Socio-cultural Diversity and Urban Buzz

Daniel Arribas-Bel[1]

Karima Kourtit[1]

Peter Nijkamp[1,2]

[1] *Faculty of Economics and Business Administration, VU University Amsterdam;*

[2] *Tinbergen Institute.*

Socio-cultural Diversity and Urban Buzz

Daniel Arribas-Bel   **Karima Kourtit**   **Peter Nijkamp***

Dept. of Spatial Economics
VU University Amsterdam
darribas@feweb.vu.nl
k.kourtit@vu.nl
p.nijkamp@vu.nl

* Tinbergen Institute, The Netherlands.

Abstract
Cities have become playing grounds for competitive behaviour and rapid economic dynamics. But in many cities (or urban agglomerations) economic growth is mainly manifested in specific geographic areas, where creative people and innovative entrepreneurs are located. This paper offers first the foundation for analysing the so-called 'urban buzz' and its interlinked primary drivers. The paper will next develop an analytical framework for testing the buzz hypothesis, with a special reference to the importance of social networks in Amsterdam. In our empirical analysis, we use a unique data set on social network connectivity and spatial concentration in a city, based on location-sharing services through the use of Foursquare. Our urban buzz model shows clearly that buzz and socio-economic (cultural) diversity are closely connected phenomena.

Keywords: cultural diversity, urban buzz, agglomeration, creativity, piazza, spatial dependence, social networks

1. The Urban Stage

In the past decades the awareness has grown that cities have become epicentres of socio-economic dynamism. New knowledge, innovations, creative lifestyles and entrepreneurial heroism find often their genesis in urban agglomerations. The rising importance of the 'urban magnetism' is clearly demonstrated in the ever increasing urbanisation rates world-wide. Clearly, there are examples of shrinking cities (such as Detroit or Leipzig), but the loss in their urban population is overshadowed by the rise in urbanisation elsewhere. The urbanized way of living – not necessarily in compact city centres, but more broadly in urban agglomerations including satellite towns and edge cities – has gradually become a dominant megatrend in our world. In several recent publications, Kourtit and Nijkamp (2013a, b, c) have offered much evidence on the emerging 'urban century' which was coined by them the '*New Urban World'*.

It should be noted however, that the '*New Urban World'* does not display a uniform settlement pattern. On the contrary, it is characterized by a great diversity in living and working patterns, in urban land use and architecture, and in urban management and governing institutions. There is an abundance of literature which traces the roots of rising urbanisation. These are mainly sought in the presence of spatial externalities (often in the form of so-called MAR – Marshall-Arrow-Romer – externalities dealing with various economies of scale in urban areas) and social capital benefits (often referred to as 'melting pot' advantages in the spirit of Jane Jacobs). Extensive treatments on these issues can be found in Nijkamp (2008) and van der Ploeg and Poelhekke (2008).

The variety of scholarly contributions on advantages of urban areas can essentially be subsumed under the heading of three driving forces, viz. economies of *density,* economies of *proximity* and economies of *connectivity*. The first category focuses essentially on joint advantages of spatial concentration of people and activities (see e.g. Glaeser et al. 1992, Nijkamp 2008). The next class addresses the benefits of physical or socio-psychological access of people and activities to each other (see e.g. Torre and Gilly 2005; Boschma 2005;Tranos et al. 2013). And finally, the last category concerns the economies generated in a city that emerge from social capital or network linkages – physical or virtual – among heterogeneous groups of people and activities (see Sahin et al. 2007), be it at a short distance or at a long distance (see Tranos et al. 2013).

These classes of external economies explain the booming character of modern cities, in contrast to small towns, villages or rural areas. This does not mean that the latter areas have no socio-economic prospect (see e.g. de Noronha Vaz et al. 2013), but in general the socio-economic future of our world tends to be determined by areas with a high degree of

urbanisation, as such areas generate a wealth of unrivalled centripetal and centrifugal forces and associated benefits. One of the resulting dynamic urban constellations that is frequently mentioned in the current literature on urbanisation advantages is '*urban buzz*', the potential to generate creative, innovative and unconventional initiatives or activities in cities or in specific urban districts. In the next section we will offer a concise introduction to this urban buzz. The aim of the paper is to investigate whether and to what extent the above mentioned economies of density, proximity and connectivity offer a significant contribution to the emergence of urban buzz. The paper will address in particular the impact of socio-economic diversity – including cultural diversity – as an intervening factor in favouring urban buzz. This conceptual model will next be operationalized and empirically tested in a case study on Amsterdam. An extensive database from the Foursquare location-sharing service will be employed to estimate econometrically the 'urban buzz equation'.

2. The Urban Buzz

The phenomenon of urban buzz has to be seen against the background of firm dynamics and innovation in urban areas. Innovative firms are change agents in a creative urban 'milieu'. The economic performance of business enterprises depends on both the firms' indigenous capabilities and the supply of resources in their flanking environment (see e.g., Barney 1991; Kramer and Diez 2012; Kourtit and Nijkamp 2012). Clearly, the growth of companies will be constrained, if there is a shortage or weakness in the available resources, or in the capability to mobilize or generate adequate resources. Reid and Garnsey (1998) distinguish between different stages of growth, ranging from achieving access to resources to the mobilization of resources, and the companies' own generation of resources. The use of the right combination of resources at the right time by young, innovative entrepreneurs enables them to undertake a jump in growth (Kourtit and Nijkamp 2012). Failing to use the right combination at the right time may cause a delay in growth and even a fall back into previous stages (Vohora et al. 2004). In the early growth stages and after a fall back to such stages, firms rely heavily on resources available in the direct environment or proximity, including the urban environment and its constituent infrastructure and suprastructure. Cities offer in many cases the creative network conditions for acquiring new knowledge and expertise. In recent years, the resource-based growth perspective has clearly extended its scope from physical resources to human and social capital resources. It is nowadays broadly accepted that regions and cities – and sometimes, urban districts in a Marshallian sense – may use their indigenous

resources and may offer unique geographic and location conditions and facilities – beyond other competitive assets – to attract talents and firms to (relatively less favoured) regions in the belief that they generate (more) positive externalities. In turn, this may bring about positive socio-economic achievements which may enhance the territorial competitive advantages. As a result, over the past decades cities – and their creative districts – all over the world have managed to reinforce their socio-economic position, be it sometimes with up and downs. This brings us to the notion of urban buzz.

