



OECD Economics Department Working Papers No. 1643

Why are some U.S. cities successful, while others are not? Empirical evidence from machine learning

**Damien Azzopardi,
Fozan Fareed,
Patrick Lenain,
Douglas Sutherland**

<https://dx.doi.org/10.1787/7f77c2e7-en>

ECONOMICS DEPARTMENT

WHY ARE SOME U.S. CITIES SUCCESSFUL, WHILE OTHERS ARE NOT? EMPIRICAL EVIDENCE FROM MACHINE LEARNING

ECONOMICS DEPARTMENT WORKING PAPERS No. 1643

By: Damien Azzopardi, Fozan Fareed, Patrick Lenain and Douglas Sutherland

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Authorised for publication by Alvaro Pereira, Director, Country Studies Branch, Economics Department.

All Economics Department Working Papers are available at www.oecd.org/eco/workingpapers.

JT03469759

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works.

Comments on Working Papers are welcomed, and may be sent to OECD Economics Department, 2 rue André Pascal, 75775 Paris Cedex 16, France, or by e-mail to eco.contact@oecd.org.

All Economics Department Working Papers are available at www.oecd.org/eco/workingpapers.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

© OECD (2020)

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. Requests for commercial use and translation rights should be submitted to PubRights@oecd.org

Abstract/Résumé**Why are some U.S. cities successful, while others are not? Empirical evidence from machine learning**

The U.S. population has become increasingly concentrated in large metropolitan areas. However, there are striking differences in between the performances of big cities: some of them have been very successful and have been able to pull away from the rest, while others have stagnated or even declined. The main objective of this paper is to characterize U.S. metropolitan areas according to their labor-market performance: which metropolitan areas are struggling and falling behind? Which ones are flourishing? Which ones are staying resilient by adapting to shocks? We rely on an unsupervised machine learning technique called Hierarchical Agglomerative Clustering (HAC) to conduct this empirical investigation. The data comes from a number of sources including the new Job-to-Job (J2J) flows dataset from the Census Bureau, which reports the near universe of job movements in and out of employment at the metropolitan level. We characterize the fate of metropolitan areas by tracking their job mobility rate, unemployment rate, income growth, population increase, net change in job-to-job mobility and GDP growth. Our results indicate that the 372 metropolitan areas under examination can be categorized into four statistically distinct groups: booming areas (67), prosperous mega metropolitan areas (99), resilient areas (149) and distressed metropolitan areas (57). The results show that areas that are doing well are predominantly located in the south and the west. The main features of their success have revolved around embracing digital technologies, adopting local regulations friendly to job mobility and business creation, avoiding strict rules on land-use and housing market, and improving the wellbeing of the city's population. These results highlight that cities adopting well-targeted policies can accelerate the return to growth after a shock.

This Working Paper relates to the 2020 OECD Economic Survey of the United States

<http://www.oecd.org/economy/united-states-economic-snapshot/>

Keywords: Labour mobility, job-to-job flows, metropolitan areas, clustering analysis, United States.

JEL Classification: E24, J11, J61, C38, O51

Pourquoi certaines villes américaines sont-elles florissantes alors que d'autres vivent ?**Données empiriques obtenues au moyen d'une technique d'apprentissage automatique**

La population américaine est de plus en plus concentrée dans les grandes agglomérations métropolitaines. Cela étant, il existe des disparités frappantes entre les grandes villes : certaines sont très florissantes et ont réussi à se détacher du lot alors que d'autres ont stagné, voire régressé. La présente étude vise principalement à classer les agglomérations métropolitaines américaines en fonction de leurs résultats sur le marché du travail : quelles sont celles qui sont en difficulté et se font distancer ? Quelles sont celles qui sont florissantes ? Quelles sont celles qui sont en mesure de faire preuve de résilience en s'adaptant aux chocs ? Nous avons utilisé une technique d'apprentissage automatique non supervisé appelée Hierarchical Agglomerative Clustering (HAC) pour mener à bien cette enquête empirique. Les données proviennent de nombreuses sources, telle que la nouvelle base de données sur les passages d'un emploi à un autre du Bureau du recensement des États-Unis, qui rend compte de l'univers proche des transitions professionnelles (entrées dans l'emploi ou le non emploi ou sorties de l'emploi ou du non emploi) dans les agglomérations urbaines aux États-Unis. Nous décrivons l'évolution des agglomérations métropolitaines en suivant leur taux de mobilité professionnelle, leur taux de chômage, la croissance des revenus, leur croissance démographique, la variation nette de la mobilité professionnelle et la croissance du PIB. Selon les résultats de l'étude, les 372 agglomérations couvertes par l'enquête peuvent être classées dans quatre catégories statistiquement distinctes : les agglomérations en plein essor (67) ; les mégapoles prospères (99), les agglomérations résilientes (149) ; et les agglomérations métropolitaines en difficulté (57). Les résultats montrent que les agglomérations qui s'en sortent bien sont principalement situées dans le sud et

l'ouest du pays. Leur succès se fonde principalement sur les facteurs suivants : l'intégration des technologies numériques, l'adoption de réglementations locales propices à la mobilité professionnelle et à la création d'entreprises, l'absence de règles restrictives relatives à l'aménagement du territoire et au marché du logement, ainsi que l'amélioration du bien-être de leurs habitants. Cette étude met en évidence que dans les villes qui ont adopté des politiques publiques bien ciblées, le retour de la croissance après un choc peut être plus rapide.

Ce document de travail est lié à l'Étude économique de l'OCDE de 2020 consacrée aux États-Unis.

<http://www.oecd.org/fr/economie/etats-unis-en-un-coup-d-oeil/>

Mots-clés : mobilité professionnelle, flux d'emploi, agglomérations métropolitaines, analyse par grappe, États-Unis.

Classification JEL : E24, J11, J61, C38, O51

Table of contents

Why are some U.S. cities successful, while others are not? Empirical evidence from machine Learning	6
Introduction	6
About the data	9
Empirical strategy	10
Results of the clustering analysis	12
Concluding remarks	15
References	15
Annex A.	17
Tables	
Table 1. Characteristics of Metropolitan Clusters in 2017	14
Table A.1. Descriptive statistics	21
Table A.2. Characteristics of Metropolitan Clusters in 2009	21
Figures	
Figure 1: The population is increasingly urban	8
Figure 2: Firms and employment are increasingly located in metro areas	8
Figure 3: Metropolitan areas are becoming more unequal	9
Figure 4: Dendrogram from the hierarchical clustering analysis	12
Figure A.1. Manufacturing outputs have changed markedly	17
Figure A.2. Regional differences persist in labor market characteristics	19
Boxes	
Box A.1. The Job-to-Job (J2J) Flows Data from the U.S. Census Bureau	18

Why are some U.S. cities successful, while others are not? Empirical evidence from machine Learning

By Damien Azzopardi, Fozan Fareed, Patrick Lenain and Douglas Sutherland¹

Introduction

1. The economic geography of the United States is changing. As technology, trade and preferences have changed so has the location of the population and economic activity (Hyatt et al., 2018). In some areas, factories have closed and jobs have been lost. Employment has grown elsewhere as new industries have developed in places that are not bound by past production networks. Moreover, over the past 50 years, the growth of services and high-tech products, the rising importance of foreign trade and integration of global value chains to the US economy have also contributed to changing locus of economic activity (OECD, 2020b). For example, the share of manufacturing in employment has declined, particularly in the old industrial heartland of the Mid-East and the Great Lakes region. This region accounted for one half of total production and employment but now only accounts for one third as manufacturing activity has shifted to the south and west.²

2. Accompanying these changes has been increasing urbanization. At the last census in 2010, around 80% of the population was urban (areas with at least 2,500 inhabitants), with the share increasing on average by 3 percentage points each decade, as shown in Figure 1 (U.S. Census Bureau, 2012). The share of the population living in metropolitan areas in the U.S. is greater than the OECD average of 64% (OECD, 2013). Moreover, metropolitan areas with populations greater than 500,000 are especially prevalent and they are important centres for employment and economic activity. Since 2000, larger cities have accounted for over three-quarters of national GDP growth (OECD, 2020b). Firm creation rates, while declining overall, have fallen by less in metropolitan areas, contributing an increasing urban concentration of firms and employment (Figure 2). Employment has fallen since the great recession in rural areas, particularly those that are not close to a metropolitan area (Arnosti and Liu, 2018). Generally, smaller urban areas have struggled during the 2010s in comparison with larger and better-connected areas.

