The role of entrepreneurship, innovation, and urbanity-diversity on growth, unemployment, and income: US State-level evidence and an fsQCA elucidation

Abstract

This study considers roles played by dimensions of entrepreneurship, innovation, and geography on United States (US) state level growth, unemployment, and income employing Fuzzy-set Qualitative Comparative Analyses (fsQCA). One important developmental feature of the analyses is the use of a novel fuzzy membership score creation process, undertaken to calibrate the considered condition and outcome variables. Moreover, fuzzy cluster analyses are undertaken, using the fuzzy c-means technique, on sets of constituent variables to produce sets of clusters interpretable to the relevant condition and outcome variables. A series of fsQCA investigations are undertaken across the different outcome variables of growth, unemployment, and income. The fsQCA results offer novel insights into variations in the US state level based outcome variables, and how dimensions of entrepreneurship, innovation, and the urbanity-diversity of the states contribute to this. The novel applied and technical developments offer expanding ideas on this area of research.

1. Introduction

The relationship between entrepreneurship and wider measures of economic prosperity is complex, widely debated, and interlinked with issues of innovation and economic geography. For example, van Stel et al. (2005) evaluated the impact of Total Entrepreneurial Activity (TEA, effectively a business start-up) on economic growth in 36 countries, identifying an inconsistent relationship where TEA is found to be significantly positive for economically wealthy countries, but significantly negative for poorer countries (particularly developing economies). They also posit that in more developed economies, the positive relationship is associated with start-up entrepreneurship linked to innovation (specifically commercialization). Colombelli et al. (2016) and Malchow-Møller et al. (2011) further posit that such business start-ups contribute to job creation (and therefore to reducing unemployment), provided that the net effect of the new entrants, namely, taking over the market shares of exiting firms, also allow for overall market growth. In contrast, van Stel et al. (2005) argue that poorer countries failing to benefit from such entrepreneurial activity is an indication of insufficient larger firms being able to generate economies of scale, technology, and learning effects from innovation.

In terms of the stock of small businesses, Carree and Thurik's (2008) study found initially direct positive effects from increasing numbers of business owners on employment, gross domestic product (GDP), and productivity growth. This was followed by negative business exit-related effects, the final overall net effect (including supply-side effects) being positive for employment and GDP growth. Thurik et al. (2008) found, however, that the overall impact on reducing unemployment rates can take up to eight years. Contrastingly, Casson (2010) notes that older entrepreneurs remaining in existing businesses might reduce predicted benefits from entrepreneurship, where it prevents entrepreneurs moving to exploit opportunities elsewhere. Casson (2010) also suggests that an economy with a good supply of entrepreneurs but serious inefficiencies in its market for entrepreneurs, may find that entrepreneurs move to exploit opportunities elsewhere. This suggests both that existing business rates may be negatively related to start-ups and that high-growth entrepreneurship and international migration into an area potentially indicates the desirability of that location from an entrepreneurial start-up, growth, and innovation perspective, when complimented with a positive entrepreneurial environment. Furthermore, Acs (2008) identified that the most significant contributions to national economies were achieved by faster-growing "gazelle" firms, while Shane (2009) also argued that efficient policy should direct resources toward high-growth firms in particular, rather than those focused only on survival.

Interactions exist between entrepreneurship, innovation, and desirability of location; these variables also impact growth, unemployment, and income (Li et al., 2016; Rupasingha, 2017). Growth, unemployment, and income conditions, however, also interact with each other, making this a particularly complex issue to analyzes, with entrepreneurial, innovation and geographical drivers likely to impact in different ways in diverse environments (Huggins et al., 2017).

Therefore, the context of this study lies within the often-contested overlapping roles of entrepreneurship, innovation, and geography in economic development as measured by growth, unemployment, and income. Specifically, the study adopts the position that the literature is fundamentally related to the alternate ways in which entrepreneurship can be measured, in terms of start-up, small-firm survival, and high-growth, as well as the links between these different entrepreneurship measures, innovation, and the geographical context in which these activities take place. Synthesising both entrepreneurship and innovation, while also introducing an economic geography component, within the regional innovation systems literature, Cooke (2003) suggests that successful regions, as measured by their relatively high-growth and income and low unemployment rates, have “entrepreneurial” innovation systems, while peripheral regions have different, more “institutional,” government-directed systems. This dynamic suggests that a heterogeneous mixture of conditions and consequent policies will be relevant depending on the region in question.

There is a lack of analyses of which combinations of entrepreneurship, innovation, and geographical context lead to which outcome, a gap which is particularly relevant at the regional level, given the importance often placed on government policy in supporting economic development. The study utilizes United States (US) state-level data, where the literature is largely nascent, in order to conduct this analyses. Fazio et al. (2016), for example, suggests that while entrepreneurship levels significantly influence the US economy, creating a foundation for economic dynamism and prosperity with high-growth start-ups contributing disproportionately to both job creation and impactful innovation, there is also a variation in entrepreneurial potential for such growth across the US. Their review of existing analyses highlights either studies conducted on a small number of US states over time, or on cities and regions, rather than across US states. It is this gap that this study seeks to address. To evaluate these issues, fuzzy-set qualitative comparative analyses (fsQCA) (Ragin, 2000; 2008) is used in the analysese of US state-level data. As a set theoretic-based analyses technique, it is closely associated with small “*n*” data-level analyses, relevant in the case of US state-level data. In addition, where fsQCA has been previously employed using US state-level data, for example, in pension movement on US old-age policy (Amenta et al., 2005) and environmental justice policy (Kim and Verweij, 2016), it has allowed heterogeneity created by having very different state-level contexts to be accounted for through the identification of multiple pathways to an outcome. As Glaeser and Gottlieb's (2009) work highlights, the heterogeneous economies that US states represent means that fsQCA is potentially pertinent where equifinality and asymmetric aspects of the analyses may also be relevant, as well as conjunctional cauzation. In a review of QCA in public policy analyses, Rihoux et al. (2011) suggested US state-level analyses would benefit from this particular form of analyses.

One technical development in this study is a novel form of data generation for use with fsQCA, where the variables, condition and outcome must be in fuzzy membership score form. Here, sets of constituent-variables are considered for each variable (condition or outcome) used in the analyses, using fsQCA. For each set of constituent-variables, a two-step clustering process is first employed (Bacher et al., 2004; Şchiopu, 2010), whereby *i*) the optimum numbers of clusters is first found with concomitant silhouette coefficients (Zhou and Gao, 2014), and then, *ii*) fuzzy clustering, using fuzzy *c*-means (Bezdek, 1980) is used to establish the subsequent sets of clusters. This approach was employed in Beynon et al. (2018) to create a single outcome variable (with fsQCA results compared with this outcome variable and those found from a factor-analyses-derived outcome variable). For all considered condition and outcome variables, two cluster solutions are identified as optimum from their constituent-variables, from which fuzzy membership scores associating cases to one of the two clusters are shown to be appropriate as the necessary fuzzy membership scores for use with fsQCA.

The study's contribution to the theory of entrepreneurship and innovation offers novel insights into how different combinations of entrepreneurship, innovation, and urbanity-diversity affect growth, unemployment and income levels across US states, consequently effectively informing entrepreneurship policy and practice regarding which policies are most effective to generate specific outcomes.

The structure of the rest of the paper is as follows: in section 2, a discussion on the position of entrepreneurship, innovation, and urbanity-diversity differences, including within the US, and their consequence on growth, unemployment, and income, is undertaken. Section 3 describes the methodology and data, with the creation of a fuzzy membership score based on condition and outcome variables (using two-stage clustering). Section 4 describes the results from the fsQCA analyses of the considered US state-level data. Section 5 interprets the fsQCA-based findings. In section 6, focused conclusions develop the discussion of the results, articulated within the extant literature, presenting implications for theory and practice, and limitations and directions for future research.

2. Interactions between entrepreneurship, innovation and urbanity-diversity with growth, unemployment, and income

2.1 The relationships between different measures of entrepreneurship, innovation, and economic geography and growth

In terms of economic growth, Wennekers and Thurik (1999), Audretsch et al. (2006) and Cumming et al. (2014), found entrepreneurial activity measured in terms of business start-up to be an engine for economic development, increased innovation, and regional development in both industrialized and developing countries. Colombelli et al. (2016), Wennekers and Thurik (1999) and Dejardin (2011) suggest business start-ups play a key role in supporting competition, enabling innovation, and assisting new business sector development. Simultaneously, government policy-makers encourage innovation activity (Hausman, 2005), Beynon et al. (2015) and Santos (2000) positing that innovation is itself a critical process in improving business growth and performance.

With regard to the types of entrepreneurship of potential importance, Wennekers and Thurik (1999) implicitly identify start-up, fast growth, and innovation-focused entrepreneurship as beneficial to economic growth, competition effects also meaning that business exits (and consequently lower levels of existing businesses) are also regarded as positive. Valliere and Peterson (2009), for more developed economies, found a significant portion of economic growth attributable to high-growth focused entrepreneurs exploiting national innovation systems and regulatory freedom, whereas in less developed economies this effect is absent. Wong et al. (2005) used cross-sectional GEM data on 37 countries and found that for high-growth potential TEA, necessity TEA, opportunity TEA, and overall TEA, only high-growth potential entrepreneurship had a significant impact on economic growth. In the US context specifically, Phillips and Kirchoff (1989) find firm survival rates positively related to firm growth. Mueller’s (2007) results also indicate that increased innovative start-up activity is more effective than increased general entrepreneurship for economic growth.

Minghao et al. (2016) found that local economic conditions favorable to high-growth firms are significantly different from new firms more generally and that such conditions differ between more urban and rural areas. Valliere and Peterson (2009) also find that regional concentration of economic activity, driven by increasing localized returns from labor markets, reduced transportation costs, and increased demand for manufactured goods, is positively related to economic growth; these aspects are more likely to be found in more urbanized areas.

Thus, these findings indicate that start-up and particularly high-growth entrepreneurship (rather than firm survival) are relevant to economic growth, and that high-growth entrepreneurship is often linked to innovation activities in these processes. In addition, there are also potential impacts from the specific geographical context in which these activities take place, including within a country where individual regions may exhibit different economic development stages and disparate performance levels, and where more urbanized regions would seem to have an advantage. This dynamic suggests that high start-up entrepreneurship, in combination with high-growth entrepreneurship, and innovation within regions displaying stronger aspects of urbanity are more likely to experience high-growth, with more rural areas being likely to require different policy combinations to achieve the same outcomes.