Urban buzz – as a result of density, proximity and connectivity externalities – has received quite some attention in the recent urban literature. Buzz areas – be it cities as a whole or urban districts – are powerhouses of innovation, creativity and unconventional lifestyles. In a study by Storper and Venables (2003), the authors refer in particular to the ease of communication and information exchange between different actors in the urban space as the source of a local buzz economy. A recent article by Rodriguez-Pose and Fitjar (2012) highlights the need for a broader interpretation of urban buzz: this concept is a container for local entrepreneurial dynamism (Acs et al. 2008), innovation access and intensity (Duranton and Puga 2001), knowledge generation and diffusion (Puga 2010), competitive cluster formation (Porter 1990), industrial districts (Pyke and Sengenberger 1992), learning areas (Morgan 1997) or spatial systems of innovation (Cooke et al. 1998). More detailed analyses of buzz phenomena can be found inter alia in Amin and Thrift (1994), Gertler (1995), Bathelt et al. (2004), McCann (2008) and Polèse (2009). Many of these contributions are pointing out important elements of dynamic urban developments, but in most cases a solid evidence-based econometric test on the underlying hypotheses is missing. The main ambition of the present study is to offer a statistical-econometric framework for examining the validity of the urban buzz hypothesis.

It should be noted that urban buzz may relate to socio-economic factors and productivity-enhancing factors. The first class relates to the economies of cultural and social diversity in urban areas (Jacobs 1961 1969; Waldinger et al. 1990; Perdikogianni and Penn 2005; Choenni 1997; Sahin et al. 2007; Kahanec and Zimmermann 2011; Longhi et al. 2010; Kourtit and Nijkamp 2012). Such social buzz factors create various advantages for the population concerned, such as a great variety in product supply, variability in skills and socio-economic capabilities (e.g., ethnic entrepreneurship) (Kloosterman and Rath 2003; Masurel et al. 2002; Zhou 2004; Kourtit and Nijkamp 2011; Ozgen et al. 2011; Sahin et al. 2012), seedbed functions for new forms of art and culture etc. The second class is more focused on business sector advantages, such as a rise in innovativeness, access to creative ideas, vicinity of institutions for higher education, use of advanced telecommunication systems and networks, etc.

Urban buzz is not a phenomenon that is uniformly spread over all population groups in the city or urban districts. There is a clearly demographic component involved with intra-urban dynamics, in particular in regard to ethnic variety in modern cities. The underlying idea is that in modern cities a great deal of socio-economic dynamics is created by various types of foreign migrants. In particular, the emergence of migrant entrepreneurship has led to an unprecedented dynamics in many cities (see e.g. Kourtit and Nijkamp 2012).

In many cases, urban buzz is not only showing up all over a city, but in dedicated or specific areas where a concentration of initiatives, innovations and an interactive expressions of life styles are taking place. This action place of urban buzz resembles the *piazza* of old Italian cities, where in the past all activities and communications in the city were concentrated. The *piazza* is essentially the spatial bundling of urban buzz and very much depends on factors relating to both the built environment as well as the socio-economic conditions. Therefore, in our research on the spatial expression and geographic projection of urban buzz, we will concentrate our efforts on the economic functioning and outreach of such *piazza*s in modern cities.

We will focus attention in this paper mainly on urban social networks, more in particular from a cultural diversity perspective on urban buzz. Clearly, there are other determinants of local dynamism as well, but we will offer in this contribution an explanatory analysis for the broader cultural diversity components in an urban buzz constituency. Thus, we will regard creative classes, migrants and social media users as major change agents in a local or urban economy, at the interface of urban buzz and cultural diversity.

3. Methodological Framework for Urban Buzz Analysis

From the previous section, we conclude that urban buzz results from a series of background factors. Some have a more physical background, like urban form or infrastructure, while others stem from characteristics of the environment associated with the people who live or spend time there, like the type of amenities available. Additionally, as this paper argues, the presence of a culturally diverse neighbourhood might be considered as an additional amenity valued by the main actors that drive urban buzz and hence induce a higher intensity of activity and interactivity. Now we will test the above propositions on the basis of a case study in the Netherlands.

Our empirical contribution is centered around the city of Amsterdam. Two main reasons explain the decision to opt for this location. In the first place, Amsterdam has a

long tradition of openness and tolerance. Historically, the city has offered shelter to various cultures and ethnicities and has shaped its character around this idea of inclusion. If cultural diversity is to have an effect on the popularity of a given area within a city, this effect is likely to appear in Amsterdam more than anywhere else. Secondly, on a more practical level, Amsterdam is sufficiently large, culturally-oriented and high-tech-savvy to induce a degree of penetration of location-sharing services that ensures meaningful results when using the dataset designed to capture urban buzz (see also Section 4.1 for an in-depth description). We will use the boundaries of the municipality of Amsterdam to delineate the spatial extent of our analysis. Furthermore, in order to study the variation of activity within the city, we use the neighbourhood ('buurt' in the Dutch terminology) level, because it combines a good spatial resolution and relevant data on migrants' presence.

In order to formally test the effect of cultural diversity on the level of buzz, we need to translate the ideas from the previous section into a framework that allows for hypothesis testing. In particular, we will use regression analysis to estimate the existence and importance of such an effect. Based on the literature reviewed in the previous section, we hypothesize that the volume of buzz in a particular area of a city may be conceptually expressed as a function of the following factors:

$$B_i = \alpha + \beta_1 F_i + \beta_2 E_i + \gamma Div_i + \varepsilon_i \tag{1}$$

where B_i is the level of buzz in neighbourhood i, α is a constant term, F_i is a set of variables relating to the amount of possibilities for buzz to occur in i, E_i is a group of variables describing different characteristics of the urban form in i, Div_i is the level of cultural diversity that characterizes i, and ε_i is a well-behaved error term. At the same time, β_1, β_2 and γ are parameters that capture the effect on the level of buzz. The particular implementation, the definition and methodology to quantify buzz in B_i, as well as the actual selection of variables that we consider to empirically represent F_i, E_i and Div_i are explained in detail in Section 4.

