Volume 7, Number 2, 2020

REGION

The Journal of ERSA
Powered by WU

Table of Contents

Articles

Risky move: New evidence on the determinants of the willingness to migrate
Geraint Johnes 1

Strengthening local economy – an example of higher education institutions' engagement in "co-creation for sustainability"
Anna Bogedain , Rüdiger Hamm 9

Non-Employed People in the Westpfalz Region (Germany) as Target Group of Continuing Higher Education
Jessica Neureuther , Dominik Böhmer 29

Flatten the Curve!
Thomas Wieland 43

Does the Crisis Change the Nature of Agglomeration Economies in Indonesia?
Khoirunurrofik 85

Resources

classification for English primary schools using open data
Stephen Clark , Nik Lomax , Mark Birkin R1

A reproducible notebook to acquire, process and analyse satellite imagery
Meixu Chen , Dominik Fahrner , Daniel Arribas-Bel , Francisco Rowe R15

Letters

Macroeconomic determinants of Port and Douro wine exports: An econometric approach
Anthony Macedo , Sofia Gouveia , João Rebelo L1

Funded by

REGION, The journal of ERSA / Powered by WU
ISSN: 2409-5370

ERSA: http://www.ersa.org
WU: http://www.wu.ac.at

Articles

REGION
The Journal of ERSA Powered by WU

Volume 7, Number 2, 2020, 1–7
DOI: 10.18335/region.v7i2.304
journal homepage: region.ersa.org

Risky Move: New Evidence on the Determnants of the Willingness to Migrate

Geraint Johnes[1]

[1] Lancaster University Management School, Lancaster, United Kingdom

Received: 6 February 2020/Accepted: 8 September 2020

Abstract. Data from a bespoke Totaljobs survey of workers in the United Kingdom are used to revisit issue of workers' willingness to migrate in order to enhance their career opportunities. Demographic variables such as age, gender, and family circumstances are found to have high explanatory power. Education is also an important cofactor, as is the individual's current income – though the latter has a highly nonlinear effect. Workers located in the north east – a region relatively remote from other large population centres, and one with a strong and distinct cultural identity – are significantly less likely to express a willingness to move. The paper is novel in two respects: in identifying the role played by individual income in mobility, and in allowing for the potential endogeneity of variables associated with attitudes to risk-taking.

JEL classification: J61

Key words: migration

1 Introduction

Imbalances between regions within a country have long been a matter of concern to economists (Gabriel et al. 1993, Bentivogli, Pagano 1999). Where some regions are operating close to capacity while others are characterised by slack, there is potential for aggregate economic outcomes to be improved by a rebalancing of regional activity (Gardiner et al. 2013). Labour migration is a key mechanism for achieving such rebalancing (Rowthorn 2010), but, given both the financial and psychic costs of mobility, and given rigidities in the housing market, it is not clear that this can be relied upon to achieve rapid convergence across regions. Indeed, evidence on the rate of convergence of regions is indicative of "significant barriers to factor mobility within countries" (Gennaioli et al. 2014).

A prerequisite for migration is that individuals should be willing to move. There are few data sources that allow this to be examined, however, and so most studies in this area concern actual rather than potential migration. In this paper, we use a novel data set that provides information about workers' willingness to move between regions in the United Kingdom. This allows us to build on previous work conducted rather a long time ago, and in particular enables application of contemporary methods of analysis. To be specific, our research question is this: what factors are associated with individuals' willingness to move, and is there any evidence to support a causal interpretation?

The remainder of the paper is structured as follows. In Section 2, we provide a brief review of the relevant literature. Section 3 introduces the data set used in the analysis, while Section 4 documents the results. Conclusions are drawn in Section 5.

2 Literature

Arguably the most direct ancestor of the work reported here is the seminal paper by Hughes, McCormick (1985). Using General Household Survey data, they find that potential migration rises with education and occupational status and falls with age. They also find results concerning housing tenure that have considerable policy importance; council tenants (locked into provision by a specific local authority) are least likely to express willingness to move, while private tenants are the most likely, with owner-occupiers being in the middle. This last finding has stimulated a literature on the role played by housing tenure in exacerbating labour market rigidities; the Oswald (1996) hypothesis of a positive relationship between owner-occupation and unemployment rates has subsequently been investigated by, inter alia, Munch et al. (2006), Battu et al. (2008), Blanchflower, Oswald (2013) and Laamanen (2017).

Another important paper that has focused on potential mobility is due to Drinkwater, Ingram (2009). In common with earlier work, this finds that mobility increases with education and declines with age. It also finds significant gender effects (with women being more reluctant to move than men), and some evidence on the importance of household composition (for example with widowed or divorced respondents being more willing to move). Many of these effects appear to be robust across more than 20 countries studied in their analysis, which uses data collected as part of the International Social Surveys Programme.

Unfortunately, more recent work on potential, as opposed to actual, migration between regions has often been frustrated by a lack of data. Moreover, neither the Hughes and McCormick nor the Drinkwater and Ingram papers make use of data on individuals' incomes, likely to be a key determinant of potential mobility. In the next section, we discuss the data used in the present study.

3 Data

Our data were collected by Opinium[1] in January 2019 as part of the "Northern Pound" study conducted by Red Consultancy[2] on behalf of jobs board, Totaljobs[3]. Some 1821 individuals located in and around 9 British cities (London, Liverpool, Manchester, Sheffield, Leeds, York, Newcastle, Glasgow and Edinburgh) were surveyed online. Respondents are screened so that all are in full-time work[4]. The survey comprises a wide range of questions concerning respondents' demographic characteristics (age, gender, marital status, household composition etc.), educational background, industry in which employed, income, housing tenure, commuting practices, and patterns of expenditure. It also gathers information about respondents' willingness to move (both in principle and for various levels of financial inducement) [5]. Finally, the survey collects data about the time spent on activities outside of work. Data on income and gender are incomplete for a small number of respondents, and this leaves 1707 observations that are used in the analyses that follow. To ensure representativeness, we weight respondents by city population data[6].

Descriptive statistics for the variables used in the analysis that follows are reported in Table 1. Willingness to move is self-reported (and coded 1 if the respondent is willing, and

[1] https://www.opinium.co.uk/

[2] http://bit.ly/2QZglaa

[3] https://recruiting.totaljobs.com/northern-pound. The data are used here with permission of Red Consultancy and Totaljobs.

[4] These workers may be more likely than others to invest in migration, since they have greater opportunity to recoup the financial costs of the investment in the move. The results reported here therefore reflect the willingness to migrate only of those who select themselves into full-time work.

[5] As dependent variable, we use a binary indicator of willingness to move. Some of those indicating such willingness may be prepared to move for no wage gain, while others would do so only for substantial wage gain.

[6] http://bit.ly/2QTpXTS

Table 1: Descriptive statistics

variable	mean	standard deviation
willing to move	0.7253	0.4465
male	0.5442	0.4982
age	43.94	11.48
income (£ pa)	36409.8	22605.0
married	0.4862	0.5000
number of children	0.6497	0.9365
education	0.6087	0.4882
suburban	0.9350	0.2466
mortgage	0.4897	0.5000
friends	0.1793	0.3837

zero otherwise), and indicates a willingness in principle; thus, respondents who declare themselves willing to move may only be prepared to do so if rewarded by a substantial (say, £20000) increase in annual earnings. This being the case, it is perhaps not surprising to find that over 70% of respondents are prepared to relocate. Nevertheless, a substantial minority are not.

The gender mix of the sample is reasonably representative, with somewhat more than half of all respondents being male. Recall that the sample comprises only full-time workers, and this likely accounts for the fact that women are less likely than men to be included. The average age is around 44 years. Mean income is around £36000, very close to the average for full-time workers reported by the Office for National Statistics (http://bit.ly/388ewhh). Just under one half of respondents are currently married, and the number of children living with the typical respondent is quite low – though the range is quite wide, with some respondents living with five children.

Education is coded 1 if the respondent has achieved education at least to the level of a certificate of higher education, and zero otherwise; on this metric, just over 60% of respondents have a high level of education. A substantial majority, some 94%, of respondents live within a 90-minute commute of their place of work, and we describe this in the table as "suburban". Almost one half have a mortgage. Finally, we find that some 18% of respondents spend 5 or more hours per week socialising with friends; we hypothesise that these may be less likely than others to be willing to move because of the roots they have in their local communities. In the next section a more formal analysis of these data is presented.

4 Analysis

The dependent variable used in our analysis is the willingness to move. Logit results (weighting observations by the population of the urban area from which they are drawn) and the corresponding marginal effects for a variety of specifications are reported in Table 2. Reading from left to right, the columns of this table report a basic model, our preferred model, and the preferred model augmented by a full set of city dummies. In the preferred model, the city dummy for Newcastle is retained because it is, or is close to being, statistically significant at conventional levels in all specifications; none of the other city dummies achieves a z value as high as one. The coefficient estimates and marginal effects are all reasonably robust across all specifications.

Males are likelier than females to be willing to move, by a large margin. This may reflect differences in risk-taking propensities across gender (see, for example, Anbarci et al. 2016; but, for an important caveat, see Booth, Nolen 2012), or may be a reflection of perceived historic gender roles. As respondents age, they become less likely to be willing to move. This may be attributable to their roots in a locality, but it may also reflect the fact that the flattening of the income-age relationship over time offers diminished pecuniary incentives to mobility as people age.

Income enters the model in nonlinear form. The turning point is at a very low level

Table 2: Logit results

variable	coeff.	marg. eff.	coeff.	marg. eff.	coeff.	marg. eff.
male	0.8587	0.1297	0.7962	0.1085	0.7901	0.1075
	(3.39)	(3.24)	(3.22)	(3.07)	(3.18)	(3.01)
age	-0.0746	-0.0106	-0.0633	-0.0081	-0.0622	-0.0080
	(6.37)	(6.76)	(5.27)	(5.07)	(4.98)	(4.62)
income ($\times 10^{-6}$)	-42.1	-5.99	-46.8	-5.99	-47.1	-6.03
	(2.15)	(2.15)	(2.36)	(2.40)	(2.37)	(2.43)
income^2 ($\times 10^{-7}$)	37.6	5.34	36.4	4.67	36.4	4.67
	(2.12)	(2.15)	(2.06)	(2.12)	(2.06)	(2.13)
number of children			-0.3608	-0.0463	-0.3643	-0.0466
			(1.13)	(1.12)	(1.14)	(1.13)
number of children			1.0046	0.1288	1.0095	0.1293
$\times$ suburban			(2.93)	(2.96)	(2.93)	(2.96)
education			0.6010	0.0859	0.5839	0.0831
			(2.24)	(2.06)	(2.18)	(1.98)
married			0.0576	0.0074	0.0654	0.0084
			(0.22)	(0.22)	(0.25)	(0.25)
mortgage			-0.3256	-0.0425	-0.3113	-0.0406
			(1.29)	(1.31)	(1.21)	(1.22)
friends			-0.4654	-0.0657	-0.4710	-0.0665
			(1.65)	(1.53)	(1.66)	(1.54)
Newcastle			-0.4806	-0.0718	-0.5314	-0.0805
			(2.19)	(1.97)	(1.85)	(1.68)
constant	4.8623		4.1495		4.1596	
	(7.39)		(6.04)		(6.03)	
pseudo R2	0.1271		0.1820		0.1829	
number of observations	1707	1707	1707	1707	1707	1707

Notes: z values in parentheses. The specification in the last two column includes a full set of city dummies (not reported for reasons of space). "coeff." = "coefficient", "marg. eff." = "marginal effect".

of income, so the interesting feature of the nonlinearity concerns the relatively modest impact of income on willingness to move at low income levels, contrasted with a much higher (positive) impact at higher levels. Higher income respondents are likely more able than others to bear the costs of migration, and may also perceive greater economic returns to that migration.

The presence of children in the household reduces the likelihood that a respondent is willing to move, but the effect is not significant at conventional levels. For those who live within a 90-minute commute of their place of work, however, having more children is associated with a greater willingness to move. This might reflect a desire to access greater living space or superior amenities.

Those educated to Certificate of Higher Education or beyond are markedly more willing to move than others, with a marginal effect of over 0.08. As with the gender variable, there may be differences between highly educated individuals and others in attitudes to risk. This issue is explored further in the sequel.

Marital status has no significant effect. Other variables that are significant only at generous levels include owner-occupation with a mortgage (those with a mortgage may be less likely to be willing to move) and time spent socialising with friends (those spending 5 hours a week or more are less likely to be willing to move).

Finally, those currently living in or around Newcastle (a city in the north east of England, fairly remote from other major centres of population in the country) are less likely to be willing to move than other respondents. This effect is significant in some specifications at the 5% level, and is borderline significant in others. Numerically, the effect is substantial, with a marginal effect of between 0.07 and 0.08.

Table 3: Average treatment effects

variable	nearest neighbour coefficient	nearest neighbour z	IPWRA coefficient	IPWRA z
educated	0.0360	1.37	0.0460	1.95
male	0.0838	3.31	0.0808	3.71
married	-0.0029	0.10	-0.0280	1.17
mortgage	-0.0383	1.52	-0.0152	0.71
friends	-0.0065	0.20	-0.0206	0.75
newcastle	-0.0755	1.63	-0.0781	2.19

As noted earlier, attitudes to risk are likely to influence the dependent variable and some of the explanatory variables. Establishing that the link between, say, education and the willingness to move is causal therefore requires further analysis. Our data lack any intervention that can be used as a discontinuity in order to establish causality, and so we appeal to matching methods[7]. In Table 3, we report on the average treatment effect (ATE) associated with education in a logistic propensity score nearest neighbour matching model with one match per observation[8]. This has a positive value, indicating that education does indeed impact positively on willingness to move; however it falls short of statistical significance, and the numerical value of the effect is somewhat less than half that of the corresponding marginal effect in the analysis reported in Table 2. It would be heroic therefore, on the basis of this evidence, to conclude that education has a causal effect on willingness to move.

For completeness, we also report in Table 3 the ATEs associated with other binary explanatory variables. In almost every case the magnitude of the effect is smaller than reported in Table 2; the exception is the Newcastle dummy. Only the gender dummy has an ATE that is significant at conventional levels.

There are many variations on the simple propensity score matching model, and so as a robustness check Table 3 reports also results obtained from an inverse probability weighted regression adjustment (IPWRA) model. The results are broadly similar, but in this case the coefficient on Newcastle is significant at better than 5%.

Typically, earlier analyses of the willingness to move have not provided a causal analysis; while recognising that matching methods do not provide a panacea for this deficiency, the weakening of the coefficients reported here nevertheless suggests that the results of these earlier analyses should be treated with caution.

5 Conclusion

In response to the promise of higher earnings, a high proportion of workers are willing to consider migrating within country. Nevertheless, this willingness varies systematically across respondents according to demographic characteristics. Furthermore, we find that income and education are important influences on the willingness to move; in the case of education, the size and the significance of the effect is weakened somewhat when estimated using matching models.

The data set used in this study is unusual in that it was conducted on behalf of a private organisation interested in researching labour markets. The sampling strategy involved a focus on a number of cities around the United Kingdom. While the statistical analysis conducted in the paper uses information on the geographical distribution of respondents to weight observations, future work would benefit from use of a truly representative sample drawn from across the country.

[7] See Stuart (2010) for a discussion of how matching models allow a causal interpretation. She notes the recent development of methods such as coarsened exact matching (Iacus et al. 2012) that can improve balancing between treated and untreated groups. Application of such methods are not considered here owing to the preponderance of binary variables in the analysis.

[8] This is estimated using the teffects psmatch command in Stata, with default values of options, and with all variables included in the preferred model of Table 2 used to predict treatment assignment.

Several of our findings have policy relevance. Gender differences in willingness to move are unsurprising, but may nonetheless cause inefficient allocation of resources across an economy. The relative reluctance of women to move may be due in part to longstanding social norms, and continued efforts to promote women's aspirations as they progress through education and into the labour market may serve to weaken this reluctance. More generally, increased education (which is borderline significant in the IPWRA estimates) may promote mobility. The finding that those located in and around Newcastle are less likely than others to contemplate relocation is particularly interesting and suggests that policies aimed at disseminating information about positive opportunities elsewhere may be needed if further mobility is to be promoted.

Potential mobility is not, of course, the same as actual mobility. Those who are willing to move may not do so for a variety of reasons, including weak financial incentives. It is nonetheless instructive to examine the factors that influence such willingness. The survey used in the present note is unusual in that it enables research on this issue. It is hoped that further data collection will allow further insights to be gained in this space.

References

Anbarci N, Lee J, Ulker A (2016) Win at all costs or lose gracefully in high-stakes competition? Gender differences in professional tennis. *Journal of Sports Economics* 17: 323–353

Battu H, Ma A, Phimister E (2008) Housing tenure, job mobility and unemployment in the UK. *Economic Journal* 118: 311–328

Bentivogli C, Pagano P (1999) Regional disparities and labour mobility: The Euro-11 versus the USA. *Labour* 13: 737–760

Blanchflower D, Oswald A (2013) Does high home ownership impair the labor market? IZA Discussion Paper, available at http://bit.ly/2QSMFMc, accessed 6 February 2020

Booth AL, Nolen P (2012) Gender differences in risk behaviour: Does nurture matter? *Economic Journal* 122: F56–F78

Drinkwater S, Ingram P (2009) How different are the British in their willingness to move? Evidence from international social survey data. *Regional Studies* 43: 287–303

Gabriel S, Shack-Marquez J, Wascher W (1993) Does migration arbitrage regional labor market differentials? *Regional Science and Urban Economics* 23: 211–233

Gardiner M, Martin R, Sunley P, Tyler P (2013) Spatially unbalanced growth in the British economy. *Journal of Economic Geography* 13: 889–928

Gennaioli N, La Porta R, Lopez De Silanes F, Shleifer A (2014) Growth in regions. *Journal of Economic Growth* 19: 259–309

Hughes G, McCormick B (1985) Migration intentions in the UK: Which households want to migrate and which succeed? *Economic Journal* 95: 113–123

Iacus S, King G, Porro G (2012) Causal inference without balance checking: Coarsened exact matching. *Political Analysis* 20: 1–24

Laamanen JP (2017) Home ownership and the labour market: Evidence from rental housing market deregulation. *Labour Economics* 48: 157–167

Munch JR, Rosholm M, Svarer M (2006) Are homeowners really more unemployed? *Economic Journal* 116: 991–1013

Oswald A (1996) A conjecture of the explanation for high unemployment in industrialised nations. Warwick university economic research paper, available at http://bit.ly/2QYBEsq, accessed 6 February 2020

Rowthorn R (2010) Combined and uneven development: Reflections on the north-south divide. *Spatial Economic Analysis* 5: 363–388

Stuart E (2010) Matching methods for causal inference: A review and a look forward. *Statistical Science* 25: 1–21

REGION
The Journal of ERSA Powered by WU

Volume 7, Number 2, 2020, 9–27
DOI: 10.18335/region.v7i2.271
journal homepage: region.ersa.org

Strengthening local economy – an example of higher education institutions' engagement in "co-creation for sustainability"

Anna Bogedain, Rüdiger Hamm[1]

[1] Hochschule Niederrhein, Krefeld, Germany

Received: 24 July 2019/Accepted: 2 November 2020

Abstract. Major societal challenges like energy efficiency, climate change and resource scarcity trigger and influence continuous change processes worldwide, nationwide, but also on all regional levels. They force regions to think about (a more) sustainable development. As the transformation processes necessary for sustainable development are complex there is a need for actors willing to engage and support sustainability transitions. Higher education institutions (HEIs) are often expected to be one of these supporters on the regional level. The central aim of this paper is to show by the use of an example, that HEIs are able to provide impulses for sustainable transformation. Following Pflitsch, Radinger-Peer (2018) HEIs can play different roles in regional sustainable transition; the authors use two dimensions to distinguish these roles – depth and autonomy:

- As to depth HEIs' roles can be "comprehensive, involving diverse actors and approaching sustainability with a holistic perspective" or "more fragmented and passive, but also more focused on specific topics".
- As to autonomy the roles can be "autonomous, the university defining its own focus and priorities through interacting with a broad range of regional actors" or "more directed, the university working on topics that are relevant from the perspective of the regional or federal-state government".

Using this rough classification the HEI in our example focuses on a "specific topic" and it is interacting with other regional respectively local actors on a topic that is not only relevant from the perspective of most German cities and their citizens but also from the national and federal government's perspective. The paper starts with a short systemisation of transfer channels and missions of HEIs. It starts with a description of transfer channels used by the two traditional missions of HEIs – education and research. Afterwards the concept of "third mission" is introduced and distinguished from a possible fourth mission of HEIs – the concept of "co-creation for sustainability". Afterwards it deals with important concepts and approaches which are characteristic elements of "co-creation for sustainability" – transformative research, participatory action research (PAR), urban living labs and student service learning. The "specific topic" that serves as an example is introduced after that: it is about the problem of local economy in urban neighbourhoods. Local economy will be defined, its problems resulting from the functional change of urban neighbourhoods are sketched and the arising necessity of strengthening local economies will be discussed. We show the methodological concept that is used to develop strategies and specific measures for strengthening local economy. The paper

shows that the elements of this concept are typical approaches of transformative sciences. Afterwards concrete examples stemming from an urban neighbourhood which is part of the city of Viersen (Northrhine-Westphalia, Germany) is used to show, how the approach works in practise. The paper ends trying to explain, why projects like this give an example of HEIs' impulses for sustainable development in their regional surrounding. Furthermore, the usefulness, but also the shortcomings and further research necessities of the approach will be discussed.