2.2 *The relationships between different measures of entrepreneurship, innovation, and economic geography, and unemployment*

Turning the focus to unemployment highlights a different combination of measures of entrepreneurship as being of more importance. Cumming et al. (2014), for example, found that entrepreneurship, in terms of business start-up, positively impacts by reducing unemployment. Moreover, Audretsch and Thurik (2000) identified that increased entrepreneurship, measured by number of business owners as a proportion of the labor force, leads to lower levels of unemployment. While Baptista and Preto (2007) suggest a more ambiguous relationship, Fritsch and Mueller (2004) identify time-dependent effects, with initial positive direct effects of job creation in new entities, then negative indirect effects of new businesses crowding out competitors and, lastly, improved supply conditions and competitiveness, peak positive impacts of new businesses on regional development reached eight years after entry.

These findings indicate that firm start up, in conjunction with firm survival, is of particular relevance to unemployment. Again, there are geographical aspects to this, with several, sometimes contradictory processes in operation, linked to firm start-up, survival, and innovation. In the US context, Acs et al. (2007) found that greater geographical concentration of businesses reduces initial new-firm formation rates but also increases survival rates of business start-ups, advantages given by geographical concentration reducing formation of short-lived firms. Conversely, while a positive geographical concentration effect for innovation can be seen to exist, due to, for example, knowledge spillover effects (Acs and Varga, 2002), greater geographical concentration can also lead to lower innovative new-firm survival rates because of the potential for short-lived imitative businesses. This highlights that the quality of business start-ups, in terms of its effects on competitiveness, is itself related to innovation, the impact of geography then occurring through processes highlighted previously by Acs et al. (2007).

In terms of unemployment, therefore, the evidence suggests both that combinations of firm survival with one or more of start-up, innovation and geography are particularly relevant but also that, because of the geographical effects of concentration (where concentration will be more strongly associated with urban areas), different combinations will be required in other geographies to generate unemployment benefits. Specifically, in urban areas, unemployment benefits from entrepreneurship are likely related to high survival rates in areas without high innovation, while in more rural areas, high survival rates related to high firm start-up and/or innovation are more likely linked to positive unemployment effects.

2.3 The Relationships between different measures of entrepreneurship, innovation, and economic geography and income

Regarding effects on income, Cumming et al. (2014) suggest that entrepreneurship, measured by new firm formation, has a significantly positive relationship with GDP per capita (i.e., income levels). This is itself, however, also linked to economic geography, which also affects the type of entrepreneurship and the importance of innovation.

At the country level, Ács et al. (2008) identify entrepreneurship as having a U-shaped relationship with economic development (for which income-per-head levels serves as a proxy). This is related to individuals undertaking entrepreneurial activity for opportunity- and necessity-based reasons (Tominc and Rebernik, 2007). In lower-income, lower-innovation, developing, factor-driven economies (which are also likely to be more rural), entrepreneurship (particularly necessity-based) activity tends to be high. It then tends to decrease as economies enter the efficiency phase, which is more dominated by manufacturing, before entrepreneurship activity (more opportunity-based) rises again during the services and innovation-driven phase for developed, higher-income, higher-innovation economies (likely also more urbanized).

This dynamic suggests therefore, that while innovation is central to high-income outcomes generally, this will be in combination with different variables dependent on economic geography. In more urbanized geographies, where high-income is more likely, high-growth entrepreneurship and innovation are typically associated with high-income outcomes. In more rural economies, however, where low incomes are prevalent, obtaining high-income outcomes will require the development of a stock of surviving, innovation-driven growth firms.

3 Methodology and data

The above discussion identifies a complex interplay among combinations of geography, entrepreneurship, and innovation-related factors and overall growth, income and unemployment outcomes. This dynamic identifies a requirement for a method able to deal with both conjunctional causation and equifinality, with issues of asymmetry in outcomes also likely to exist, and a study of US states requires a method capable of dealing with a small *n* dataset. In summary, the literature review reveals a requirement for further research to enhance our understanding of the associations among entrepreneurial activity, innovation, and different aggregate-level outcomes in terms of growth, unemployment, and income. This is particularly pertinent in the context of comparative US states analyses, where the literature is nascent. This review therefore presents a critical perspective on the literature currently available, particularly in the US context. This section outlines the fsQCA methodology, provides a description of the data, and describes the creation of fuzzy membership score values for use in fsQCA.

FsQCA

FsQCA is a set-theoretic-based data analyses technique introduced in Ragin (2000; 2008). As a development of the original crisp QCA (Ragin, 1987), where crisp sets are binary, fsQCA can utilize fuzzy sets, sets in which membership can be expressed in degrees. To elucidate fsQCA against a more traditional quantitative approach, here regression, the more traditional regression-based approach investigates the effect of a condition variable on an outcome variable, the orientation of fsQCA is on what conditions lead to an outcome (Elliott, 2013).

FsQCA as a configurational comparative approach offers two practical dimensions in its analyses, as described in Fiss et al. (2013), "…*allows for equifinality (different configurations leading to the same outcome) and asymmetric causality (absence of causal conditions associated with an outcome not leading to absence of the outcome)*" (p. 192). Since its introduction, fsQCA has been employed in a range of business management disciplines, including in the area of entrepreneurship and innovation (Beynon et al., 2016a, b, 2018; Mallon et al., 2018). For further details on fsQCA, including its particular usage within the entrepreneurship and innovation discipline, see Kraus et al. (2018).

Data

The data considered here are with regard to US state-level analyses, and following the intended employment of fsQCA, are broken down into condition and outcome variables. We present the variables included in the study in addition to the reasons to include them instead of others.

<u>Condition variables</u>

This sub-section describes the construction of five condition variables. Three of these are derived from the most recent (2016) data contained within The Kauffman Index of Entrepreneurship series (Kauffman, 2017), which is an umbrella of annual reports that measure US entrepreneurship activities. Because this dataset provided a consistent range of data related to firm start-up, existing small-firm levels, and high-growth activity, it was determined that the data be used to construct the entrepreneurship-related variables. Specifically, three types of state-level entrepreneurship behavior are taken from this series and considered here:

1. The Start-up Activity Index focuses on the beginnings of entrepreneurship, specifically new business creation (Rate of New Entrepreneurs - RoNE), market opportunity (Opportunity Share of New Entrepreneur - OSoNE), and start-up density (Start-up Density - SD).
2. The Main Street Index focuses on the prevalence of local, small business ownership (Rate of Business Growth - RoBG; Survival Rate of Firms - SRoF; Established Small Business Density - ESBD).
3. The Growth Entrepreneurship Index focuses on growing companies (Rate of Start Growth - RoSG; Share of Scaleups - SoS; High-Growth Company Density - H-GCD).

While in this study these are used as condition variables, these areas are often considered as outcome variables, particularly with regards to start-up (Beynon et al., 2016a), and high-growth (Lee, 2014). In this study, and motivated by the intended use in fsQCA, a novel analyses is undertaken through the notion of clustering. Namely, the study considers the association of each US state to the area derivatives of low and high, Start-up Activity, Main Street and Growth Entrepreneurship.

In addition, a process of identifying three relevant constituent-variables from which to create further condition variables was also used for Innovation and for Urbanity-diversity, the specific data chosen on the basis of being the most recently available from official government sources and available for all US states. The three constituent-variables chosen for innovation and urbanity-diversity were also selected on the basis of complementarity in providing different while related measures of innovation and urbanity-diversity. We therefore believe they provide a reasonable basis for the fsQCA analyses subsequently conducted.

4. The Innovation Index focuses on patents (2015 Patents per 1000 population from the US Patent Office (2018), Higher Education Qualifications (2015 Percentage of Post-18 Population with Bachelor's Degree or Higher from the United States Census Bureau, (2018), and Research and Development spending (R&D spending as a percentage of State GDP from the latest available 2012 Bureau Of Economic Analyses (2018) data).
5. The Urbanity-Diversity Index focuses on the concentration of the state population in urban areas (Percentage of Population) living in urban areas with more than 50,000 people from the 2010 Census (United States Census, 2010), the density of those areas (Urban Area Population Density score from the 2010 Census (United States Census, 2010), and the diversity of the state population (in terms of the percentage of the state population that is foreign-born according to the 2015 American Community Survey (United States Census Bureau, 2018).

Moreover, following the approach adopted in Beynon et al. (2018), fuzzy c-means (FCM) clustering forms the clustering process employed (Bezdek, 1980), post-establishment of the optimum number of clusters using the first part of the often-employed two-step approach (Bacher et al., 2004; Şchiopu, 2010).

For each condition variable, the three identified sub-variables are used in the cluster analyses to form cluster solutions, which, since using FCM, means grades of membership are evaluated on the association of each case (US state) to the interpreted clusters. This is pertinent in the construction of variables for fsQCA, when there are two clusters established, and the concomitant grades of membership are limited based on the derivatives of high and low outcome.

<u>Outcome variables</u>

We chose three relevant outcome variables partly for reasons of consistency with the fuzzy-clustering-based creation process of the condition variables, available US government data being used for growth, unemployment, and income-per-head data, for specific variables where the data existed for all US states. Unlike the condition variables, however, a single definition (one for growth, unemployment, and income per head) was used because of the more specific nature of the outcome variables, with the last three years of annually available data being used in each case to reduce the potential for a single year of data to skew the data. Again, while other definitions for each of the variables could have been chosen, we believe they provide a reasonable basis for the fsQCA analyses subsequently conducted as follows:

1. US state Annual Growth rates as measured by Annual Percentage Change in Real GDP (in chained US dollars) for 2013-14, 2014-15, and 2015-16 from the Bureau of Economic Analyses (2018).
2. US state per capita personal Income for 2014, 2015 and 2016 from the Bureau of Economic Analyses (2018).
3. US state unemployment rates (U6 definition) for 2014, 2015 and 2016 from the Bureau of Labor Statistics (2018).