Ultimately, what we are interested in is the sign and magnitude of γ, which will allow us to state whether cultural diversity does indeed have an effect on buzz at an intra-urban level. A positive sign will point to a positive impact, meaning that the more cultural diversity, the higher the buzz; the opposite will apply if the sign is negative; an insignificant coefficient would suggest no relationship between the two. However, this analysis can also shed some light on other determinant factors. Since the idea of quantitatively measuring buzz in a city is one of the novel contributions of our work, the

5

question of which aspects of a neighbourhood matter for the level of local activity is also of interest in this context. This means that we will also pay attention to the sign and magnitude of the estimates of β_1 and β_2. They will help us better understand the nature of buzz and what type of human activities and economic functions trigger this phenomenon.

The distribution of buzz across Amsterdam is likely to have a marked spatial dimension. Some parts of the city, such as the historical old town, will most likely have much higher levels of activity than other more peripheral and residential ones. At the same time, this pattern is not likely to perfectly match administrative boundaries of the neighbourhoods. While these entities do capture to some extent socio-economic characteristics that drive the outcome, they remain fairly stable over time and hence cannot accommodate the more dynamic nature of buzz. If this is the case, spatial autocorrelation will be present in the data that we observe since the spatial unit we use (i.e. neighborhood, see Section 4) does not perfectly capture the extent of the phenomenon we are studying, that is urban buzz. The issue of spatial autocorrelation, although a technical econometric problem, may have important consequences for the final conclusions drawn from the analysis. In particular, failing to account for the spatial correlation of the dependent variable in an econometric model when necessary makes the estimates biased and inefficient (Anselin 1988). For that reason, we decided to expand the baseline model including what is commonly known as "spatial lag": an explanatory variable that expresses the value of the dependent variable in the surrounding neighborhood. In particular, the extended model may be formulated as:

$$B_i = \alpha + \rho \, \Sigma^N_j \, w_{ij} B_j + \beta_1 F_i + \beta_2 E_i + \gamma Div_i + \varepsilon_i \tag{2}$$

where ρ is a parameter and w_{ij} is the value if row i and column j of the spatial weights matrix W. This NxN matrix contains a formal representation of the spatial relationships between all the observations in the sample; if i is defined as a spatial neighbour of j, $w_{ij} > 0$; otherwise, the weight assigned to such relationship is zero. When W is row-standardized, the spatial lag of B_i effectively becomes the average value of B in i's surrounding locations.

The justification for the introduction of a spatial lag in this context is akin to the filtering of temporal correlation in the time series literature[1]. If we consider equation (2) in matrix form:

---

[1] For a similar discussion in the context of housing prices, we refer to Anselin and Lozano-Gracia (2008).

$$B = \alpha + \rho W B + \beta_1 F + \beta_2 E + \gamma Div + \varepsilon \tag{3}$$

and rearrange it, we obtain:

$$B - \rho W B = B(I - \rho W) = \alpha + \beta_1 F + \beta_2 E + \gamma Div + \varepsilon \tag{4}$$

which reflects the correction of the spatial scale mismatch present in B by the operator $(I - \rho W)$. This introduction however has clear implications for the estimation method required. A spatial lag of the dependent variable introduces endogeneity in the model that must be accounted for and corrected when the model is estimated. To that end, we use OLS for the baseline equation but adopt a maximum likelihood (ML) approach as suggested in Anselin (1988) for the spatial lag model.

4. Database

4.1 Measuring buzz with location-sharing services

This paper adopts a novel approach to measure buzz within an urban environment. We take advantage of a new phenomenon that is spreading quickly among creative residents of cities: location-sharing services. These are online applications with which users, empowered by a location-aware device connected to the internet such as a smartphone or tablet, can share their geographical position at a given point in time with their friends and broadcast it to the internet. To offer a sense about their degree of penetration, in 2010 Foursquare, one of the main leaders in this industry, claimed to be processing one million posts per day (Grove 2010). Since then, its popularity has even increased, particularly among the young and skilled stratum of the population. A key piece in these services is what has come to be known as checkin. Whenever a user finds him/herself in a place and wants to share it with his/her social network (and potentially with the whole internet), this kind of mobile applications can be used to "check in". The place where the checkin occurs, or venue, can vary greatly in its nature: from a bar or restaurant to a train station or even a park; most urban spaces can be "checked in". A checkin is not only defined by the place/venue where the user finds him/herself at a moment, but also by its accurate coordinates (provided by the digital device carried) and a time stamp. This degree of detail automatically translates into a dataset with very high granularity in both space and time.

The dataset we employ for this application is comprised by more than 70,000 checkins extracted by the original database presented in Cheng et al. (2011). Throughout

most of 2010 and beginning of 2011, Cheng et al. (2011) crawled the social network Twitter to compile checkins from Foursquare broadcast, not only to the immediate social network but to the whole internet. This yielded more than 20 million observations that they later released[2] in an open format and which has already been included in some studies by other researchers (e.g. Cranshaw et al., 2012). Although the observations are scattered all around the world, because they are geo-referenced, we can subset them to consider only those that occurred within the Amsterdam city boundary limits. In its basic form, each observation includes the latitude and longitude of the checkin, a time stamp and some text from the tweet, which we use to link it to the venue and then are able to obtain more characteristics (see Section 4.2 for a more detailed explanation). These data, although appealing, are not without drawbacks. One in particular stands out: it is not possible to know details about the characteristics of the users since they are anonymized. This means that we cannot trace back the origins of buzz (who are *producing* it) and thus we cannot be totally sure about who it comes from. This certainly limits some of the applications the dataset could be used for, since it prevents us to delve deeper into the actual nature of the buzz, but not all of them. In particular, the focus of this paper, to what extent cultural diversity and other socio-economic factors influence the volume of buzz in a neighborhood (whoever this activity comes from), can still be tested.

Figure 1 displays the distribution of the checkins, both at the point level (points have been made extremely small and transparent in order not to overcrowd the figure) and aggregated at the neighbourhood, which is the unit we will be using for the analysis. The checkins closely match the spatial structure of Amsterdam. Smaller, more central and denser neighbourhoods of the historical center capture most of the checkins and peripheral, larger areas devoted to residential and industrial uses (e.g. the harbour in the NW area) barely attract any activity. An exception is a large polygon in the SE of the city, which shows a large volume of checkins, considered its location and neighbors. This corresponds to the Bijlmer area, where a whole entertainment center was developed, including the soccer team (Ajax) stadium, many movie theaters and a large concert hall. By looking at the map, one can easily conclude that the dataset is appropriate to measure urban buzz at a very detailed level.