1 Introduction

Major sustainability topics like energy systems, climate change, pollution, resource scarcity and economic inequality trigger and influence continuous change processes worldwide, nationwide, and at the regional level. They force regional decision makers to think about (a more) sustainable development. As the transformation processes necessary for sustainable development are complex, there is a need for actors willing to engage in and support sustainability transitions.[1] Higher education institutions (HEIs) are often expected to be one of these supporters on the regional level. The central aims of the following paper are to show by use of a case-study that HEIs are able to provide impulses for sustainable transition as well as to describe the research approach and methods used to generate these impulses, comparing the experiences of the case-study on hand with findings from other case studies.

The next section (Section 2) gives a short historical outline of the research interests regional sciences had and still have in the context of regional effects of HEIs. It starts with a description of transfer channels used by the two traditional missions of HEIs – education and research. Moreover, the concept of "third mission" is introduced and distinguished from a possible fourth mission of HEIs – the "co-creation for sustainability". Section 3 deals with important methodological concepts and approaches which are characteristic elements of "co-creation for sustainability" – transformative research, participatory action research (PAR), urban living labs and student service learning. A specific example about the problem of "local economy" in urban neighbourhoods is introduced in Section 4. Local economy is defined; its problems resulting from the functional change of urban neighbourhoods are sketched, and the arising necessity of strengthening local economies is discussed. Section 5 starts with a description of a methodological concept used to develop strategies and specific measures for strengthening local economies. The paper shows that the elements of this concept are typical approaches of transformative sciences. Afterwards concrete examples stemming from an urban neighbourhood in the city of Viersen (North-Rhine Westphalia, Germany) are used to show how the approach works in practice. The paper ends with a concluding summary (Section 6) explaining why projects like this illustrate an example of HEIs' impulses for sustainable development at the regional level. Furthermore, the usefulness, but also the shortcomings of and further research necessities for the approach are discussed.

2 Regional Effects of HEIs – From Demand-Side Effects to Co-Creation for Sustainability[2]

For about fifty years regional sciences have been dealing with research on regional (economic) effects of HEIs. HEIs' missions and the channels through which they can provide impulses to their respective regions have been focal points of this research. Three research directions were pursued consecutively. The first line focused on effects HEIs

[1] As in the literature, the terms "transition" and "transformation" are used as synonyms in this paper "to express the ambition to shift from analysing and understanding problems towards identifying pathways and solutions for desirable environmental and societal change." (Hölscher et al. 2018, 1). The authors further argue (Hölscher et al. 2018, 1-3) that "transition" is used by the sustainability research community when talking about fundamental social, technological, institutional and economic change, but the term is mainly employed to analyse changes in societal sub-subsystems (e.g. energy, mobility, cities). Researchers concerned with global environmental change normally use transformation to refer to large-scale changes in whole societies, which can be global, national or local.

[2] The following considerations are partly based on Hamm, Koschatzki (2020)

achieve through their own economic activities; research concentrated almost exclusively on impulses a university gives to its regional environment solely through its existence (overviews can be found in Voß 2004, Stoetzer, Krähmer 2007). These studies showed that universities have a considerable impact on regional employment and income. Meanwhile, the scientific interest in this pure estimation of so-called demand-side effects generated by HEIs' employees, students, expenses and investments has declined considerably, as the methods and limits of such analyses are widely known.

The second line of research led to a broadening of the perspective. The demand-oriented approach was extended to include the effects of HEIs via supply-side connections. This shift was caused by the increasing diversity of universities' tasks, the orientation towards the "Entrepreneurial University" transfer model (outlined by Clark 1998, Gibbs 2001) and the development of the Triple-Helix model (Abramson et al. 1997, Etzkowitz, Leydesdorff 1995). In this context, academic spin-offs (Koschatzky, Hemer 2009, Stahlecker 2006) and scientific analyses of the research and innovation system played an important role. In the meantime, a multitude of analyses, dealing with supply-side effects of HEIs theoretically and empirically, have considerably improved the state of knowledge on their regional economic effects.

The third line of research represents a further broadening of the perspective. So far, the focus has been exclusively on the regional economic effects of HEIs. In accordance with the model of the "engaged university", all kinds of impulses stemming from HEIs and having an impact on society at the regional level are considered. The term "transfer" has therefore been extended to include not only the economically relevant transfers, but any kind of regional transfer from HEIs into society. This can be seen in analyses of regional innovation systems (e.g. Asheim, Gertler 2005, Cooke 1992, 2002), which particularly address the role of the research sector. The societal dimension of regional transfer activities reflects actions that are geared to a broad regional commitment of universities and their members (employees and students). Such activities are often carried out without particular cooperation partners, but aim at specific target groups or general contributions to social life in a region (see Koschatzky et al. 2013); examples include students' consultancy projects, universities for children, pupils and senior citizens, as well as open knowledge transfer events. Thus, regional social commitment becomes another significant transfer channel for universities. It complements transfer activities of the "entrepreneurial university" focused on education, research and economic transfer with facets that go beyond the previous university commitment and are reflected in the "engaged university" (see Breznitz, Feldman 2012), meaning the global civic engagement of higher education institutions and their efforts to use intellectual resources to tackle societal issues.

Three main questions arise as a result of an extension related to this third line of research. The first research question is concerned with which socially relevant effects universities might have on their regional surrounding as by-products of fulfilling their traditional core tasks – research and education. In fact, in addition to economic effects, HEIs can trigger a multitude of socially relevant effects on their regional environment (see Henke et al. 2016). The second questions deals with whether HEIs have a third core task in addition to education and research. Meanwhile HEIs' "third mission" is generally accepted in literature. It deals with reciprocal interactions between the university and the particular region where it is located (Roessler et al. 2015). Furthermore, Henke et al. (2016) stress co-operation with non-academic partners aiming at societal interests as other necessary elements of the "third mission". The "third mission" bundles all kind of services that lead to a beneficial integration of the higher education institution into its extramural environment through mutual interactions.

The identification of social and economic impulses running from HEIs to their regional environments and considerations for optimising these impulses by the "engaged university type" finally led to the third follow-up research question: Are universities also able to provide impulses for sustainable transformation processes in their regional environments? There is little doubt that HEIs can increase the flexibility of regions to adapt to structural economic change and thus provide an impetus for economic transformation (Pinheiro et al. 2015, Zomer, Benneworth 2011, Zawdie 2010). This question is new, however, due to its regard of transformation as change towards ecological, economic and social sustainability.

The said question specifically examines whether universities can provide important impulses to support sustainable transformation via their three missions (education, research and the so called "third mission") or whether they can act as a change agent for regional sustainable development in the first place. According to Stephens et al. (2008), all challenges related to sustainable transformation can be assigned to ecological, social and technological change. To meet the sustainability challenges, Stephens et al. (2008) call for a transition to more sustainable practices and lifestyles. In this process HEIs must be viewed from two sides: on one hand, they are themselves objects of transformation, but on the other hand, they can also take over the function of a driving force or even a change agent. In the latter case, HEIs are expected to have a considerable solution potential at the strategic, tactical and operational levels. At the strategic level, they can participate in the development of long-term social visions. At the tactical level, they can initiate and strengthen cooperation among the regional stakeholders. Moreover, they can advance the desired transition at the operational level by changing the orientation of education, research and transfer and by internal sustainability efforts of their own. Stephens et al. (2008) distinguish four concrete operational categories in which universities can support a transformation towards sustainability:

1. HEIs can help to ensure that they themselves are perceived as integrated, transdisciplinary agents.

2. HEIs can carry out problem-driven research projects to overcome urgent sustainability challenges.

3. HEIs can develop sustainable solution-oriented practices for society and promote their implementation.

4. HEIs can provide participants with skills for holistic thinking and for coping with sustainability challenges in their study programs.

Trencher et al. (2014) also investigate the increasing involvement of HEIs in promoting the process of sustainable transformation in their regional environments in cooperation with governments and civil society actors. They call this commitment "co-creation for sustainability" and define it "as a role where the university collaborates with diverse social actors to create societal transformations with the goal of materialising sustainable development in a specific location, region or societal sub-sector" (Trencher et al. 2014, 152). It must be stressed that "co-creation for sustainability" clearly differs from the "third mission". While the latter contributes to economic and social development through transfer, "co-creation for sustainability" supports society in its quest for sustainable development (Trencher et al. 2014). Trencher et al. (2014) therefore suggest viewing "co-creation for sustainability" as an evolving new mission of HEIs. However, "co-creation for sustainability" should not become the sole focus of HEIs. As the three missions (education, research and transfer) already exist side by side in an "entrepreneurial university", reinforcing each other, they could coexist with a fourth one, namely the "co-creation for sustainability", in a "Transformative University" (Trencher et al. 2014).

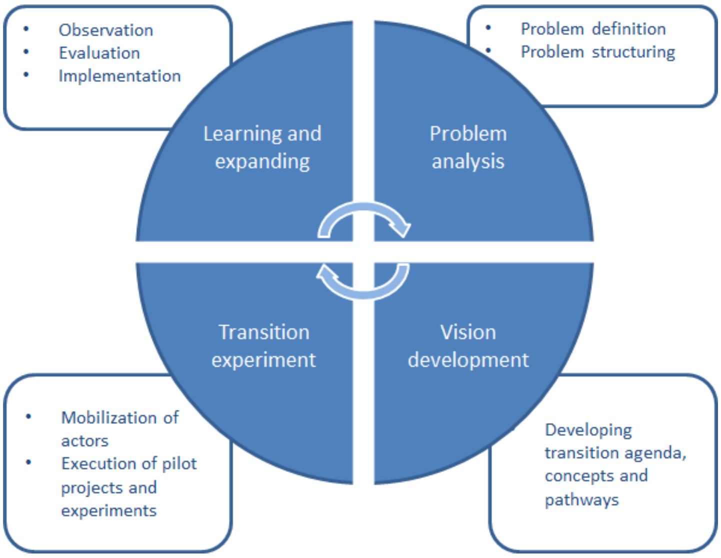
Partly based on Goldstein et al. (1995) and Uyarra (2010), the explanations given so far can be summarised into four groups of regional impacts also connected with the missions of HEIs:

1. Regional demand effects (resulting from HEIs' roles as economic actors)

2. Regional human capital effects and regional knowledge effects (resulting from HEIs' education and research activities).

3. Regional social and political effects (resulting from HEIs' "Third Mission")

4. Regional co-creation for sustainability effects (resulting from a possible "Fourth Mission").



Source: Own diagram following Loorbach (2010), Schneidewind, Scheck (2012), Schäfer, Scheele (2014).

Figure 1: The Transition-Cycle Model

3 Methods of co-creation for sustainability

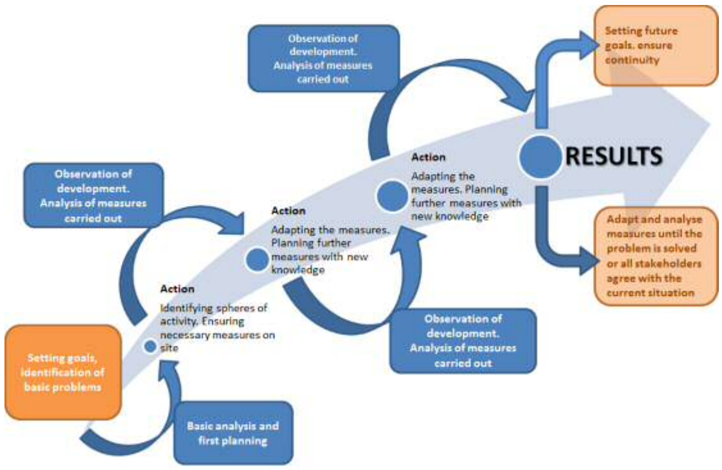
To implement co-creation for sustainability, new forms of science and knowledge production become necessary (Trencher et al. 2014); research agendas should have a greater societal relevance (Gibbons 1999), and scientific knowledge should be increasingly produced and applied in cooperation with extramural stakeholders. This explains why the methods necessary in transformation sciences and for the sustainable transformation process via co-creation for sustainability differ from those used in "traditional sciences". Following Trencher et al. (2014), this is illustrated by the move from 'mode 1' – knowledge production that is characterized by theory building and testing within one discipline – towards 'mode 2', where knowledge is produced for application. These new forms of knowledge production require inter- and transdisciplinarity, a high level of participation, a mutual interaction of scientific analysis and work on site, as well as a continuous reflection of the chosen measures followed by necessary improvements. These requirements can be fulfilled by methodological approaches such as transformative research, participatory action research, real-world laboratories and student service learning. These methodological approaches are presented in some more detail below.

3.1 Transformative Research

The German Advisory Council on Global Change defines transformative research as follows: "Transformative research supports transformation processes in practical terms through the development of solutions and technical as well as social innovations, including economic and social diffusion processes and the possibility of their acceleration, and demands, at least in part, a systemic perspective and inter- and cross-disciplinary methods, including stakeholder participation" (WBGU 2011, 322).

Transformative research is mostly based on the transition-cycle-model (Figure 1), which is characterized by four different phases (Loorbach 2010, Singer-Brodowski, Schneidewind 2014, Schneidewind, Scheck 2012, Schäfer, Scheele 2014):

- Problem analysis: The cycle starts with analysing the question: What is the current situation? Existing knowledge from different disciplines helps to form a "snapshot" of the socio-cultural, economic, institutional and ecological situations. System knowledge is necessary in this first phase.
- Development of transition agenda, visions and pathways: As there is a need to



Source: Own diagram following Walter (2009).

Figure 2: The PAR-Cycle

develop targets and future visions, the results of the first phase are used to answer the question: "What should the situation look like in the future?". To come up with adequate visions, all relevant stakeholders must be integrated. Future visions are generated depending on the requirements. With the help of transformative science, the chance for long lasting changes increases.

- Initiation and execution of transition experiments: Based on these visions, a third phase of experiments must follow. The central question is: Which measures should be taken to reach the desired outcome in the future? Arrangements are implemented and projects are carried out.

- Evaluation, learning and expanding: Finally, in the fourth phase of the cycle, learning effects should arise because the entire outlined process is continuously reflected on by supporters, critically evaluated and, if necessary, readjusted.

Transformative research uses research approaches focusing on collaborative and experimental learning by scientists and laypersons. One of these approaches is Participatory Action Research (PAR), which is described in the following section.

3.2 Participatory Action Research

The origins of Participatory Action Research can be traced to the work of the Prussian psychologist Kurt Lewin (1944), who is considered the founder of action research. Attwood (1997, cited from MaxDonald 2012) explains that PAR's philosophy embodies "the concept that people have a right to determine their own development and recognises the need for local people to participate meaningfully in the process of analysing their own solutions, over which they have (or share, as some would argue) power and control, in order to lead to sustainable development".

PAR means that activities should take the interests of people concerned into account. The activation of the citizens via e.g. activating surveys, initiation of networks and empowerment are essential elements of its operating principles. PAR is a qualitative research methodology that fosters collaboration among participants and researchers (MacDonald 2012). It is a concept that tries to find problem solutions through the interaction of participation and action. Those affected by the research results participate in the research. Research questions do not come from "outside" but are articulated by the affected people who also participate in the search for problem solutions. The aim is to bring about positive changes and to achieve the research objective by a bottom-up approach.

Figure 2 demonstrates that PAR (Greenwood, Levin 2007, Walter 2009) uses a multi-level, cyclical approach; it encompasses a "cyclical process of fact finding, action, reflection, leading to further inquiry and action for change" (MacDonald 2012, 37). The detailed steps of the cyclical approach are as follows:

- A problem is identified by the persons involved.
- The involved community cooperates with researchers. Involved supporters and researchers act jointly, developing ideas to solve the identified problem and creating a plan.
- The plan is implemented.
- The results of the implementation are monitored by the researchers and the community.
- The final stage of the first cycle is the reflection on the results. If the actors involved are satisfied with these results, the described process of planning, action, observation and reflection will continue, building on the successful outcome. If the first action is deemed ineffective, the evaluation will influence the actions planned for the next cycle.
- This process is repeated or continued as often as necessary until the problem at hand is solved or the desired aims are achieved.

3.3 Urban Living Labs as an example of Real-World Labs

Real-world labs are another methodological element of co-creation for sustainability. A real-world lab refers to a social context in which researchers carry out "real experiments" in order to learn more about social dynamics and processes (Schneidewind 2014). Real-world labs are places of learning with different types of impact (Schneidewind et al. 2016):

- they create solutions for actual problems,
- they serve as a testing ground for the created solutions and
- they can facilitate transferability of solutions to other contexts.

If HEIs are engaged in real-world labs, they should develop place-based problem solutions and help test them in a real world environment, cooperating with society instead of working in the scientific ivory tower.

In the context of this paper, urban living labs (ULLs) are a relevant example of these real-world labs. Cities are of particular interest as places for real-world labs for at least three reasons (Schneidewind 2014). First, social experiments have a long tradition in urban science. Secondly, cities are interesting objects of reference and experiments as the entire socio-technical structure of a modern society can be found there. However, cities are less complex and therefore easier to analyse than countries. And lastly, cities are often the starting points for all kind of changes. This probably explains why real-world lab research approaches (urban living labs) are increasingly used in urban research. Three levels of urban real-world laboratories can be distinguished (Schneidewind 2014): the household level (households or blocks of flats), the neighbourhood level (urban neighbourhoods) and the city level (entire city). In the special case described in this paper, ULLs work on the level of neighbourhoods; this means the examination room remains manageable for the research process.

ULLs are seen as a new form of intervention responding to the social, economic and environmental challenges in the urban context, thus contributing to the achievement of their sustainability goals. They are defined as panels "for innovation, applied to the development of new products, systems, services and processes, employing working methods to integrate people into the entire development process as users and co-creators, to explore, examine, experiment, test and evaluate new ideas, scenarios, processes, systems, concepts and creative solutions in complex and real contexts" (Bulkeley et al. 2017, 13).

Universities and researchers using ULLs are often initiators of sustainable development in disadvantaged neighbourhoods. This means that researchers carry out experiments in order to learn about social dynamics and processes. Research institutes collaborate with politicians, the private sector, and civil society groups in this approach. ULL approaches are always "place-based", and they aim at strengthening relevant stakeholders through actions and activities.

3.4 Student Service Learning

Student service learning is the last methodological element of co-creation for sustainability to be introduced here. It engages students in active, relevant and collaborative learning (Bringle, Hatcher 2000). Bringle, Hatcher (1995, 112) "consider service-learning to be a course-based, credit-bearing educational experience in which students (a) participate in an organized service activity that meets identified community needs and (b) reflect on the service activity in such a way as to gain further understanding of course content, a broader appreciation of the discipline, and an enhanced sense of civic responsibility." Service learning focuses on the service being provided and the learning that is occurring. Accordingly, service learning is designed in such a way that both service aspects enhance learning, while learning processes enhance service in an integrated way. On the one hand, service learning enables students to gain new knowledge and competencies as active service providers, and on the other, the outcomes of the service activity facilitate changes towards sustainability (Adomßent et al. 2014). If the topic of student service learning is connected to sustainability in the context of a municipality or neighbourhood, it can also be seen as an example of HEIs' increasing engagement in community outreach and aspects of societal transformation.

4 Functional Change of Urban Neighbourhoods and Consequences for Local Economy

After the introduction of relevant transformative science methods, the following considerations help describe the example of the paper on hand. Migration, demographical changes, changes in mobility behaviour and digitalization are important global trends causing and influencing adaption processes on all regional levels. This also applies to cities, urban districts and urban neighbourhoods. During these processes urban neighbourhoods often lose their previous functions in city structures. Finding new functions is often problematic and can become a long-lasting process. Economic and social problems are often a consequence for the residents and the local economy in those neighbourhoods.

The local economy in this context means all economic activities related to the development of a certain urban neighbourhood (Birkhölzer 2000). The firms behind the local economy are primarily small enterprises, and they consist of retailers, bars and restaurants, handicrafts, as well as social and household-oriented services. They are placed in and highly connected to the local neighbourhood. They fulfil different functions for their neighbourhood: they supply people living in the urban neighbourhood with everyday commodities; they offer opportunities for work; they are a place of communication and integration; and they upgrade the neighbourhood's living conditions (Jakubowski, Koch 2009, Henn 2013). As the local economy does not only have a pure economic function for the neighbourhood and the people who live there, but also fulfils social functions for the residents, securing the provision of basic services, it should be preserved even in times of increasing digitalization and online trade – perhaps in a modified form. A process of social and economic transformation becomes necessary.

Many German cities and their urban neighbourhoods face this problem. The German government has recognised this and initiated programmes to support urban neighbourhoods suffering from problems like this. BIWAQ (stands for "Education, Economy and Work in Neighbourhoods") is one of these programmes. It is funded by the European Social Fund (ESF) and the German Ministry of the Interior. The aims of BIWAQ are twofold – to improve job perspectives of long-term unemployed persons and to strengthen the local economy (see BMUB 2016). In case of BIWAQ-funded projects, municipalities must apply for the funding. In case of a successful application, municipalities usually cooperate with different partners and institutions (e.g. non-profit organizations and sometimes HEIs) and define subprojects with concrete objectives and measures. The institutions involved in the implementation of BIWAQ-funded projects face the challenge of developing a strategy and appropriate measures as well as actions to strengthen the local economy. Thus, they support a process of social and economic transformation or the preservation and strengthening of existing structures.