¹ Damien Azzopardi is a Statistician in the OECD Economics Department. Fozan Fareed is an OECD Consultant and an Adjunct Lecturer at Université Paris-Est, Patrick Lenain is an Assistant Director in the OECD Economics Department; and Douglas Sutherland is a Senior Economist in the same department. The authors thank Duncan MacDonald and Mikkel Hermansen for their comments and valuable suggestions. Editorial assistance from Stephanie Henry is also gratefully acknowledged.

² Annex 1 provides recent statistics on how manufacturing output has changed markedly across different regions over time.

3. The COVID-19 pandemic has triggered severe recessions around the world. Beyond this short-term impact, long-lasting changes are also likely to ensue (Barrero et al., 2020). Furthermore, many U.S. cities had not fully recovered from the global financial crisis when they were hit by the COVID-19 pandemic.

4. The aim of this paper is to analyse the labour market performance of metropolitan areas and investigate why some areas remain depressed while others are prosperous. Using an unsupervised machine learning technique (known as Hierarchical Agglomerative Clustering), our analysis identifies which metropolitan areas are doing relatively well, which ones are struggling and falling behind, and which ones are staying resilient by adapting to shocks. Our results indicate that the main features of success revolve around embracing digital technologies, adopting local regulations friendly to job mobility and business creation, avoiding strict rules on land-use and housing market, and improving the wellbeing of the city's population. Based on these results, we draw lessons for the post COVID-19 era and add to the recent literature that has critically analysed the U.S. labour market (Autor et al., 2020; Azar et al., 2020; Barrero et al., 2020, among others).

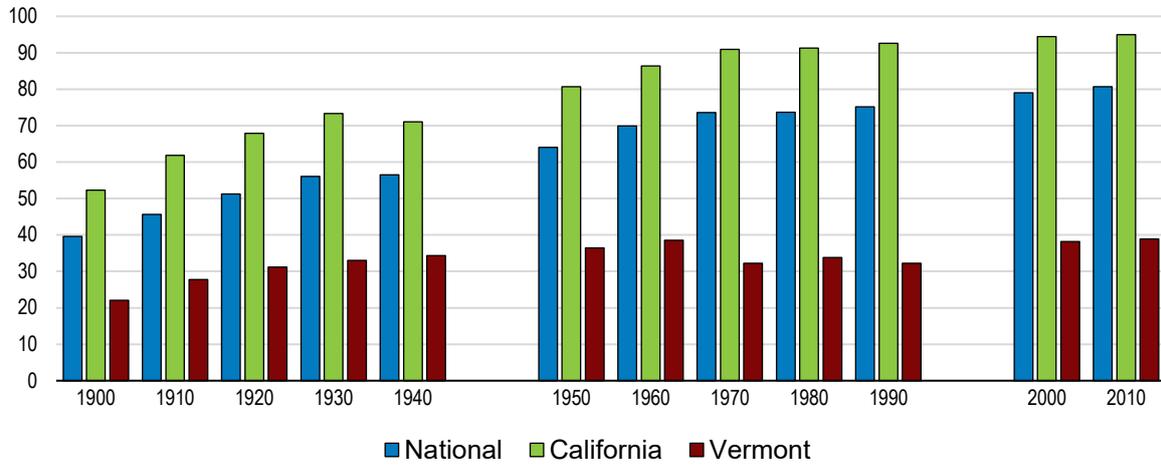
5. While the plight of states gives an insight into the differences across the United States, developments within metropolitan areas are more relevant for a number of reasons. Productivity growth seems to be increasingly concentrated in metropolitan areas, and differences between metropolitan areas are emerging as better performing cities pull away from the rest (Figure 3, Panel A). Higher income metropolitan areas have enjoyed stronger real output growth since the turn of the century. In addition, larger metropolitan areas are experiencing faster employment growth, although average output growth since 2001 has been slightly below that of smaller metropolitan areas (Figure 3, panel B). The outcomes of small metropolitan areas are quite heterogeneous. A few, such as Midland, Texas, have been growing rapidly due to the expansion of shale oil production, but other smaller cities are falling behind and losing employment as output contracts. Overall these developments suggest a pattern of returns to scale benefitting larger cities, but constraints, such as regulatory impediments or land-use restrictions, damping growth for others (Rappaport, 2018).

6. The divergences between metropolitan areas create social challenges because jobs are being created in places far away from the places where old jobs are lost. Having jobs created in areas where the population is located is vital to boost productivity and help people remain active (Bambalaitė et al., 2020; Foster and Pöschl, 2016). Moving from job to job can help workers to avoid spells of joblessness, remain productive and benefit from higher earnings (Haltiwanger et al., 2018). There is also literature that finds that job-to-job moves can help workers with positive assortative matching where they find a better match for their skills (Foster and Pöschl, 2016). However, the U.S. population has become less mobile: the share of the population moving each year has fallen from around 20% in the 1970s to under 10% more recently, with moves across state and metropolitan boundaries or moves to look for work also having been reduced (OECD, 2020b). Therefore, in order to devise better policies to address the economic challenges of the 21st century, it is important to understand why some different metropolitan areas have dynamic labor markets, while others do not.

7. The rest of the paper is organized as follows. Section 2 provides an overview of the data. Section 3 describes our empirical strategy. It elaborates the details of the clustering analysis that is being used to identify homogenous clusters of metropolitan areas. Section 4 presents the results of the clustering analysis. It provides a detailed discussion on how the different clusters of metropolitan areas vary from each other in terms of their labour market characteristics and economic performance. Section 5 concludes.

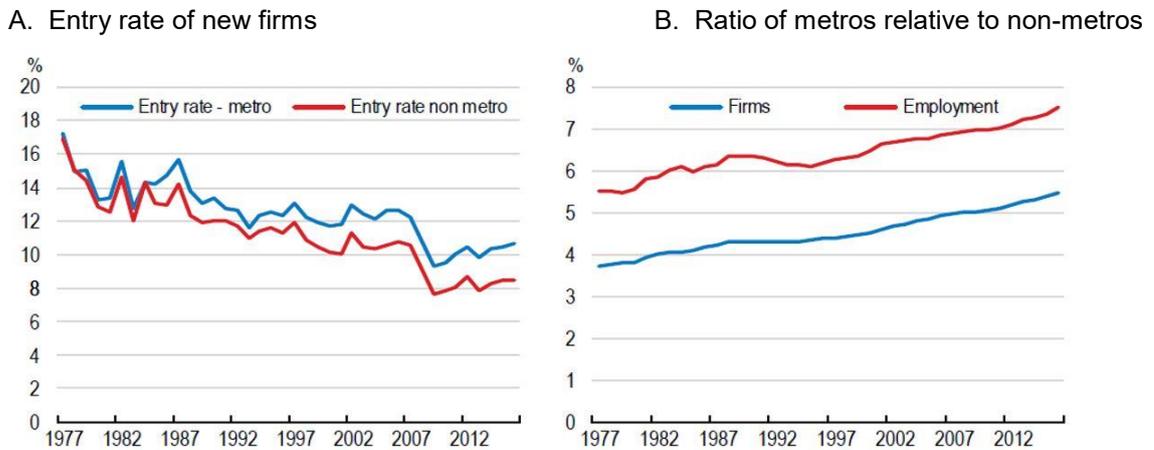
Figure 1: The population is increasingly urban

Per cent of population that lives in urban areas in the United States and selected states.