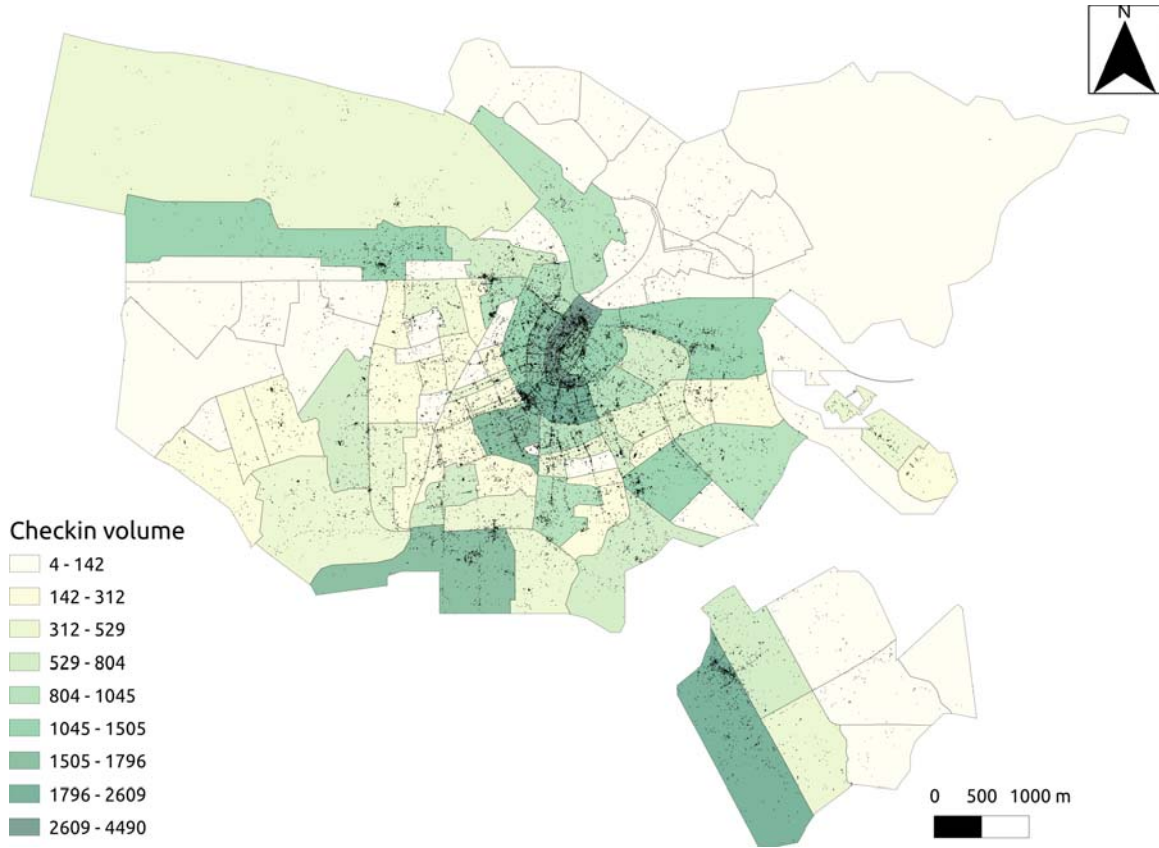[2] See http://infolab.tamu.edu/data for more information.

Figure 1. Distribution of the volume of Foursquare checkins

4.2 Cultural diversity and land use data

Once we have presented our measure of buzz, the effect of cultural diversity can only be examined by adopting an index that captures in a quantitative way its presence. Although the literature on how to measure this phenomenon is extensive and rich (e.g. e.g. Sassen 1994; Alesina et al. 1999; Alesina et al. 2004; Boeri and Brücker 2005; Musterd and Deurloo 2006; Ottaviano and Peri 2006a,b; Evans 2009; World Bank 2008; Nijkamp 2008; Kourtit and Nijkamp 2011; Ozgen et al. 2011b), in this case we decide to take a traditional approach and adopt a well-established index. In particular, we use the index of fractionalization, presented in Mauro (1995) and widely employed in other studies on the effects of cultural diversity, such as Ottaviano and Peri (2006). Its formal expression is:

$$div_r = 1 - \Sigma^{M}_{i=1} \ (CoB^r)^2 \tag{5}$$

where div_r is the diversity index in area r, M is the total number of different cultural origins and $C\,oBr$ is the share of the population with cultural origin i in area r. The index

9

is bounded between *0* and (*1 − 1/M*) and accounts for two aspects of cultural diversity: the richness, or how many different groups there are; and the evenness, or how the population is distributed across those groups. In an extreme case where everyone in a neighbourhood belonged to the same group, the probability of picking up two different groups at random is non- existent and, accordingly, *div* = 0. As a neighbourhood includes more variety (*M* increases) and the distribution across them stays even or proportionate, the probability of randomly selecting two different persons increases and so does the index. To calculate this measure in the Amsterdam neighbourhoods, we use data from the Dutch Bureau of Statistics (CBS for its Dutch acronym). At that level of resolution, we can access the shares of immigrants in the following groups: Western migrants, Moroccans, Surinamese, Turkish, from Antilles and other origins. In this context, a person is considered to be a migrant if at least one of the parents was born outside the Netherlands. Equally, a migrant is considered to be Westerner if he/she comes from Europe, North America, Japan, Indonesia and Oceania. This implies that, in the context of equation (5), *M = 7* and thus the maximum degree of diversity we can reach in our dataset is 0.86.

In addition to cultural diversity, we include other variables to explain the volume of checkins in a neighbourhood. These are introduced to control for the number of possibilities available to users and for the economic function of the area. The first set of controls is extracted from the Foursquare, Inc. (2012) database using their Venue platform. It contains information about all the venues where the users checked in during the time the checkins dataset was compiled[3]. In order to use it in conjunction with the cultural diversity index in a regression framework, we aggregate the data to the neighbourhood level and simplify it to include only major venue categories. The result is reflected in the eight following variables: total number of venues as well as percentage of venues in "Arts & Entertainment", "College & University", "Food", "Out- doors & recreation", "Professional & Other places", "Travel & transport" and "Other". Some other minor categories were initially considered but dropped later due to multicollinearity.

