due to the concentration of people.

Employment density is an important factor in wage differences as demonstrated in agglomeration literature. The three biggest cities of Turkey show the highest employment density, as expected. The employment density in the Eastern cities continues to decline possibly showing migration from these cities into larger Western cities which host a larger number of industries. The overall density levels have risen over time but Eastern cities continue to have a lower number of employees when compared to the Western cities.

Graphic : Comparison of Density between 2014 and 2019 across Turkey





Two-stage Regression

To estimate the difference in wages that are not explained by the employee characteristics, a two-stage regression is utilized. In the first stage, the mincer-equation is estimated with an interaction dummy term for the sector, region, and year.

$$\log(w) = \alpha + \beta_1 age + \beta_2 age^2 + \beta_3 D_{gender} + \sum_{occ=1}^{9} \beta_4 , occ D_{occ} + \sum_i \sum_r \sum_t \varphi_{irt} D_j D_r D_t + \varepsilon$$

The interaction dummy term will take the coefficient of the particular industry within a region at time t. These coefficients give us the differences that are occurring in each industry and region over time. In the second stage of the regression, we will regress the different characteristics of each region that we had calculated earlier.

$$\varphi_{jrt} = \alpha + \beta_1 Spec + \beta_2 Div + \beta_3 Comp + \beta_4 Den + \beta_5 Area + \beta_6 D_j + \beta_7 D_t + \varepsilon_{irt}$$

In the second stage of the regression, we regress the different agglomeration variables estimated earlier on the interaction term coefficients. φ_{irt} is the interaction term that is estimated in the first stage of regression, $\beta_1 Spec$ represents specialization for each region, $\beta_2 Div$ represents diversity measured by Shannon's entropy, $\beta_3 Comp$ represent competition measured using the Hirschman-Herfindahl index, $\beta_4 Den$ represents employee density, and $\beta_5 Area$ represents the log of region's area. In addition, sector and time dummy variables are used to check for their fixed effects.

To obtain robust standard errors, we have used clustered standard errors at the regional level. This is in line with the approach used by Groot et al. (2014) in their study of the Netherlands micro data. Some results shown below have these regional clustered standard errors where specified.

Results

Cross-section results for years 2009 and 2021 are presented in Table 3., The firm size dummy results are also displayed because it can explain wage differences since larger firms may be more productive and pay higher wages. Taking micro-enterprises as a point of reference, we observe higher wages as the firm size increases. At each level between small and medium enterprises, and between medium enterprises and above SMEs, the difference is more than double. Individuals looking to earn higher wages should aim for larger firms where they can expect to be compensated better.

Table 3:District-level cross-section results for the Mincerian regression from 2009-2021

	2009	2014	2019	2021
	(1)	(2)	(3)	(4)
Variables	Ln(w)	Ln(w)	Ln(w)	Ln(w)
Age	0.0420***	0.0368***	0.0338***	0.0376***
	(0.000116)	(7.64e-05)	(7.25e-05)	(7.24e-05)
Age squared	-0.000491***	-0.000436***	-0.000382***	-0.000424***
	(1.64e-06)	(1.06e-06)	(9.92e-07)	(9.85e-07)
Female	-0.0224***	-0.0953***	-0.0960***	-0.106***
	(0.000382)	(0.000268)	(0.000238)	(0.000235)
Small Enterprise	0.119***	0.120***	0.0976***	0.113***
	(0.000444)	(0.000325)	(0.000307)	(0.000330)
Medium Enterprise	0.299***	0.272***	0.234***	0.262***
	(0.000493)	(0.000351)	(0.000331)	(0.000347)
>SMEs	0.539***	0.473***	0.431***	0.489***
	(0.000480)	(0.000345)	(0.000315)	(0.000328)
Constant	2.383***	3.036***	3.889***	4.149***
	(0.00200)	(0.00135)	(0.00129)	(0.00130)
Observations	7,522,589	11,444,253	11,501,701	14,059,330
R-squared	0.325	0.431	0.448	0.404
Occupation Dummy	No	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes	Yes
Region Dummy	Yes	Yes	Yes	Yes
Firm Size Dummy	Yes	Yes	Yes	Yes

*Micro-enterprises are taken as a reference point for the Firm size dummy.