Note: Definitions were changed in 1950 and 2000. For example, in 2010, all urbanized areas and urban clusters were delineated solely on population density, without reference to place boundaries³.

Figure 2: Firms and employment are increasingly located in metro areas

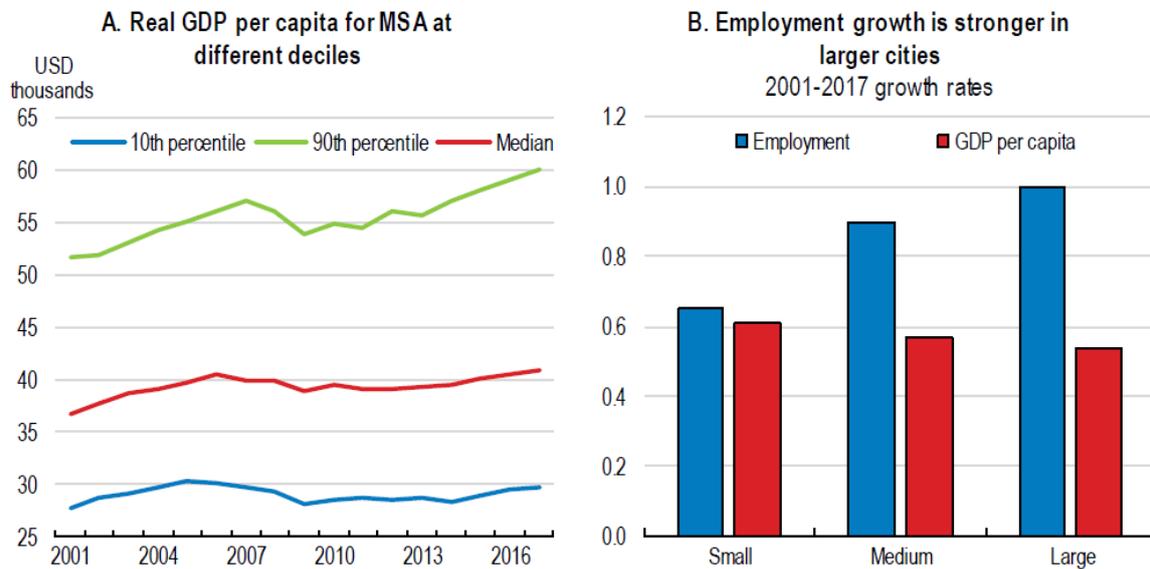


Source: Census Bureau, BDS database

³ For a detailed description of the changes in the definitions of “urban” and “rural” areas, please see: https://www.census.gov/history/www/programs/geography/urban_and_rural_areas.html.

Figure 3: Metropolitan areas are becoming more unequal

Real GDP per capita and employment in Metropolitan Statistical Areas (MSA).



Note: City size is the smallest, middle and upper third of the MSA distribution in 2001.

Source: Bureau of Economic Analysis

About the data

8. The data that we use for the purpose of this analysis is at the metropolitan level.⁴ It is obtained after combining information from several data sources. First, we use the new Job-to-Job (J2J) flows dataset from the U.S. Census bureau to get information on the job-to-job mobility rate and the net change in job-to-job flows at the metropolitan level. In addition to the J2J data, we rely on Bureau of Labour Statistics (BLS), Bureau of Economic Analysis (BEA) and the U.S. Census Bureau to obtain metropolitan level data on unemployment rate, income growth, population size, and GDP growth rate for the year 2017. The overall sample includes 372 metropolitan areas.⁵

9. The J2J data is constructed from a linked employer-employee database (Longitudinal Employer Household Dynamics, LEHD data), which is collected by the state unemployment insurance programs. It provides detailed information on hires and separations in the United States, with a focus on worker reallocation across employers, on a quarterly basis. The dataset has a nearly universal coverage of job transitions and movements into and out of employment. The coverage of private sector employment is almost complete and most public sector employees are included (missing employment are the self-employed, contract workers, federal government jobs and others not covered by state unemployment

⁴ The United States Office of Management and Budget (OMB) delineates metropolitan statistical areas according to published standards that are applied to Census Bureau data. According to the broad definition, “the general concept of a metropolitan or micropolitan statistical area is that of a core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core”. For a more detailed description, please see <https://www.census.gov/programs-surveys/metro-micro/about.html>.

⁵ Due to the lack of availability of data, 12 metropolitan areas are excluded from the analysis. For example, data is not available for metropolitan areas in the state of *Alaska*. Other excluded metro areas include Springfield (MA), Boston-Cambridge-Newton (MA-NH), Manchester-Nashua (NH), New Haven-Milford (CT) and Norwich-New London, (CT).

insurance). For example, the data counts an employment stock of 131 million workers and almost 15 million job hires in 2018 Q1, of which about half are job-to-job hires with available information on the previous job⁶.

10. A job-to-job hire is defined as a job hire following a separation from another firm with no or only a brief nonemployment spell (i.e. change of employer within the same quarter or with the new employer in the subsequent quarter), whereas the job-to-job mobility rate is measured by dividing the total number of job-to-job hires by the total number of employed people in that time period. A brief description of these definitions and how they are measured in the J2J data is provided in Annex 2.

11. Job-to-job hires account for about 50% of all job hires (Hermansen, 2019). Moreover, the J2J data also provide information on the origin and the destination of the job-to-job moves. This information enables us to aggregate data about the difference in the number of job-to-job flows that move into a metropolitan area as compared to job-to-job flows outside a metropolitan area. This difference highlights the net change in job-to-job flows in a metropolitan area. For example, the metropolitan area *Abilene, Texas* has a negative net change in job-to-job flows in 2017 i.e. the number of job-to-job hires that left *Abilene* in 2017 were greater than the job-to-job hires that moved into *Abilene*. Annex 3 provides metropolitan-level maps highlighting the differences in job-to-job mobility, unemployment rate and net change in job mobility across different areas. Descriptive statistics are provided in Annex 4.