The last source we use in this study introduces land-use data. This comes from the Dutch register of addresses and buildings (BAG, for its Dutch initials[4]), which keeps track of the land-use category at the unit level, even within a building. In the municipality of Amsterdam, this means almost half a million records. Similarly to the

---

[3] The process on linking checkins with the venue where they had occurred consisted of extracting the link of the venue from the tweet text to later query the Foursquare venue platform for further information.
[4] For more information, see http://bag.vrom.nl/.

venue data, we need to bring this to a spatial unit at which they can be linked to the degree of cultural diversity. This is possible by aggregating the amount of area devoted to each use in every neighbourhood and obtaining the percentage it represents of the total. For this analysis, we use the percentages of use in the following categories: industrial, office space, sports and retail. Additionally, we include the number of units, which is correlated with the building density and urban form of the area.

5. Empirical Results

In this section, we present the results and main interpretation of applying the methodology outlined in Section 2 to the data described in Section 4. In the first place, as a benchmark, we adopt a non-spatial approach and estimate equation (1) through traditional OLS, which allows us to obtain a set of baseline results. In a second stage, we include a spatial lag into the model and estimate the results via Maximum Likelihood (ML henceforth) as expressed in equation (2). Finally, we present an extension in which we "stretch" the data into a pseudo-panel, taking advantage of the temporal granularity of the Foursquare data, in order to break the results into different times of the day.

The first column of results in Table 1 shows the coefficients estimated for the non-spatial model using traditional OLS. Focusing first on the control variables, we can see they show the expected signs. Higher density of buildings, measured by the total number of units (within buildings) in the neighbourhood, has a large positive impact; that is the more crowded and denser the area, the more Foursquare activity it receives. This is a highly anticipated result that nevertheless can be taken as a sanity check on the validity of the dataset. Intuitively, the presence of more possibilities to "checkin", measured by the amount of venues, leads to higher total volumes of activity. In terms of shares, larger proportions of all the categories included, with the exception of "College & University", are associated with more activity as well. The strongest effect, unsurprisingly, comes from the presence of venues in the category "Arts & Entertainment", in which most cultural amenities such as museums and cinemas are included as well as some bars[5]. This is in line with the consumption-amenity nature of Foursquare checkin data and reinforces the argument of them being a good index of urban buzz. Different proportions of land-use in a neighbourhood shape a different profile in terms of its socio-economic role

---

[5] The coding scheme that determines what category a venue is in is not totally deterministic in that it depends on the judgement of users. Because of this, for example, some bars are coded as "Arts & Entertainment" while other are part of "Food".

within the larger context of the city. This is reflected in the coefficients for the shares of land-use types. Presence of office and retail space, both functions that attract daylight activity, significantly increases the number of checkins an area receives. Somehow counter-intuitively, a significant and positive coefficient is associated with industrial use. This might be driven by areas that used to be industrial (and are still coded that way) but that in the last years have undergone regeneration processes which have attracted many cultural activities (e.g. Northern docks across from the Central Station), strong drivers of "checkins". The model seems to offer a plausible explanation of the data and capture correctly most of the anticipated effects on the levels of Foursquare activity. In addition, it also captures a good deal of the variation in the data, reaching a R^2 of slightly over 0.73, which gives us confidence in the specification and in the subsequent interpretation of the effect of diversity.

The estimated effect of diversity is 0.0152 (0.1518 x 10, due to rescaling) and the coefficient is significant at the 10% level. This implies that, according to our model, a unity increase in the degree of cultural diversity present in the neighbourhood induces a boost in the number of "checkins" of approximately 1.5%. If that is the case, the evidence from this study suggests that Foursquare users positively value cultural diversity and, everything else equal, show a preference for diverse enclaves when they check-in. If we assume this is a good proxy to measure where activity (i.e. urban buzz) locates within the city at a given point in time, we are thus stating that cultural diversity in fact does have a positive and significant impact on urban buzz, at least for the case of Amsterdam.

The results shown so far relate to a model that does not take space explicitly into account. In that context, observations are assumed to be independent of each other, regardless of where they are located in space. However, as we have noted in Section 3, if that is not the case and the spatial scale we are using does not completely encapsulate the values of urban buzz, we might be incurring in a bias when estimating the coefficients (Anselin 1988). In order to properly account for these spatial effects, we estimate[6] the spatial lag model in equation (2) via ML. There are two details we first have to take care of before delving into the results. The first one is the nature of the spatial relationships we introduce in the model through W . The choice of one or another type of matrix has to be made ex-ante and, because of that, the use of W has been criticized in the literature (e.g. Gibbons and Overman 2012). In this study, we adopt a pragmatic approach and test several different specifications to make sure our results and conclusions are not

---

[6] All the computations relating to spatial methods in this paper were performed using the Python library PySAL (Rey and Anselin 2010).

exclusively tied to the choice of W. In particular, we tried K-nearest neighbours ($k = 6$), a distance threshold and contiguity weights based on the queen criterion (i and j are neighbours if they share at least an edge or vertex). Because estimates are very similar and conclusions do not change, we only report results for the latter. In addition, we always row standardize the matrix so the spatial lag can be interpreted as the average value of an observation in its vicinity. The second issue relates to the assumption of normality implicit in ML estimation, as reliable estimates are only to be obtained if this condition holds. The bottom of Table 1 shows the Jarque-Bera test of normality, which takes the null of a normal distribution in the OLS residuals and, if rejected, it can be taken as a sign of a problem. We obtain 1.711, with which we cannot reject, leading us to conclude that ML is an appropriate estimation method for our model. Finally, in order to reinforce our choice of a spatial lag specification, the very bottom of the table shows the results of the LM diagnostics of spatial dependence, which offer guidance on the spatial data generating process (either a lag or an error, see Anselin, 1988 for more details). When using their more reliable robust version, the indication is clear towards a spatial lag process.