Large differences were not expected between the city and district-level regressions for most variables since one is a more macro representation of another. Here 2009 results are included to get an idea of the relationship between wages and independent variables across time, 2009 to 2021 gives us a large window to observe the variables. However, not having occupational data is a large limitation for the dataset of

2009. This can be observed by the rise in R-squared when an occupation dummy is added in years 2014 onwards. One noticeable result is the jump in the coefficient for female employees from 2009 to 2014, according to the model there was far less wage inequality between the genders in 2009. However, that coefficient remains fairly consistent from 2014 to 2021, therefore we can postulate that the results in 2009 may be affected by the lack of occupational or educational data.

Table 4: City-level cross-section results for the Mincer regression from 2009-2021

	2009	2014	2019	2021
	(1)	(2)	(3)	(4)
Variables	Ln(w)	Ln(w)	Ln(w)	Ln(w)
Age	0.0436***	0.0377***	0.0348***	0.0388***
	(0.000118)	(7.77e-05)	(7.39e-05)	(7.38e-05)
Age squared	-0.000512***	-0.000449***	-0.000396***	-0.000439***
	(1.67e-06)	(1.08e-06)	(1.01e-06)	(1.00e-06)
Female	-0.0241***	-0.100***	-0.100***	-0.109***
	(0.000389)	(0.000272)	(0.000242)	(0.000239)
Small Enterprise	0.124***	0.124***	0.100***	0.117***
	(0.000451)	(0.000330)	(0.000312)	(0.000335)
Medium Enterprise	0.307***	0.284***	0.241***	0.270***
	(0.000495)	(0.000352)	(0.000333)	(0.000351)
>SMEs	0.572***	0.502***	0.454***	0.512***
	(0.000472)	(0.000339)	(0.000308)	(0.000323)
Constant	2.344***	3.010***	3.863***	4.119***
	(0.00204)	(0.00137)	(0.00131)	(0.00132)
Observations	7,522,590	11,444,253	11,501,701	14,059,331
R-squared	0.295	0.411	0.425	0.380
Occupation Dummy	No	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes	Yes
Region Dummy	Yes	Yes	Yes	Yes
Firm Size Dummy	Yes	Yes	Yes	Yes

*Micro-enterprises are taken as a reference point for the Firm size dummy.

The panel results show a lower coefficient for the female workers and continue to decrease as we add more years to the dataset. As observed in the cross-section results, the district results are slightly lower when compared to the city results. The panel results both on district and city-level are tested both with standard errors and then with clustered standard errors (clustering at district and city-level respectively). The coefficients for the variables show slight changes between the two different regressions (SE and clustered SE), with the errors increasing quite a bit when regional clustering is applied. The standard errors are larger at the city-level, as the geographical size increases the errors tend to increase with it. Cameron and Miller (2015) suggest clustering at the highest level (i.e., broader level), which in this case is city level. though it is useful to see the results for standard errors on both district and city levels. In both cases, the results are significant and only change very slightly.

Considering the highest occupation category (Managers) as the reference point, we find that as we move down the list, the wages decrease. This is expected as on average one would expect there to be higher wages since higher job positions are generally associated with higher experience and education.

Table 5: District and city-level panel results for the Mincer regression from 2014-2019