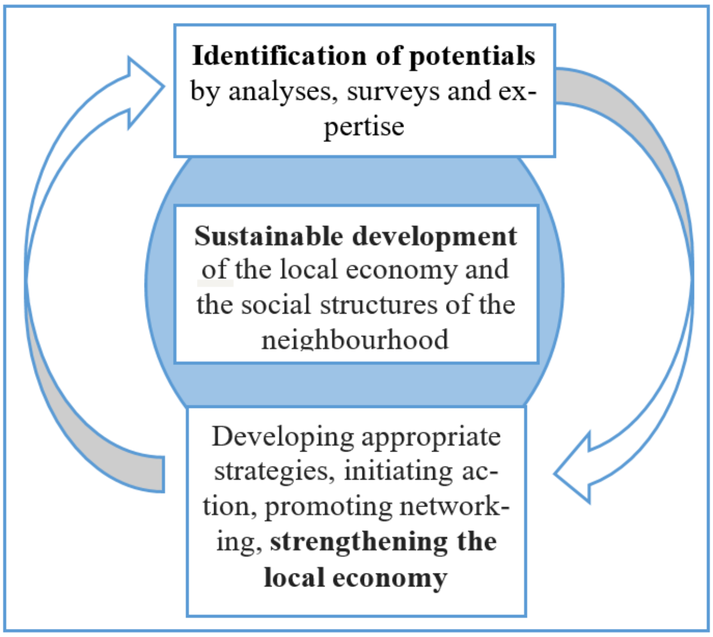
5 Case Study: Strengthening Local Economy in the "Viersener Südstadt"

5.1 The Scientific Approach in Use

Between 2011 and 2018, NIERS and SoCon – two research institutes of Niederrhein University of Applied Sciences (NUAS) – have been engaged in three BIWAQ-funded projects. These projects aimed at strengthening the local economy in disadvantaged urban neighbourhoods of North-Rhine Westphalian cities (Viersen, Solingen and Leverkusen). In addition to the heads of projects (a political scientist and an economist), two research assistants (economic geographers, social scientists and economists) were working on each of these BIWAQ-funded projects. One of these two researchers was mainly responsible for carrying out analytical tasks necessary for developing strategies and place-based measures (e.g. stocktaking by socio-economic indicators, surveys, guideline interviews, mapping). The other researcher mainly worked as a "business manager" in the neighbourhood; he visited and advised companies. Both researchers worked closely together, being in constant contact with the project management; each researcher was always able to take over the tasks of the other. If more than one project was being worked on at the same time, the researchers were also in constant exchange with one another. The authors of this paper had been engaged in these projects as research assistant and head of project, respectively.

When the first project started, the NUAS-researchers realized that supporting economic and social transformation processes of urban neighbourhoods by strengthening local economies suggests the use of a new and different scientific approach called "transformative science". Schneidewind et al. (2016) define transformative science as "a specific type of science that does not only observe and describe societal transformation processes, but rather initiates and catalyzes them. Transformative science aims to improve our understanding of transformation processes and to simultaneously increase societal capacity to reflect on them". Starting from this definition, a general scientific approach was developed, fulfilling some requirements:

- Neighbourhoods with a need to strengthen local economy often are "multiple burdened" – economic and social problems overlap and occur simultaneously; there are also health and ecological problems in many cases. Therefore, an interdisciplinary approach had to be chosen.
- Every neighbourhood is different. The specific problems of each district require innovative ideas that must be appropriate to the special situation of the quarter on hand. But in many cases the existing knowledge and information about the urban districts' situation is not sufficient to develop place-based perspectives, visions and measures for strengthening the local economy. A necessary condition for initiating concrete measures in the neighbourhoods is therefore learning more about the local economy's structure, about its strengths and weaknesses, as well as about the everyday problems of the respective neighbourhood. This requires analyses before action; scientific support ensures innovative measures adequate to the causes of existing problems.
- As suggested by the PAR-approach, the necessary research and measures to be implemented should not be determined by scientists from the outside; instead a practice-oriented approach is necessary. Research should get its questions from local actors (residents and firms). On-site cooperation with local actors should help scientists identify additional research questions and further needs for information to be analysed. Results should provide impulses to the districts and its actors by



Source: Own illustration.

Figure 3: The project approach

stimulating, discussing, possibly modifying, and finally initiating and implementing projects as well as concrete measures.

- Measures should be developed in participatory processes that take local needs into account and are supported by local actors, residents and entrepreneurs. Effects of measures should be communicated to and discussed with local actors who should be regarded as experts on "their" neighbourhood. This participation-oriented approach increases the likelihood of achieving long-term effects.

Based on these requirements and the theoretical pillars introduced in Section 3, an approach (Figure 3) similar to the transition cycle model in transformative research (see Figure 1 above) was used.

- In the first step, a baseline study of the urban district on hand was elaborated. The aim of the said study was to get more information on the local economic structures, on the residents' socio-economic situations as well as on the neighbourhood's image and conditions. This study was mainly based on secondary statistical data and document analyses. Activating firm and passers-by surveys as well as questionnaires for local stakeholders and external experts were used to provide further information. This information was used to identify strengths and weaknesses which are the basis for developing ideas, strategies and measures for strengthening the local economy. It was important to involve the local actors and their networks in this process. On one hand, they acted as experts of their neighbourhood, and on the other, they guaranteed that measures were in line with the wishes of local residents and entrepreneurs.

- The second step encompassed the development, coordination and implementation of concrete measures and projects for the local economy in the neighbourhood. A "commercial neighbourhood manager" was working on-site in order to initiate the developed projects and measures. The "commercial neighbourhood manager" specifically addressed the local actors and supported their networking activities. Other stakeholders (local authorities, chamber of commerce and industry, regional development agencies etc.) were integrated where necessary. Furthermore, the "commercial neighbourhood manager" formed the "link" between the different local actors and the researchers responsible for the scientific analysis, thus promoting the mutual interaction of analysis and on-site work in the neighbourhood.

The description shows that PAR is a basis of this scientific approach. The connection of the approach to the concept of urban living labs becomes clear when looking at some of the requirements mentioned by Schneidewind (2014):

Interdisciplinarity : in the case on hand research institutes from social sciences and regional economics are cooperating to strengthen local economy. The project team consisted of social scientists, business economists, regional economists, economic geographers, regional planners and political scientists.

Transdisciplinarity is guaranteed by continuous cooperation of the research team and local actors during the whole process of research and development.

Co-creation and co-production between the research team and the civil society is guaranteed in the research process via the participatory and activating elements.

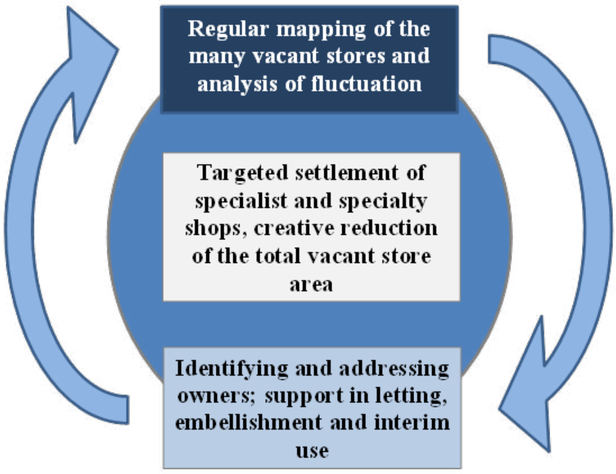
Continuous reflection of scientific methods and selected measures is ensured by discussions between the research team and local actors but also by discussions with external experts coming from different disciplines.

5.2 Examples of Concrete Action

The city of Viersen is located in the Central Lower Rhine Area (CLRA) in the west of the river Rhine, close to the Dutch-German border. With around 76,000 inhabitants (cf. City of Viersen 2016), Viersen is the capital of the Viersen district. The southern part of the inner city forms a neighbourhood with numerous stationary retailers having their shops on the ground floors; this neighbourhood is called "Viersener Südstadt". As a former part of the inner-city centre, "Viersener Südstadt" lost its importance for retail after the pedestrian zone development in the northern part of the inner city. Since then, the area has become a peripheral inner-city location that no longer belongs to the central supply area (Hagemann et al. 2011). In 2013, a total of 112 shops were located in the streets of this neighbourhood. Cityscape and population structure clearly show traces of the changing social conditions in the neighbourhood. As in many other German cities, the functional change becomes evident, e.g. due to the relocation of retail and services, increasing vacancies and investment backlogs as well as the refurbishment of buildings. (MWEBWV NRW 2014). The vacancy rate on the main shopping streets of "Viersener Südstadt" varied between 20 and 25% during the project period. The shops in the Südstadt often do not have barrier-free access and almost exclusively have less than 100 sqm of commercial space (Busch 2013).

In the course of the project, the service providers and retail stores in "Viersen Süd" were identified as the core target group. The vast majority of these shops were owner-managed and were specialised in particular, often somewhat unusual products. During the first step, information on the economic situation of the firms was ascertained with the help of an activating firm survey (problem analysis). The firms' opinions on the location quality factors in the neighbourhood, as well as on their future prospects, concrete needs for support and expectations connected with the BIWAQ-project were surveyed. A SWOT analysis (showing strengths, weaknesses, opportunities, and threats) and reflections of the survey results with experts from scientific and on-site practice formed the basis for developing measures to stabilise stationary retailing. Furthermore, measures were taken to raise the profile of the "Viersener Südstadt" as a location for specialized retail stores (vision development). All this was supported by another focal point of the project work: the implementation of a business-oriented on-site neighbourhood management, crosslinking (traditional) scientists using a new research-approach and practitioners from the neighbourhood.

The owners of the specialized retail stores as well as other (locally relevant) actors from existing networks of professional district developers from the urban environment – e.g. social planning, urban development, district management, city management – were



Source: Own illustration.

Figure 4: Examples of transformative methods in the project "Strengthening Local Economy in the Viersener Südstadt"

sensitized to the problems and visions of the local economy and mobilized to jointly advertise for the southern part of Viersen in a location community of specialty shops (experiments). Joint advertising measures were intensified and supported by targeted public relations work on the advantages of the location at the levels of the city administration, associations and the regional press. The left side of Fig. 4 illustrates the described interaction of analysis and work on-site in the concrete example.

The accompanying vacancy management (see right side of Figure 4) also followed this transformative research approach. Regular theme-based mapping (see Figure 5) revealed the need for action as well as documented fluctuations and categorised them into long-term and short-term vacancies (problem analysis). Having the local vision for retail trade described above in mind, an assessment of the development perspectives of the vacant retail store was made (vision development). In concrete terms, vacancy owners were contacted by the business-oriented on-site neighbourhood management in order to provide better information on the vacant stores to (specialist) shop-owners interested in settling there. Moreover, the possibilities to rededicate stores to housing were discussed with the city administration (experiment). Changes such as the settlement of further shops, the improvement of existing vacancies and the conversion of apartments were again observed, evaluated and controlled (learning and expansion).

Finally, student service learning became an integrated element of the project approach as well. Students from master courses at NUAS supported the team of researchers by conducting different types of surveys. These surveys, oriented towards specific problems of Viersener Südstadt and the existing lack of information at the local level, examined the image of the neighbourhood and – later on in the project – whether the neighbourhood had succeeded in positioning as a location for special shops (learning and expansion). On one hand, the master's students had to provide a relevant service job for the city of Viersen, facilitating changes towards sustainability. On the other hand, the students gained new knowledge and competencies – about empirical methods, surveys and the interpretation of their results regarding the economic and social situation of the Viersener Südstadt as well as the general social and economic consequences of functional change of inner urban areas.

6 Discussions and Conclusions

The concluding section has two objectives. First, it aims to discuss the project as a possible contribution of a university to sustainable regional development of its regional environment. Secondly, the substantive results of the project is summarised and evaluated.

Source: Own illustration following (Busch 2013, 14)

Figure 5: Size, structure and barrier-free accessibility of shops in Südstadt

In this context, the following aspects are discussed in more detail:

1. Does the project presented in the paper focus on sustainability (focus)?

2. Is the project an example of co-creation for sustainability (co-creation)?

3. How can the HEIs' "sustainability impulse" be classified (classification)?

4. Which factors enabled the implementation of the project and was the project complicated by obstacles often typical in a sustainability transition process (enablers and obstacles)?

5. How can the substantive results be assessed (results)?

Focus

General trends – e.g. migration, demographical changes, changing mobility behaviour and digitalization – force structural adaption processes not only at the national level, but also at the level of regions, cities and urban districts. Neighbourhoods lose their previous functions in city structures, and the process of finding new functions is usually long-lasting and difficult. In many urban districts this process is accompanied by economic and social problems; ecological and health problems also come up often. Urban neighbourhoods in many German cities face similar problems; they are "multi-burdened" urban districts, for economic, social and ecological burdens often occur simultaneously. Efforts to support social and economic (sustainable) transformation become necessary. Strengthening the local economy is an important part of these efforts. By pursuing this objective, the projects aim to simultaneously achieve social stabilization in the neighbourhood. In the case study presented here, a university of applied sciences helps strengthen local economy and stabilize the social situation in urban districts, thus supporting transformation processes for a more sustainable development at the local level. With regard to ecology, "neutrality" has been implicitly assumed in the project. This is certainly a weakness of the project as the assumption was neither questioned nor analysed. Two dimensions of sustainability, however, have been explicitly mentioned.

Co-creation

In the presented case study, the university collaborates with social actors of urban districts to create transformation processes for sustainable development at the local level. Knowledge is produced for direct application by the use of research approaches and forms of knowledge production highly related to sustainable transition processes:

- transformative research
- participatory action research
- the idea of urban living labs and
- the concept of student service learning.

With these practises the project fulfils the necessities of co-creation explained above.

Classification

According Pflitsch, Radinger-Peer (2018), HEIs can play different roles in regional sustainable transition; they use two dimensions to distinguish these roles – depth and autonomy:

- When it comes to depth, HEIs' roles can be "comprehensive, involving diverse actors and approaching sustainability with a holistic perspective" or "more fragmented and passive, but also more focused on specific topics".
- As for autonomy, the roles can be "autonomous, the university defining its own focus and priorities through interacting with a broad range of regional actors" or "more directed, the university working on topics that are relevant from the perspective of the regional or federal-state government".

Using this rough classification, the HEI in our example apparently does not approach sustainability with a holistic strategy. Instead of that two research institutes of this university of applied sciences interact with local actors and work on a "specific topic" that is not only relevant from the perspective of most German cities and their citizens but also from the national and federal governments' perspectives.

Thus, the case study presented here clearly differs from those presented and discussed in detail by Radinger-Peer et al. (2020). Their examples (Augsburg, Freiburg, Linz and Darmstadt) can all be classified as "comprehensive, involving diverse actors and approaching sustainability with a holistic perspective". In these four cities, an awareness of sustainability was already politically and civically anchored when the HEIs became active (Radinger-Peer et al. 2020). Moreover, the HEIs also have implemented the necessary institutional and organisational changes (Radinger-Peer et al. 2020). Both hardly hold true for the NUAS and its regional environment.

Enablers and obstacles

In the case study on hand, two aspects were important for the realization of the projects:

- Application oriented scientists – which are typical for German universities of applied sciences – with a personal research interest in the special problem of local economies in disadvantaged urban neighbourhoods and sustainable development initiated the project.
- These researchers succeeded in raising funds for their topic in close collaboration with local actors from the cities concerned. This aspect was decisive for the realisation of the project as German universities of applied sciences have almost no research budget of their own.

Thus, the combination of scientific curiosity of the participating scientists, their willingness of to change the situation in co-operation with local actors, the openness to new approaches on part of the local actors and the possibility to raise public funding has paved the way for the project implementation. On the one side, this supports the opinion of Radinger-Peer et al. (2020) that special features of the higher education system (e.g. freedom of research and teaching, loose coupling of units and flat hierarchies) allow for a high level of bottom-up and niche activities. But on the other side. one must confess that the paper deals with a small number of projects due to the individual commitment of staff members, i.e. only about a "pin-brick" compared to the potentials universities have in supporting sustainable development in their regional environments. Although Radinger-Peer et al. (2020) emphasise that individual actors – like those in the example presented here – sometimes have provided an impetus for the universities' developments towards sustainability, they also mention (Radinger-Peer et al. 2020) that projects like this are often unconnected, lacking far-reaching visibility, and that, as a consequence of this, some of them might disappear without durable effects. In their view (Radinger-Peer et al. 2020) efforts to achieve a regional sustainability transition can only be attained through the support of the university management and the establishment of interdisciplinary institutes. This, however, requires institutional and organisational changes within the university.

The already discussed distinction between a university-wide strategic orientation towards a regional sustainability transition and individual projects aiming at regional sustainability helps explain why challenges and obstacles that often arise in connection with the new role of universities (Aleixo et al. 2018, Radinger-Peer et al. 2020, Verhulst, Lambrechts 2015) are rather irrelevant in the described example:

- Conflicts between and discussions about the relevance of different missions of universities did not occur, because the acting researchers had the freedom to decide and put the research focus on sustainability issues.

- The same reasoning explains why a modest willingness for change or a lack of acceptance of the concept of sustainability on the side of the university stakeholders have not been relevant.

- The concept of sustainability is often seen as an abstract and complex topic. This might lead to a different understanding of sustainability between scientists in the university and practitioners in the region. Different definitions of sustainability from actors with different disciplinary backgrounds sometimes make project work even more difficult. In addition, sustainability transitions are sometiems highly conflictual processes accompanied by long-lasting negotiations. The projects in the example did not pose these difficulties, as they did not aim at a general transition, but worked on a sustainability topic in a limited niche. In the example on hand university stakeholders and regional practitioners had a common interest. They jointly applied for funding and declared their willingness to cooperate. Accordingly, the project did not begin with theoretical discussions about definitions of sustainability, but with pragmatic work in the neighbourhood.

- Finally, there was no lack of financial resources as the researchers successfully applied for the necessary governmental and European funds. This gave them financial freedom.

Results

The last point of these concluding remarks deals with an evaluation of concrete project results achieved for the urban district on hand. This evaluation is not a scientifically based analysis, and its authors were personally involved in the projects. Nevertheless, the following aspects should be mentioned:

- The continuous interaction of local participants with an inter-disciplinary research team has proven to be a promising concept. The concept encouraged participation from local actors. They could be motivated to engage and participate in activities for the sustainable local development. The elaboration of measures was a result of co-design and co-production of researchers and local actors in the research process. The taken measures were not set by a "top-down" arrangement but were rather elaborated and initiated by a research team and local actors through a participatory "bottom-up procedure".

- Information and knowledge about the neighbourhood could be improved by the continuous combination of analysis and activity. These improvements allowed for the development of evidence-based measures appropriate for dealing with the causes of problems at hand.
- Neighbourhoods differ from each other and so do their problems. This means there is a need for place-based solutions fitting to the specific problems of the neighbourhood. The concept led to place-based measure recommendations.
- Continuous reflection on the scientific methods and selected measures was ensured by discussions between the research team and local actors but also by discussions with external experts coming from different disciplines.
- Finally, according to the opinion of the authors partly supported by own evaluations, the neighbourhood project was able
  - to improve the initial situation (e.g. by implementing a community of entrepreneurs, locating specialist shops and making the cityscape more attractive by reducing vacancies),
  - to counteract the negative image of the neighbourhood,
  - to strengthen the neighbourhood's economic potential,
  - to create to create a more optimistic mood in the neighbourhood
  - to identify realistic functions and objectives for the future of the neighbourhood.
- As a consequence, all this contributed to improving the social situation in the neighbourhood, giving people new perspectives.

The example clearly shows that universities can give impulses for sustainable development in their regional environments not only by a consequent university-wide orientation towards sustainability, but also by being perceived as a "sustainable". They can also give these impulses through their involvement and engagement in single research projects with regional partners. Universities' contribution to sustainable regional transition often can be seen as a side product of research. In this case, universities do not play the role of a change agent for sustainable regional development but rather give pinpricks of change, and the higher the number or intensity of these pinpricks – the higher the effects for regional sustainability.

Despite this positive conclusion, there are some questions to be dealt with in future research:

- First of all, long-term external evaluation would be necessary to properly assess the success achieved by the projects.
- Secondly, the concrete, decisive factors for the success of the described case study remains uknown. The methodological approach used here seems to be a key success factor. Nevertheless, further research is necessary for a better understanding of the determinants for success or failure; governance and local socio-economic structures as well as the types of involved actors could all be important factors determining the outcome.
- Another very policy-relevant research question is concerned with whether the improvements for the neighbourhoods will be durable even after the funding ends and with how it will be possible to perpetuate the cooperation of local actors as well as their participation and engagement.

References

Abramson HN, Encarnação J, Reid PP, Schmoch U (1997) Technology transfer systems in the United States and Germany – lessons and perspectives. Washington, DC

Adomßent M, Barth M, Fischer D, Richter S, Rieckmann M (2014) Learning to change universities from within: A service-learning perspective on promoting sustainable consumption in higher education. *Journal of Cleaner Production* 62: 72–81. CrossRef.

Aleixo AM, Leal S, Azeiteiro UM (2018) Conceptualization of sustainable higher education institutions, roles, barriers, and challenges for sustainability: An exploratory study in Portugal. *Journal of Cleaner Production* 172: 1664–1673. CrossRef.

Asheim BT, Gertler MS (2005) The geography of innovation: Regional innovation systems. In: Fagerberg J, Mowery DC, Nelson RR (eds), *The Oxford Handbook of Innovation*. 291-317, New York. CrossRef.

Attwood H (1997) An overview of issues around the use of participatory approaches by post-graduate students. IDS, Participatory Research, IDS PRATopic Pack (February, 1997). Brighton: IDS, University of Sussex

Birkhölzer K (2000) Formen und Reichweite lokaler Ökonomien. In: Ihmig H (ed), *Wochenmarkt und Weltmarkt. Kommunale Alternativen zum globalen Kapital*. USP International, Bielefeld, 56–64

BMUB – Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (2016) Städtebauförderung 2016. Anwenderhinweise zu den Förderprogrammen. Berlin

Breznitz SM, Feldman MP (2012) The engaged university. *The Journal of Technology Transfer* 37: 139–157

Bringle RG, Hatcher JA (1995) A service-learning curriculum for faculty. *Michigan Journal of Community Service Learning* 2: 112–122

Bringle RG, Hatcher JA (2000) Institutionalization of service learning in higher education. *The Journal of Higher Education* 71: 273–290. CrossRef.

Bulkeley H, Coenen L, Frantzeskaki N, Hartmann C, Kronsell A, Mai L, Marvin S, McCormick K, van Stenbergen F, Voytenko-Palgan J (2017) Urban living labs: Governing urban sustainability transitions. *Current Opinion in Environmental Sustainability* 22: 13–17. CrossRef.