The spatial lag results are reported in the second column of Table 1. The general conclusions of the model point in the same direction as for the OLS case since no sign or level of significance is greatly affected. However, it can be seen that coefficients are in their vast majority (with the exception of one significant variable) overestimated: when space is accounted for in the model, their magnitude is more limited. So is the case with the coefficient of diversity, which is downwardly corrected from 0.15 to 0.13, but remains significant at the 10% level, pointing to a significant impact of a more diverse neighbourhood in its level of buzz. The OLS upward bias is a typical result when positive spatial autocorrelation is present and highlights the importance of these methods to properly deal with spatial effects and obtain correct estimates. The ρ parameter, associated with the spatial lag of the dependent variable, is slightly larger than 0.4 and highly significant. This is indication of a process of positive spatial autocorrelation by which neighboring observations tend to have similar values.

Table 1. Results for aggregated volume of *checkins*

	OLS	ML-Lag
Foursquare - Arts & Entertainment	0.0285***	0.0158**
	(0.008)	(0.007)
Foursquare - College & University	-0.0104	-0.0063
	(0.012)	(0.010)
Foursquare – Food	0.0208***	0.0132***
	(0.006)	(0.005)
Foursquare – Other	0.0121***	0.0128***
	(0.003)	(0.003)
Foursquare - Outdoors & Recreation	0.0171**	0.0093
	(0.008)	(0.007)
Foursquare - Professional & Other places	0.0129**	0.0099**
	(0.005)	(0.004)
Foursquare - Travel & Transport	0.0138**	0.0134***
	(0.006)	(0.005)
Industrial use	0.0184**	0.0217***
	(0.008)	(0.006)
Office use	0.0339***	0.0224***
	(0.008)	(0.006)
Sports use	-0.0271	-0.0213
	(0.030)	(0.024)
Retail use	0.0187	0.0408
	(0.032)	(0.026)
Total venues	0.0209***	0.0138***
	(0.006)	(0.005)
Total units	0.1853***	0.1737***
	(0.036)	(0.030)
Diversity	0.1518*	0.1267*
	(0.080)	(0.065)
Constant	2.1549***	-0.0511
	(0.524)	(0.670)
P		0.4399***

R^2 0.735
Jarque-Bera 1.711

LM Lag Robust	LM Lag	LM Error	20.8748***
	Robust	LM Error	12.0798***
			8.9422***
			0.1472

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, insignificant otherwise. Standard errors are shown in parethesis below the estimates.
"Total units" and "Diversity" are rescaled to avoid multicolinearity issues by pre-multiplying them by 0.001 and 10 respectively.

The last piece of the analysis comprises an exploration into the granular temporal dimension of the Foursquare data. We set out to study how the impact of cultural

diversity varies across time of the week. For every neighbourhood, we disaggregate by time of the week the number of checkins and re-run different models of equation (2) using the volume of activity in each time slot as the dependent variable. The following times of the day are considered: morning (5am to noon), weekday afternoon (noon to 6pm), evening (6pm to 10pm) and night (10pm to 5am). Each of those is further divided between weekday (Monday through Friday) and weekend (Saturday and Sunday), resulting in eight different times of the week. Figure 2 displays the ML estimates[7] for the coefficient of diversity on each time of the week, along with the 95% confidence intervals. The effect is lowest on a weekday morning and grows over the day to reach its peak during the night of a week day. Over the weekend, the effect decreases a little bit in the morning but picks up again in the evening and night. Except for the weekday morning, they are all significant at the 5% level. This pattern is very much consistent with the idea of cultural diversity valued by Foursquare users as a consumption amenity: in typical work hours, its effect is negligible, but as the use of time shifts more into leisure and entertainment, the effect becomes more relevant, highlighting that users particularly prefer diverse neighbourhoods for activities they do during the end of the day or in the weekend.
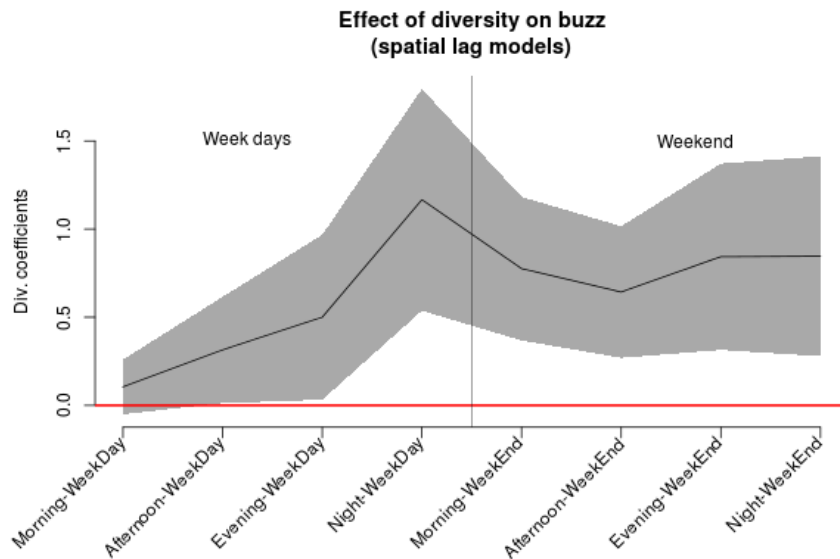

Figure 2. Coefficient estimates and 95% confidence intervals

[7] Although we only present ML estimates from the spatial model, we tested several alternative specifications, including a pooled regression, a regime regression using time of the week as the regime variable and spatial versions of those. Since all the main conclusions remain unchanged, we only show the ML models for consistency with the previous part of the analysis. Additional results are available from the authors at request. Equally, although we only show the estimates for diversity, full model results are available.

6. A Policy Perspective on Urban Buzz

This paper adopts a novel dataset and approach to quantitatively measure urban buzz and to study its main determinants, with a particular focus on the effect of cultural diversity. Using data from the online location-sharing service Foursquare, we quantitatively define urban buzz as "checkin" volume per neighborhood and present an empirical model for the city of Amsterdam. This includes not only a measure of cultural diversity, but also information on the availability of venues to "check in" as well as on land use, and properly accounts for the presence of spatial autocorrelation in the buzz variable. The main results suggest a positive and significant effect of cultural diversity on the level of buzz activity that occurs in a neighborhood. This implies that, given the same economic functions and availability to check in, greater level of cultural diversity is associated with a larger volume of checkins.