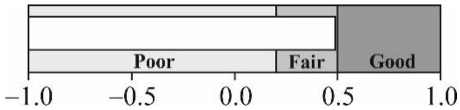
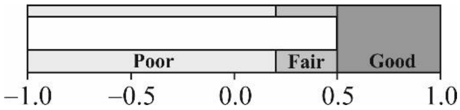
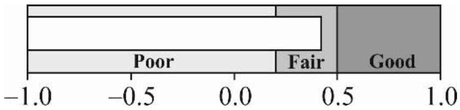
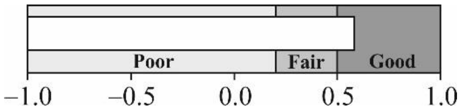
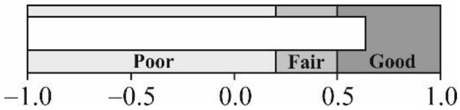
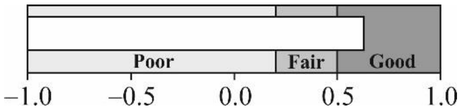
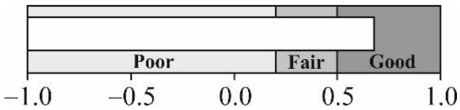
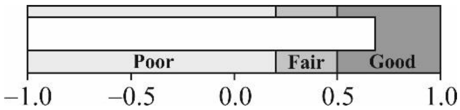
<u>Establishment of optimum number of clusters</u>

Following the discussion of the considered condition and outcome variables, and in particular the constituent-variables, the creation of the fsQCA useable variables is considered next. With a view to performing fuzzy clustering on the separate sets of constituent-variables, this section outlines the identification of the optimum number of clusters associated with each set of constituent-variables, for future condition and outcome variables appropriate for use in fsQCA.

This optimum (number of clusters) issue is the first part of the two-step clustering approach available in the Statistical Package for the Social Sciences used in this analyses (SPSS Inc., 2001); for independent evaluations of this approach, see Bacher et al. (2004) and Şchiopu (2010). In technical terms, the number of clusters can be automatically determined using a two-phase estimator, first using Bayesian Information Criterion, then by examining the ratio change in distance between clusters (Bacher et al., 2004). To quantify the quality of the chosen number of clusters for each set of constituent-variables, the concomitant silhouette coefficient plots are found (Zhou and Gao, 2014). The silhouette coefficient is a measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation), and ranges from −1 to +1 (Rousseeuw, 1987), with subranges noting different qualitative grades of appropriateness (using the SPSS two-step approach), namely, bad, fair and good (noting also if most objects have a good value, then the clustering configuration is appropriate, SPSS Inc., 2001).

The optimum number of clusters and concomitant silhouette coefficient plots for each set of constituent variables are shown in Table 1.

Table 1. Optimum number of clusters and concomitant silhouette coefficient plots for each set of constituent variables

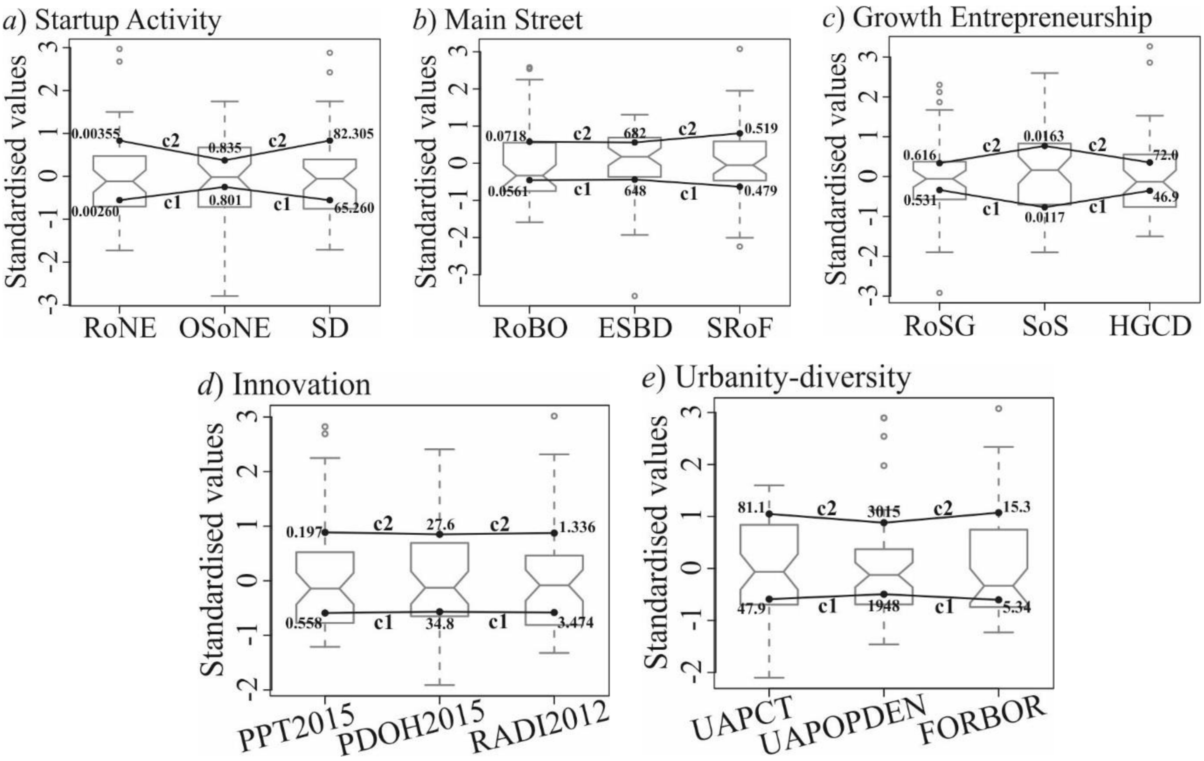
Variable (No. clusters) Silhouette plot	Variable (No. clusters) Silhouette plot
Start-up Activity (2 clusters) 	Main Street (2 clusters) 
Growth Entrepreneurship (2 clusters) 	Innovation (2 clusters) 
Urbanity-diversity (2 clusters) 	Growth (2 clusters) 
Unemployment (2 clusters) 	Income (2 clusters) 

In Table 1, each set of constituent-variables is shown to be optimally considered in terms of two clusters, the first stage of this approach being to 'grades of membership' based variable construction. The quality of these suggested optimizations of cluster numbers is shown by the specific silhouette plots for each set of constituent-variables, the plots shown (in terms of horizontal white bar) indicating a predominance of high-fair to good quality. We next move onto the fuzzy clustering part of this process, using FCM on each set of three constituent variables to establish two clusters.

<u>Fuzzy clustering of sets of constituent variables</u>

Following the approach in Beynon et al. (2018), which similarly used two clusters, FCM was employed to identify the grouping of US states to the two clusters shown to be optimum, for each of the areas, Start-up Activity, Main Street, Growth Entrepreneurship, and Urbanity-diversity. The FCM cluster results for each condition variable set of constituent-variables are graphically shown in Figure 1.

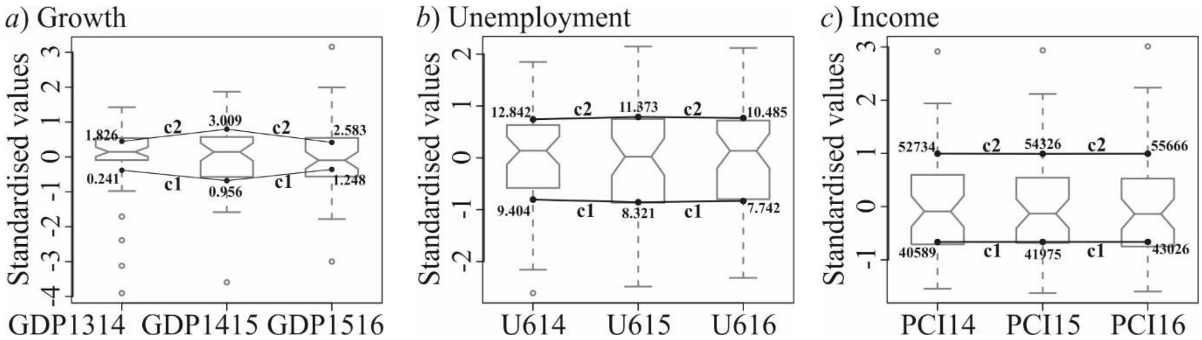
Figure 1. Constituent cluster variable means for the sets of three sub-variables across the five condition variables, Start-up Activity, Main Street, Growth Entrepreneurship, and Urbanity-diversity, found using FCM



In each graph shown, the points depict the constituent means for the sets of variable values associated with US states found to be most associated to a cluster. The lines between points enable the indication of sets of constituent means describing a cluster. While the *y*-axis indicates standardized values of the relevant scales (as used in FCM), the values shown in the plot, of the cluster constituent means, are given in their non-standardized values (for ease of interpretation).

Inspection of the sets of constituent means for the pairs of clusters shows there is one cluster with consistently higher mean values than the others. The indication being that the clusters, across each condition variable, can be consistently considered the concomitant derivatives of low and high Start-up Activity (C1 and C2 in Figure 1*a*), low and high Main Street (C1 and C2 in Figure 1*b*), low and high-Growth Entrepreneurship (C1 and C2 in Figure 1*c*), low and high Innovation (C1 and C2 in Figure 1*d*), and low and high Urbanity-diversity (C1 and C2 in Figure 1*e*). A similar approach is taken with the three considered outcome variables Growth, Unemployment, and Income (Figure 2).

Figure 2. Constituent cluster variable means for the sets of three sub-variables across the three outcome variables constructed, Growth, Unemployment and Income, found using the FCM



As with the condition variables, the reported constituent means across each set of constituent-variables discerns the clusters in terms of low and high-Growth (C1 and C2 in Figure 2*a*), low and high Unemployment (C1 and C2 in Figure 2*b*) and low and high Income (C1 and C2 in Figure 2*c*). Across all the cluster results for the considered condition and outcome variables, importantly, C2 is associated with the high derivative. In general, letting μ_1 and μ_2 be the grades of membership associating a US state to the C1 and C2 clusters, respectively, found using FCM, it follows with two clusters, $\mu_1 + \mu_2 = 1$, so $\mu_1 = 1 - \mu_2$. Hence, the μ_2 value can be used for each area to measure the grade of membership to 'high outcome', and $1 - \mu_2$ for grades of membership to 'low outcome'.

Description of variable fuzzy membership scores

Based on the μ_2 grades of membership (to C2 clusters), each US state is now represented as a fuzzy membership score (grades of membership), appropriate for use in fsQCA. Figures 3 and 4 report the spreads of grades of membership values across the condition (Figure 3) and outcome (Figure 4) variables.