Year	Panel 2014-2019		Panel 2014-2019 w/Clustered SE	
	District	City	District	City
Variables	Ln(w)	Ln(w)	Ln(w)	Ln(w)
Age	0.0359***	0.0364***	0.0360***	0.0366***
	(2.70e-05)	(2.74e-05)	(0.00175)	(0.00479)
Age squared	-0.000413***	-0.000421***	-0.000416***	-0.000424***
	(3.60e-07)	(3.65e-07)	(2.05e-05)	(5.63e-05)
Female	-0.0852***	-0.0894***	-0.0838***	-0.0880***
	(0.000106)	(0.000108)	(0.00428)	(0.00447)
Small Enterprise	0.111***	0.115***	0.110***	0.114***
	(0.000135)	(0.000137)	(0.00592)	(0.0181)
Medium Enterprise	0.251***	0.262***	0.251***	0.262***
	(0.000144)	(0.000145)	(0.00931)	(0.0321)
>SMEs	0.440***	0.469***	0.441***	0.470***
	(0.000140)	(0.000137)	(0.0123)	(0.0290)
Professionals	-0.102***	-0.103***	-0.100***	-0.102***
	(0.000318)	(0.000323)	(0.0100)	(0.0101)
Technicians and Associate Professionals	-0.271***	-0.279***	-0.271***	-0.279***
	(0.000314)	(0.000319)	(0.0148)	(0.0220)
Clerical Support Workers	-0.318***	-0.326***	-0.319***	-0.326***
	(0.000319)	(0.000324)	(0.0135)	(0.0311)
Services and Sales Workers	-0.439***	-0.451***	-0.438***	-0.450***
	(0.000298)	(0.000302)	(0.0190)	(0.0431)
Skilled Agricultural, Forestry and Fishery Workers	-0.469***	-0.482***	-0.468***	-0.481***
	(0.000685)	(0.000695)	(0.0208)	(0.0405)
Craft and Related Trades Workers	-0.457***	-0.474***	-0.457***	-0.474***
	(0.000302)	(0.000307)	(0.0173)	(0.0501)
Plant and Machine Operators, and Assemblers	-0.453***	-0.475***	-0.454***	-0.475***
	(0.000306)	(0.000310)	(0.0162)	(0.0492)
Elementary Occupations	-0.476***	-0.491***	-0.475***	-0.490***

	(0.000295)	(0.000300)	(0.0176)	(0.0500)
Constant	3.808***	3.799***	3.805***	3.686***
	(0.000573)	(0.000582)	(0.0284)	(0.0747)
Observations	73,628,279	73,628,279	73,265,359	73,265,359
R-squared	0.516	0.500	0.516	0.481
Occupation Dummy	Yes	Yes	Yes	Yes
Industry Dummy	Yes	Yes	Yes	Yes
Region Dummy	Yes	Yes	Yes	Yes
Firm Size Dummy	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes

*Micro-enterprises are taken as a reference point for the Firm size dummy, managers are taken as a reference point for the occupation dummy.

The employee characteristics (age, gender, occupation) utilized in this model alongside the industry, region, firm size, and year dummies give us a good idea of how the wages of individuals are determined in Turkey. Still yet, there is a large proportion of wages that are not explained by the employee characteristics. In order to explain those differences, we follow a two-stage model proposed by Groot et al. (2011) and further utilized by many other studies especially when studying the case of emerging economies. Equation (x) is estimated to get the following results presented in Table 6

Table 6: Explaining regional wage differences not explained by employee characteristics

Variables	2014-2019	2014-2019	2014-2019 Clustered SE
Specialization	0.00492	-0.0625	-0.0625
	(0.0652)	(0.0620)	(0.120)
HHI Index	-0.155***	-0.132***	-0.132***
	(0.0139)	(0.0133)	(0.0420)
Diversity	0.769***	0.844***	0.844***
	(0.0724)	(0.0690)	(0.190)
Density	0.0534***	0.0527***	0.0527***
	(0.00162)	(0.00154)	(0.00501)
Log Area	0.0390***	0.0379***	0.0379***
	(0.00294)	(0.00279)	(0.00790)
Constant	-0.382***	-0.395***	-0.395***
	(0.0273)	(0.0260)	(0.0771)
Observations	8,802	8,802	8,802
R-squared	0.422	0.478	0.478

Industry Dummy	Yes	Yes	Yes
Year Dummy	No	Yes	Yes

The results we find are similar to those that have been found in previous studies.(here cite some papers) A significant positive effect of density is observed with and without fixed year effects, 5.27% and 5.34% respectively. That is, the doubling of employment density in the region would have a positive effect on wages to the extent of over 5%. Additionally, we find a positive significant effect of diversity and area. Competition, however, as observed from the results, has a significant negative effect on the overall wages. The negative effects of competition may be offset by the positive effects of diversity where knowledge spillovers between industries lead to improved results of agglomeration economies. The results show no significant effect of specialization, whereas diversity is observed to have the largest effect on wages in Turkey. Therefore, in the case of Turkey, employees' wages benefit the most in more industrially diverse cities where there is a large variety of industries instead of a concentration of specific industries.