Empirical strategy

12. We rely on an unsupervised machine learning technique known as Hierarchical Agglomerative Clustering to conduct this analysis. Broadly speaking, machine learning algorithms take advantage of high computational power of computers to learn from the data. Models that fall within the confines of machine learning are generally non-parametric in nature, as compared to econometric models that are mainly parametrical.⁷

13. Parametric models rely on a functional form that link the outcome variable with the regressors based on certain assumptions. Besides the main advantages of parametrical models such as their ease of interpretability, fast computational speed, less reliance on large datasets, some of their main weaknesses include their limited predictive power and specification risks (James et al., 2013). Nonparametric models, however, make the minimum amount of assumptions regarding the functional form of the relationship in order to have more flexibility to learn from the data (Azzopardi et al., 2019). Their main strength is associated with their flexibility and high predictive power, whereas one main shortcoming of these model is the risk of overfitting and their reliance on a stable underlying structure to the data (Bazarbash, 2019; Jordan and Mitchell, 2015).⁸

14. When it comes to machine learning, it is important to distinguish between two different types: supervised and unsupervised machine learning. In case of supervised learning, the model is trained using data which is “labelled”. In this case, the data on the outcome variable is available and the supervised algorithm learns from this “labelled” data to predict the outcomes for previously unseen data.⁹ However, in

⁶ Annex 2 provides an overview of the dataset. The J2J statistics are quarterly and available from 2000 onwards, currently until Quarter 1 of 2018. For more details about this dataset, please see <https://lehd.ces.census.gov/data/#j2j>.

⁷ There are some exceptions such as Least Absolute Selection and Shrinkage Operator, to name a few.

⁸ See Hastie, Tibshirani and Friedman (2016) for a detailed discussion on the comparison between econometric modelling and machine learning.

⁹ Examples of some of the most commonly used machine learning techniques include the random forest, gradient boosting decision trees, neural networks, and the support vector machine algorithm (Bazarbash, 2019).

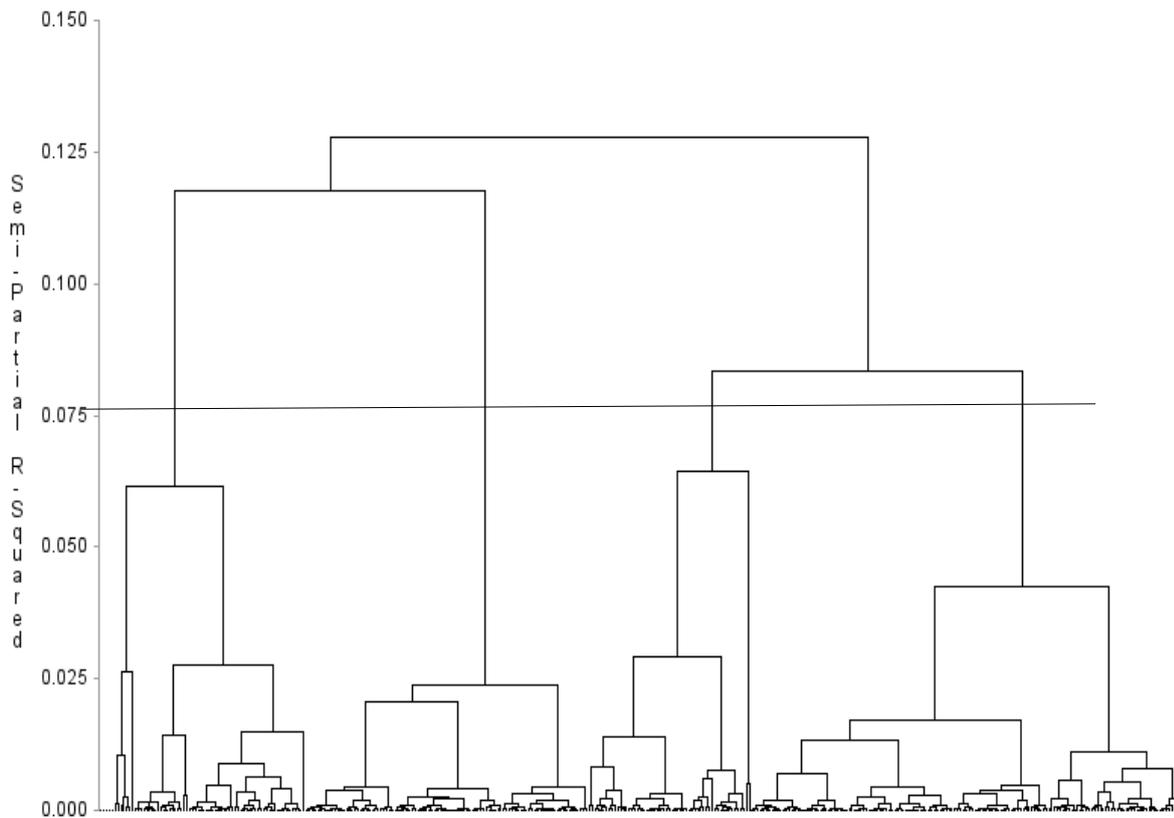
the case of unsupervised machine learning models, the outcome variable is not available, and the models basically deal with “unlabeled data”. The goal of the algorithm in this case is to find similarities among different features and group characteristics based on a set of identified variables. Since the outcome variable is not directly available in this case, we rely on unsupervised machine learning. Unsupervised models can cluster metropolitan areas into different groups with similar features, based on a set of relevant characteristics that explain labour market performance. One common form of unsupervised machine learning models is the hierarchical clustering analysis which we use to conduct this analysis.¹⁰

15. A partitioning method hierarchical agglomerative clustering is applied on the sample of 372 metropolitan areas in the U.S. to obtain homogeneous clusters of metropolitan areas. The variables used to carry out this analysis comprise of labor market indicators including measures of job-to-job mobility, income growth, unemployment rate, and population size and also economic performance indicator measured by metropolitan GDP growth. This algorithm makes it possible to break down the set of observations into several subsets that are statistically homogeneous in their characteristics. This ascending approach, or agglomeration, starts with an observation in each class, then successively merges the two closest classes, and stops when there is only one class containing all the observations. The distance measure used here is based on Euclidean distance.

16. We use the data from 372 metropolitan areas, covering about 86% of the population, for the year 2017, to carry out this analysis. The results of the clustering algorithm suggest that there are four statistically distinct groupings of the U.S. metropolitan areas. The benefit of the hierarchical clustering technique is the tree-based representation of the observations on a figure known as the dendrogram which is particular easy to interpret. Figure 4 represents the results of the hierarchical clustering. The results show a clear separation of four distinct clusters of metropolitan areas on the dendrogram, with the semi-partial R-squared (SPRSQ) represented by the vertical line. The semi-partial R-squared (SPRSQ) is a measure of the homogeneity of merged clusters and it provides a measure to read the dendrogram. The SPRSQ is basically considered as the loss of homogeneity due to combining two groups or clusters to form a new group or cluster. The value of SPRSQ should be small implying that, after merging, members of the groups are homogenous.

¹⁰ Another option would have been to rely on K-means clustering technique. K-means algorithm also partitions the full set of observation into K numbers of clusters, however, the number of total clusters must be defined prior to running the algorithm, which is why this technique is not considered here. Hierarchical Agglomerative Clustering does not impose any such condition prior to running the algorithm.