Figure 3. Probability density functions and spreads of grades of membership values (x-axis) for five condition variables

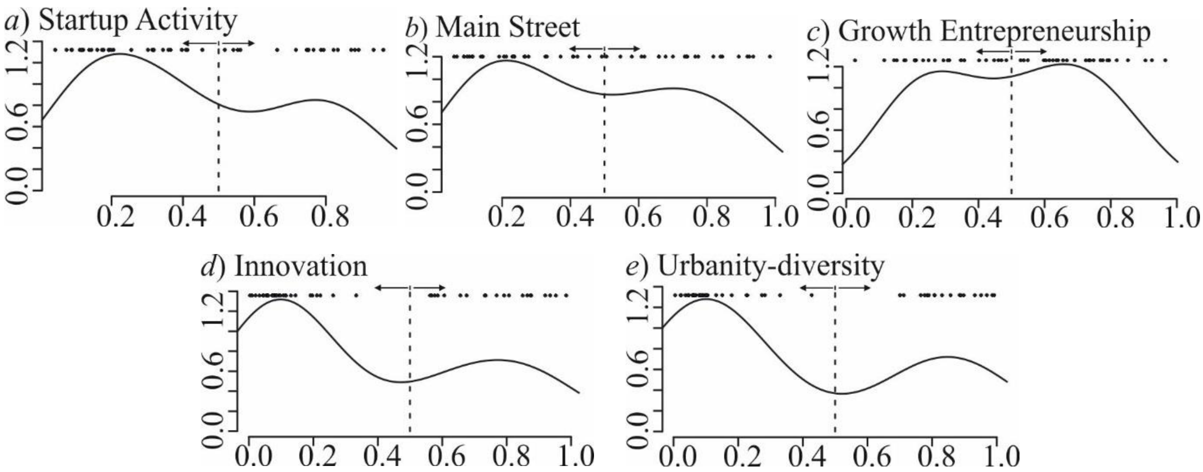
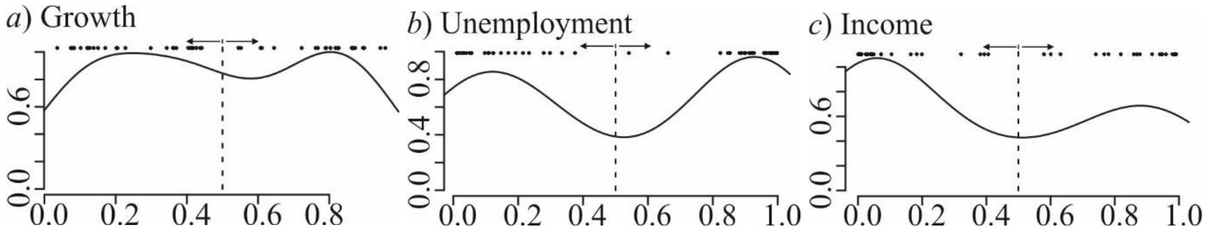


Figure 4. Probability density functions and spreads of grades of membership values (x-axis) for three outcome variables



In each graph in Figures 3 and 4, a probability density function (pdf) representing the spread of the previously created grades of membership values for each variable is shown, to elucidate their distributions across the respective 0.0 – 1.0 domains. Above each pdf are the specific values the 50 US states take for a variable; furthermore, the arrows near the base of each graph show the number of US states above and below the 0.5 grades of membership value (of note in the fsQCA undertaken – in terms of strong membership terms – see Ragin, 2008). As noted, this is a novel form of calibration of variables to a form of variables appropriate for use in fsQCA.

4. FsQCA analyses

This section evaluates the fsQCA analyses of the constructed US state-level data set (using the fuzzy membership scores values found previously). The analyses are deconstructed into three parts: *i*) necessity analyses of all the condition variables across the different outcome variables, *ii*) elucidation of truth table and concomitant frequency and consistency thresholds used in fsQCA, and *iii*) sufficiency analyses across the different outcome variables in the creation of concomitant causal recipes.

Necessity Analyses

The necessity analyses undertaken here are to investigate if there exists necessary conditions in relation to an outcome (Ragin, 2008) in relation to the respective membership scores of a condition variable (X_i), and outcome variable (Y_i), and a measure is employed (measuring the level that the respective Y_i values are less than the respective X_i values) (Table 2). The analyses are performed with respect to the presence and absence (~) of the condition and outcome variables.

Table 2. Analyses of necessity results for Growth (GRTH and ~GRTH), Unemployment (UNPT and ~UNPT), and Income (INCM and ~INCM) (Cons - Consistency and Cov - Coverage)

Variable	Var	Growth				Unemployment				Income			
		GRTH		~GRTH		UNPT		~UNPT		INCM		~INCM	
		Cons	Cov	Cons	Cov	Cons	Cov	Cons	Cov	Cons	Cov	Cons	Cov
Start-up Activity	SA	0.671	0.752	0.557	0.642	0.516	0.628	0.546	0.577	0.561	0.507	0.515	0.706
	~SA	0.680	0.599	0.784	0.710	0.652	0.623	0.648	0.538	0.674	0.479	0.640	0.688
Main Street	MS	0.606	0.651	0.639	0.705	0.425	0.496	0.703	0.711	0.626	0.543	0.482	0.632
	~MS	0.725	0.661	0.684	0.641	0.753	0.745	0.502	0.431	0.575	0.423	0.651	0.725
Growth Entrepreneurship	GE	0.726	0.719	0.648	0.659	0.576	0.619	0.599	0.559	0.612	0.488	0.576	0.696
	~GE	0.656	0.644	0.723	0.731	0.589	0.629	0.591	0.548	0.619	0.491	0.576	0.692
Innovation	In	0.516	0.689	0.451	0.619	0.430	0.623	0.450	0.566	0.652	0.701	0.293	0.478
	~In	0.714	0.559	0.773	0.622	0.701	0.595	0.700	0.516	0.515	0.325	0.817	0.780
Urbanity-diversity	Ud	0.542	0.700	0.443	0.587	0.501	0.701	0.384	0.467	0.631	0.656	0.324	0.511
	~Ud	0.680	0.543	0.774	0.635	0.619	0.537	0.754	0.567	0.530	0.341	0.782	0.762
Stats	Min	0.516	0.543	0.443	0.587	0.425	0.496	0.384	0.431	0.515	0.325	0.293	0.478
	Max	0.726	0.752	0.784	0.731	0.753	0.745	0.703	0.711	0.674	0.701	0.817	0.780

The results in Table 2 show the necessity measures, Consistency (Cons), and Coverage (Cov), of each condition variable (the Var and ~Var versions) against each outcome variable (absence and presence of versions). Inspection of the bottom Stats rows shows the min and max of the values for each outcome variable; importantly, no value goes above 0.817 in terms of the consistency values, less that the often-employed threshold of approximately 0.9 (see Ragin, 2008; Young and Park, 2013), which would signify a necessary variable. Hence, no condition variable derivative was considered necessary for an outcome variable derivative in this study.

Truth table

With combinations of condition variables now the focus of investigation (no one condition variable deemed necessary), the associated truth tables formed from the five condition variables to each outcome are next explored; see Table 3. The truth table, which is partly a list of the logically possible combinations of conditions, is used to synthesize the results of fuzzy-set analyses of the logically possible configurations of a given set of conditions (Ragin, 2008).

For a detailed description of the details of a truth table, such as that in Table 3, see Ragin (2008). Each row of the truth table describes a logically possible configuration of condition variables (combinations of 0 s and 1 s under the condition variable columns – with five condition variables there are $2^5 = 32$ possible configurations). The number column notes the number of US states with ‘strong membership’-based association to the respective configurations (following Greckhamer, 2015, this is determined by assigning a 1 to cases with a set membership score >0.5 and a 0 to cases with a set membership score <0.5), with the US state column listing the actual states with this association. The remaining columns provide the relevant consistency values associating each configuration to the presence of and absence of derivatives of each outcome variable. Similar to the necessity-based analyses undertaken previously, the consistency scores denote the degree to which cases sharing a given condition or combination of conditions agree in displaying the considered outcome.

Thereafter, the first of two threshold terms must be considered, namely, frequency threshold, particular to the number of US states associated with a configuration for it to be considered in the later sufficiency analyses. With 32 possible configurations to consider, and the spread of the States across them, it was decided that at least two US states were necessary to be associated with a configuration for it to be further considered (across all three outcome variables). Thus, in Table 3, those configurations with zero or one US state associated with them are struck through, indicating not further considered (more specifically, considered as remainders later; see Ragin, 2008).

Table 3. Truth Table: All configurations existing from data, based on five condition variables