Busch R (2013) Perspektiven des Einzelhandelsstandortes "Viersen-Südstadt". Dortmund

Clark B (1998) Creating entrepreneurial universities: Organizational pathways of transformation. Bingley

Cooke P (1992) Regional innovation systems: Competitive regulation in the new Europe. *Geoforum* 23: 365–382. CrossRef.

Cooke P (2002) Regional innovation systems: General findings and some new evidence from biotechnology clusters. *Journal of Technology Transfer* 27: 133–145. CrossRef.

Etzkowitz H, Leydesdorff L (1995) The triple helix-university-industry-government relations: A laboratory for knowledge-based economic development. *EASST Review* 14: 14–19

Gibbons M (1999) Science's new social contract with society. *Nature* 402: C81–C84. CrossRef.

Gibbs P (2001) Higher education as a market: A problem or solution? *Studies in Higher Education* 26: 85–94. CrossRef.

Goldstein HA, Maier G, Luger M (1995) The university as an instrument for economic and business development: Us and european comparisons. In: Dill D, Sporn B (eds), *Emerging patterns of social demand and university reform: Through a glass darkly*. Elmsford, 105–133

Greenwood DJ, Levin M (2007) *Introduction to Action Research: Social Research for Social Change* (2 ed.). Sage, Thousand Oaks

Hagemann J, Kruse S, Mayer A (2011) Einzelhandelskonzept für die Stadt Viersen. Fortschreibung, im Auftrag der Stadt Viersen. Endbericht. Dortmund

Hamm R, Koschatzky K (2020) Kanäle, Determinanten und Hemmnisse des regionalen Transfers aus Hochschulen. Hochschulen und nachhaltige Regionalentwicklung. Arbeits- und Sitzungsberichte der Akademie für Raumforschung und Landesplanung

Henke J, Pasternack P, Schmid S (2016) Third Mission bilanzieren – Die dritte Aufgabe der Hochschulen und ihre öffentliche Kommunikation. HoF-Handreichungen 8, Beiheft zu "die hochschule", Halle-Wittenberg

Henn S (2013) Die Rolle der lokalen Ökonomie für Stadtteile mit besonderem Entwicklungsbedarf. Presentation on Nov. 12, 2013. Leipzig

Hölscher K, Wittmayer JM, Loorbach D (2018) Transition versus transformation: What's the difference? *Environmental Innovation and Societal Transitions* 27: 1–3. CrossRef.

Jakubowski P, Koch A (2009) Stärkung lokaler Ökonomie – eine stadtentwicklungspolitische Aufgabe. *Informationen zur Raumentwicklung* 3/4.2009: 241–253

Koschatzky K, Dornbusch F, Hufnagl M, Kroll H, Schnabl E, Schulze N, Teufel B (2013) Regionale Aktivitäten von Hochschulen. Fallstudien von neun deutschen Universitäten und Fachhochschulen. Karlsruhe: Fraunhofer ISI, http://www.isi.fraunhofer.de/isi-wAssets/docs/p/de/projektberichte/Regionale-Aktivitaeten-von-HO_Fallstudien.pdf

Koschatzky K, Hemer J (2009) Firm formation and economic development: What drives academic spin-offs to success or failure? In: Varga A (ed), *Universities, Knowledge Transfer and Regional Development*. Edward Elgar, 191-218

Loorbach D (2010) Transition management for sustainable development: A prescriptive, complexity-based governance framework. *Governance* 23: 161–183. CrossRef.

MacDonald C (2012) Understanding participatory action research: A qualitative research methodology option. *Canadian Journal of Action Research* 13: 34–50

MWEBWV NRW – Ministerium für Wirtschaft, Energie, Bauen, Wohnen und Verkehr des Landes Nordrhein-Westfalen, Städtenetz Soziale Stadt NRW (2014) Stadtteilprofil Viersen Südstadt. http://www.soziale-stadt.nrw.de/stadtteile_projekte/profil.php?.st=-viersen-sued. (Sep. 9, 2016)

Pflitsch G, Radinger-Peer V (2018) Developing boundary-spanning capacity for regional sustainability transitions – A comparative case study of the universities of Augsburg (Germany) and Linz (Austria). *Sustainability* 10: 918–918. CrossRef.

Pinheiro R, Langa P, Pausits A (2015) One and two equals three? The third mission of higher education institutions. *European Journal of Higher Education*: 232–249. CrossRef.

Radinger-Peer V, Pflitsch G, Freytag T, Döring T (2020) Transformationsprozesse im Hochschulsystem in Richtung nachhaltige Regionalentwicklung ("Empirische Illustrationen"). In: Postlep RD, Roessler I, Duong S, Hachmeister CD (eds), *Forschungsbericht der ARL 11*. Verlag der ARL, Hannover, 177–207

Roessler I, Duong S, Hachmeister CD (2015) Welche Missionen haben Hochschulen? Third Mission als Leistung der Fachhochschulen für die und mit der Gesellschaft. Arbeitspapier Nr. 182, Centrum für Hochschulentwicklung. http://che-ranking.de/-downloads/CHE_AP_182_Third_Mission_an_Fachhochschulen.pdf. (May 3, 2019)

Schäfer E, Scheele U (2014) Bridging the transformation gap with "living labs"? Conference paper: Urban regions under change: Towards social-ecological resilience

Schneidewind U (2014) Urbane Reallabore – Ein Blick in die aktuelle Forschungswerkstatt. *pnd-online, www.planung-neu-denken.de*: 1–7. Heft III. Aachen

Schneidewind U, Scheck H (2012) Zur Transformation des Energiesektors – Ein Blick aus der Perspektive der Transition-Forschung. In: Servatius HG, Schneidewind U, Rohlfing D (eds), *Smart Energy – Wandel zu einem nachhaltigen Energiesystem*. Springer, Berlin, Heidelberg, 45–61

Schneidewind U, Singer-Brodowski M, Augenstein K, Stelzer F (2016) Pledge for a transformative science – A conceptual framework. 191 Wuppertal Paper. Wuppertal

Singer-Brodowski M, Schneidewind U (2014) Transformative Literacy: gesellschaftliche Veränderungsprozesse verstehen und gestalten. In: Umweltdachverband GmbH (ed), *Krisen- und Transformationsszenarios: Frühkindpädagogik, Resilienz & Weltaktionsprogramm*. Bildung für nachhaltige Entwicklung, Jahrbuch, 131–140

Stahlecker T (2006) Regionale Bindungen im Gründungs- und Entwicklungsprozess wissensintensiver Dienstleistungsunternehmen - Dargestellt am Beispiel der Regionen Bremen und Stuttgart. Berlin

Stephens JC, Hernandez ME, Román M, Graham AC, Scholz RW (2008) Higher education as a change agent for sustainability in different cultures and contexts. *International Journal of Sustainability in Higher Education* 9: 317–338. CrossRef.

Stoetzer MW, Krähmer C (2007) Regionale Nachfrageeffekte der Hochschulen – Methoden, Probleme und Ergebnisse empirischer Untersuchungen für die Bundesrepublik Deutschland. Jenaer Beiträge zur Wirtschaftsforschung 6

Trencher G, Bai X, Evans J, McCormick K, Yarime M (2014) University partnerships for co-designing and co-producing urban sustainability. *Global Environmental Change* 28: 153–165. CrossRef.

Trencher G, Yarime M, McCormick KB, Doll CNH, Kraines SB (2014) Beyond the third mission: Exploring the emerging university function of co-creation for sustainability. *Science and Public Policy* 41: 151–179. CrossRef.

Uyarra E (2010) Conceptualizing the regional roles of universities. Implications and contradictions. *European Planning Studies* 18: 1227–1246. CrossRef.

Verhulst E, Lambrechts W (2015) Fostering the incorporation of sustainable development in higher education: Lessons learned from a change management perspective. *Journal of Cleaner Production* 106: 189–204. CrossRef.

Voß R (2004) Regionale Wirksamkeit von Hochschulen – ein Konzept zur Analyse und Bewertung. In: Präsident der Technischen Fachhochschule Wildau (ed), *Wissenschaftliche Beiträge*. Forschung, Lehre, Technologietransfer. Wildau, 103-113

Walter M (2009) Participatory action research. social research methods. In: Bryman A (ed), *Social Research Methods*. The Falmer Press, London, 151–158

WBGU – German Advisory Council on Global Change (2011) World in transition – A social contract for sustainability. Berlin: WBGU. http://www.wbgu.de/en/flagship-reports/fr-2011-asocial-contract/

Zawdie G (2010) Knowledge exchange and the third mission of universities. *Industry & Higher Education* 24: 151–155. CrossRef.

Zomer A, Benneworth P (2011) The rise of the university's third mission. In: Eders J, de Boer HF, Westerhejiden DF (eds), *Reform of Higher Education in Europe*. Sense, Rotterdam. CrossRef.

REGION
The Journal of ERSA
Powered by WU

Volume 7, Number 2, 2020, 29–42
DOI: 10.18335/region.v7i2.269

journal homepage: region.ersa.org

Non-Employed People in the Westpfalz Region (Germany) as Target Group of Continuing Higher Education – Potentials of a regional demand-orientation

Jessica Neureuther, Dominik Böhmer[1]

[1] Technische Universität Kaiserslautern, Kaiserslautern, Germany

Received: 5 July 2019/Accepted: 10 November 2020

Abstract. It is stated that Higher Education Institutions (HEIs) and also continuing higher education has an influence on the development of a region. By focusing on the challenges and demands of the regional economy as well as its population, these data help to develop study programmes that meet these needs. In further education there is research suggesting that the identification and consideration of the demands of target groups helps to increase the attractiveness and participation in continuing higher education at HEIs. In this paper, the Westpfalz region (Germany) is considered as an example because it is known as structurally weak; particularly, the participation rate on further education in general, and the unemployment rate are below the national average there. For this purpose, a representative regional population survey was conducted. It describes the population of the Westpfalz region, their educational demands and it provides clues to the development of study programmes that meet their needs. The group of non-employed people stood out in this survey. In comparison with the employed group, it is astonishing that the non-employed people have a higher willingness to participate in continuing higher education. They also have more time to invest in study programmes. The paper shows interesting results of the population survey concerning the non-employed group. It also presents how these results can take up in the development of (regional) study programmes. At the end of the paper, it is reflected what potentials the demand-oriented development of continuing higher education programmes has for the regional development, which are comparable to the Westpfalz region.

Key words: continuing higher education; target groups; non-employed people; regional development

1 Introduction

In times of a globalised and digitalised division of labour, the knowledge economy has set a new agenda in the field of regional development. A sole focus on economic structures is not enough and neglects the regional population. This is where the educational potential of a region comes into play. In general, further education contributes to the improvement of job opportunities and social participation, thus having a positive effect in the development of a region (Nuissl 2000, Fritsch 2009). In this context, universities are also gaining more and more importance (Rohs et al. 2015). Due to the research-orientation, Higher Education Institutions (HEIs) have the possibility to assume the function as "innovation driver"

(Fritsch et al. 2008) by contributing profound knowledge about the educational needs of the regional economy as well as its population (Rohs, Steinmüller 2020). Therefore, they can react by developing study programmes that meet the needs of a region. This results in both economic and social development for a region. Referring to this, the project E^B "Education as exponent of individual and regional[1]" wanted to identify new target groups for higher education, i.e. to enlarge the social participation and to increase the level of education in the Westpfalz region (English: West Palatinate) in Germany. This region is generally considered to be structurally weak. It is therefore an aim to contribute to regional development. The overall aim is to make the participating universities into places of lifelong learning (Longworth 2006). To learn more about the region and its educational needs, the demand-oriented approach is applied by the project. In addition to many other surveys in the context of the project (for more information see Schwikal, Steinmüller 2017), the regional population survey is highlighted in this paper. The group of non-employed people has been identified as a potential new (regional) target group for continuing higher education. The following section examines how the demand-oriented development of study programmes promotes the achievement of this target group and what potentials arise for the regional development in the Westpfalz.

This paper primarily describes the Westpfalz region, followed by a brief introduction into the relationship between further education in general and continuing higher education; in particular, demand-orientation and region. After that, the representative regional population survey and its results are presented. Thus, the characteristics and educational demands of non-employed people for continuing higher education are contrasted to the traditional target group of working people. Finally, it is discussed how continuing higher education can address this target group and what potentials this has for the Westpfalz region and other comparable regions.

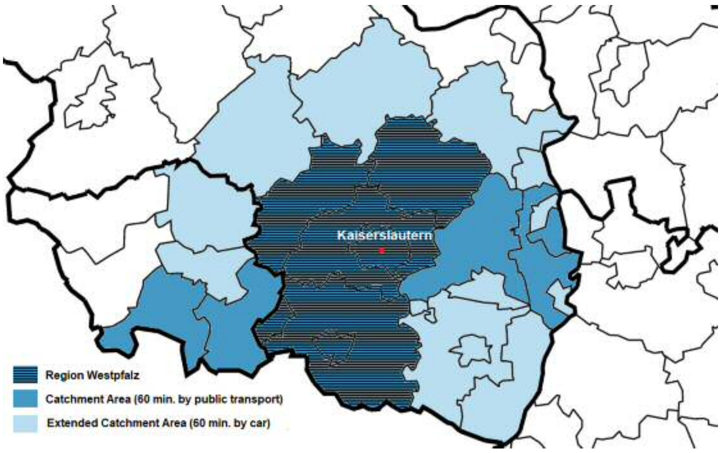
2 Description of the Westpfalz Region

In the following discussion, a delimitation of the Westpfalz region is made and special characteristics of this region are worked out. The Westpfalz region is a part of the federal state Rheinland-Pfalz in Germany. Different aspects like the administrative affiliation (see the dark blue shaded area, Figure 1), functional interrelation, homogeneity or sociocultural factors (Bernhard 2017, Dobischat et al. 2006) are possibilities for defining a region. Figure 1 shows also the aspects of defining the region in the project E^B. It shows the functional interrelation of the catchment area with public transport (see the middle blue shaded area) and by car (see the light blue shaded area) which is reachable within a duration of 60 minutes. The administrative affiliation – here: districts (see the dark blue shaded area) is also considered. More specifically, this is the so-called "extended" region of the Westpfalz or the "region E^B" (Marks 2015). Figure 1 shows the location and the relevant areas, which were central for the population survey.

By adding all relevant districts, in the region E^B has a population about 2.5 million (Statistisches Amt Saarland 2018, Statistisches Landesamt Rheinland-Pfalz 2017).

The Westpfalz region, with its centre of Kaiserslautern, limited by the administrative affiliation, is known as structurally weak. Following the statistics of the Bundesagentur für Arbeit (2019), it is characterised by a permanently higher rate of unemployment compared to the national average (5.6 percent in December 2019; total Germany: 4.9 percent). If one considers the unemployment rate among the group of those with an academic degree, this averaged only 2.2 percent (total Germany: 2.2 percent) in the 2018 reporting year (Bundesagentur für Arbeit 2019). This goes hand-in-hand with low economic growth, which is the consequence of the closure of many industrial plants (e.g. "Pfaff" – sewing machines) or relocation of production and growing competition in the Asian as well as Southern and Eastern European regions, in the wake of globalisation (e.g. companies of shoe industry in Pirmasens) since the 1990s. Since then, the region has suffered from an increasingly ageing and rapidly declining population (Ludewig et al. 2007). According to

[1] The project was a joint project between the Kaiserslautern University of Applied Sciences, the Ludwigshafen University of Business and Society and the Technische Universität Kaiserslautern. It ended in July 2020. For more information see https://www.e-hoch-b.de/e-hoch-b/



Source: Marks (2015, p. 14)
Notes: own editing

Figure 1: The Region E[B]

the "Prognos Future Atlas 2019" (Prognos AG 2019), which examines the development opportunities and risks for all German regions, three districts (Kaiserslautern, Stadt; Kaiserslautern, Landkreis and Donnersbergkreis) of the Westpfalz region (dark blue shaded area) are among the regions showing the worst development in the past few years (Prognos AG 2019). In particular, it shows very large social inequalities (Prognos AG 2016). In addition, a study by the Berlin Institute for Population and Development, which evaluates the demographic sustainability of Germany's regions on the basis of a system of indicators, tends to give the Westpfalz region poor marks, especially the Südwestpfalz (English: Southwest Palatinate) (Slupina et al. 2019). Among other things, one of the lowest gross domestic products per capita in Germany as an indicator of economic strength and the highest municipal debt per inhabitant are reported. Furthermore, the "Atlas of further education in Germany" (German: Weiterbildungsatlas) reveals that the participation rate in further education, in general, is below national average (Bürmann, Frick 2016). With regard to the participation rate in continuing higher education in Germany, there is insufficient data, due to the absence of systematic recording, to date (Widany et al. 2020).

Nevertheless, the situation in the Westpfalz region has improved slightly in recent years. A forecast study of the region of Kaiserslautern indicated that the economy has changed in the last two decades. Before, there were more manufacturing companies and now the region is in a process of transition to a service, science and IT location (Kujath 2015). The next chapter briefly discusses the relationship between region and education. Besides that, it shows how the demand-oriented approach can be supportive in the development of study programmes.

3 Region, continuing higher education and demand-orientation

Against the background of globalisation and digitalisation, in general, there are increasing labour market challenges. Therefore, the demand for qualified specialists is becoming more important. Due to the demographic and structural changes, regional and social challenges are posed, e.g. ageing of the population or poor economic development opportunities (for more detail see Rohs, Steinmüller 2020). Especially structurally weak regions like the Westpfalz region are challenged to meet these requirements and to be particularly attractive and innovative in order not to succumb completely to demographic change. Therefore, in the last few years, the importance of the region for further education in general has moved into focus (Benneworth, Hospers 2007, Martin et al. 2015). The connection between learning and region was recognised in the 1970s, when the UNESCO, OECD or EU (Eckert, Tippelt 2017, Kallen, Bengtsson 1973) promoted the paradigm of lifelong learning. Generally, the goal of lifelong-learning is the improvement of the economic and environmental competitiveness of each state and to force "education for a more highly skilled workforce; personal development leading to a more rewarding life; and the creation of a stronger and more inclusive society" (Aspin et al. 2001, p. 21). Thus, HEIs as a location for lifelong-learning moved more into the focus. They also contribute to the social, cultural and economic development of a region (Schäfer 1988). Therefore, the universities are assigned a third mission in addition to research and teaching. This is defined as:

> "Third stream activities are therefore concerned with the generation, use, application and exploitation of knowledge and other university capabilities outside academic environments. In other words, the Third Stream is about the interactions between universities and the rest of society." (Molas-Gallart et al. 2002, p. iii)

This mission is divided into three activities: Continuing (Higher) Education, Technology Transfer and Innovation, and Social Engagement (Carrión et al. 2012). In this paper, the focus is mainly on the aspect of continuing (higher) education, whereby the expansion of education is also closely linked to social engagement. Carrión et al. (2012) see educational outreach as an indicator of social engagement. According to Pasternack, Zierold (2014), universities contribute to improving the quality of life, providing public services and infrastructure or strengthening civil society.

Therefore, continuing higher education has an influence, which should not be underestimated. Rohs, Steinmüller (2020) point out that the role of continuing higher education has received little attention so far – especially in Germany. One reason is that continuing higher education in Germany is challenged to be part of the higher education and science system and of the further education market – the so called "doppelte Systembindung" (double systemic commitment) (Wolter 2005). As part of the science system, continuing higher education has access to the latest research findings and thus contributes to the transfer to society. In addition, it is subject to market structures, which results in a stronger demand-oriented development of educational offerings in the last years (Seitter et al. 2015, Wolter 2015).

In order for continuing higher education to contribute to regional development, it is necessary to know the demands (concrete and active) and needs (passive state descriptions) in the region. These educational demands are able to point out educational potentials. Information on target groups and their demands is therefore important for the accuracy of fit of study programmes. Finally, this can be transferred locally into tailor-made study programmes and ultimately helps to increase the attractiveness of educational opportunities at universities. Therefore, participation in continuing higher education increases, which contributes to regional development (Rohs, Steinmüller 2020).

In order to be able to draw conclusions for continuing higher education, it is essential to consider the demands and needs of the Westpfalz region. In this context, the project E^B conducted the population survey to get more information about the population and their educational demands or needs.

4 The Regional Population Survey

4.1 Methodical Approach

As mentioned above, a representative regional population survey was conducted (for an overview of general results see Schwikal, Steinmüller 2017). The intention was to describe the population of the region E^B and their educational needs, thereby identifying new target groups for study programmes. The aim was to draw conclusions about the demands these target groups articulate in terms of learning interests and conditions and the formats of continuing (higher) education programmes. Moreover, the data provide information on motivations and barriers to participation in further education in general, which can support the development of regionally effective educational concepts.

The survey gives a representative database of the region E^B, which means a characteristics-specific representativeness with the usual probability of error in social scientific surveys ($p = 0.05$). In other words, there is a confidence interval of 95 percent. According to the German Census (2015), the population of the E^B region comprises about 2.5 million residents. The sample size of the survey included 521 people (at least 400 respondents were needed for representativeness) and was drawn in a two-step random selection: First, the RLD method (randomised last digit) was used to select a household in the region (household level). On the individual level, the respondent of the household was selected with the Last-Birthday method. The sample was quoted by gender distribution and age limits analogous to the Census Data (2015). Only people between 17 and 64 years were included in the survey in order to ensure that they could potentially pursue employment. The upper age limit was set at 64 years to ensure a suitable time distance to retirement. On behalf of the project E^B, the Umfragezentrum Bonn (UZBonn) conducted the survey as Computer Assisted Telephone Interview (CATI) in November and December 2016 in German. The used questionnaire was based on already existing survey instruments and items (e.g. of Adult Education Survey or Mikrozensus) to ensure reliability and validity of the questions, as well as the comparability of the results.