Figure 4: Dendrogram from the hierarchical clustering analysis

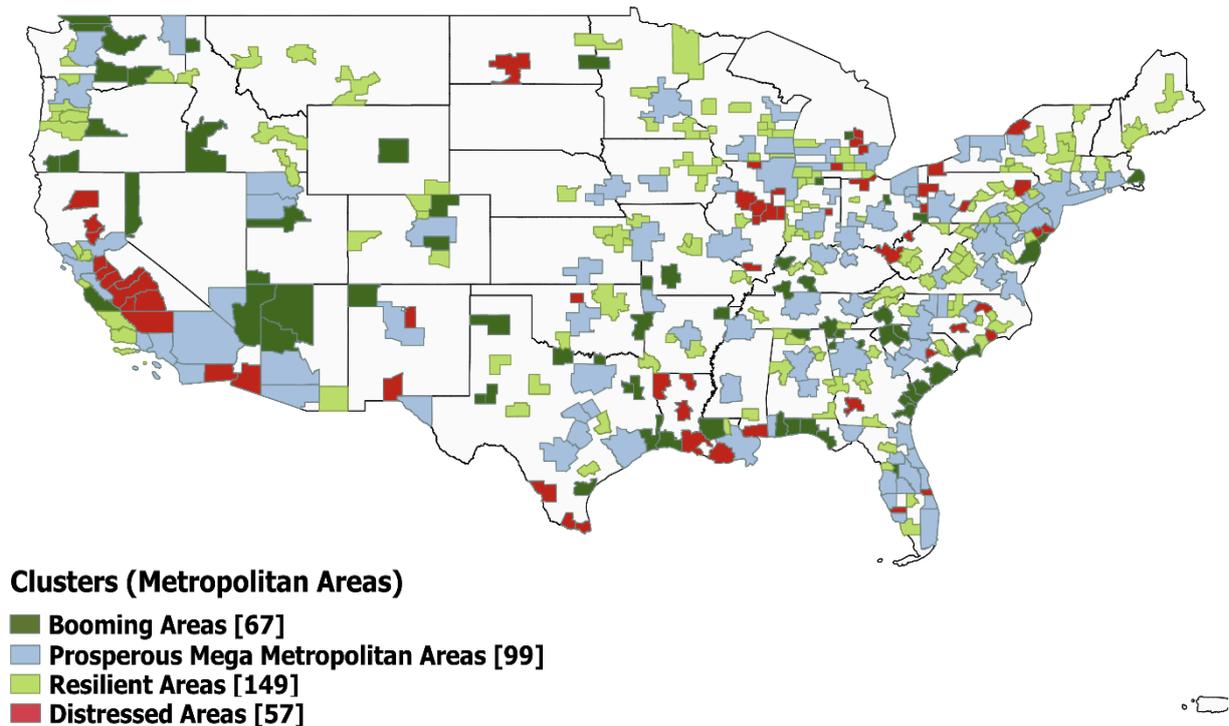


Results of the clustering analysis

17. The results of the clustering analysis suggest that metropolitan areas in the U.S. can be categorized into four distinct groups. The results reveal that some metropolitan areas are doing well, some are struggling and falling behind, but others are nonetheless adapting to shocks. This approach makes it possible to break down very diverse areas into different subsets that are statistically similar with also differences in the indicators of labor market fluidity. Given below are the details of the four clusters. These clusters can be characterized as: Booming areas, prosperous metropolitan areas, resilient areas, and distressed areas. Figure 5 depicts these clusters on a map, whereas Table 1 provides a summary of main characteristics associated with these clusters.

Figure 5: Booming metropolitan areas are mostly in the West and South

Results of the clustering analysis conducted on metropolitan areas (2017)



Source: OECD staff calculations.

- **Booming Metropolitan Areas:** These 67 metropolitan areas, home to about 7% of urban population, have enjoyed very fast growth of GDP per capita. They have often found success thanks to fast-growing industries, often with technology clusters - Midland, Austin, and Colorado Springs are examples. Other cities have found prosperity by becoming retirement destinations – most notably cities in Florida (The Villages, Pensacola area, Panama City). Cities in this group have become magnets for people looking for good jobs, high quality of life and comparatively low cost of living. For example, in 2017, about 305,000 workers were attracted by cities in Texas, many having decided to leave California and Louisiana. About 260,000 workers left states such as Georgia and New York and moved to Florida.
- **Prosperous Mega Metropolitan Areas:** This cluster is the largest one in terms of population: it includes 99 metropolitan areas and about three quarters of the U.S. urban population resides here. These are very large cities, with an average population size of 2 million, which can take advantage of agglomeration externalities. This category includes some of the largest U.S. metropolitan areas such as New York, Los Angeles, Chicago, Dallas, Houston, Washington DC and Miami. They have remained buoyant in the face of shocks and have benefited from low unemployment rates, average job mobility rate, and a high income per capita as compared to other clusters. However, rising inequality is a challenge here, and their future will depend on improving housing affordability and transportation.
- **Resilient Metropolitan Areas:** 149 metropolitan areas are part of this cluster and account for about 11% of the urban population. This cluster is mainly composed of relatively smaller areas such as Lewiston, ID-WA, Great Falls, MT, Columbus, IN and Kokomo, IN. Neither booming nor in

distress, these areas are generally classified by relatively low job mobility. However, they have a comparatively higher income per capita growth rate, and a number of these areas seem to be on an upward trajectory. The average population size of this group is the lowest among all clusters.

- **Distressed Metropolitan Areas:** This cluster includes 57 metropolitan areas that appear to be struggling. Home to about 6% of the total urban population, these areas are classified by a low job mobility rate, high unemployment rates, and a significantly lower GDP and income per capita growth rate, as compared to all other clusters. This group includes many trailing cities and old industrial areas. They can be found in North Dakota (Bismarck), Illinois (Bloomington, Champaign-Urbana) and Southern California (El Centro). Many metropolitan areas in central California are also in this cluster. In 2017, more than one-quarter million job-to-job moves went out of California to other states. The highest number of these jobs went to Texas (about 33,000) followed by Arizona (about 25,000) and Washington (about 24,000). Another major reason behind these moves seems to be the high cost of living and the high housing prices in some of these metropolitan areas.

Characteristics of Metropolitan Clusters in 2017

Table 1. Characteristics of Metropolitan Clusters in 2017

Clusters (metropolitan areas)	Main characteristics	Average J2J mobility <i>(in % of employment)</i>	Average GDP growth <i>(2016- 17 annual rate, %)</i>	Average unemployment rate <i>(2016-17, %)</i>	Average income per capita <i>(US\$, 2017)</i>
Booming areas	Very high mobility, net job gainers, high GDP growth	7.0	3.1	4.5	44,301
Prosperous mega areas	Average mobility, high income per capita, low unemployment, very highly populated	5.8	2.0	4.1	50,843
Resilient areas	Low mobility, high income growth, low unemployment	5.6	1.5	4.1	44,076
Distressed areas	Low mobility, lowest income	5.4	-0.2	6.5	40,952

	growth, high unemployment				
All		5.9	1.7	4.5	45,619

Source: OECD analysis based on data from BEA, BLS and Census Bureau

18. A comparison of these clusters for the year 2009 is also provided in Annex 5. It indicates how some clusters have performed better (booming areas, resilient areas) than other clusters (distressed areas), in terms of different economic indicators, since the financial crisis.

Concluding remarks

19. This paper reports that labor market fluidity in the U.S. has declined over time and substantial differences remain across metropolitan areas. We construct a unique measure of labor market fluidity at the metropolitan level by relying on a hierarchical agglomerative clustering algorithm. The results indicate that about 15% of the metropolitan areas can be classified as distressed, 18% as booming, 40% as resilient areas, and the remaining 27% are prosperous areas. The analysis also seems to highlight that economic growth is increasingly concentrated in certain areas, and differences between metropolitan areas are evolving as better performing metros pull away from others.