Urban buzz reflects the wealth-creating potential of urban areas as a result of density, connectivity and proximity advantages among heterogeneous groups – including distinct migrant groups – that reside in modern cities. Creativity and diversity appear to be key factors in the dynamic performance conditions of such areas. What can policy do to favour such urban buzz phenomena? Urban buzz finds its genesis in urban agglomeration advantages of various kinds, and it seems plausible that policy efforts to exploit the benefits of urban territorial capital (including social, creative, entrepreneurial and innovative capital) may concentrate on the provision of conditions that favour such capital. Land use policy, educational policy and entrepreneurial stimuli may then offer promising strategic directions. Moreover, since the results presented in the previous sections point to a positive and significant effect of cultural diversity on this phenomenon, a focus from policy makers on protecting and even stimulating such characteristic of cities and neighborhoods is an additional recommendation that derives from this study.

Especially in a European setting – where most cities house a wealth of cultural heritage which acts often as an attraction force for innovative business – due policy attention for the exploitation of historic-cultural resources as a source of economic progress would be meaningful. There is an increasing awareness in Europe that cultural heritage is not meant to craft a city in stone, but to use the past as an engine for economic progress, by attracting visitors, business and residents.

This study represents a first experiment to explore the usefulness of new data sources originating from the web and how they can be employed to answer questions about the

16

physical world. In this sense, the positive results obtained regarding the main question asked in the beginning of the paper (how does cultural diversity affect urban buzz?) should prompt more activity in the future rather than be taken as a road end. In particular, and without being exhaustive, we will conclude by suggesting three avenues that we consider especially relevant. Further research could extend the present results by looking at the components of diversity in relation to urban buzz (what particular socio-economic groups have a larger participation?) and trying to disentangle potential identification issues by including information on wages and wage-gaps between groups. An aspect that has been assumed throughout the paper but that would certainly be interesting to empirically test too, is to what extent the effect of urban buzz influences the socio-economic development of a neighborhood. Finally, a deeper analysis on the nature of Foursquare data, trying to delve into the intricacies and characteristics of its users, as well as their socio-demographics, would also be illuminating. For all those reasons, this study should be taken as an inspiring starting point rather than as a conclusive end.

References

Acs, Z. J., Bosma, N., and R. Sternberg (2008). *The Entrepreneurial Advantage of World Cities: Evidence from Global Entrepreneurship Monitor Data*, Working Paper SCALES, Utrecht: University of Utrecht.

Alesina, A., R. Baqir and W. Easterly (1999). Public Goods and Ethnic Division, *Quarterly Journal of Economics*, 111 (4) : 1243-1284.

Alesina, A., R. Baqir and C. Hoxby (2004). Political Jurisdictions in Heterogenous Communities, *Journal of Political Economy,* 112 (2): 384-39

Amin, A., and N. Thrift (1994). *Globalization, Institutions and Regional Development in Europe,* Oxford, New York: Oxford University Press.

Anselin, L. (1988). *Spatial Econometrics*. Dordrecht: Kluwer Academic Publishers.

Anselin, L. and N. Lozano-Gracia (2008). Errors in Variables and Spatial Effects in Hedonic House Price Models of Ambient Air Quality, *Empirical Economics* 34 (1): 5-34.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage, *Journal of Management* 17:99-120.

Bathelt, H., Malmberg, A. and P. Maskell (2004). Clusters and Knowledge: Local Buzz, Global Pipelines and the Process of Knowledge Creation, *Progress in Human Geography* 28:31-56.

Boer, J. de, and J. Beekmans (2012). Check-in-Stedenbouw, Rooilijn, 45 (6), 404-409.

Boeri, T. and H. Brücker (2005). Why are Europeans so Tough on Migrants?, *Economic Policy,* 20(44): 629-703.

Boschma, R.A. (2005). Proximity and Innovation: a Critical Assessment. *Regional Studies* 39:61-74.

Cheng, Z., Caverlee, J., Lee, K., D.Z. Sui (2011). *Exploring Millions of Footprints in Location Sharing Services*. In: Proceeding of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM), Barcelona.

Choenni, A. (1997). *Veelsoortig Assortiment*, Amsterdam: Het Spinhuis.

Cooke, P., Uranga, M.G. and G. Etxebarria, (1998). Regional Systems of Innovation: an Evolutionary Perspective, *Environment and Planning A* 30: 1563-1584.

Cranshaw, J., Schwartz, R., Hong, J. and N. Sadeh (2012). *The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City*, in: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media, ICWSM, volume 12.

Duranton, G. and D. Puga, (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products, *American Economic Review* 91(5):1454-1477.

Evans, G.L. (2009). Creative Cities, Creative Spaces and Urban Policy, *Urban Studies,* 46(5-6): 1003-1040.

Foursquare, Inc. (2012). Accessed Sept. 5, 2012, http://foursquare.com.

Gertler, M.S. (1995). 'Being there': Proximity, Organization, and Culture in the Development Adoption of Advanced Manufacturing Technologies, *Economic Geography* 71:1-26.

Gibbons, S. and H.G. Overman (2012). Mostly Pointless Spatial Econometrics?, *Journal of Regional Science* 52 (2): 172-191.

Glaeser, E., Kallal, H. D., Schinkmann, J. A. and A. Shleifer (1992). Growth in Cities, *Journal of Political Economy* 100: 1126-1152.

Grove, J. (2010). *Foursquare Nearing 1 Million Checkins per Day*, Mashable, http://mashable.com/2010/05/28/foursquare-checkins/.

Jacobs, J., (1961). *The Death and Life of Great American Cities*, New York: Vintage Books.

Jacobs, J., (1969). *The Economy of Cities*, New York: Random House.

Kahanec, M. and K. Zimmermann (eds) (2011). *Ethnic Diversity in European Labor Markets*, Cheltenham: Edward Elgar.

Kloosterman, R. and J. Rath (2003). *Immigrant Entrepreneurs: Venturing Abroad in the Age of Globalization*, Berg Publishers: Oxford.

Kourtit, K. and P. Nijkamp (2012). Strangers on the Move: Ethnic Entrepreneurs as Urban Change Actors, *European Review*, Vol. 20 (3), pp 376- 402.