Config	Condition Variables					Number	US State	Raw Consistency Values					
								GRTH		UNPT		INCM	
	Start-up Activity	Main Street	Growth Entrepreneurship	Innovation	Urbanity-diversity			Presence	Absence	Presence	Absence	Presence	Absence
1	0	0	0	0	0	3	Arkansas, Kentucky, Mississippi	0.7574	0.8889	**0.7862**	0.6017	0.4232	**0.8658**
~~2~~	~~0~~	~~0~~	~~0~~	~~0~~	~~1~~	~~0~~		~~0.9232~~	~~0.8804~~	~~0.8469~~	~~0.6511~~	~~0.6661~~	~~0.8081~~
~~3~~	~~0~~	~~0~~	~~0~~	~~1~~	~~0~~	~~1~~	Michigan	~~0.9452~~	~~0.8817~~	~~0.8307~~	~~0.6795~~	~~0.5552~~	~~0.8931~~
4	0	0	0	1	1	2	Illinois, Washington	0.8697	0.892	**0.8856**	0.5916	0.7473	0.6812
5	0	0	1	0	0	4	Alabama, South Carolina, Tennessee, West Virginia	0.7896	0.8115	**0.8379**	0.5292	0.3930	**0.8898**
~~6~~	~~0~~	~~0~~	~~1~~	~~0~~	~~1~~	~~0~~		~~0.8928~~	~~0.8816~~	~~0.8528~~	~~0.6415~~	~~0.6932~~	~~0.7857~~
7	0	0	1	1	0	2	Delaware, New Mexico	**0.9285**	0.9034	**0.8077**	0.7169	0.6044	**0.8947**
8	0	0	1	1	1	2	Rhode Island, Virginia	0.8243	**0.9294**	**0.8295**	0.6712	**0.7748**	0.6898
9	0	1	0	0	0	5	Indiana, Iowa, Maine, Nebraska, Wisconsin	0.7592	0.8941	0.6079	**0.7678**	0.4739	**0.8790**
~~10~~	~~0~~	~~1~~	~~0~~	~~0~~	~~1~~	~~0~~		~~0.9219~~	~~0.9008~~	~~0.8099~~	~~0.6945~~	~~0.7027~~	~~0.8092~~
~~11~~	~~0~~	~~1~~	~~0~~	~~1~~	~~0~~	~~1~~	Vermont	~~0.8788~~	~~0.8874~~	~~0.6643~~	~~0.7478~~	~~0.6671~~	~~0.7695~~
12	0	1	0	1	1	2	Connecticut, Oregon	0.8521	0.8870	**0.8578**	0.5802	0.7180	0.7052
13	0	1	1	0	0	4	Kansas, Louisiana, Ohio, Pennsylvania	0.7986	**0.9066**	0.7182	0.6926	0.5355	**0.8574**
~~14~~	~~0~~	~~1~~	~~1~~	~~0~~	~~1~~	~~0~~		~~0.8995~~	~~0.8949~~	~~0.8113~~	~~0.6870~~	~~0.7236~~	~~0.7885~~
15	0	1	1	1	0	2	Minnesota, New Hampshire	**0.9365**	0.8900	0.6529	**0.8035**	**0.7555**	0.7418
16	0	1	1	1	1	2	Maryland, Massachusetts	0.8513	0.8527	0.7283	0.7232	**0.8120**	0.6205
17	1	0	0	0	0	3	Alaska, Georgia, Missouri	0.8306	0.8922	**0.7353**	0.6891	0.5161	**0.8035**
~~18~~	~~1~~	~~0~~	~~0~~	~~0~~	~~1~~	~~1~~	Florida	~~0.9016~~	~~0.8198~~	~~0.8462~~	~~0.5920~~	~~0.5985~~	~~0.8301~~
~~19~~	~~1~~	~~0~~	~~0~~	~~1~~	~~0~~	~~1~~	Idaho	~~0.9149~~	~~0.8860~~	~~0.7677~~	~~0.7478~~	~~0.5866~~	~~0.8626~~
20	1	0	0	1	1	3	California, New Jersey, New York	0.7859	0.8480	**0.8647**	0.5167	**0.7765**	0.5839
~~21~~	~~1~~	~~0~~	~~1~~	~~0~~	~~0~~	~~1~~	North Carolina	~~0.8973~~	~~0.8529~~	~~0.8100~~	~~0.6253~~	~~0.5115~~	~~0.8310~~
22	1	0	1	0	1	3	Arizona, Nevada, Texas	0.8841	0.7778	**0.7842**	0.6210	0.5555	**0.8535**
~~23~~	~~1~~	~~0~~	~~1~~	~~1~~	~~0~~	~~0~~		~~0.9307~~	~~0.9266~~	~~0.8225~~	~~0.7341~~	~~0.6754~~	~~0.8570~~
24	1	0	1	1	1	2	Colorado, Utah	**0.9115**	0.8107	**0.7611**	0.6887	0.7222	0.7260
25	1	1	0	0	0	3	Montana, South Dakota, Wyoming	0.8944	0.8867	0.6177	**0.8242**	0.5694	**0.8319**
~~26~~	~~1~~	~~1~~	~~0~~	~~0~~	~~1~~	~~0~~		~~0.9351~~	~~0.9193~~	~~0.8022~~	~~0.7108~~	~~0.6780~~	~~0.8569~~
~~27~~	~~1~~	~~1~~	~~0~~	~~1~~	~~0~~	~~0~~		~~0.9103~~	~~0.9161~~	~~0.7655~~	~~0.7759~~	~~0.6430~~	~~0.8663~~
~~28~~	~~1~~	~~1~~	~~0~~	~~1~~	~~1~~	~~0~~		~~0.9262~~	~~0.9189~~	~~0.8262~~	~~0.6918~~	~~0.7141~~	~~0.7874~~
29	1	1	1	0	0	2	North Dakota, Oklahoma	0.8781	**0.9386**	0.6763	**0.7827**	0.5992	**0.8216**
~~30~~	~~1~~	~~1~~	~~1~~	~~0~~	~~1~~	~~1~~	Hawaii	~~0.9245~~	~~0.9101~~	~~0.7969~~	~~0.7068~~	~~0.6954~~	~~0.8441~~
~~31~~	~~1~~	~~1~~	~~1~~	~~1~~	~~0~~	~~0~~		~~0.9378~~	~~0.9521~~	~~0.8046~~	~~0.7876~~	~~0.7162~~	~~0.8536~~
~~32~~	~~1~~	~~1~~	~~1~~	~~1~~	~~1~~	~~0~~		~~0.9112~~	~~0.9172~~	~~0.8083~~	~~0.7208~~	~~0.7653~~	~~0.7576~~
Frequency Threshold								> 1		> 1		> 1	
Consistency Threshold								> 0.904		> 0.750		> 0.750	
Number of configurations (US states)								(3 (6), 3 (8)) [6 (14)]		(9 (23), 4 (12)) [13 (35)]		(4 (9), 9 (31)) [14 (40)]	

Consideration then turns to the consistency threshold to adopt, over which configurations may be associated with either the presence or absence of derivatives of an outcome. The different outcome variables were considered separately on this issue, and the criteria expressed in Andrews et al. (2016) and Beynon et al. (2016a) initially adopted here, namely, the lowest possible consistency threshold to adopt while not allowing the same configuration to be then associated with both the presence and absence of derivative forms of an outcome, with a further check for when an evaluated threshold value is below the often-employed minimum standard of 0.750; in such cases the consistency threshold is raised to this 0.750 value. The respective consistency thresholds for each outcome are (see Table 3), namely, 0.904 for Growth, 0.750 for Unemployment, and 0.750 for Income (for Unemployment and Income these values have been found from raising the initially identified values). The last row displays the number of configurations, based on consistency thresholds, associated with each presence or absence outcome derivative (as well as subsequent number of US states). As noted, this phase can also mean for an outcome variable, and there may be further configurations now considered remainders, since they lack one of the two concomitant consistency values above the concomitant consistency threshold. The explained details in the adapted truth table in Table 3 (for all three outcome variables) are next considered in terms of relevant sufficiency analyses.

Sufficiency analyses

In providing a rigorous analyses and presentation of the main results under discussion, the sufficiency analyses undertaken here first considers those conditions, or sets of conditions, which lead to the outcome (but a set of sufficient conditions may not be the only conditions that lead to the outcome). For each outcome variable, the sufficiency analyses is exposited using an amended form of the Ragin and Fiss (2008) circle notation, employed in Andrews et al. (2016). This notation system highlights the thinking on how to consider those configurations identified as remainders (see discussion around Table 3).

It follows that the analyses distinguishes between complex (discerning between those configurations considered in the truth table – not remainders) and parsimonious solutions (discerning between those in the truth table – including possible remainders); see Ragin (2008), with the elucidation of both these solutions advocated by Wagemann and Schneider (2010). Only discerning against those other configurations still considered in Table 3 is termed the complex solution (avoid using any remainders to simplify the truth table), while incorporating the remainders that yields the most parsimonious solutions is termed the parsimonious solution.

Sufficiency analyses are presented for each of the three outcome variables, Growth, Unemployment, and Income. Furthermore, maps are given showing collections of US states associated with causal recipes from the parsimonious solutions (limited space restricted full discussion of some complex solutions: causal recipes). Emphasis is given to causal recipes with associated consistency equal to or above 0.750, with details of others given for completeness.

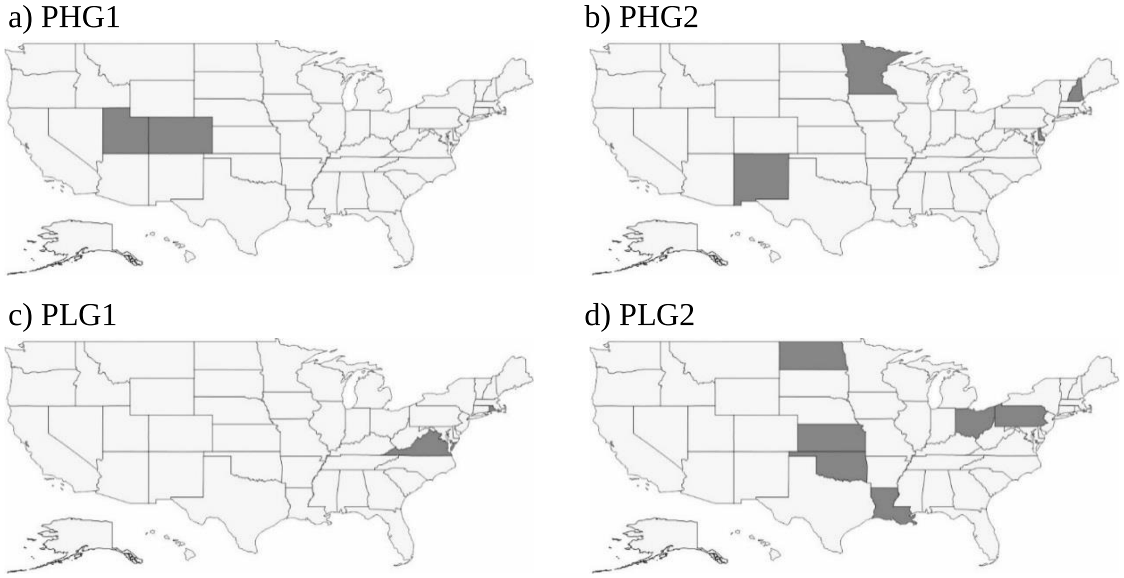
Growth

Sufficiency analyses for the Growth presence of (GRTH) and absence of (~GRTH) outcome variables are shown in Table 4 (and maps shown in Figure 5).