20. Diverging trends between cities create social challenges because new jobs are being created in places far away from the places where old jobs are lost. In order to address the economic and social challenges that the ongoing COVID-19 pandemic has brought to the fore, cities need to act now to avoid long periods of economic downturn (OECD, 2020a). With drastic changes happening in the urban ecosystem (Barrero et al., 2020), it has become more important than ever to focus on housing and land zoning rules, and other restrictions to mobility, notably occupational licensing (Bambalaite et al., 2020; Hermansen, 2019). With a major reallocation coming up, cities that address these regulatory barriers would be in a better position to benefit from new opportunities and attract businesses and talents looking for a new home.¹¹

References

- Arnosti, N., and Liu, A. (2018). "Why rural America needs cities". S.I.:Brookings Institution.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). "The fall of the labor share and the rise of superstar firms". *The Quarterly Journal of Economics*, 135(2), 645-709.
- Azar, J., Marinescu, I., Steinbaum, M., & Taska, B. (2020). "Concentration in US labor markets: Evidence from online vacancy data". *Labour Economics*, 66, 101886.
- Azzopardi, D., Fareed, F., Lenain, P. and Sutherland, D. (2019), "Assessing Household Financial Vulnerability: Empirical Evidence from the U.S. using Machine Learning", in OECD Economic Survey of the United States: Key Research Findings, OECD Publishing, Paris, 123- 145.
- Bambalaite, I., G. Nicoletti and C. von Rueden (2020), "Occupational entry regulations and their effects on productivity in services: Firm-level evidence", OECD Economics Department Working Papers, No. 1605, OECD Publishing, Paris.

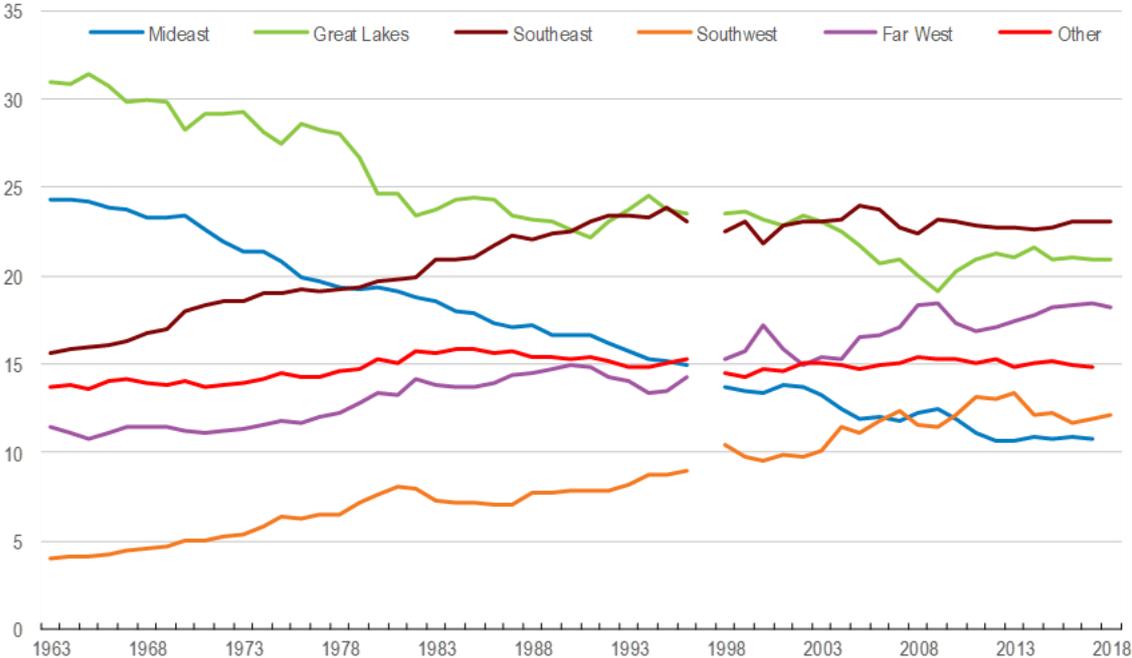
¹¹ Barrero et al. (2020) show that COVID-19 is also a reallocation shock and 2 out of 5 jobs lost during the pandemic in the United States may not come back.

- Barrero, J. M., Bloom, N., and Davis, S. J. (2020). "Covid-19 is also a reallocation shock". National Bureau of Economic Research Working Paper, No. w27137.
- Decker, R., J. Haltiwanger, R. Jarmin and J. Miranda (2016), "Declining Business Dynamism: What We Know and the Way Forward", *American Economic Review: Papers & Proceedings*, Vol. 106/5, pp. 203-207.
- Bazarbash, M. (2019). "FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk". International Monetary Fund.
- Foster-McGregor, N., and Pöschl, J. (2016). "Productivity effects of knowledge transfers through labour mobility". *Journal of Productivity Analysis*, 46(2-3), 169-184.
- Haltiwanger, J., Hyatt, H., and McEntarfer, E. (2018). "Who moves up the job ladder?". *Journal of Labor Economics*, 36(S1), S301-S336.
- Hastie, T., R. Tibshirani, and J. Friedman (2016). "The elements of statistical learning". New York: Springer.
- Hermansen, M. (2019). "Occupational licensing and job mobility in the United States". OECD Economics Department Working Papers, No. 1585, OECD Publishing, Paris
- Hyatt, H., E. McEntarfer, K. Ueda, and Zhang, A. (2018). "Interstate Migration and Employer-to-Employer Transitions in the United States: New Evidence from Administrative Records Data", *Demography*, Vol. 55 (6), 2161-2180.
- Hyatt, H. R., and Spletzer, J. R. (2013). "The recent decline in employment dynamics". *IZA Journal of Labor Economics*, 2(1), 5.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). "An introduction to statistical learning". Vol. 112. New York: springer.
- Jordan, M. and T. Mitchell (2015). "Machine learning: Trends, perspectives, and prospects". *Science*, 349(6245), 255-260.
- Molloy, R., Smith, C. L., and Wozniak, A. (2011). "Internal migration in the United States". *Journal of Economic Perspectives*, 25(3), 173-96.
- Molloy, R., Trezzi, R., Smith, C. L., and Wozniak, A. (2016). "Understanding declining fluidity in the US labor market". *Brookings Papers on Economic Activity*, 2016(1), 183-259.
- OCDE (2020a), "Cities Policy Responses", COVID-19 notes on policy responses.
- OECD (2020b). "[OECD Economic Surveys: United States 2020](#)", OECD Publishing, Paris.
- OECD (2018a). "[OECD Economic Surveys: United States 2018](#)", OECD Publishing, Paris.
- OECD (2018b), *Rethinking Urban Sprawl: Moving Towards Sustainable Cities*, OECD Publishing, Paris,
- OECD (2013), "Metropolitan areas", OECD Regional Statistics (database), <http://dx.doi.org/10.1787/data-00531-en>.
- Pugsley, B., and Sahin, A. (2015). "Grown-Up Business Cycles." Staff Report no. 707. Federal Reserve Bank of New York.
- Rappaport, J. (2018). "The Faster Growth of Larger, Less Crowded Locations". The Federal Reserve Bank of Kansas City Economic Review.
- U.S. Census Bureau (2012). "2010 Census of Population and Housing, Population and Housing Unit Counts, United States Summary" U.S. Government Printing Office, Washington, DC.

Annex A.

Figure A.1. Manufacturing outputs have changed markedly

Gross value added in manufacturing as per cent of total manufacturing output.



Note: Before 1997 the data are based on SIC and after that date on NAICS.
Source: Bureau of Economic Analysis.