Before carrying out a specific data evaluation to answer the research interest mentioned above, some general socio-demographic descriptive figures from the survey are outlined. 521 people took part in the survey. Their average age was 45 years with a standard deviation of 13 years. 54 percent of the respondents were female and 46 percent were male. Almost two-thirds (63.3 percent) lived in a household with their partners, 40 percent with their children and only about one fifth of the respondents (18.1 percent) lived in a single-person household (for more details see Schwikal, Steinmüller 2017).

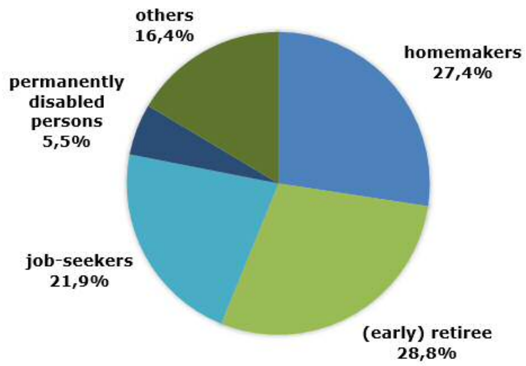
4.2 Empirical Results

4.2.1 The Group of Non-Employed People

In the survey, the question “Are you currently employed?” was included. It was noticeable that 19.8 percent ($n = 103$) answered “no”. This group is composed of respondents who answered that they are neither currently working nor on parental leave. Because of the size of this group, further analyses were carried out.

A closer look on the non-employed people show that 70.9 percent ($n = 73$) of them would have the entrance qualification for participating in continuing higher education (e.g. because of graduation, previous academic degree or vocational qualification). The access requirements in Germany are usually a first academic degree or a vocational qualification and work experience (Wolter 2015). In contrast, the data show that 93.1 percent ($n = 389$) of the 418 employed people would have this entrance qualification. This represents a larger share than in the group of non-employed people (+22.2 percentage points). In the following, the groups of persons with a possible access to continuing higher education are considered. Due to the size of the non-employed group, it is interesting to analyse them as a new target group for continuing higher education. Non-employment is caused by different personal life situations: Most respondents (28.8 percent) were already (early) retired. 27.4 percent were homemakers and 21.9 percent were currently seeking employment. There were also people who are permanently disabled and unable to work (5.5 percent). All other persons and their reasons (16.4 percent) were grouped into the category “others”, because they could not be assigned to any of the other groups (see Figure 2). The majority of this group of “others” were undergraduate students still in the process of completing their studies. They were potentially interested in continuing higher education, as well as persons in partial retirement, private individuals and persons undergoing restructuring.

Looking at the gender distribution, the data show that 58 percent of the non-employed people were female and 42 percent were male. The group of employed people ($n = 389$) is “more male” (47 percent) than the group of non-employed. Within the group of non-employed people, the average age was around 48 years and the majority of those surveyed, living together with their partner (61.6 percent), followed by living in a household with children (31.5 percent) and living alone (27.4 percent). For this question, it was possible to give multiple answers. Those who lived together with their children mostly had one child (60.9 percent). 21.7 percent of the respondents had two children and only 17.4 percent



Notes: in percent; $n = 73$

Figure 2: Composition of Non-Employed People in the population survey

had three children and more. Furthermore, the employed people were younger (average of 46.8 years), more often parents (+11.9 percentage points) and lived less alone (-10.5 percentage points). In addition to that, the majority of those who lived together with their children had two children (compared with mostly one child of the non-employed).

Regarding the participation in further education in the last ten years of the non-employed group, 65.8 percent have not participated in any courses lasting several weeks (two weeks to a maximum of eight weeks) or months (longer than two months). Moreover, greater than half of this group had not participated in any short-term trainings (58.9 percent). The percentages of non-participation for long-term trainings do not differ much from those of the employed group (65.3 percent). They only differ in terms of short-term trainings. Merely 28.1 percent did not take part in short-term trainings. The employed people participated more frequently in short-term trainings.

It is interesting to note that 20.5 percent of the non-employed people are considering participation in continuing higher education at a HEI in the next five years. The same percentage could not yet assess it. In comparison with the employed group, the data show a stronger rejection of participation in continuing higher education at a university over the next five years (-5.8 percentage points) and a higher proportion of indecision (+6.5 percentage points). There is a statistically significant difference between the two groups in terms of the proportion of those who can imagine attending continuing higher education at a university in the next five years is statistically significant[2].

In addition, the population was asked about their motives for participation in the past. The respondents rated on a six-step Likert scale (from 1-6) to what extent their participation motive was work-related or personal-motivated (1: attended entirely for work-related reasons to 6: entirely personal interest). Comparing the mean values of the group of non-employed ($M = 2.71$) and employed people ($M = 2.01$) shows that there is a difference between the two groups. The non-employed people were more likely to participate in further education out of personal interests. The difference in mean values were tested by the procedure of independent t-tests[3]. The t-test shows that there is a significant difference for both ($T = -2,243$; $p(t) = 0.030$). For this item, the differences between the two groups using the Cohens d method have an effect strength of $d = 0.49$. This corresponds to a small-to-medium effect.

With regard to attitudes towards further education in general, the question arises whether participants have undergone further education in order to carry out their professional activity better and to advance professionally. A six-step Likert scale (from 1: do not agree at all to 6: strongly agree) was used – the group of employed people agreed almost fully ($M = 5.39$) while the group of non-employed people simply agreed ($M = 4.02$). The

[2]At one percent significance level, based on an unpaired t-test. This indicates that the difference is not only random in the sample of this study, but that these groups also differ in the population.

[3]First of all, the necessary prerequisites must be met of variance homogeneity and normal distribution. The variance homogeneity is checked by means of the Levene-test. This test shows variance heterogeneity.

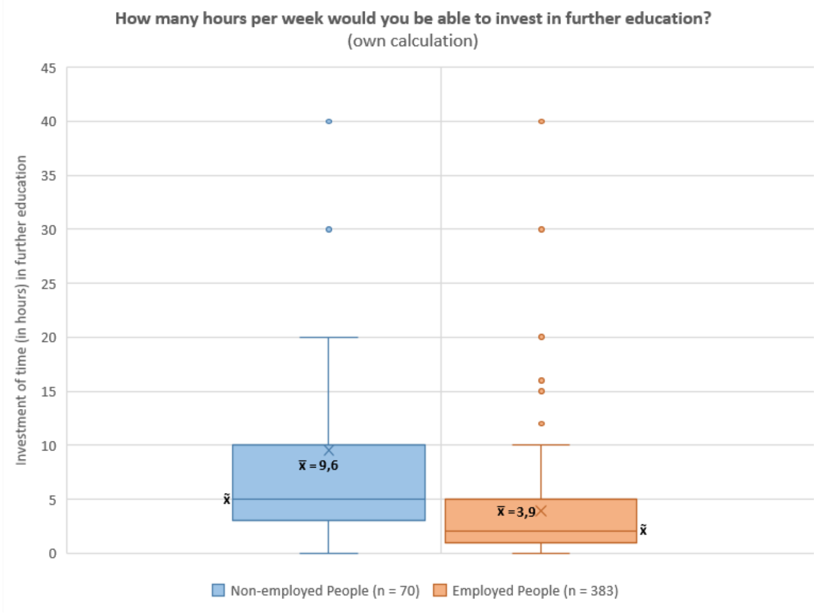


Figure 3: Boxplot of investment of time in further education by non-employed and employed people in the population survey

t-test shows that there is a significant difference ($T = 4.483$; $p(t) : 0.000$). The differences between the groups have an effect strength of $d = 0.26$ by using the Cohens d method, which means a weak effect.

Almost three-quarters of the group of non-employed people (70.8 percent) agreed that they would participate in further education if it could be more easily integrated into everyday life. This can be a barrier for participating in further education.

Furthermore, almost half of the respondents (49.3 percent) saw further education as a natural part of their personal and vocational development. There is a difference with the employed respondents (75.1 percent), who agree more strongly with this statement. 60.3 percent of the non-employed respondents rejected the statement that participation in further education is uninteresting and serves only to increase the number of job opportunities on the labour market. The rejection of this statement was about twelve percentage points higher among the working population.

A remarkable difference between the two groups concerning the investment of available time in further education is shown (see Figure 3). Employed people would be willing to spend approximately four hours per week on further education. In comparison, the non-employed group would be willing to invest at least nine and a half hours per week on further education. In particular, the subgroup of job-seekers (15.3 hours per week) stand out.

Regarding the willingness to bear the costs of further education programmes, more than half of the non-employed people (53.2 percent) would pay up to 3,000 Euros. 5,000 Euros or 10,000 Euros would be paid in each case by 23.4 percent of the respondents. Among the group of employed people was a greater willingness to pay more than 10,000 Euros for further education programmes. However, a higher proportion of non-employed people would pay 5,000 Euros up to 10,000 Euros (+9.9 percentage points compared to the group of employees).

As the format of further education programmes and the mode of learning were concerned, the non-employed preferred programmes which last over a long duration and can be completed in smaller sections (54.1 percent). Flexible entry points (58.6 percent) and opportunities to communicate with other participants (66.7 percent) were also important for them. Another aspect is relevant here: The non-employed people favoured short arrival routes (66.7 percent) and the receipt of a certificate (73.9 percent). Compared to the working population, there are only minimal differences. These results

Table 1: Further education topics in the population survey

Further Education Topics	Number of Mentions
Job and career	24
Languages	15
Information technology	14
Art, culture and history	7
Politics, economy and society	7
Environment	6
Health and nutrition	6
Natural science and technology	5
Others	7
Total:	91

can also be applied to continuing higher education, because these design requirements are also important here.

The non-employed people were also asked on which topics they would like to receive further training. Altogether many topics were named. 49 persons from the non-employed group answered this question. Thus, 91 statements were collected (see Table 1). The statements were categorised according to subject areas. There were 24 mentions in the subject area of (specific) career and profession – e.g. in the fields of geriatric care, business administration, warehouse logistics or mechatronics. Languages like English, Spanish or Italian (15 mentions) were mentioned second-most frequently and IT topics third-most frequently. There were seven mentions each for the subject areas "art, culture and history" and "politics, economy and society"; six mentions each for "health and nutrition" and "natural science and technology". They specifically mentioned scientific interests like scientific education and research results. The remaining statements were summarised under the category "others". However, it is not always clear whether the subjects were named for professional reasons or for personal interest.

In the following section, these results are discussed and what effect a demand-oriented development of study programmes may have for the regional development of the Westpfalz region.

5 Discussion of the Results

The Westpfalz region in particular, as a structurally weak region with special challenges, needs new approaches to meet them. The regional population survey helps to identify further education demands by planning (regional) study programmes. Continuing higher education in Germany has so far mainly focused on traditional target groups of those who have a first academic degree, especially working academics (Faulstich et al. 2008, Wolter 2015). Thus, it is particularly important to address new target groups. In the population survey, the group of those who are not employed was identified as such a new target group. Because of the higher rate of unemployment and lower rate of participation in further education in the Westpfalz region, it is interesting to consider this group in greater detail.

In some points, the survey responses show that both groups – non-employed and employed people – have similar expectations concerning continuing higher education. The vast majority of the two groups prefer a certificate to a bachelor's or master's degree. Regarding the mode of learning, there is also a similarity between the two groups. The importance of communication with other participants, flexible entry points and the desire to complete an offer in smaller sections over a longer period were emphasised. These aspects show that there is an understanding – independent of the group of persons – on what continuing higher education has to provide. Comparing the difference between the employed and the non-employed group, it is astonishing that the non-employed participants have a higher willingness to participate in continuing higher education. One reason could be that they have more time to spend on further education than the employed people do. Another aspect could be that, despite the unemployment, they still want to participate in social life and civil society activities. A further indication of this is also seen in the more highly valued importance of social interaction and exchange, as it is mentioned in the statement that they prefer opportunities to communicate with other participants (66.7 percent). A third aspect is that the factor of social desirability of the responses may have also contributed to the group difference.

The results of different national studies show that persons who were involved in civil society organisations, participated in educational activities significantly more frequently (Kaufmann-Kuchta, Kuper 2017, Smith 1994). If the participation rate could be increased, a higher involvement in civil society organisations could be forced in the Westpfalz region, and thus have a positive effect on the common social life.

The population survey shows that more than 61.6 percent of the group of non-employed has no regular "traditional" university entrance qualification, which means that they have left school without Abitur by the first or second educational route. However, it is interesting that continuing higher education at HEIs is taken into consideration. In Rheinland-Pfalz in particular, the higher education regulations have been adapted for non-traditional students (e.g. persons without an academic background) to promote an increase in the educational level in the region. One conclusion would be that persons without an academic background are not yet sufficiently informed about the further education possibilities at universities. These potentials can be used by universities in cooperating with regional stakeholders to better reach this target group, e.g. through targeted marketing measures.

Another aspect could be that further education offers do not yet meet individual requirements. This is also seen in the context of the high level of agreement with the statement that they would participate in further education if it could be integrated into their everyday life more easily. The group of non-employed people had fewer family and work responsibilities than the employed people, but programmes should nevertheless be flexible to reach this target group, e.g. low threshold and short-term offers like Massive Open Online Courses (MOOCs) or web-based offers as an entry into study programmes. MOOCs are particularly useful here because their flexibility in terms of time and space make them easier to integrate into everyday life. Furthermore, the results demonstrate that further education was perceived as interesting and as a natural part of one's life, even though only half of the respondents agreed. One explanation might be that this item also includes vocational development, which is not relevant for all persons of the group.

A distinguishing difference between the groups, is that the non-employed population preferred short arrival routes. This may be caused by being short of money and supports the fact that mature students often choose programmes in close proximity to their place of residence (Gibbons, Vignoles 2012, Harker et al. 2001). A national study also concluded that good accessibility is particularly important for groups with a lower educational background (without university entrance qualification) to enable participation in further education (Stöhr, Baur 2018). Therefore, the university benefits as a location here, especially in light of the fact that these people are interested in scientific topics such as research results in general, history, or computer science. Regarding the third-most frequently named topic of information technologies, it is interesting that the surveyed companies in the Westpfalz region stated that they have the greatest need for further education in the area of information technology and business administration (Steinmüller 2018). If this interest is met by appropriate study programmes or courses (especially for unemployed people), this would also benefit the companies.

If the attractiveness of study programmes at universities and the participation rate will be increased for this target group in particular, e.g. through flexible programmes, this can have a positive effect on regional development. These programmes can contribute to updating professional skills, which is helpful in the context of economic development dynamics, innovations and structural changes. In general, they can contribute to raising the level of education and educational opportunities by enabling or maintaining social participation for those who are not working or who are unable to work. It is precisely through demand-oriented approach that social challenges can be identified and linked to research results. For example, the results of the population survey show that also the non-employed participants were interested in research results or IT themes. By establishing more study courses in this area, this can accelerate the further development of the Westpfalz region into a service, science and IT location, as stated above.

In programmes, a critical and reflective examination of social changes could be supported. Further education is strengthened, which increases the attractiveness for continuing higher education and leads to a higher participation rate. If this regional target group of non-employed people is better motivated and integrated into study programmes, they will participate more on continuing higher education programmes, therefore enabling greater social participation, which could lead to more life satisfaction for them. Ultimately, this has an overall impact on society in the region. Beyond that, it helps to counteract the demographic change in the Westpfalz by increasing further education participation rate and reducing the unemployment rate. Finally, this may produce a suction effect. Eventually, more people move to the region and the settlement of companies and non-university research institution is favoured, which will also benefit regional development of the Westpfalz region (Rohs, Steinmüller 2020).

6 Conclusions

Traditionally, the group of working academics are focused on continuing higher education. In the last years, however, the higher education system in Germany has been opened up to promote access to different groups, especially those with vocational qualifications (Wolter 2015, Wolter, Kerst 2015). This is also reflected in higher education regulation (e.g. Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland 2009, Landesrecht Rheinland-Pfalz 2018). It allows persons without a traditional higher education entrance qualification to study at a university in Germany. Due to economic interest, the group of employed persons is the main target group. The aim was to show to what extent other target groups are also interesting, especially regarding the challenges in the Westpfalz region, such as the demographic change and the low level of participation in further education.

The results of the regional population survey suggest that it is worth considering non-employed people as a new target group for continuing higher education. The demands for further education on the part of this target group was identified and has shown among other things that they have a higher willingness to participate than the employed group. These results also give hints for other regions with a high rate of non-employed people who are basically fit for work, or for regions that are confronted with social problems, but where non-employed people are interested in research findings or scientific topics. This could help to arouse their interest in social engagement, which could have an overall positive impact for region with social problems. In this contribution, the demands of the target groups were mainly considered, but these are always in relation to social and entrepreneurial demands.

The opportunity to base the development of study programmes on such data not only help to make continuing higher education at HEIs more attractive, it also allows a wider focus on new target groups and possibilities to influence regional development through continuing higher education. In addition, such a demand-oriented development of continuing higher education programmes should be forced to comply with the responsibility of HEIs. Thereby, a rise in the attractiveness and competitiveness of the region as a business location is entailed, the impending shortage of skilled workers is counteracted and social participation is improved (Teichler 1991, Wolter 2011). This demand-orientation is advantageous in many respects, also for other regions that have similar regional challenges. Even it involves a lot of effort.

It has been shown that the group of non-employed could also contribute to regional development if the needs of this target group are given greater consideration in the development of study programmes in continuing higher education. Here it is exciting to carry out further research into the extent to which a demand-oriented approach really increases the participation of the corresponding target group.

Consequently, universities and continuing higher education, in particular, have a special responsibility in regards to regional development by increasing attractiveness and prosperity of a region, particularly for the structural weak Westpfalz region and other regions with similar characteristics. If the development and design of study programmes will be geared to the needs of the different regional target groups, there is at least an opportunity to meet the social and economic challenges of a region. Therefore, demand-oriented development of study programmes can serve as a catalyst for regional development. In this way, the universities also fulfil their tasks as described in the third mission. Ultimately, universities really become places of lifelong-learning.

It is assumed that further education in general, and continuing higher education in particular is of outstanding importance, has a demonstrable (potential) significance for the course of regional developments, and thus also represents a resource that could be drawn upon in the context of efforts (Nuissl 2000) – not only for the Westpfalz region. It would be worthwhile to have more specific, empirical data on this target group of non-employed people from other regions, especially with regard to their expectations by participation in continuing higher education. Previously, it would also be interesting to seek out other new region-specific target groups for continuing higher education in order to increase the attractiveness and the participation rate. In the long term, it is then necessary to investigate whether a demand-oriented approach really have a positive effect to regional development.

In the future, it is important to identify the further education demands of the regional population and companies. In addition, for universities and regional players it will be essential to cooperate more closely in order to collect these data and therefore develop programmes that meet regional challenges.

References

Aspin D, Chapman J, Hatton M, Sawano Y (2001) Introduction and overview. In: Aspin D, Chapman J, Hatton M, Sawano Y (eds), *International Handbook of Lifelong Learning*. Springer, Dordrecht. CrossRef.

Benneworth P, Hospers GJ (2007) The new economic geography of old industrial regions: Universities as global – local pipelines. *Environment and Planning C: Government and Policy* 25: 779–802. CrossRef.

Bernhard C (2017) *Erwachsenenbildung und Region: Eine empirische Analyse in Grenzräumen*. Theorie und Praxis der Erwachsenenbildung. W. Bertelsmann Verlag, Bielefeld

Bürmann M, Frick F (2016) Deutscher Weiterbildungsatlas: Teilnahme und Angebot in Kreisen und kreisfreien Städten. Bertelsmann Stiftung, Gütersloh

Bundesagentur für Arbeit (2019) Arbeitsmarkt im Überblick. Retrieved from https://statistik.arbeitsagentur.de/Navigation/Statistik/Statistik-nach-Regionen/BA-Gebietsstruktur/Rheinland-Pfalz-Saarland/Kaiserslautern-Pirmasens-Nav.html

Carrión A, García-Gutiérrez VR, Bas MC, Carot JM (2012) A new methodology for measuring third mission activities of universities. Technical University of Valencia, Valencia

Dobischat R, Düsseldorf C, Nuissl E, Stuhldreier J (2006) Lernende Regionen – begriffliche Grundlagen. In: Nuissl E (ed), *Theorie und Praxis der Erwachsenenbildung. Regionale Bildungsnetze: Ergebnisse zur Halbzeit des Programms "Lernende Regionen – Förderung von Netzwerken"*. Bertelsmann, Bielefeld

Eckert T, Tippelt R (2017) Learning Regions – Learning Cities – Learning Communities: Auf dem Weg zur Gestaltung regionaler Bildungsräume? In: Eckert T, Gniewosz B (eds), *Bildungsgerechtigkeit*. Verlag für Sozialwissenschaften, Wiesbaden. CrossRef.

Faulstich P, Graeßner G, Schäfer E (2008) Weiterbildung an Hochschulen: Daten zu Entwicklungen im Kontext des Bologna-Prozesses. *REPORT – Zeitschrift für Weiterbildungsforschung* 31: 9–18. https://www.die-bonn.de/doks/faulstich0801.pdf

Fritsch M (2009) Was können Hochschulen zur regionalen Entwicklung beitragen? *Die Hochschule* 1: 39–52. http://ids.hof.uni-halle.de/documents/t1868.pdf

Fritsch M, Henning T, Slavtchev V, Steigenberger N (2008) Hochschulen als regionaler Innovationsmotor? Innovationstransfer aus Hochschulen und seine Bedeutung für die regionale Entwicklung. Arbeitspapier 158, Hans-Böckler-Stiftung, Düsseldorf

Gibbons S, Vignoles A (2012) Geography, choice and participation in higher education in England. *Regional Science and Urban Economics* 42: 98–113. CrossRef.