Table 4. Sufficiency Analyses Results for Growth Outcomes

Conditions	Growth			
	GRTH		~GRTH	
Start-up Activity	●	⊝	⊖	
Main Street	⊝		⊖	●
Growth Entrepreneurship	●	•	●	●
Innovation	●	●	•	⊖
Urbanity-diversity	•	⊖	●	⊝
Complex	**CHG1**	**CHG2**	**CLG1**	**CLG2**
Configurations	24	7, 15	8	13,29
Consistency	0.911	0.923	0.929	0.895
Raw Coverage	0.239	0.264	0.224	0.386
Unique Coverage	0.074	0.100	0.081	0.243
Solution Consistency	0.889		0.897	
Solution Coverage	0.339		0.467	
Parsimonious	**PHG1**	**PHG2**	**PLG1**	**PLG2**
Configurations	24	7, 15	8	13, 29
Consistency	0.900	0.820	0.876	0.878
Raw Coverage	0.278	0.330	0.266	0.449
Unique Coverage	0.074	0.125	0.041	0.224
Solution Consistency	0.822		0.867	
Solution Coverage	0.403		0.490	

Figure 5. Maps showing US states associated with the identified Parsimonious solution-based causal recipes for Growth



The parsimonious and complex solutions for presence and absence of Growth indicate equifinality and asymmetry as well as conjunctional cauzation in the relationships. Specifically, CHG1 (covering Colorado, and Utah) shows that US state growth is present where there is the presence of Start-up, Growth Entrepreneurship, Innovation and Urbanity-diversity, and the absence of Main Street (suggesting strong competitive effects), this recipe in many ways showing an “ideal” scenario for entrepreneurial processes and their impact on growth. However, CLG1 (Rhode Island, and Virginia) notes State growth being absent where Urbanity-diversity is present even where Entrepreneurial Growth (core) and Innovation are also present, if there is also the (core) absence of Start-up and Main Street activity. This suggests that Start-up Activity is key to growth in high-entrepreneurial, high-growth urban areas, where strong competitive effects lead to an absence of Main Street entrepreneurship, because of the way in which such Start-up Activity influences competition.

CHG2 (Delaware, New Mexico, Minnesota, and New Hampshire) highlights that the presence of State growth is also possible where Urbanity-diversity is absent, where Innovation (core) and Entrepreneurial growth is also present but start-up is absent. This may suggest that where Urbanity-diversity and the competitive effects related to it are absent, it is a focus on quality of firms in terms of high-growth innovative businesses, rather than quantity in terms of start-up, that is of even more importance.

In addition, CLG2 (Kansas, Louisiana, North Dakota, Ohio, Oklahoma, and Pennsylvania) illustrates that State growth can be absent where Urbanity-diversity is absent, even where Growth Entrepreneurship is present, where Innovation is absent and Main Street is present (suggesting weak competitive effects). This suggests that where there is an absence of Urbanity-diversity, innovative activity per se is potentially more important in driving overall growth than Growth Entrepreneurship that may thus manifest itself in terms of firms numbers (presence of Main Street), rather than presence of growth more generally.

Overall, this indicates that Growth Entrepreneurship and Innovation are necessary, but not always sufficient, to obtain economic growth, and additional supporting entrepreneurial activities and processes also necessary, differing depending on the economic geography of the particular States in question. This illustrates the complex relationships among entrepreneurship, innovation, and State Urbanity-diversity and growth, indicated by the relatively low numbers of States encompassed by the analyses. The analyses, therefore, only partially supports previous literature (Wennekers and Thurik, 1999), particularly for Start-up and Growth Entrepreneurship being relevant to economic growth when linked to Innovation activities in these processes. There is specific support for high start-up entrepreneurship, in combination with Growth Entrepreneurship, and Innovation, within regions displaying stronger aspects of Urbanity-diversity, being more likely to experience the presence of growth. The evidence also suggests that where Startup is absent from this combination, growth is also absent, but that more rural areas experience very different policy combinations to achieve the same outcomes. Geographical context, therefore, including through its relationship with the presence or absence of Main Street entrepreneurship, through processes of competition also appears to be of importance in determining whether growth is present or absent. The results showing that geography per se does not determine the presence or absence of growth is also important, suggesting that (albeit different) entrepreneurship-innovation-related policy combinations may be relevant.

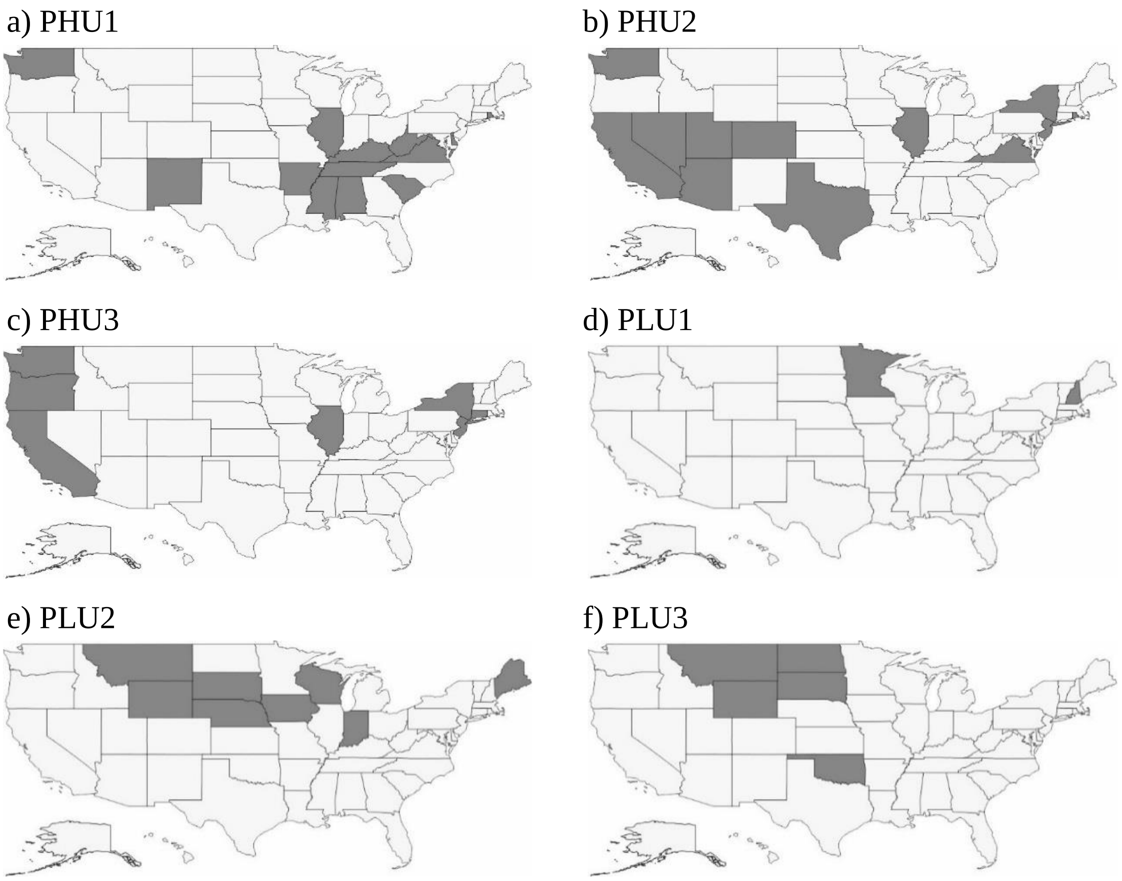
Unemployment

Sufficiency analyses for the Unemployment presence of (UNPT) and absence of (~UNPT) derivative outcome variables are shown in Table 5 (maps in Figure 6).

Table 5. Sufficiency Analyses Results for Unemployment Outcomes

Conditions	Unemployment							
	UNPT					~UNPT		
Start-up Activity	⊖	⊖	•		⊝	⊝		⬤
Main Street	⊖	⊖	⊖	⊖		⬤	⬤	⬤
Growth Entrepreneurship		•	•		⊖	•	⊖	
Innovation	⊝			•	•	⬤	⊝	⊝
Urbanity-diversity	⊝	⊝	⬤	⬤	⬤	⊖	⊖	⊝
Complex	**CHU1**	**CHU2**	**CHU3**	**CHU4**	**CHU5**	**CLU1**	**CLU2**	**CLU3**
Configurations	1	5, 7	22, 24	4, 8, 20, 24	4, 12	15	9, 25	25, 29
Consistency	0.814	0.818	**0.730**	0.807	0.874	0.804	0.796	0.812
Raw Coverage	0.408	0.371	0.255	0.304	0.233	0.214	0.409	0.343
Unique Coverage	0.047	0.006	0.040	0.077	0.042	0.053	0.090	0.028
Solution Consistency	0.772					0.795		
Solution Coverage	0.668					0.498		
Parsimonious	**PHU1**		**PHU2**		**PHU3**	**PLU1**	**PLU2**	**PLU3**
Configurations	1, 4, 5, 7, 8		4, 8, 20, 22, 24		4, 12, 20	15	9, 25	25, 29
Consistency	0.798		0.784		0.852	0.787	0.774	0.751
Raw Coverage	0.412		0.412		0.384	0.279	0.467	0.401
Unique Coverage	0.383		0.044		0.045	0.039	0.090	0.073
Solution Consistency	0.774					0.756		
Solution Coverage	0.714					0.587		

Figure 6. Maps showing US states associated with the identified Parsimonious solution-based causal recipes for Unemployment



For unemployment, a larger number of US states were considered in the final analyses. Specifically, supporting the work of Wennekers and Thurik (1999), Main Street was consistently present for absence of unemployment causal recipes and absent in four of the five presence of unemployment causal recipes, suggesting that it a key ingredient in the recipes, though in conjunction with different sets of factors depending on whether urbanity-diversity was present or absent. Urbanity-diversity was also consistently absent in the absence of unemployment causal recipes, while being present in three of the five presence of unemployment complex causal recipes.

Combining these two results suggests that it is states where Urbanity-diversity is present where unemployment is concentrated (although not exclusively so) and where policies to support Main Street activity may potentially be beneficial (as indicated by the first set of parsimonious solutions). However, there is also diversity in the recipes where urbanity-diversity is present. Where Main Street is absent, the presence of other entrepreneurship-innovation processes (Acs et al., 2007) may be feeding into competition (and potentially in some cases broader economic growth) rather than reducing unemployment. Where Main Street is neither absent nor present in the recipe, however, the Innovation activity taking place is doing so in the absence of Start-up Activity, suggesting it is this lack of complementarity that is contributing to the presence of unemployment (in the case of the two configurations covered by this recipe, neither growth nor income outcomes reaching the required level of consistency to be included in the analyses of these outcomes).

For more rural areas, where there is also the presence of high unemployment, this was consistently associated with the absence of Main Street entrepreneurship. In one recipe (CHU2), this was also linked to the absence of Start-up Activity and Innovation, in the case of the configuration covered by this recipe suggesting the near-complete absence of entrepreneurship as well as Innovation activity, suggesting a relative lack of resources for these activities that may also be detrimental to income levels. In the other recipe, the absence of Main Street is in conjunction with the absence of Start-up Activity but also the presence of Growth Entrepreneurship, potentially suggesting a focus on a narrow range of firms that may have improved access to the financial resources available (because of the lack of competition indicated by the absence of start-up activity), leading to more capital-intensive Growth Entrepreneurship that is not beneficial to reducing unemployment or income levels, though likely more beneficial to broader growth.