Our findings are similar to those of other emerging economies such as Indonesia Ridhwan (2021), Ecuador (Torres-Gutiérrez et al., 2019), China, India, and Brazil (Chauvin et al., 2017), Brazil (Silva & Azzoni, 2021), and Colombia (Duranton, 2015). These studies all provide support for a positive effect of density on wage differences. For example using Indonesian data Ridhwan (2021) estimates a density effect of 6% compared to our 5% estimate. Our results are also comparable to Torres-Gutiérrez et al., (2019) where density is defined as employment with respect to area in EcuadorEcuador (). Yet these estimates remain lower than those found for other developing countries where density has been defined as population density (Prasertsoon and Puttanapong (2022)).

Similar to density, literature in developing economies also finds mixed results for diversity. Ridhwan (2021) finds a significant negative effect between 3% to 7% (depending on their model specifications) in the case of Indonesia. On the other hand, Barufi (2015) explains that the negative or positive sign for diversity may be due to the way it is measured. Measuring diversity in three different ways they find a positive effect when using Shannon entropy and a negative effect when using the share of 5 largest sectors as a proxy for diversity. This could also be a result of the sectors of study; therefore, the results must be interpreted with care. In the case of Turkey, this study finds a significant positive effect of diversity and we also utilize Shannon's entropy to measure diversity.

Conclusion

This study is the first to use individual level worker and firm data to measure agglomeration externalities in Turkey. Using a novel dataset over a 6-year period, this study finds that employment density plays an important role in wage premiums across Turkish cities. Furthermore, diversity is found to have a significant positive effect on wages whereas competition is found to have a significant negative effect. This study contributes to literature on agglomeration economies in developing countries By providing important evidence for Turkiye.

In the case of Turkey, we found that agglomeration has a positive effect on wage premiums both trough the increasing of density (number of employees within a region) or diversity (the number of different firms and industries within a region). Our results also reveal a significant negative effect of competition on wage premiums, however, that should be interpreted with caution since competition was measured among different industries and not within industry among firms. We also found that employees gain higher wages as they improve their occupations, which explains how education affects the ability of workers to earn wages. The strong relationship found between density, diversity, and labor wages shows the importance of urbanization in developing economies.

Turkey is an emerging country with both a growing population and economy. As cities grow rapidly across Turkey and other developing countries, the importance of agglomeration and its effects on labor wages will augment. This study sets out to contribute to this literature with micro data from Turkey which will further the debate on agglomeration externalities and bring about more studies in developing countries along with Turkey.

This study, like all others, has limitations and should be interpreted keeping those in mind. Though we show the effects of agglomeration economies on wage premiums across Turkey, we cannot be certain they are causal. However, as found in plethora of literature, agglomeration externalities do exist and have evidence in both developed and developing countries. In such studies, instrumental variables (IVs) have been used in the past, especially the use of historical population data to show that in fact cities have been growing as a historical phenomenon and not due to the benefit of higher wage premiums. However, we do not apply any IVs in this study though we plan to utilize them in the future to further our work on this topic within Turkey. Future studies may also study specific sectors across Turkey (in addition to the Manufacturing sector which has garnered attention in the past). In addition, migration and wage premiums shall also be of interest in future studies.

References

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of human resources*, *50*(2), 317-372.

Kent, O. (2015). *Agglomeration economies and geographic distribution of manufacturing industries in Turkey* (Doctoral dissertation). http://www.dspace.yildiz.edu.tr/xmlui/handle/1/12198

Özgüzel, C. (2023). Agglomeration effects in a developing economy: evidence from Turkey. *Journal of Economic Geography*, *23*(4), 823-846.