Box A.1. The Job-to-Job (J2J) Flows Data from the U.S. Census Bureau

The U.S. Census Bureau provides detailed statistics on job mobility and transitions in and out of employment (freely available at https://lehd.ces.census.gov/data/j2j_beta.html). The J2J statistics are quarterly and available from 2000 Q2 onwards, currently until 2018 Q1. The data is constructed from a linked employer-employee database (Longitudinal Employer Household Dynamics, LEHD data), which is collected by state unemployment insurance programs. The coverage of private sector employment is almost complete and most public sector employees are included (missing employment compose self-employed, contract workers, federal government jobs and others not covered by state UI). For example, the data counts an employment stock of 131 million workers and almost 15 million job hires in 2018 Q1, of which half are job-to-job hires with available information on the previous job. In addition, statistics are available on average earnings for (stable) job stayers each quarter and for job movers by the quarter before and after a job-to-job move.

Measures from the J2J Data applied in this paper

- *Job hire*: A worker is employed in a firm by the end of the quarter (i.e. has the largest combined earnings from this employer) and did not receive earnings from the same firm in the previous quarter.
- *Job-to-job hire*: A job hire following a separation from another firm with no or only a brief nonemployment spell (i.e. change of employer within the same quarter or with the new employer in the subsequent quarter).
- *Nonemployment hire*: A job hire after no main job for at least a quarter (i.e. no earnings from the same employer in both the beginning and end of a quarter). Referred to as persistent nonemployment in the J2J Data.
- *Earnings*: Average earnings are calculated by dividing the sum of earnings received in a quarter by the count of workers for selected job histories. Information on hours worked or e.g. bonuses included are not available.

Structure of the data and level of aggregation

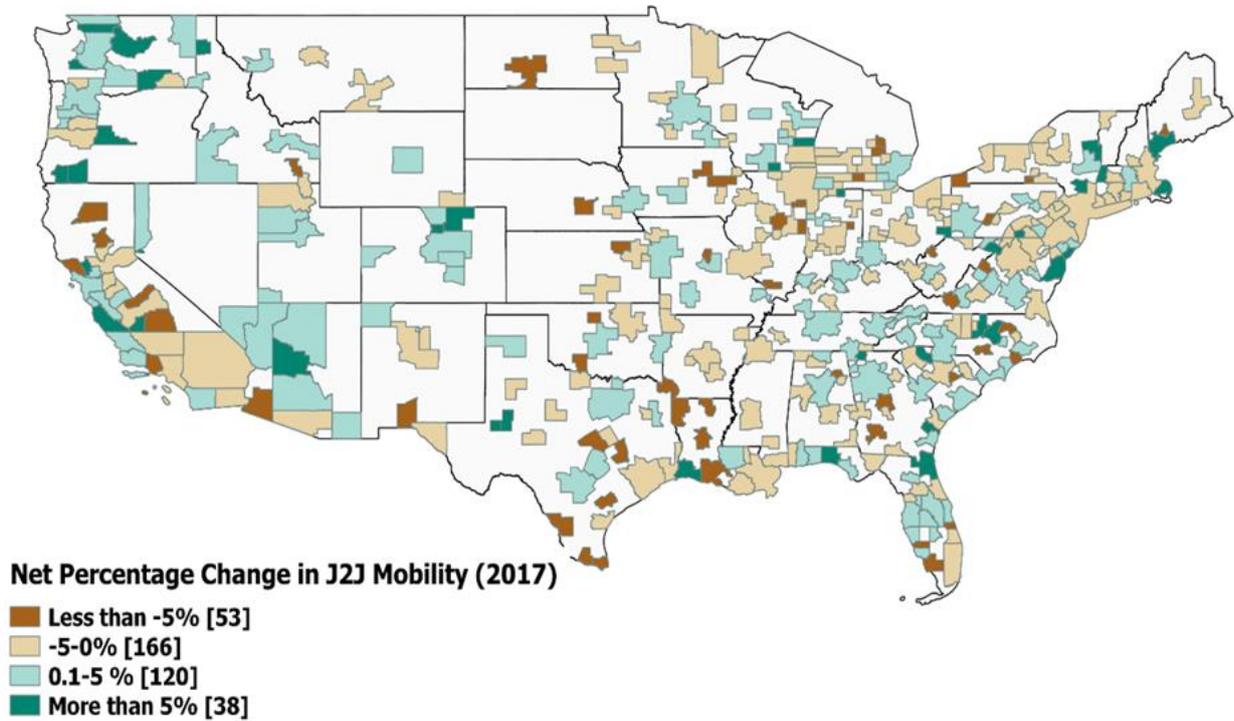
The J2J Data is released by semi-aggregate tabulations. This means that the number of hires and separations is reported in cells by metropolitan area, state, industry, worker and firm characteristics (sex, age group, race/ethnicity, education, firm age and firm size). Specifically, the data is made available in two different forms:

- *J2J Counts*: Includes counts of individuals with a specified job history in each quarter and details about either the destination job (hires) or the origin job (separations). For example, 1350 job-to-job hires is counted for women aged 25-34 in the management industry in New York in 2017 Q2. Average earnings of USD 16,485 is reported for job stayers in the same cell.
- *J2J Origin-Destination*: Includes counts of job-to-job changes only and provides details about both the origin and destination job. For example, three men aged 45-54 made a job-to-job move from the manufacturing industry in Michigan to retail trade in California in 2017 Q2. Their average earnings declined from USD 52,315 in the quarter prior to the move to USD 42,686 in the quarter after the move.

Source: Source: Hermansen (2019).

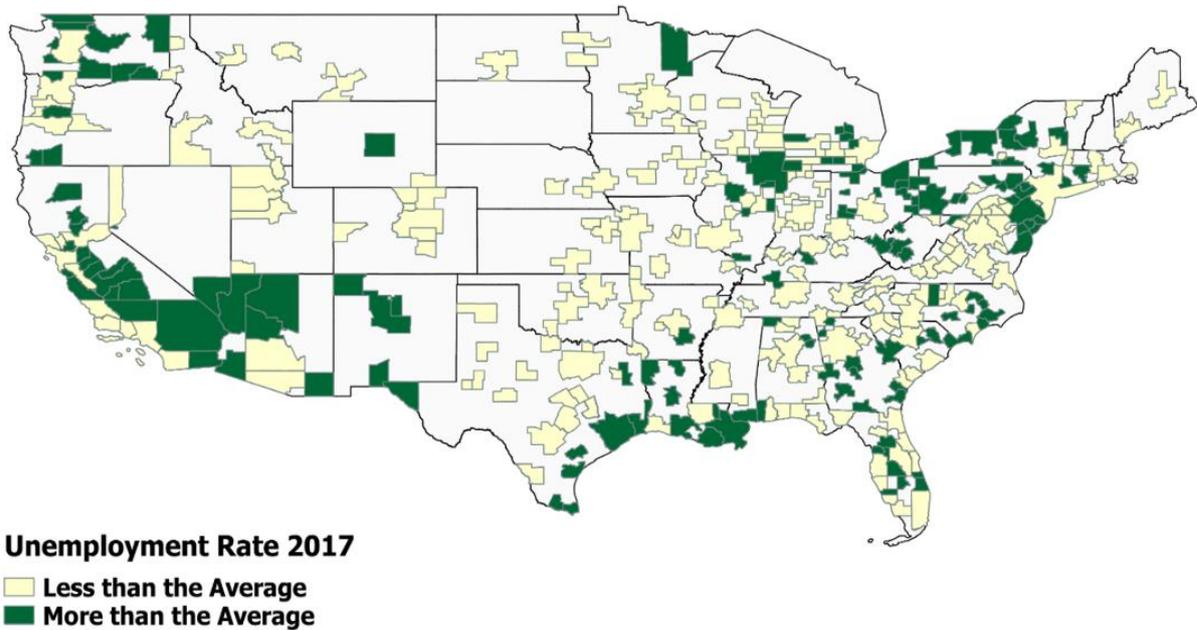
Figure A.2. Regional differences persist in labor market characteristics