Kourtit, K. and P. Nijkamp (2013a), In Praise of Megacities in a Global World, *Regional Science Policy and Practice* (forthcoming).

Kourtit, K. and P. Nijkamp (2013b) The Future of the New Urban World, *International Planning Studies* (forthcoming).

Kourtit, K. and P. Nijkamp (2013c), Challenges of the New Urban World, *Applied Spatial Analysis and Policy* (forthcoming).

Kourtit, K. and P. Nijkamp (2011). Strategic Choice Analysis by Expert Panes for Migration Impact Assessment, *International Journal of Business and Globalisation* 7(2):166-194.

Kramer, J.-Ph., Diez, J.R. (2012). Catching the Local Buzz by Embedding?, *Regional Studies* 46 (10): 1302-1317.

Longhi, S., Nijkamp, P. and J. Poot (2010). Joint Impacts of Immigration on Wages and Employment: Review and Meta-Analysis, *Journal of Geographical Systems* 12 (4):355-388.

Masurel E., Nijkamp, P., Tastan, M. and G. Vindigni (2002). Motivations and Performance Conditions for Ethnic Entrepreneurship, *Growth & Change* 33(2):238-260.

Mauro, P. (1995). Corruption and Growth, *The Quarterly Journal of Economics* 110 (3): 681-712.

McCann, P. (2008). Globalization and Economic Geography: the World is Curved, not Flat, *Cambridge Journal of Regions, Economics and Society* 1:351-370.

Morgan, K. (1997). The learning Region: Institutions, Innovation and Regional Renewal, *Regional Studies* 31:491-503.

Musterd, S. and R. Deurloo (2006) Amsterdam and the Preconditions for a Creative Knowledge City, *Tijdschrift voor Economische en Sociale Geographie,* 91(1): 80-93.

Nijkamp, P. (2008). XXQ Factors for Sustainable Urban Development: A Systems Economics View, *Romanian Journal of Regional Science* 2 (1): 1-34

Noronha Vaz, M.T. de, Leeuwen, E. van and P. Nijkamp (eds) (2013). *Towns in a Rural World,* Aldershot: Ashgate.

Ottaviano, G. and G. Peri (2006). The Economic Value of Cultural Diversity: Evidence from US Cities, *Journal of Economic Geography* 6 (1): 9-44.

Ozgen, C., Nijkamp, P. and J. Poot (2011a). *Immigration and Innovation in European Regions*, IZA Discussion Paper no. 5676, Bonn: IZA.

Ozgen, C., Nijkamp, P. and J. Poot (2011b). *The Impact of Cultural Diversity on Innovation: Evidence from Dutch Firm Level Data*. IZA DP No. 6000. Bonn: Institute for the Study of Labor (IZA).

Perdikogianni, I. and A. Penn (2005). *Measuring Diversity: a Multi-Variate Analysis of Land Use and Temporal Patterning in Clerkenwell*, Proceedings of the 5th International Space Syntax Conference, Delft.

Ploeg, F. van der, Poelhekke, S. (2008). Globalization and the Rise of Mega-Cities in the Developing World, *Cambridge Journal of Regions, Economy and Society* 1(3):477-501.

Polèse, M. (2009). *The Wealth and Poverty of Regions: Why Cities Matter,* Chicago, IL: University of Chicago Press.

Porter, M.E. (1990). *The Competitive Advantage of Nations,* London: Macmillan.

Puga, D. (2010). The Magnitude and Causes of Agglomeration Economies, *Journal of Regional Science* 50:203-219.

Pyke, F. and W. Sengenberger (1992). Industrial Districts and Local Economic Regeneration: Research and Policy Issues, in: F. Pyke & W. Sengenberger (eds.) *Industrial Districts and Local Economic Integration,* Geneva: International Labour Organization, pp. 1-30.

Reid, S. and E. Garnsey (1998). Incubation Policy and Resource Provision: Meeting the Needs of Young, Innovative Firms, in: R. Oakey and W. During (eds.), *New Technology Based Firms in the 1990s*, London: Paul Chapman, pp. 67-80.

Rey, S. J. and L. Anselin (2010). PySAL: A Python Library of Spatial Analytical Methods, in: Fischer, M. M. and Getis, A. (eds), *Handbook of Applied Spatial Analysis,* Heidelberg/Berlin: Springer, pp. 175-193.

Rodriguez Pose, A. and R.D. Fitjar (2012). Buzz, Archipelago Economies and the Future of Intermediate and Peripheral Areas in a Spiky World, *European Planning Studies*, online first publication, http://dx.doi.org/10.1080/09654313.2012.716246

Sahin, M., M. Rietdijk and P. Nijkamp (2007). Is Cultural Diversity a Socio-Economic Opportunity? – A Case of Urban Migrant Entrepreneurs, *Innovation ─ The European Journal of Social Science Research* 22(3): 251-281.

Sahin, M., Todiras, A. and P. Nijkamp (2012). An Explanatory Model for the Economic Performance of Migrant Entrepreneurs in the High-tech Sector, in: P. Nijkamp, J. Poot and M. Sahin (eds), *Migration Impact Analysis: New Horizons*, Cheltenham: Edward Elgar, pp. 227-260.

Sassen, S. (1994). *Cities in a World Economy*. Thousand Oaks, California: Pine Forge/ Sage Press.

Storper, M. and A.J. Venables (2004). Buzz: Face-to-Face Contact and the Urban Economy, *Journal of Economic Geography* 4(4): 351-370.

Torre, A. and J.P. Gilly (2005). Proximity and Localization, *Regional Studies* 39: 47-59.

Tranos, E., Reggiani, A. and P. Nijkamp (2013). Accessibility of Cities in the Digital Economy, *Cities* 30: 59-67.

Vohora, A., Wright, M., Lockett, A. (2004). Critical Junctures in the Development of University High-tech Spinout Companies, *Research Policy* 33:147-175.

Waldinger, R., Aldrich, W. and R. Ward (eds) (1990). *Ethnic Entrepreneurs,* Newbury Park CA: Sage Publishers.

World Bank (2006). *Global Economic Prospects: Economic Implications of Remittances and Migration*, World Bank, Washington, D.C .

Zhou, M. (2004). Revisiting Ethnic Entrepreneurship: Convergencies, Controversies, and Conceptual Advancements, *The International Migration Review* 38(3):1040.