Harker D, Slade P, Harker M (2001) Exploring the decision process of school leavers' and 'mature students' in university choice. *Journal of Marketing for Higher Education* 11: 1–20. CrossRef.

Kallen D, Bengtsson J (1973) Recurrent education: A strategy for lifelong learning. Organisation for Economic Co-operation and Development (OECD), Paris

Kaufmann-Kuchta K, Kuper H (2017) Informelles Lernen und soziale Teilhabe. In: Bilger F, Behringer F, Kuper H, Schrader J (eds), *Weiterbildungsverhalten in Deutschland 2016: Ergebnisse des Adult Education Survey (AES)*. Wbv, Bielefeld, 185–201

Kujath HJ (2015) Wissensgesellschaftliche Raumdifferenzierung in Deutschland. In: Fritsch M, Pasternack P, Titze M (eds), *Schrumpfende Regionen – dynamische Hochschulen: Hochschulstrategien im demografischen Wandel*. Verlag für Sozialwissenschaften, Wiesbaden. CrossRef.

Landesrecht Rheinland-Pfalz (2018) Hochschulgesetz in der Fassung vom 19. November 2010. http://landesrecht.rlp.de/jportal/portal/t/n4s/page/bsrlpprod.psml/js_pane/Dokumentanzeige#focuspoint

Longworth N (2006) *Learning cities, learning regions, learning communities: Lifelong learning and local government*. Routledge, London. CrossRef.

Ludewig O, Otto A, Stabler J (2007) Arbeitsmarkt Westpfalz: Sektoraler Strukturwandel in den Agenturbezirken Kaiserslautern und Pirmasens. IAB-Regional Rheinland-Pfalz-Saarland, Nürnberg

Marks S (2015) Region als Bezugsraum für Hochschulentwicklung: Regionsdefinition für das Projekt E-hoch-B. Arbeits- und Forschungsbericht aus dem Projekt E-hoch-B – Bildung als Exponent individueller und regionaler Entwicklung No. 1. Kaiserslautern und Ludwigshafen. https://kluedo.ub.uni-kl.de/frontdoor/index/index/docId/4111

Martin A, Schömann K, Schrader J, Kuper H (2015) Ausgewählte Ergebnisse: Die Wiederentdeckung der Bedeutung der Region. In: Martin A, Schömann K, Schrader J, Kuper H (eds), *Deutscher Weiterbildungsatlas*. wbv, Bielefeld. CrossRef.

Molas-Gallart J, Salter A, Patel P, Scott A, Duran X (2002) Measuring third stream activities. Final report to the Russel group of universities. Science and technology policy research unit university of Sussex, Brighton

Nuissl H (2000) Weiterbildung und "regionale Lernprozesse". *Raumforschung und Raumordnung* 58: 467–476. CrossRef.

Pasternack P, Zierold S (2014) Überregional basierte Regionalität: Hochschulbeiträge zur Entwicklung demografisch herausgeforderter Regionen (Special issue). HoF-Handreichungen 4. Beiheft zu "die Hochschule". Institut für Hochschulforschung (HoF), Halle-Wittenberg. Prognos AG

Prognos AG (2016) Prognos Zukunftsatlas – Das Ranking für Deutschlands Regionen. Berlin. retrieved from https://www.prognos.com/fileadmin/images/publikationen/Zukunftsatlas2016/Prognos_Zukunftsatlas_2016_Auf_einen_Blick.pdf

Prognos AG (2019) Prognos Zukunftsatlas – Das Ranking für Deutschlands Regionen. Berlin. retrieved from https://www.prognos.com/publikationen/zukunftsatlas-r-regionen/zukunftsatlas-2019/

Rohs M, Steinmüller B (2020) Wissenschaftliche Weiterbildung und Region. In: Jütte W, Rohs M (eds), *Handbuch Wissenschaftliche Weiterbildung*. Verlag für Sozialwissenschaften, Wiesbaden. CrossRef.

Rohs M, Vogel C, Marks S (2015) From supply-driven to demand-oriented academic education: Evidence-based development of study courses to match regional skill shortage with new student groups. Submission for the 12^{th} PASCAL International Observatory Conference. Retrieved from https://www.researchgate.net/profile/Matthias_Rohs2/research

Schäfer E (1988) *Wissenschaftliche Weiterbildung als Transformationsprozeß: Theoretische, konzeptionelle und empirische Aspekte*. Verlag für Sozialwissenschaften, Wiesbaden. CrossRef.

Schwikal A, Steinmüller B (2017) Die Bedarfsanalyse im Projekt E-hoch-B: Das Forschungsdesign. Arbeits- und Forschungsbericht aus dem Projekt E-hoch-B – Bildung als Exponent individueller und regionaler Entwicklung No. 14. Kaiserslautern und Ludwigshafen am Rhein. Retrieved from https://kluedo.ub.uni-kl.de/frontdoor/index/index/docId/4695

Seitter W, Schemmann M, Vossebein U (2015) Bedarf – Potenzial – Akzeptanz. Integrierende Zusammenschau. In: Seitter W, Schemmann M, Vossebein U (eds), *Zielgruppen in der wissenschaftlichen Weiterbildung*. Verlag für Sozialwissenschaften, Wiesbaden. CrossRef.

Slupina M, Dähner S, Klingholz R, Reibstein L, Scholz J (2019) *Die demografische Lage der Nation. Wie zukunftsfähig Deutschlands Regionen sind*. Berlin-Institut für Bevölkerung und Entwicklung, Berlin

Smith DH (1994) Determinants of voluntary association participation and volunteering: A literature review. *Nonprofit and voluntary sector quarterly* 23: 243–263. CrossRef.

Statistisches Amt Saarland (2018) Statistische Berichte: Bevölkerungsentwicklung im 4. Vierteljahr 2014. Bevölkerungsstand am 31. Dezember 2014. Retrieved from https://www.saarland.de/dokumente/thema_statistik/STALA_BER_AI1-VJ4-14.pdf

Statistisches Landesamt Rheinland-Pfalz (2017) Kreisfreie Städte und Landkreise in Rheinland-Pfalz: Ein Vergleich in Zahlen. Bad Ems. Retrieved from https://www.statistik.rlp.de/fileadmin/dokumente/kreisuebersichten/Kreisuebersichten_2016.pdf

Steinmüller B (2018) Fachkräfte- und Weiterbildungsbedarfe von Unternehmen in der Region Westpfalz. Arbeits- und Forschungsbericht aus dem Projekt E-hoch-B – Bildung als Exponent individueller und regionaler Entwicklung No. 21. Kaiserslautern und Ludwigshafen am Rhein. Retrieved from https://kluedo.ub.uni-kl.de/frontdoor/index/index/docId/5217

Stöhr I, Baur HR (2018) Zu weit weg? Lokale Angebotsstruktur als Zugangsbedingung für Weiterbildung. *Erwachsenenbildung.at. Das Fachmedium Für Forschung, Praxis Und Diskurs* 34: 1–14. https://erwachsenenbildung.at/magazin/artikel.php?aid=12462#12462

Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland (2009) Hochschulzugang für beruflich qualifizierte Bewerber ohne schulische Hochschulzugangsberechtigung: Beschluss der Kultusministerkonferenz vom 06.03.2009. Retrieved from http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2009/2009_03_06-Hochschulzugang-erful-qualifizierte-Bewerber.pdf

Teichler U (1991) Towards a highly educated society. *Higher Education Policy* 4: 11–20. CrossRef.

Widany S, Wolter A, Dollhausen K (2020) Monitoring wissenschaftlicher Weiterbildung: Status quo und Perspektiven. In: Jütte W, Rohs M (eds), *Handbuch Wissenschaftliche Weiterbildung*. Verlag für Sozialwissenschaften, Wiesbaden. CrossRef.

Wolter A (2005) Profilbildung und universitäre Weiterbildung. In: Jütte W (ed), *Kontexte wissenschaftlicher Weiterbildung: Entstehung und Dynamik von Weiterbildung im universitären Raum*. Waxmann, Münster

Wolter A (2011) Die Entwicklung wissenschaftlicher Weiterbildung in Deutschland: Von der postgradualen Weiterbildung zum Lebenslanges Lernen. *Beitrag zur Hochschulforschung* 33: 8–35

Wolter A (2015) Opening up higher education for new target groups in Germany: A case study for the development of university lifelong learning. In: Zgaga P, Teichler U, Schuetze HG, Wolter A (eds), *Higher Education Reform: Looking Back – Looking Forward*. Peter Lang, Frankfurt. CrossRef.

Wolter A, Kerst C (2015) The 'academization' of the German qualification system: Recent developments in the relationships between vocational training and higher education in Germany. *Research in Comparative and International Education* 10: 510–524. CrossRef.

REGION
The Journal of ERSA
Powered by WU

Volume 7, Number 2, 2020, 43–83
DOI: 10.18335/region.v7i2.324
journal homepage: region.ersa.org

Flatten the Curve! Modeling SARS-CoV-2/COVID-19 Growth in Germany at the County Level

Thomas Wieland[1]

[1] Karlsruhe Institute of Technology, Karlsruhe, Germany

Received: 14 May 2020/Accepted: 2 November 2020

Abstract. Since the emerging of the "novel coronavirus" SARS-CoV-2 and the corresponding respiratory disease COVID-19, the virus has spread all over the world. Being one of the most affected countries in Europe, in March 2020, Germany established several nonpharmaceutical interventions to contain the virus spread, including the closure of schools and child day care facilities (March 16-18, 2020) as well as a full "lockdown" with forced social distancing and closures of "nonessential" services (March 23, 2020). The present study attempts to analyze whether these governmental interventions had an impact on the declared aim of "flattening the curve", referring to the epidemic curve of new infections. This analysis is conducted from a regional perspective. On the level of the 412 German counties, logistic growth models were estimated based on daily infections (estimated from reported cases), aiming at determining the regional growth rate of infections and the point of inflection where infection rates begin to decrease and the curve flattens. All German counties exceeded the peak of new infections between the beginning of March and the middle of April. In a large majority of German counties, the epidemic curve has flattened before the "lockdown" was established. In a minority of counties, the peak was already exceeded before school closures. The growth rates of infections vary spatially depending on the time the virus emerged. Counties belonging to states which established an additional curfew show no significant improvement with respect to growth rates and mortality. Furthermore, mortality varies strongly across German counties, which can be attributed to infections of people belonging to the "risk group", especially residents of retirement homes. The decline of infections in absence of the "lockdown" measures could be explained by 1) earlier governmental interventions (e.g., cancellation of mass events, domestic quarantine), 2) voluntary behavior changes (e.g., physical distancing and hygiene), 3) seasonality of the virus, and 4) a rising but undiscovered level of immunity within the population. The results raise the question whether formal contact bans and curfews really contribute to curve flattening within a pandemic.

1 Background

The "novel coronavirus" SARS-CoV-2 ("Severe Acute Respiratory Syndrome Coronavirus 2") and the corresponding respiratory disease COVID-19 ("Coronavirus Disease 2019") caused by the virus initially appeared in December 2019 in Wuhan, Province Hubei, China. Since its emergence, the virus has spread over nearly all countries across the world. On March 12, 2020, the World Health Organization (WHO) declared the SARS-CoV-2/COVID-19 outbreak a global pandemic (Lai et al. 2020, World Health Organization 2020b). As of May 10, 2020, 3,986,119 cases and 278,814 deaths had been reported worldwide. In Europe, the most affected countries are Spain, Italy, United Kingdom and Germany (European Centre for Disease Prevention and Control 2020).

The virus is transmitted between humans via droplets or through direct contact (Lai et al. 2020). In a very influential simulation study from March 2020, the Imperial College COVID-19 Response Team (Ferguson et al. 2020) suggested a series of public health measures aimed at slowing or stopping the transmission of the virus in absence of a vaccine or a successful therapy. These so-called *nonpharmaceutical interventions* (NPI) aim at reducing contact rates in the population, including social distancing and closures of schools and universities as well as the quarantine of infected persons. The Chinese government had imposed containment measures in the Provice Hubei already at the end of January 2020. This "lockdown" included a quarantine of the most affected city Wuhan and movement restrictions for the population as well as school closures (CNN 2020). In March 2020, nearly all European countries have introduced measures against the spread of Coronavirus. These measures range from appeals to voluntary behaviour changes in Sweden to strict curfews, e.g. in France and Spain (Deutsche Welle 2020a). The public health strategy to contain the virus spread is commonly known as "flatten the curve", which refers to the epidemic curve of the number of infections: "Flattening the curve involves reducing the number of new COVID-19 cases from one day to the next. This helps prevent healthcare systems from becoming overwhelmed. When a country has fewer new COVID-19 cases emerging today than it did on a previous day, that's a sign that the country is flattening the curve" (Johns Hopkins University 2020).

In Germany, due to the federal political system, measures to "flatten the curve" were introduced on the national as well as the state level. As the German "lockdown" has no single date, we distinguish here between four phases of NPIs, of which the main interventions were the closures of schools, child day care centers and most retail shops etc. in calendar week 12 (phase 2), and the nationwide establishment of a contact ban (attributed to phase 3), including forced social distancing and a ban of gatherings of all types, on March 23, 2020. The German states Bavaria, Saarland, and Saxony established additional curfews (see Table 1). Occasionally, these governmental interventions were criticized because of the social, psychological and economic impacts of a "lockdown" and/or the lack of its necessity (Capital 2020, Süddeutsche Zeitung 2020a, Tagesspiegel 2020a, Welt online 2020a). Apart from the economic impacts emerging from a worldwide recession (The Guardian 2020), the psychosocial consequences of movement restrictions and social isolation (resulting from NPIs) have also become apparent now in terms of an increase of several mental health illnesses (Carvalho Aguiar Melo, de Sousa Soares 2020, Mucci et al. 2020, Williams et al. 2020). The effects of (forced) isolation as well as school and child day care closures are also visible through a worldwide increase in domestic abuse (New York Times 2020), reported in Germany as well (Stuttgarter Zeitung 2020, Süddeutsche Zeitung 2020b).

It is therefore all the more important to know whether these restrictions really contributed to the flattening of the epidemic curve of Coronavirus in Germany (RKI 2020a). This question should be addressed from a regional perspective for two reasons.

1. In May 2020, the competences for the measures in Germany have shifted from the national to the state and regional (county) level. In the future, counties with more than 50 new infections per 100,000 in one week are expected to implement regional measures (see Table 1).

2. A spatial perspective allows the impact of the German measures of March 2020 to be identified.

In his statistical study, the mathematician Ben-Israel (2020) compares the epidemic curves of Israel, the USA and several European countries. These curves demonstrate a decline of new infections, regardless of the national measures to contain the virus spread. Furthermore, the study reveals the trend that the peak of infections is typically reached in the sixth week after the first reported case, while a decline of the curve starts in week eight. This occurs in all assessed countries on the national level, no matter whether a "lockdown" was established (e.g. Italy) or not (e.g. Sweden).

Table 1: Main governmental nonpharmaceutical interventions with respect to COVID-19 pandemic in Germany

Phase	Measure	Entry into force	Competence /level
1	First quarantines of infected persons and suspected cases	February 2020	nationwide
up to CW	Minister of health Spahn recommends cancellation of large events ($\geq$ 1,000 participants)	(March 8, 2020)	
10/11	Bundesliga games behind closed doors ("ghost games")	March 11, 2020	nationwide
	Speeches of chancellor Merkel and president Steinmeier, recommondation to avoid social contacts and large events	(March 12, 2020)	
2	Closure of schools, child day care centers and universities	March 16-18, 2020	states
CW 12	Closure of retail facilities (except for basic supply), bars and leisure facilities	March 17-19, 2020	states
	Travel restrictions	March 17, 2020	nationwide
3	Curfew in Bavaria, Saarland and Saxony	March 21-23, 2020	states
CW 12/13 ("Lock-down")	Contact ban: ban of gatherings > 2 people (including political and religious gatherings), forced social distancing (distance $\geq$ 1.5 m), closure of "nonessential" services (e.g., gastronomy, hairdressers)	March 23, 2020	nationwide
4	Reopening of several retail facilities and services	April 20, 2020	states
CW	Mandatory face masks in public transport and shops	April 22-29, 2020	states
17	Further liberalizations; implementation of an "emergency brake": lockdowns on the county level on condition of 50 new infections per 100,000 in one week	May 6, 2020	nationwide

Source: own compilation based on an der Heiden, Hamouda (2020), Deutsche Welle (2020a,b), Tagesschau.de (2020a,b).

The focus of the present study is on the main nonpharmaceutical interventions with respect to the SARS-CoV-2/COVID-19 pandemic in Germany. This means the concrete "lockdown" measures affecting the social and economic life of the whole society (distinguishing from measures taken in most cases of infectious dieseases, such as quarantine of affected persons). In the terminology of the present study, these are the phase 2 and 3 measures, denoted in Table 1. Building upon the discrepancy outlined by Ben-Israel (2020), the present study addresses the following research questions:

- Pandemic or epidemic growth has a regional component due to regional infection hotspots or other behavorial or spatial factors. Thus, growth rates of infections may differ between regions in the same country (Chowell et al. 2014). In Germany, the prevalence of SARS-CoV-2/COVID-19 differs among the 16 German states and 412 counties, clearly showing "hotspots" in South German counties belonging to Baden-Wuerttemberg and Bavaria (RKI 2020a). Thus, the first question to be answered is: *How does the growth rate of SARS-CoV-2/COVID-19 vary across the 412 German counties?*

- The German measures to contain the pandemic entered into force nearly at the same time, especially in terms of closures of schools, childcare infrastructure and retailing (starting March 16/17, 2020) as well as the nationwide contact ban (starting March 23, 2020). Ben-Israel (2020) found a decline of new infection cases on the national level regardless of the Corona measures. To examine the effect of the German measures, we need to estimate the time of the peak and the declining of the curves of infection cases, respectively: *At which date(s) did the epidemic curves of SARS-CoV-2/COVID-19 flatten in the 412 German counties?*

- Regional prevalence and growth, as well as the mortality of SARS-CoV-2/COVID-19, are attributed within media discussions to several spatial factors, including population density or demographic structure of the regions (Welt online 2020b). Furthermore, the German measures differ on the state level, as three states – Bavaria, Saarland and Saxony – established additional curfews supplementing the other interventions (see Table 1). Focusing on growth rate and mortality, and addressing these regional differences, the third research question is: *Which indicators explain the regional differences of SARS-CoV-2/COVID-19 growth rate and mortality on the level of the 412 German counties?*

2 Methodology

2.1 Logistic growth model

According to Li (2018), in simple terms, an infectious disease spread (pandemic or epidemic) can be summarized as follows: At the beginning, one or more infectious individuals are introduced into a population of *susceptibles* (non-infected/healthy individuals). As the pathogen (e.g., virus) is transmitted from one individual to another, the number of *infected* individuals increases over time. Depending on the regarded pathogen/disease, infected individuals *recover* due to medical interventions and/or reactions of the individuals' immune system and, in many cases, gain partial or full immunity against the pathogen (e.g., through the development of antibodies against a virus). In other cases, infected people may also die from the disease. In all aforementioned cases and on condition of a stationary population, the number of susceptibles decreases and, thus, the number of new infections decreases as well. As a consequence, the pandemic/epidemic slows down and ends. The disease spread may also be contained by vaccination and/or other control and preventive measures. Note that, technically, one must distinguish between an *infection* and the *disease* which is (or may be) caused by the pathogen: "Disease is not the same as infection. Infection is said to have occurred when an organism successfully avoids innate defense mechanisms and stably colonizes a niche in the body. To establish an infection, the invader must first penetrate the anatomic and physiological barriers that guard the skin and mucosal surfaces of the host. Secondly, the organism must be able to survive in the host cellular milieu long enough to reproduce. This replication may or may not cause visible, clinical damage to the host tissues, symptoms that we call 'disease"' (Mak, Saunders 2006).

Analyzing the transmission and spreading process of infectious diseases involves the utilization of mathematical models. Pandemic growth can be modeled by deterministic models such as the SIR (susceptible-infected-recovered) model and its extensions, or by stochastic, phenomenological models such as the exponential or the logistic growth model. The former type of model does not depend on large empirical data on disease cases but requires additional information about the disease and the transmission process. The latter type of model is based on linear or nonlinear regression. Only empirical data of infections and/or confirmed cases of disease (or death) is required to estimate such models (Batista 2020a,b, Chowell et al. 2014, 2015, Li 2018, Ma 2020, Pell et al. 2018). Recently, there have already been several attempts to model the SARS-CoV-2 pandemic on the country (or even world) level, by using either the original or extended SIR model (Batista 2020b), the logistic growth model (Batista 2020a, Vasconcelos et al. 2020, Wu et al. 2020), or both (Zhou et al. 2020).

In this paper, we regard the spread of the Coronavirus primarily as an empirical phenomenon over space and time and ignore its epidemiological characteristics. We focus on 1) the regional growth speed of the pandemic and 2) the time when exponential growth ends and the infection rate decreases again. Apart from that, only infection cases and some further information are available, but not additional epidemiological information. Thus, the method of choice is a phenomenological regression model. In an early phase of an epidemic, when the number of infected individuals growths exponentially, an exponential function could be utilized for the phenomenological analysis (Ma 2020). However, officially reported SARS-CoV-2 infections in Germany (measured by the time of onset of symptoms) declined from mid-March. The corresponding reproduction number was estimated at $R = 0.71$ based on the case reports as of May 6, 2020 (RKI 2020a). This indicates that the phase of exponential growth was exceeded at this time. Thus, a logistic growth model is used for the analysis of SARS-CoV-2 growth in the German counties.