A. Net Change in Job-to-Job Mobility Across Metropolitan Areas, 2017



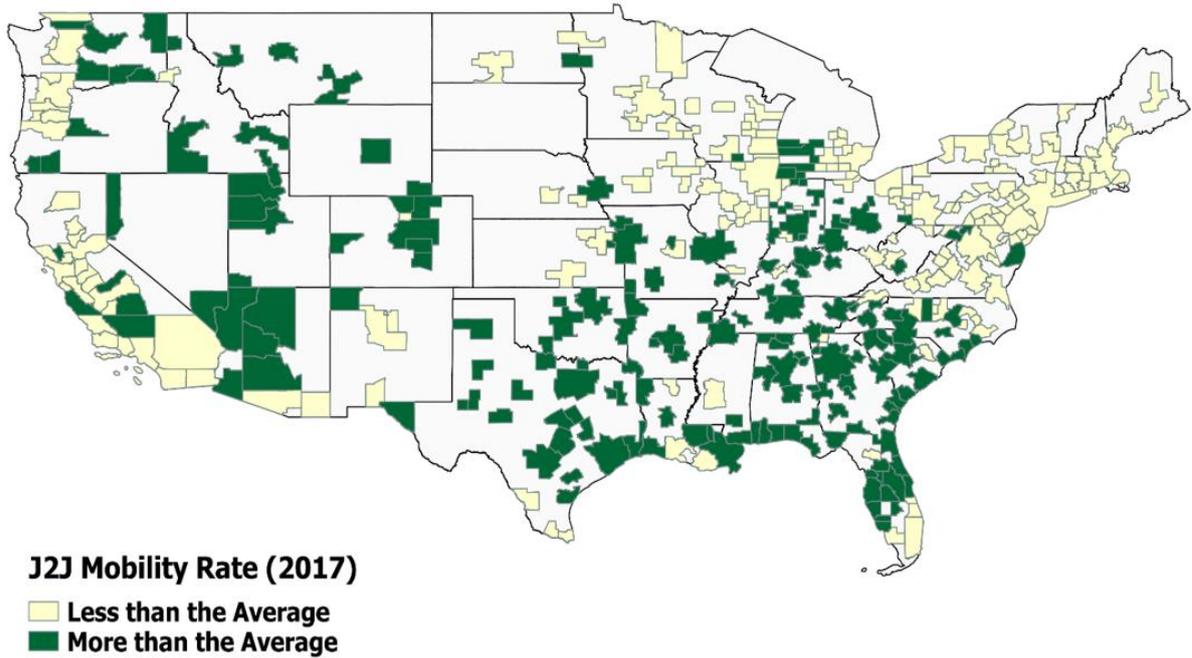
Source: Source: J2J Data from Census Bureau, 2017.

B. Unemployment Rate Across Metropolitan Areas, 2017



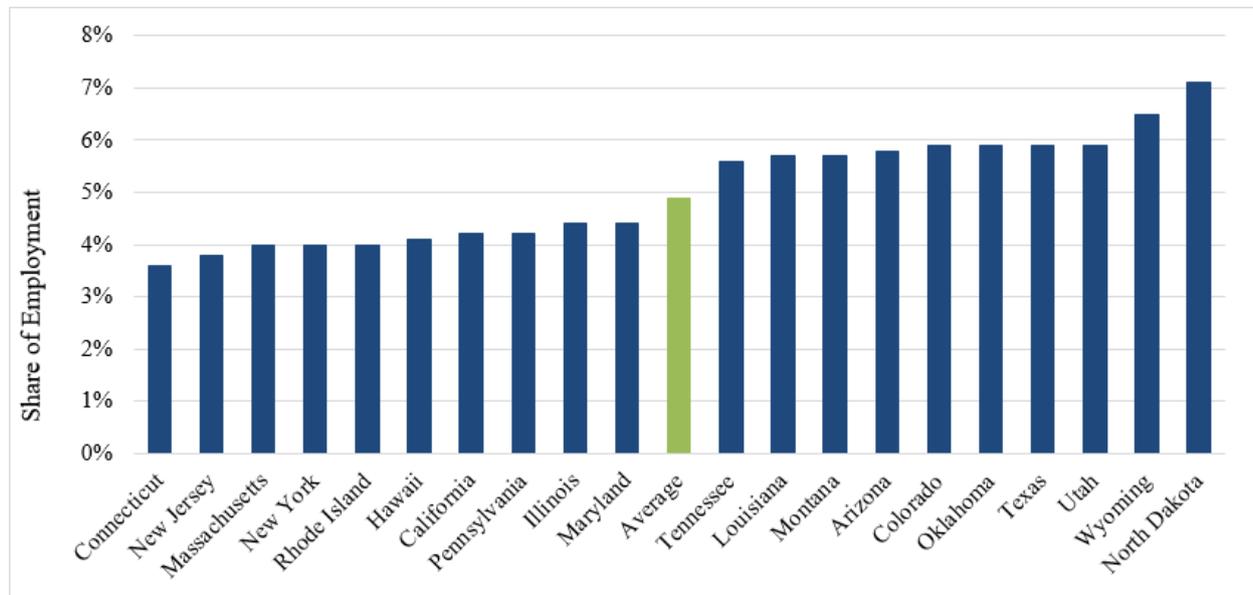
Source: Bureau of Labor Statistics and Census Bureau, 2017.

C. Job-to-Job Mobility Rate Across Metropolitan Areas, 2017



Source: J2J Data from Census Bureau, 2017.

D. Job-to-Job Mobility Rate Across States, 2010-2017 Average



Source: J2J Data from Census Bureau, 2017

Note: The charts presents ten states with the lowest rate of J2J mobility and ten states with the highest rate

Table A.1. Descriptive statistics

Variable Name	# of Obs.	Mean	S.D.	Min	Max
J2J Mobility Rate (<i>in % of employment</i>)	372	5.88	0.90	3.99	9.82
Income Growth (2016-17 annual rate, %)	372	3.28	1.46	-2.70	16.6
Unemployment (%)	372	4.53	1.60	1.98	19.63
Net Change in J2J Mobility Rate	372	-0.58	4.83	-18.10	23.80
Population	372	728646	1661776	54608	20000000
GDP Growth Rate 2016-17 annual rate %)	372	1.65	2.32	-7.80	12.1
Population Growth 2016-17 annual rate %)	372	0.26	0.60	0.02	7.19

Source: Census Bureau, BLS and BEA, 2017

Table A.2. Characteristics of Metropolitan Clusters in 2009

Clusters (metropolitan areas)	Average J2J mobility (<i>in % of employment 2009</i>)	Average GDP growth (2009-10 annual rate, %)	Average unemployment rate (2009 %)	Average income per capita (US\$, 2009)
Booming areas	4.4	2.4	4.4	34,241
Prosperous mega areas	3.7	1.7	3.6	39,260
Resilient areas	3.7	2.9	3.8	34,546
Distressed areas	3.8	1.8	5.9	32,634
All	3.8	2.3	4.2	35,595

Source: OECD analysis based on data from BEA, BLS and Census Bureau