In terms of unemployment specifically, therefore, the evidence suggests both that combinations of firm survival with one or more of Start-up, Innovation, and Urbanity-diversity are particularly relevant but also that, specifically because of the geographical effects of concentration (where concentration will be more associated with urban areas), different combinations will likely be required in different geographies to generate unemployment benefits. It was not possible to specifically explore the relationship identified in the literature that in more urban areas an absence of unemployment is more likely where there are high firm survival rates in areas without high innovation because an absence of unemployment was only related to recipes where there was an absence of Urbanity-diversity.

In more rural areas, an absence of unemployment was linked to firm survival (i.e., presence of Main Street entrepreneurship) in conjunction with either high firm Start-up or Innovation, in two of the recipes. In the other absence of unemployment recipe, however, it is linked to the presence of Main Street with an absence of Innovation and absence of Growth Entrepreneurship, suggesting a relative stagnancy in these economies that may be suggestive of a trade-off with income levels.

More broadly, the evidence suggests that different sets of entrepreneurship-innovation conditions in combination with different geographies, possibly because of their different effects on competition and resources, may be leading to different complementary/substituting relationships among growth, unemployment, and income outcomes. This implies both that different entrepreneurial processes (and hence policies) may be required to support presence of growth (linked to Entrepreneurial growth and Innovation) and absence of unemployment (linked to the presence of Main Street) outcomes but also that policies that produce complementary outcomes for growth and unemployment for one economic geography may generate substituting outcomes in another.

<u>Income</u>

Sufficiency analyses for the Income presence of (INCM) and absence of (~INCM) derivative outcome variables are shown in Table 6 (maps in Figure 7).

Table 6. Sufficiency Analyses Results for Income Outcomes

Conditions	Income						
	INCM			~INCM			
Start-up Activity	●	⊖	⊝	⊝	•		
Main Street	⊝		●	⊖	⊝		•
Growth Entrepreneurship	⊖	●	●	•	•	⊝	
Innovation	•	•	●		⊖	⊖	⊖
Urbanity-diversity	●	●		⊖	•	⊝	⊝
Complex	**CHI1**	**CHI2**	**CHI3**	**CLI1**	**CLI2**	**CLI3**	**CLI4**
Configurations	20	8, 16	15, 16	5, 7	22	1, 9, 17, 25	9, 13, 25, 29
Consistency	0.777	0.817	0.806	0.892	0.853	0.835	0.809
Raw Coverage	0.274	0.310	0.342	0.360	0.227	0.438	0.396
Unique Coverage	0.108	0.033	0.071	0.074	0.084	0.051	0.041
Solution Consistency	0.820			0.831			
Solution Coverage	0.489			0.659			
Parsimonious	**PHI1**	**PHI2**	**PHI3**	**PLI1**	**PLI2**		
Configurations	20	8, 16	15, 16	5, 7	1, 9, 13, 17, 22, 25, 29		
Consistency	**0.678**	0.770	0.807	0.870	0.780		
Raw Coverage	0.328	0.365	0.367	0.526	0.817		
Unique Coverage	0.099	0.053	0.077	0.053	0.344		
Solution Consistency	0.750			0.789			
Solution Coverage	0.560			0.870			

For Income, 40 states were covered by the analyses. Innovation was the most consistent in its presence or absence (in the complex solutions at least), being present in all the presence of income causal-recipes and absent in three of the four absence of income causal recipes. Economic geography could also be seen to play a role here (again for the complex recipes), with Urbanity-diversity present in two of the three presence of income causal recipes and absent in three of the four absence of income causal recipes (though being present in the other).

The literature review identified that innovation is central to high-income outcomes generally, the results here supporting this, as well as the presence of income outcomes being much more prevalent in urban geographies. The review also suggested, however, that the role of innovation will be in combination with different variables dependent on the economic geography, in more urbanized geographies, with the combination of high-growth entrepreneurship and innovation most likely to be linked to high-income outcomes, while in more rural economies high-income outcomes also requiring building a stock of surviving, innovation-driven growth firms. The results broadly support this but also identify more specifically the complexity in these relationships than is implied in the literature.

Figure 7. Maps showing US states associated with the identified Parsimonious solution-based causal recipes for Income

a) PHI1

b) PHI2

c) PHI3

d) PLI1

e) PLI2

In terms of the absence of income, the absence (or at least non-presence) of Innovation does appear to be of central relevance for both urban and rural geographies, albeit with different combinations of relevance to urban and rural areas, as indicated by PLI1 and PLI2. In terms of presence of income, the recipe CHI2 also broadly supports the existing literature. Here, the presence of income is related to the presence of Urbanity-diversity (encompassing Maryland, Massachusetts, Rhode Island, and Virginia) with (albeit non-core) Innovation, presence of Growth Entrepreneurship and core absence of Start-up. In this case, absence of Start-up Entrepreneurship potentially suggests that Urbanity-Diversity reduces initial start-up rates (in line with Acs et al., 2007), but that the deleterious effects of concentration on high-innovation entrepreneurship through imitation are outweighed by benefits in terms of supporting Growth Entrepreneurship. For CHI1, however, while the presence of income causal-recipe with that of Urbanity-diversity (California, New Jersey, and New York) is associated with presence of Innovation, this is not core, and along with the absence of Main Street, there is also the core absence of Growth Entrepreneurship, with the (non-core) presence of Start-up Activity. For the states covered by this recipe, therefore, broader (start-up) processes, rather than those related to high-Growth Entrepreneurship specifically (as measured by the variables here at any rate) appear to be of more importance in driving high-income levels. This may, however, still be consistent with the geographical concentration effect for innovation from knowledge spillover effects (Acs and Varga, 2002), where this greater geographical concentration leads to lower innovative new-firm survival (and firm growth) rates because of greater short-lived, imitative, businesses being created.

For CHI3, the third presence of income causal-recipe, Urbanity-diversity is neither absent nor present, the recipe including presence of Growth Entrepreneurship and Innovation, as well as presence of Main Street and (non-core) absence of Start-up. This result is consistent with the literature related to more rural national economies, this recipe including configurations covering (absent Urbanity-diversity) Minnesota, and New Hampshire (which benefit from presence of growth and absence of unemployment) but also with (present Urbanity-diversity) Maryland, and Massachusetts, who are associated with presence of unemployment (with outcomes associated with both presence and absence of growth above the threshold level and therefore not included in the analyses).

<u>Summary</u>

Overall, this might suggest that for non-urban-diverse states, such as Minnesota and New Hampshire, the presence of Innovation, Growth Entrepreneurship and Main Street, and relative absence of Start-up Activity (itself indicating a quantity versus quality issue for entrepreneurship policy), is of most relevance in supporting positive growth, unemployment and income outcomes, that do not appear to need to be traded off against one another. For more urban-diverse states, however, this combination of factors is less effective in generating positive outcomes across all three outcomes. Specifically, there are clear trade-offs in growth, unemployment, and income outcomes in more complex urban-diverse economies, likely because of the ways in which competition (the outcomes of which are indicated through Main Street entrepreneurship) is driven by geographical concentration interacting with Innovation in combination with Start-up and Growth Entrepreneurship activities.

5. Conclusions

This study offers a novel theoretical contribution to further the understanding of the complexity of the effects on growth, unemployment, and income of interactions among innovation and different entrepreneurial behaviors within the different economic geographies represented by individual US states. Specifically, it broadens the debate on the impact of entrepreneurship in countries in different states of development, to explore these issues within a highly developed economy, while also indicating that it is combinations of innovation and specific entrepreneurship activities within specific geographies that create growth, unemployment, and income outcomes. There are also trade-offs among these outcomes depending on the specific economic geography in question.

FsQCA has allowed this more nuanced understanding of these complex relationships linking entrepreneurship, innovation, and geography (here urbanity-diversity) with growth, unemployment, and income outcomes, at the US state-level. An additional, technical development in this study, which may also assist in such future research, has been the employment of fuzzy clustering to formulate the considered condition and outcome variables, and importantly the concomitant fuzzy membership scores required. Without loss of generality, three constituent-variables are considered to model each condition and outcome variable (less or more could be included), and a two-step approach employed to identify two clusters of cases across each set of constituent-variables. Using the evaluated grades of membership values for the relevant cluster, they are in the correct scale for use in fsQCA (fuzzy membership scores). It is worth noting that this approach has done away with previously considered awkward efforts required for such calibration processes, such as using the direct method approach (Ragin, 2008).

The study also offers implications for both policy and practice. In terms of policy, the study offers novel insights for US state-level decision-makers in allowing the identification of the combinations of specific types of entrepreneurial activity and innovation activity that are required to promote growth, unemployment, and income outcomes, in specific geographies. In addition, it clearly indicates the configurations of US states where there is also a requirement to understand the trade-offs among growth, unemployment, and income that may result from such policies. Thus, decision-makers can use this evidence to mirror effective practice within benchmark states toward improved entrepreneurial practice and seek to reduce the negative impacts of unemployment, low growth, and/or low income, depending on their policy priorities. If such a strategic approach could be implemented, the small business community would benefit in turn through increased, more specifically focused support and legislation, which should thereby increase firm efficiency and practices.

The findings suggest there is no “one size fits all” policy that can be suggested for the use of entrepreneurial activities to promote beneficial growth, unemployment, and income outcomes, particularly for more urban-diverse US states, where there is a clear requirement for further research. This study also recognizes, however, its limitations in terms of drawing meaningful conclusions from a single, essentially cross-sectional study. Generally, a longitudinal study is required, as well as further in-depth research, to analyzes in more detail effective entrepreneurial behavior within individual US states to identify best practices, given their diversity and the requirement to further inform policy. There is also a need to further explore the issue of effective entrepreneurial behavior within urban-diverse US states.

References

Ács, Z. (2008). Foundations of high impact entrepreneurship. *Foundations and Trends in Entrepreneurship*, 4, 535-620.

Ács, Z., & Varga, A. (2002). Geography, endogenous growth and innovation. *International Regional Science Review*, 25(1), 132-148.

Ács, Z., Armington, C., & Zhang, T. (2007). The determinants of new-firm survival across regional economies: The role of human capital stock and knowledge spillover. *Regional Science*, 86(3), 367-391.

Ács, Z.J., Desai, S., & Hessels, J. (2008). Entrepreneurship, economic development and institutions. *Small Business Economics,* 31(3), 219-234.