The following representations of the logistic growth model are adopted from Batista

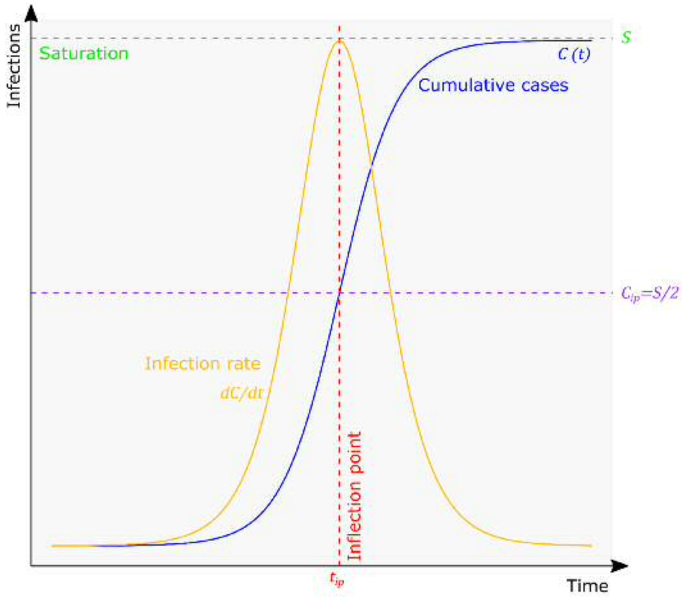


Figure 1: Logistic growth of an epidemic

(2020a), Chowell et al. (2014) and Tsoularis (2002). Unlike exponential growth, logistic growth includes two stages, allowing for a saturation effect. The first stage is characterized by an exponential growth of infections due to an unregulated spreading of the disease. As more infections accumulate, the number of at-risk susceptible persons decreases because of immunization, death, or behavioral changes as well as public health interventions. After the inflection point of the infection curve, when the infection rate is at its maximum, the growth decreases and the cumulative number of infections approximate its theoretical maximum, which is the saturation value (see Figure 1).

In the logistic growth model, the cumulative number of infected or diseased persons at time t, $C(t)$ is a function of time:

$$C(t) = \frac{C_0 S}{C_0 + (S - C_0)\exp(-rSt)} \tag{1}$$

where C_0 is the initial value of C at time 0, r is the intrinsic growth rate, and S is the saturation value.

The infection rate is the first derivative:

$$\frac{dC}{dt} = rC\left(1 - \frac{C}{S}\right) \tag{2}$$

The inflection point of the logistic curve indicates the maximal infection rate before the growth declines, which means a flattening of the cumulative infection curve. The inflection point, ip, is equal to:

$$ip = \frac{S}{2} \tag{3}$$

at time

$$t_{ip} = \frac{c}{rS} \tag{4}$$

where:

$$c = \ln\frac{C_0}{C - C_0} \tag{5}$$

When empirical data (here: time series of cumulative infections) is available, the three model parameters r, S and C_0 can be estimated empirically.

We fit the models in a three-step estimation procedure including both OLS (Ordinary Least Squares) and NLS (Nonlinear Least Squares) estimation. The former is used for generating initial values for the iterative NLS estimation, making use of the linearization and stepwise parametrization of the logistic function. Following Engel (2010), the nonlinear logistic model (Equ. 1) can be transformed into a linear model (on condition that the saturation value is known) by taking the reciprocal on both sides, taking natural logarithms and rearranging the function:

$$\ln(\frac{1}{C(t)} - \frac{1}{S}) = \ln(\frac{S - C_0}{SC_0}) - rSt \tag{6}$$

The transformed dependent variable, y_i^*, can be expressed by a linear relationship with two parameters, the intercept ($\hat{b}$) and slope ($\hat{m}$):

$$y_i^* = \ln(\frac{1}{C(t)} - \frac{1}{S}) \tag{7}$$

$$\hat{y}^* = \hat{b} + \hat{m}t \tag{8}$$

In step 1, an approximation of the saturation value is estimated, which is necessary for the linear transformation of the model. Transforming the empirical values $C(t)$ according to Equ. (7), we have a linear regression model (Equ. 8). By utilizing bisection (Kaw et al. 2011), the best value for S is searched minimizing the sum of squared residuals. The bisection procedure consists of 10 iterations, while the start values are set around the current maximal value of $C(t)$ (Interval: $[\max(C(t)) + 1;\ \max(C(t)) * 1.2]$).

The resulting preliminary start value for the saturation parameter, $\hat{S}_{\text{start}}$, is used in step 2. We transform the observed $C(t)$ using Equ. (7) with the preliminary value of $\hat{S}$ from step 1, $\hat{S}_{\text{start}}$. Another OLS model is estimated (Equ. 8). The estimated coefficients are used for calculating the start values of $\hat{r}$ and $\hat{C}_0$ for the nonlinear estimation (Engel 2010):

$$\hat{r}_{\text{start}} = -\frac{\hat{m}}{\hat{S}_{\text{start}}} \tag{9}$$

and

$$\hat{C}_{0_{\text{start}}} = \frac{\hat{S}_{\text{start}}}{1 + \hat{S}_{\text{start}} \exp{(\hat{b})}} \tag{10}$$

In step 3, the final model fitting is done using Nonlinear Least Squares (NLS), while inserting the values from steps 1 and 2, $\hat{S}_{\text{start}}$, $\hat{C}_{0_{\text{start}}}$ and $\hat{r}_{\text{start}}$, as start values for the iterative process. The NLS fitting uses the default Gauss-Newton algorithm (Ritz, Streibig 2008) with a maximum of 500 iterations.

Using the estimated parameters $\hat{r}$, $\hat{C}_0$ and $\hat{S}$, the inflection point of each curve is calculated via equations (3) to (5). The inflection point t_{ip} is of unit time (here: days) and assigned to the respective date $t_{ip_{\text{date}}}$ (YYYY-MM-DD). Based on this date, the following day $t_{ip_{\text{date}+1}}$ is the first day after the inflection point at which time the infection rate has decreased again. For graphical visualization, the infection rate is also computed using Equ. (2).

2.2 Estimating the dates of infection

In the present study, we use the daily updated data on confirmed SARS-CoV-2/COVID-19 cases, provided by federal authorities, the German Robert Koch Institute (RKI) (RKI 2020b). This dataset includes all persons who have been tested positive on the SARS-CoV-2 virus using a PCR (*polymerase chain reaction*) test and reported from local health authorities to the RKI. However, one must consider that neither the volume of tests nor the criteria for conducting a test are constant over time: Up to and including May 2020, almost exclusively people with acute respiratory symptoms were tested for SARS-CoV-2, as, with few exceptions, the presence of relevant symptoms is an exclusion criterion for testing in the RKI guidelines for medical doctors (IBBS 2020). In other words, this testing policy is targeted at the *disease* (COVID-19), not the *virus* (SARS-CoV-2). Thus, most of the cases in the present data are COVID-19 sufferers, whilst asymptomatic infected people and individuals with milder course are underrepresented. In the vast majority of cases, the date of onset of symptoms is reported in the dataset as well (an der Heiden, Hamouda 2020). The test volume was increased heavily from calendar week 11 (127,457 tests) to 12 (348,619 tests) but remains in the same order of magnitude until calendar week 18 (300,000-400,000 tests per week) (RKI 2020c).

The dataset used here is from May 5, 2020 and includes 163,798 cases. This data includes information about age group, sex, the related place of residence (county) and the date of report (Variable *Meldedatum*). The reference date in the dataset (Variable *Refdatum*) is either the day the disease started, which means the onset of symptoms, or the date of report (an der Heiden, Hamouda 2020). The date of onset of symptons is reported in the majority of cases (108,875 and 66.47%, respectively).

The date of infection, which is of interest here, is either unknown or not included in the official dataset. Thus, it is necessary to estimate the approximate date of infection dependent on two time periods: the time between the infection and the onset of symptons (incubation period) and the delay between onset of symptoms and official report (reporting delay). From the 108,875 cases where the onset of the symptoms is known, we can calculate the mean reporting delay as 6.84 days. Additionally, we assume an incubation period of five days. This is a rather conservative assumption (which means a relatively short time period) referring to the current epidemiological estimates (see Table 2). In their model-based scenario analysis towards the total number of diseases and deaths, the RKI also assumes an average incubation period of five days (an der Heiden, Buchholz 2020). Taking into account incubation period and reporting delay, there is an average all-over delay between infection and reporting of about 12 days (see Figure 2).

But this is just one side of the coin. As an inspection of the case data reveals, the delay differs by case characteristics (age group, sex) and counties. In their current prognosis, the RKI estimates the dates of onset of symptoms by Bayesian nowcasting based on the reporting date, but not taking into account the incubation time. The RKI nowcasting model incorporates delays of reporting depending on age group and sex, but not including spatial (county-specific) effects (an der Heiden, Hamouda 2020). Exploring the dataset used here, we see obvious differences in the reporting delay with respect to age groups and sex. There seems to be a tendency of lower reporting delays for young children and older infected individuals (see Table 3). Taking a look at the delays between onset of symptoms and reporting date on the level of the 412 counties (not shown in table), the values range between 2.39 days (Würzburg city) and 17.0 days (Würzburg county).

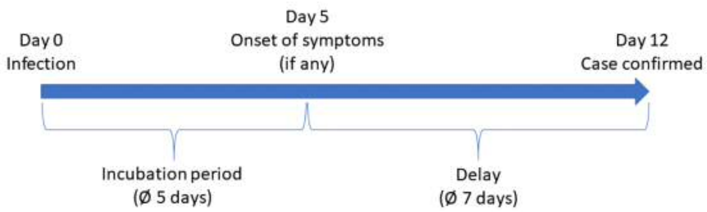


Figure 2: Time between infection and reporting of case

Table 2: Studys estimating the incubation period of SARS-CoV-2/COVID-19

Study	n	Distribution	Mean (CI95)	SD (CI95)	Median (CI95)	Min.	Max.
Backer et al. (2020)	88	Weibull	NA	2.3 (1.7, 3.7)	6.4 (5.6, 7.7)	NA	NA
Lauer et al. (2020)	181	Lognormal	5.5	1.52 (1.3, 1.7)	5.1 (4.5, 5.8)	NA	NA
Leung (2020)	175 (a)	Weibull	1.8 (1.0, 2.7)	NA	NA	NA	NA
	175 (b)	Weibull	7.2 (6.1, 8.4)	NA	NA	NA	NA
Li et al. (2020)	10	Lognormal	5.2 (4.1, 7.0)	NA	NA	NA	NA
Linton et al. (2020)	158	Lognormal	5.6 (5.0, 6.3)	2.8 (2.2, 3.6)	5.0 (4.4, 5.6)	2	14
Sun et al. (2020)	33	NA	4.5	NA	NA	NA	NA
Xia et al. (2020)	124	Weibull	4.9 (4.4, 5.4)	NA	NA	NA	NA

Notes: (a) = Travelers to Hubei, (b) = Non-Travelers. *Source*: own compilation.

Table 3: Delay between onset of symptoms and official report by age group and sex

Age group	Sex	Delay between onset of symptoms and reporting date [days] Mean	SD
A00-A04	female	5.82	5.68
A05-A14	female	6.09	5.34
A15-A34	female	6.82	5.59
A35-A59	female	7.00	5.98
A60-A79	female	7.08	6.22
A80+	female	5.10	5.83
unknown	female	8.71	9.79
A00-A04	male	5.93	5.98
A05-A14	male	6.04	5.12
A15-A34	male	6.78	5.52
A35-A59	male	7.18	5.90
A60-A79	male	7.20	6.14
A80+	male	5.70	5.84
unknown	male	9.86	8.42
A00-A04	unknown/diverse	3.50	2.39
A05-A14	unknown/diverse	4.00	5.20
A15-A34	unknown/diverse	6.44	4.87
A35-A59	unknown/diverse	6.95	5.51
A60-A79	unknown/diverse	7.36	5.87
A80+	unknown/diverse	6.50	11.40
unknown	unknown/diverse	9.60	3.58
all-over		6.84	5.90

Source: own calculation based on data from RKI (2020b). *Note*: The date of onset of symptoms is known for 108,875 (66.47%) of 163,798 cases in the dataset.

For the estimation of the dates of infection, it is necessary to distinguish between the cases where the date of symptom onset is known or not. In the former case, no assumption must be made towards the delay between onset of symptoms and date report. The calculation is simply:

$$\hat{di}_i = do_i - incp \tag{11}$$

where $\hat{di}_i$ is the estimated date of infection of case i, do_i is the date of onset of symptoms reported in the RKI dataset and $incp$ is the average incubation period equal to five (days).

For the 54,923 cases without information about onset of symptoms, we estimate this delay based on the 108,875 cases with known delays. As the reporting delay differs between age group, sex and county, the following dummy variable regression model is estimated (stochastic disturbance term is not shown):

$$\hat{ds}_{asc} = \alpha + \sum_{a}^{A-1} \beta_a D_{\text{agegroup}_a} + \sum_{s}^{S-1} \gamma_s D_{\text{sex}_s} + \sum_{c}^{C-1} \delta_c D_{\text{county}_c} \tag{12}$$

where $\hat{ds}_{asc}$ is the estimated delay between onset of symptoms and report depending on age group a, sex s and county c, D_{agegroup_a} is a dummy variable indicating age group a, D_{sex_s} is a dummy variable indicating sex s, D_{county_c} is a dummy variable indicating county c, A is the number of age groups, S is the number of sex classifications, C is the number of counties and α, β, γ and δ are the regression coefficients to be estimated.

Taking into account the delay estimation, if the onset of symptoms is unknown, the date of infection of case i is estimated via:

$$\hat{di}_i = dr_i - \hat{ds}_{asc} - incp \tag{13}$$

where dr_i is the date of report in the RKI dataset.

2.3 Models of regional growth rate and mortality

To test which variables predict the intrinsic growth rate and the regional mortality of SARS-CoV-2/COVID-19, respectively, two regression models were estimated. In the first model with the intrinsic growth rate r as dependent variable, we include the following predictors:

- In the media coverage about regional differences with respect to COVID-19 cases in Germany, several experts argue that a lower population density and a higher share of older population reduce the spread of the virus, with the latter effect being due to a lower average mobility (Welt online 2020b). It is well known that human mobility potentially increases the spread of an infectious disease. Also work-related commuting and tourism are considered as drivers of virus transmission (Charaudeau et al. 2014, Dalziel et al. 2014, Findlater, Bogoch 2018). To test these effects, four variables are included into the model: 1) The population density ($POPDENS$), 2) the share of population of at least 65 years ($POPS65$), 3) an indicator for the intensity of commuting (CMI) formulated by Guth et al. (2010), and 4) the number of annual tourist arrivals per capita ($TOUR$) for each county. All variables were calculated based on official statistics for the most recent year (2018/2019) (Destatis 2020a,b,c).

- In the media coverage, the lower prevalence in East Germany is also explained by 1) a different vaccination policy in the former German Democratic Republic and 2) a lower affinity towards carnival events as well as 3) less travelling to ski resorts due to lower incomes (Welt online 2020b). Thus, a dummy variable (1/0) for East Germany is included in the model ($EAST$).

- We test for the influence of different governmental interventions by including dummy variables for the states ("Länder") Bavaria (BV), Saarland (SL), Saxony (SX) and North Rhine Westphalia (NRW), as well as Baden-Wuerttemberg (BW). Unlike the other 13 German states, the first three states established a curfew additional to the other measures at the time of phase 3, as is identified in the present study. Like Bavaria, North Rhine Westphalia and Baden-Wuerttemberg belong to the "hotspots" in Germany, with the latter state having a prevalence similar to Bavaria. Saxony has a prevalence below the national average (RKI 2020a).

- Apart from any interventions, when a disease spreads over time, also the susceptible population *must* decrease over time. As more and more individuals get infected (maybe causing temporal or lifelong immunization or, in other cases, death), there are continually fewer healthy people to get infected (Li 2018) (see also Section 2.1). Consequently, regional growth *must* decrease with increasing regional prevalence and over time (and vice versa). In the specific case of SARSCoV-2/COVID-19, the outbreak differs between German counties (starting with "hotspots" like Heinsberg or Tirschenreuth county). Differences in growth may be due to different periods of time the virus is present and differences in the corresponding prevalence. Thus, two control variables are included in the model, the county-specific prevalence (PRV) and the number of days since the first (estimated) infection ($DAYS$).

In the second model for the explanation of regional mortality (MRT), five more independent variables have to be incorporated:

- From the epidemiological point of view, the "risk group" of COVID-19 for severe courses (and even deaths) is defined as people of 60 years and older. The arithmetic mean of deceased attributed to COVID-19 is equal to 81 years (median: 82 years). Out of 6,831 reported deaths on May 5 2020, 6,524 were of age 60 or older (95.51%). This is, inter alia, because of outbreaks in residential homes for the elderly (RKI 2020a). Thus, the raw data from the RKI (RKI 2020b) was used to calculate the share of confirmed infected individuals of age 60 or older in all infected persons for each county ($INFS60$), which is included into the regression model for regional mortality.

- Several health-specific variables are found to influence the mortality risk (as well as the risk of severe course) of COVID-19. These individual-specific risk factors include, inter alia, diabetes, obesity, other respiratory diseases, or smoking (Engina et al. 2020, Selvan 2020). There are several possible health indicators which are unfortunately not available for German counties. Thus, the average health situation is captured by incorporating the average regional life expectancy into the model (*LEXP*), which is made available by the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR 2020).

- On the regional level, air pollution was found to be a contributing factor to COVID-19 fatality (Ogen 2020, Wu et al. 2020). According to Wu et al. (2020), the regional air pollution with respect to particulate matter (annual mean of daily PM_{10} values, unit: $\mu g/m^3$) is included into the model (*PM10*). Since Ogen (2020) shows a correlation between nitrogen dioxide concentration and COVID-19 fatality, this type of air pollution (annual mean of daily NO_2 values, unit: $\mu g/m^3$) is incorporated into the model as well (*NO2*). Both air quality indicators are made available by the German Environment Agency (UBA 2020a) on the level of single monitoring stations. These stations are available geocoded (UBA 2020b) and have been assigned to the German counties via a nearest neighbor join. Thus, the county-level values of both indicators equal the values of the nearest monitoring station.

- The intrinsic growth rate of each county is incorporated into the model as well. Considering the chronology of an infectious disease spread, there must be a reciprocal relationship between growth speed and mortality: The more individuals die in the context of the regarded disease, the fewer susceptibles are left to be infected, resulting in a deceleration of the pandemic spread (Li 2018) (see also Section 2.1). Thus, there *must* be a negative correlation between mortality and growth rate, all other things being equal. Consequently, the county-specific intrinsic growth rate (r) is included as control variable.

See Table 4 for all variables included into the models. All continuous variables, including the dependent variables (r and *MRT*, respectively), were transformed via natural logarithm in the regression analysis. This leads to an interpretation of the regression coefficients in terms of elasticities and semi-elasticities (Greene 2012). Two variants were estimated for the growth rate model (with and without dummy variables) and three for the mortality model (with and without growth rate as well as a third model including both growth rate and dummy variables). The minimum significance level was set to $p \leq 0.1$. In the first step, the regression models were estimated using an Ordinary Least Squares (OLS) approach and tested with respect to multicollinearity using variance inflation factors (*VIF*) with a critical value equal to five (Greene 2012).

However, SARS-CoV-2/COVID-19 cases are obviously not evenly distributed across all German counties as the disease spread started in a few "hotspots" in Bavaria, Baden-Wuerttemberg and North Rhine Westphalia (RKI 2020a, Tagesspiegel 2020b). Of course, an infectious disease can be transmitted across county borders, in particular, by contact between residents of one region and a nearby region. As a consequence, it is to be expected that indicators of disease spread – such as the regarded variables growth rate and mortality – are similar between nearby regions. Thus, further model-based analyses require considering possible spatial autocorrelation in the dependent variables (Griffith 2009). Consequently, both dependent variables were tested for spatial autocorrelation using Moran's I-statistic and the model estimation was repeated using a spatial lag model. In this type of regression model, spatial autocorrelation is modeled by a linear relationship between the dependent variable and the associated spatially lagged variable, which is a spatially weighted average value of the nearby objects. The influence of spatial autocorrelation is captured by adding a further parameter, ρ, to the regression equation, which is also tested for significance. Spatial linear regression models are not fitted by OLS but by Maximum Likelihood (ML) estimation. Both Moran's I and the spatial lag model require a weighting matrix to define the proximity of the regarded spatial object to nearby objects (Chi, Zhu 2008, Rusche 2008). Here, the weighting matrix for the spatial object (county i) was defined as all adjacent counties.

Table 4: Variables in the regression models for growth rate and mortality

Variable		Calculation/unit	Data source
r	Intrinsic growth rate	see Section 2.1	own calculation based on RKI (2020b)
MRT	Mortality cumulative (Porta 2008)	$\frac{D_i}{pop_i} * 100000$	own calculation based on RKI (2020b), Destatis (2020b)
PRV	Prevalence cumulative (Porta 2008)	$\frac{C_i}{pop_i} * 100000$	own calculation based on RKI (2020b), Destatis (2020b)
$DAYS$	Time since first infection	days (discrete)	own calculation based on RKI (2020b)
$POPDENS$	Population density	$\frac{pop_i}{A_i}$	own calculation based on Destatis (2020b)
$POPS65$	Share of population age 65 or older	$\frac{pop_{65+_i}}{pop_i} * 100$	own calculation based on Destatis (2020b)
CMI	Intensity of commuting (Guth et al. 2010)	$\frac{CM_{out_i}+CM_{in_i}}{L_{res_i}+L_{work_i}}$	own calculation based on Destatis (2020c)
$TOUR$	Tourist density	$\frac{T_i}{pop_i} * 1000$	own calculation based on Destatis (2020a,b)
$INFS60$	Share of infected age 60 or older in all infected persons	$\frac{C_{60+_i}}{C_i} * 100$	own calculation based on RKI (2020b)
$LEXP$	Life expectancy	years (mean)	BBSR (2020)
$PM10$	Air pollution PM_{10}	$\mu g/m^3$ (annual mean)	UBA (2020a,b)
$NO2$	Air pollution NO_2	$\mu g/m^3$ (annual mean)	UBA (2020a,b)
$EAST$	Dummy for East Germany	1=East Germany, else 0	
BV	Dummy for Bavaria	1=Bavaria, else 0	
SL	Dummy for Saarland	1=Saarland, else 0	
SX	Dummy for Saxony	1=Saxony, else 0	
NRW	Dummy for North Rhine Westphalia	1=North Rine Westphalia, else 0	
BW	Dummy for Baden-Wuerttemberg	1=Baden-Wuerttemberg, else 0	

Note: C_i is the cumulative number of reported SARS-CoV-2/COVID-19 cases in county i, D_i is the cumulative number of reported deaths attributed to COVID-19 in county i, C_{60+_i} is the cumulative number of reported infected persons of age 60 or older in county i, emphpop$_i$ is the population of county i, A_i is the area of county i, emphpop$_{65+_i}$ is the number of inhabitants of county i of age 65 or older, emphCM$_{\text{emphout}_i}$ and CM_{in_i} is the number of commuters from and to county i, respectively, L_{emphres_i} and L_{emphwork_i} is the number of employees whose place of residence is county i and whose place of work is county i, respectively.