Amenta, E., Caren, N., & Olasky, S.J. (2005). Age for leisure? Political mediation and the impact of the pension movement on US old-age policy. *American Sociological Review*, 70(3), 516-538.

Andrews, R., Beynon, M.J., & McDermott, A.M. (2016). Organizational capability in the public sector: a configurational approach. *Journal of Public Administration Research and Theory*, *26*(2), 239-258.

Audretsch, D.B., & Thurik, A.R. (2000). Capitalism and democracy in the 21st Century: from the managed to the entrepreneurial economy. *Journal of Evolutionary Economics*, 10(1–2), 17–34.

Audretsch, D.B., Keilbach, M.C., & Lehmann, E.E. (2006). *Entrepreneurship and Economic Growth*. Oxford: Oxford University Press.

Bacher, J., Wenzig, K., & Vogler, M. (2004). SPSS TwoStep Cluster – A First Evaluation, http://www.statisticalinnovations.com/products/twostep.pdf Accessed January 2018.

Baptista, R., & Preto, M.T. (2007). The dynamics of causality between entrepreneurship and unemployment. *International Journal of Technology, Policy and Management*, 7(3), 215-224.

Beynon, M.J., Jones, P., Pickernell, D., & Packham, G. (2015). A NCaRBS analyses of SME Intended Innovation: Learning about the Don't Knows. *Omega*, 59(Part A), 97–112.

Beynon, M.J., Jones, P., & Pickernell, D. (2016a). Country-based comparison analyses using fsQCA investigating entrepreneurial attitudes and activity. *Journal of Business Research*, 69(4), 1271-1276.

Beynon, M.J., Jones, P., & Pickernell, D. (2016b). Country-level investigation of innovation investment in manufacturing: Paired fsQCA of two models. *Journal of Business Research*, 69(11), 5401-5407.

Beynon, M.J., Jones, P., & Pickernell, D. (2018). Entrepreneurial Climate and self-perceptions about entrepreneurship: A country comparison using fsQCA with dual outcomes. *Journal of Business Research*, 89, 418-428.

Bezdek, J.C. (1980). A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms. *IEEE Transactions on Pattern Analyses & Machine Intelligence*, 2(1), 1-8.

Bureau of Economic Analyses (2018). https://www.bea.gov/ Accessed 7th February 2018.

Bureau of Labour Statistics (2018). https://www.bls.gov/cps/tables.htm Accessed 7th February 2018.

Carree, M.A., & Thurik, A.R. (2008). The lag structure of the impact of business ownership on economic performance in OECD countries. *Small Business Economics*, 30(1), 101–110.

Casson, M. (2010). Entrepreneurship, business culture and the theory of the firm. In Z. Acs, & D. Audretsch (Eds.), *Handbook of Entrepreneurship Research* (pp. 225-245). The Netherlands: Kluwer Academic Publishers.

Colombelli, A., Krafft, J., & Vivarelli, M. (2016). To be born is not enough: the key role of innovative start-ups. *Small Business Economics*, 47(2), 277–291.

Cooke, P. (2003). The regional innovation systems in Wales: Evolution or eclipse. In: P. Cooke, M. Heidenreich, & H. Braczyk (Eds.) *Regional Innovation Systems* (2nd Edn.), London: Routledge.

Cumming, D., Johan, S., & Zhang, M. (2014). The economic impact of entrepreneurship: comparing international datasets. *Corporate Governance: an International Review*, 22(2), 162-178.

Dejardin, M. (2011). Linking net entry to regional economic growth. *Small Business Economics*, 36(4), 443-460.

Fazio, C., Guzman, J., Murray, F., & Stern, S. (2016). A New View of the Skew: A Quantitative Assessment of the Quality of American Entrepreneurship. MIT Innovation Initiative, Laboratory for Innovation Science and Policy, Cambridge, MA. http://innovation.mit.edu/assets/A-New-View_Final-Report_5.4.16.pdf Accessed 5th March 2018.

Elliott, T. (2013). *Fuzzy set qualitative comparative analyses: an introduction*. Research Notes. Statistics Group: UCI.

Fiss, P.C., Sharapov, D., & Cronqvist, L. (2013). Opposites attract? Opportunities and challenges for integrating large-N QCA and econometric analyses. *Political Research Quarterly*, 66(1), 191-198.

Fritsch, M. & Mueller, P. (2004). Effects of new business formation on regional development over time. *Regional Studies*, 38(8), 961-975.

Glaeser, E.L., & Gottlieb, J.D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, 47(4), 983-1028.

Greckhamer, T. (2015). Qualitative Comparative Analysis. In G.B. Dagnino, and M.C. Cinici, (Eds.) *Research Methods for Strategic Management* (pp. 229-252). Milton Park, Abingdon: Routledge.

Hausman, A. (2005). Innovativeness among small businesses: theory and propositions for future research. *Industrial Marketing Management*, 34(2), 773-782.

Huggins, R., Prokop, D., & Thompson, P. (2017). Entrepreneurship and the determinants of firm survival within regions: human capital, growth motivation and locational conditions. *Entrepreneurship & Regional Development*, 29(3/4), 357-389.

Kauffman (2017). http://www.kauffman.org/kauffman-index/about/about

Accessed 10th March 2018.

Kim, Y., & Verweij, S. (2016). Two effective causal paths that explain the adoption of US state environmental justice policy. *Policy Sciences*, 49(4), 505-523.

Kraus, S., Ribeiro-Soriano, D., & Schüssler, M. (2018). Fuzzy-set qualitative comparative analysis (fsQCA) in entrepreneurship and innovation research–the rise of a method. *International Entrepreneurship and Management Journal*, 14(1), 15-33.

Lee, N. (2014). What holds back high-growth firms? Evidence from UK SMEs. *Small Business Economics*, 43(1), 183-195.

Li, M., Goetz, S., Partridge, M., & Fleming, D. (2016). Location determinants of high-growth firms. *Entrepreneurship & Regional Development*, 28(1-2), 97-125.

Mallon, M., Lanivich, S., & Klinger, R. (2018). Resource configurations for new family venture growth. *International Journal of Entrepreneurial Behavior & Research*. 24(2), 521-537.

Malchow-Møller, N., Schjerning, B., & Sørensen, A. (2011). Entrepreneurship, job creation and wage growth. *Small Business Economics*, 36(1), 15–32.

Minghao, L., Goetz, S.J., Partridge, M., & Fleming, D.A. (2016). Location determinants of high-growth firms. *Entrepreneurship & Regional Development*, 28(1-2), 97-125.

Mueller, P. (2007). Exploiting entrepreneurial opportunities: the impact of entrepreneurship on growth. *Small Business Economics*, 28(4), 355-362.

Phillips, B.D. & Kirchhoff, B.A. (1989). Formation, growth and survival; small firm dynamics in the US economy. *Small Business Economics*, 1(1), 65-74.

Ragin, C.C. (1987). *The comparative method: moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.

Ragin, C.C. (2000). *Fuzzy set social science*. Chicago, IL: University of Chicago Press.

Ragin, C.C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago, IL: University of Chicago Press. 2008.

Ragin, C.C., & Fiss, P. (2008). Net effects versus configurations: an empirical demonstration. In C.C. Ragin (Ed.), *Redesigning social inquiry: Fuzzy sets and beyond* (pp. 190–212). Chicago, IL: University of Chicago Press.

Rihoux, B., Rezsöhazy, I., & Bol, D. (2011). Qualitative comparative analyses (QCA) in public policy analyses: an extensive review. *German Policy Studies*, 7(3), 9-82.

Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analyses. *Computational and Applied Mathematics*, 20, 53–65.

Rupasingha, A. (2017). Local business ownership and local economic performance: evidence from US counties. *Regional Studies*, 51(5), 659-673.

Santos, D. (2000). Innovation and Territory: Which Strategies to Promote Regional Innovation Systems in Portugal? *European Urban and Regional Studies*, 7(2), 147–157.

Şchiopu, D. (2010). Applying TwoStep cluster analyses for identifying bank customers' profile. *Bulletin*, 62(3), 66-75.

Shane, S. (2009). Why encouraging more people to become entrepreneurs is bad public policy? *Small Business Economics*, 33(2), 141-149.

SPSS Inc. (2001). The SPSS TwoStep Cluster Component: A scalable component enabling more efficient customer segmentation, Technical report, Chicago. https://www.spss.ch/upload/1122644952_The%20SPSS%20TwoStep%20Cluster%20Component.pdf Accessed 14th March 2018.

Thurik, A.R., Carree, M.A., van Stel, A., & Audretsch, D.B. (2008). Does self-employment reduce unemployment? *Journal of Business Venturing*, 23(6), 673–686.

Tominc, R., & Rebernik, M. (2007). Growth aspirations and cultural support for entrepreneurship: A comparison of post-socialist countries. *Small Business Economics*, 28(2/3), 239–255.

United States Census (2010). https://www.census.gov/2010census/data/apportionment-dens-text.php Accessed 7th February 2018.

United States Census Bureau (2018). https://www.census.gov/econ/geography.html Accessed 7th February 2018).

United States Patent Office (2018). https://www.uspto.gov/web/offices/ac/ido/oeip/taf/cst_utl.htm Accessed 7th February 2018.

Valliere, D., & Peterson, R. (2009). Entrepreneurship and economic growth: evidence from emerging and developed countries. *Entrepreneurship & Regional Development*, 21(5-6), 459-480.

Van Stel, A., Carree, M., & Thurik, R. (2005). The effect of entrepreneurial activity on national economic growth. *Small Business Economics*, 24(3), 311-321.

Wagemann, C., & Schneider, C.Q. (2010). Qualitative comparative analyses (QCA) and fuzzy-sets: agenda for a research approach and a data analyses technique. *Comparative Sociology*, 9(3), 376-396.

Wennekers, S., & Thurik, R. (1999). Linking entrepreneurship and economic growth. *Small Business Economics*, 13(1), 27-56.

Wong, P.K., Ho, Y.P., & Autio, E. (2005). Entrepreneurship, innovation and economic growth: evidence from GEM data. *Small Business Economics*, 24(3), 335-350.

Young, K.L., & Park, S.H. (2013). Regulatory opportunism: cross-national patterns in national banking regulatory responses following the global financial crisis. *Public Administration*, 91(3), 561-581.

Zhou, H.B., & Gao, J.T. (2014). Automatic method for determining cluster number based on silhouette coefficient. *Advanced Materials Research*, 951, 227-230.