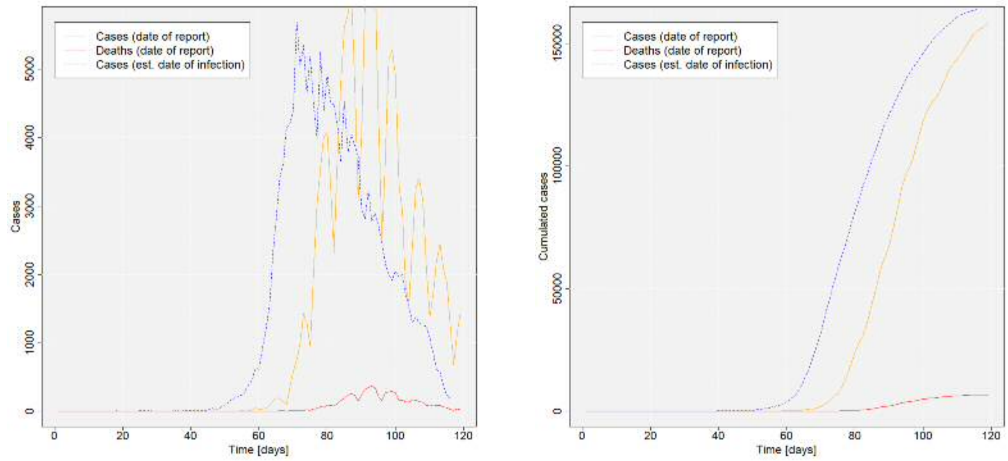
2.4 Software

The analysis in this study was executed in `R` (R Core Team 2019), version 3.6.2. The parametrization of logistic growth models was done using own functions for the OLS estimation based on the description in Engel (2010) and the `nls()` function for the final NLS estimation. For the steps of the regression analyses and presentation of results, the packages `car` (Fox, Weisberg 2019), `REAT` (Wieland 2019), `spdep` (Bivand et al. 2013), and `stargazer` (Hlavac 2018) were used. For creating maps, `QGIS` (QGIS Development Team 2019), version 3.8, was used, including the plugin `NNJoin` (Tveite 2019) for one further analysis.

3 Results

3.1 Estimation of infection dates and national inflection point

Figure 3 shows the estimated dates of infection and dates of report of confirmed cases and deaths for Germany. The curves are not shifted exactly by the average delay period



(a) Daily cases (b) Cumulative cases

Figure 3: Reported SARS-CoV-2 infections in Germany over time (dates of report vs. estimated dates of infection)

Source: own illustration.
Data source: own calculations based on RKI (2020b)

Table 5: Date of inflection point depending on assumed incubation period

Study	Median of incubation time (CI-95)	Date of inflection point Lower	Median	Upper
Linton et al. (2020)	5.0 (4.4, 5.6)	2020-03-20	2020-03-20	2020-03-19
Backer et al. (2020)	6.4 (5.6, 7.7)	2020-03-20	2020-03-19	2020-03-17

Source: own calculation based on data from RKI (2020b).

because of the different delay times with respect to case characteristics and county. The average time interval between estimated infection and case reporting is $\bar{x} = 11.92$ [days] ($SD = 5.21$). When applying the logistic growth model to the estimated dates of infection in Germany, the inflection point for Germany as a whole is on March 20, 2020.

Before switching to the regional level, we take into account the statistical uncertainty resulting from the estimation of the infection dates. About one third of the delay values for the time between onset of symptoms and case reporting was estimated by a stochastic model. Furthermore, the estimates of SARS-CoV-2/COVID-19 incubation period differ from study to study. This is why in the present case a conservative – which means a small – value of five days was assumed. Thus, we compare the results when including 1) the 95% confidence intervals of the response from the model in Equ. (12), and 2) the 95% confidence intervals of the incubation period as estimated by Linton et al. (2020). Figure 4 shows three different modeling scenarios, the mean estimation and the lower and upper bound of incubation period and delay time, respectively. The lower bound variant incorporates the lower bound of both incubation period and delay time, resulting in smaller delay between infection and case reporting and, thus, a later inflection point. The upper bound shows the counterpart. On the basis of the upper bound, the inflection point is already on March 19. Using the higher values of incubation period estimated by Backer et al. (2020), the upper bound results in an inflection point on March 17, while the lower bound variant leads to the turn on March 20. Considering confidence intervals of incubation period and delay time, the inflection point for the whole of Germany can be estimated between March 17 and March 20, 2020 (see Table 5).

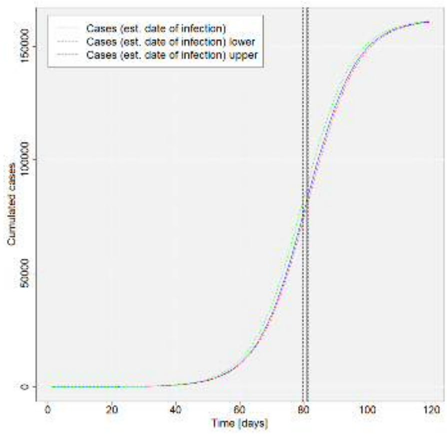


Figure 4: Estimated logistic growth model (including inflection point) for cumulative SARS-CoV-2 infections in Germany (based on estimated dates of infection) incorporating upper and lower bounds (95%-CI)
Source: own illustration. Data source: own calculations based on RKI (2020b)

3.2 Estimation of growth rates and inflection points on the county level

Figure 5 shows the estimated intrinsic growth rates (r) for the 412 counties. Figure 6 provides six examples of the logistic growth curves with respect to four counties identified as "hotspots" (Tirschenreuth, Heinsberg, Greiz and Rosenheim) and two counties with a low prevalence (Flensburg and Uckermark). There are obvious differences in the growth rates, following a spatial trend: The highest growth rates can be found in counties in North Germany (especially Lower Saxony and Schleswig-Holstein) and East Germany (especially Mecklenburg-Western Pomerania, Thuringia and Saxony). To the contrary, the growth rates in Baden-Wuerttemberg and North Rine Westphalia appear to be quite low. Taking a look at the time since the first estimated infection date in the German counties (see Figure 7a), the growth rates tend be much smaller the longer since the disease appeared in the county.

The regional inflection points indicate the day with the local maximum of infection rate. From this day forth, the exponential disease growth turns into degressive growth. In Figure 7b, the dates of the first day after the regional inflection point are displayed. The dates are categorized according to the coming into force of relevant nonpharmaceutical interventions (see Table 1). Table 6 summarizes the number of counties and the corresponding population shares by these categories. Figure 8 shows the intrinsic growth rate (y axis) and the day after the inflection point (colored points) against time (x axis). Figure 9 shows the same information against regional prevalence (x axis).

In 255 of 412 counties (61.89%) with 54.58 million inhabitants (65.66% of the national population), the SARS-CoV-2/COVID-19 infections had already decreased before phase 3 of measures came into force on March 23, 2020. In a minority of counties (51, 12.38%),

Table 6: German counties by first day after inflection point

First day after inflection point	Counties [no.]	Counties [%]	Population [Mill.]	Population [%]
Before March 13	6	1.46	0.98	1.17
March 13 to March 16	45	10.92	8.27	9.95
March 17 to March 20	138	33.50	31.03	37.33
March 21 to March 22	66	16.02	14.30	17.20
March 23 to April 19	157	38.11	28.54	34.34
Sum	412	100	83.13	100

Source: own illustration. Data source: own calculations based on RKI (2020b)

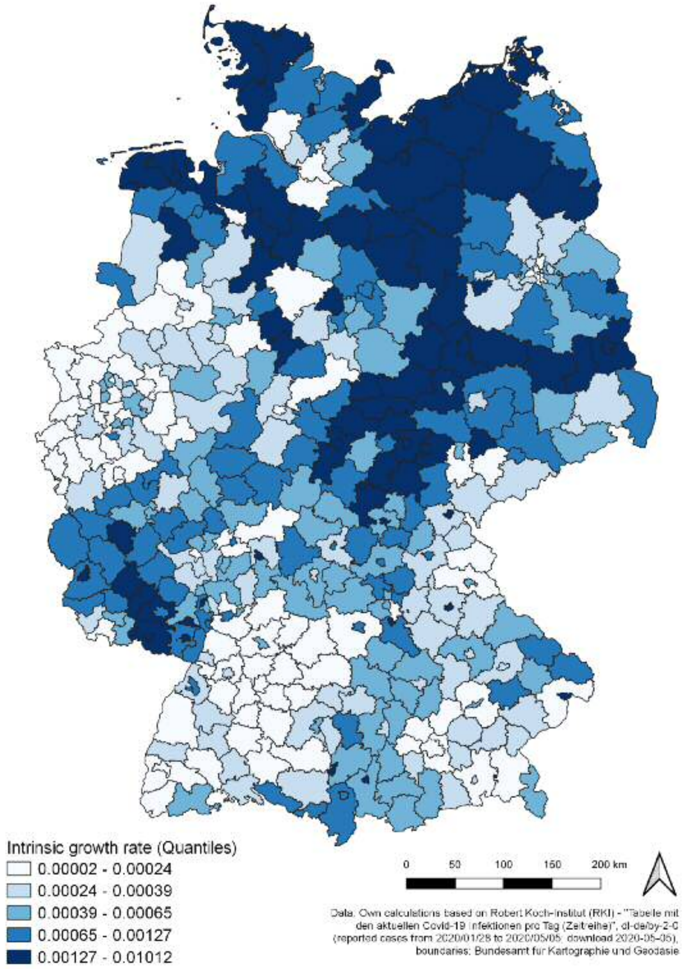


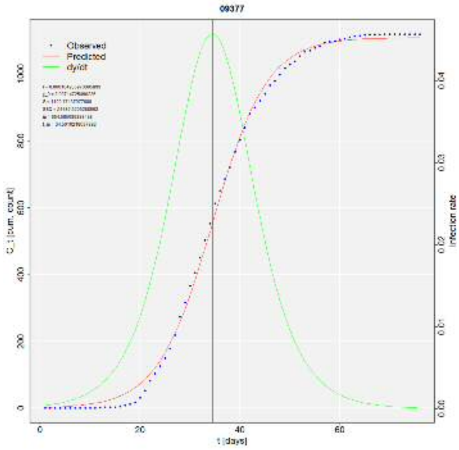
Figure 5: Intrinsic growth rate by county

the curve already flattened before the closing of schools and child day care centers (March 16-18, 2020). Six of them exceeded the peak of new infections even before March 13. This category refers to the appeals of chancellor Merkel and president Steinmeier on March 12. In 157 counties (38.11%) with a population of 28.54 million people (34.34% of the national population), the decrease of infections took place within the period of strict regulations towards social distancing and ban of gatherings.

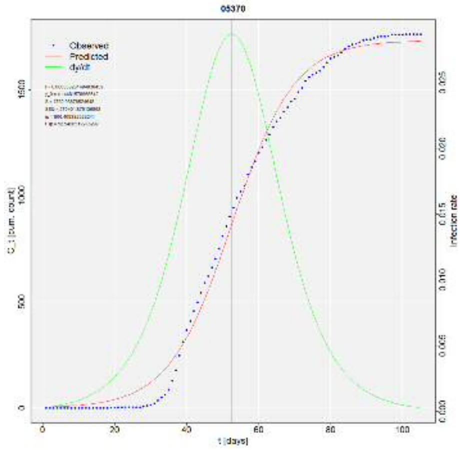
The average time interval between the first estimated infection and the respective inflection point of the county is $\bar{x} = 30.32$ [days]. However, the time until inflection point is characterized by a large variance ($SD = 11.92$), but this may be explained partially by the (de facto unknown) variance in the incubation period and the variance in the delay between onset of symptoms and reporting date.

In all counties, the inflection points lie between March 6 and April 18, 2020, which means a time period of 43 days between the first and the last flattening of a county's epidemic curve. The first regional decrease can be found in Heinsberg county (North Rhine Westphalia; 254,322 inhabitants), which was one of the first Corona "hotspots" in Germany. The estimated inflection point here took place at March 6, 2020, leading to a date of the first day after the inflection point of March 7. The latest estimated inflection point (April 18, 2020) took place in Steinburg county (Schleswig-Holstein; 131,347 inhabitants).

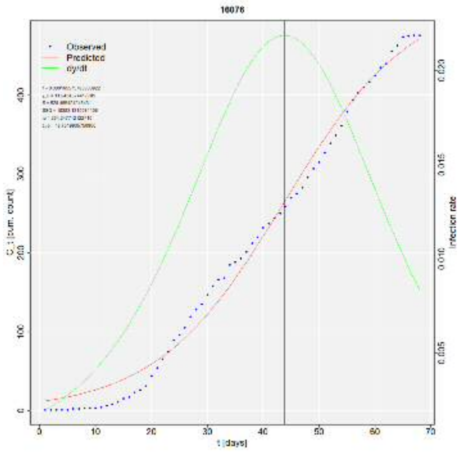
As Figure 8 shows, the intrinsic growth rates, which indicate an average growth level over time, and inflection points of the logistic models are linked. Growth speed declines



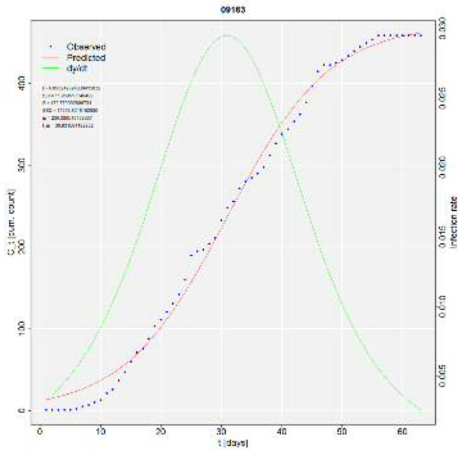
(a) Tirschenreuth county (Bavaria)



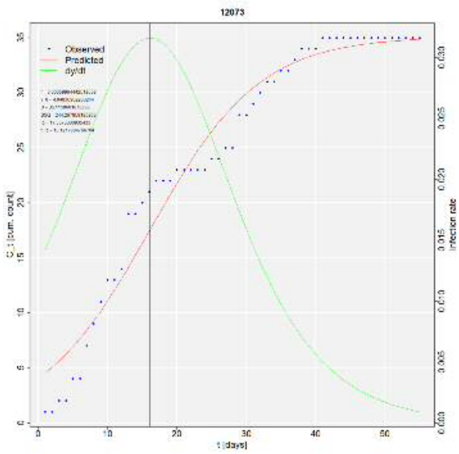
(b) Heinsberg county (North Rhine Westphalia)



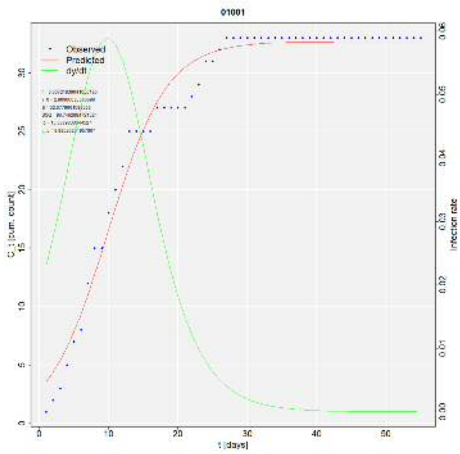
(c) Greiz county (Thuringia)



(d) Rosenheim county (Bavaria)



(e) Uckermark county (Brandenburg)



(f) Flensburg (Schleswig-Holstein)

Figure 6: Cumulative SARS-CoV-2 infections (based on estimated dates of infection) and estimated logistic growth models (including infection rate and inflection point) in six German counties

Source: own illustration. Data source: own calculations based on RKI (2020b)

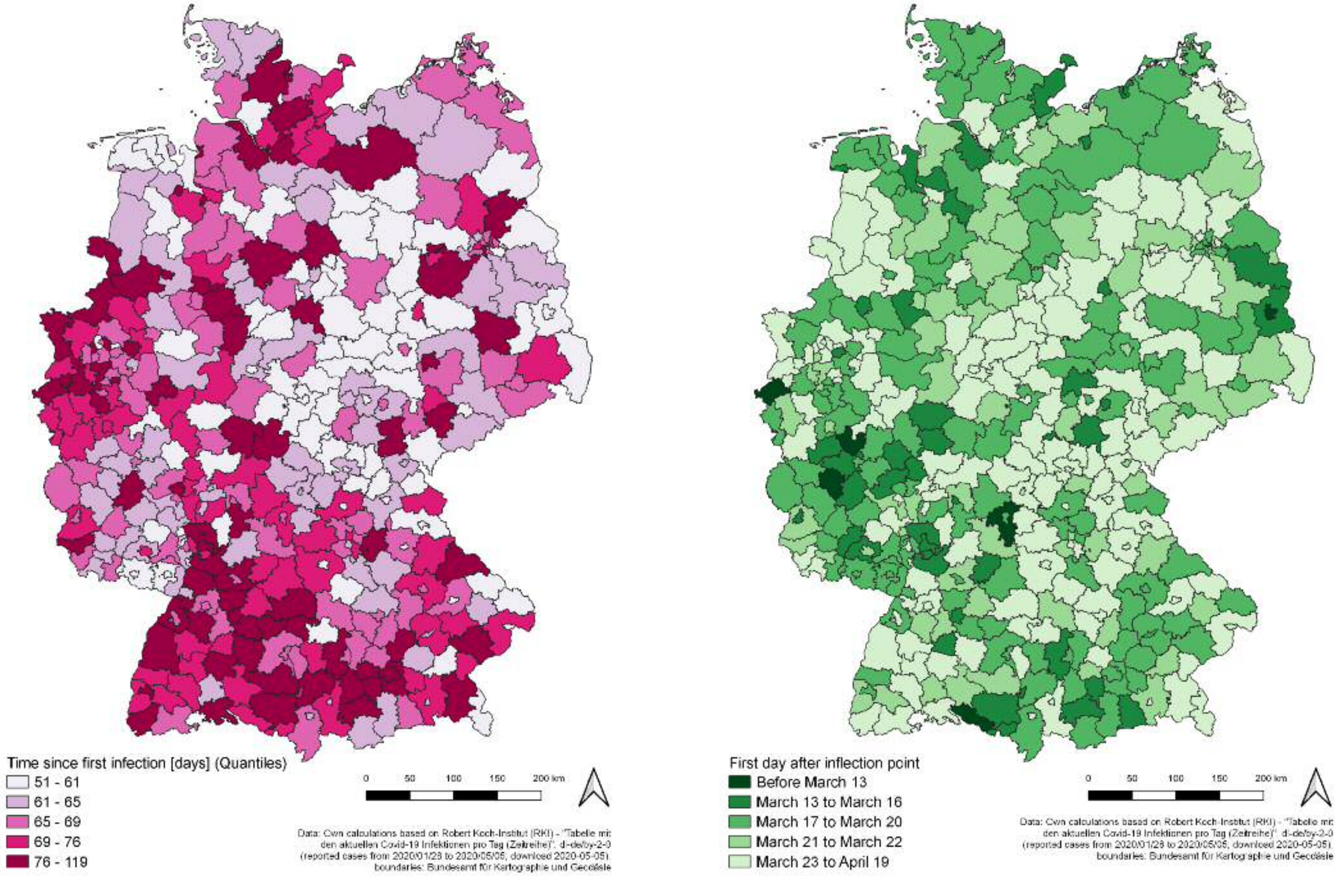


Figure 7: Time since first infection by county (left) and First day after inflection point by county (right)

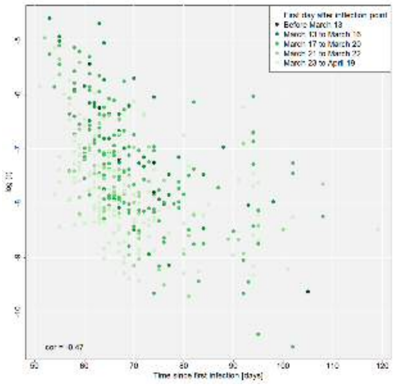


Figure 8: Growth rate and first day after inflection point vs. time
Source: own illustration. Data source: own calculations based on RKI (2020b)

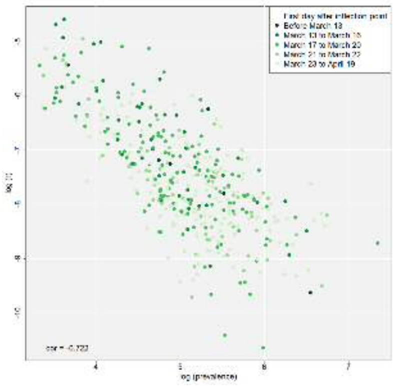


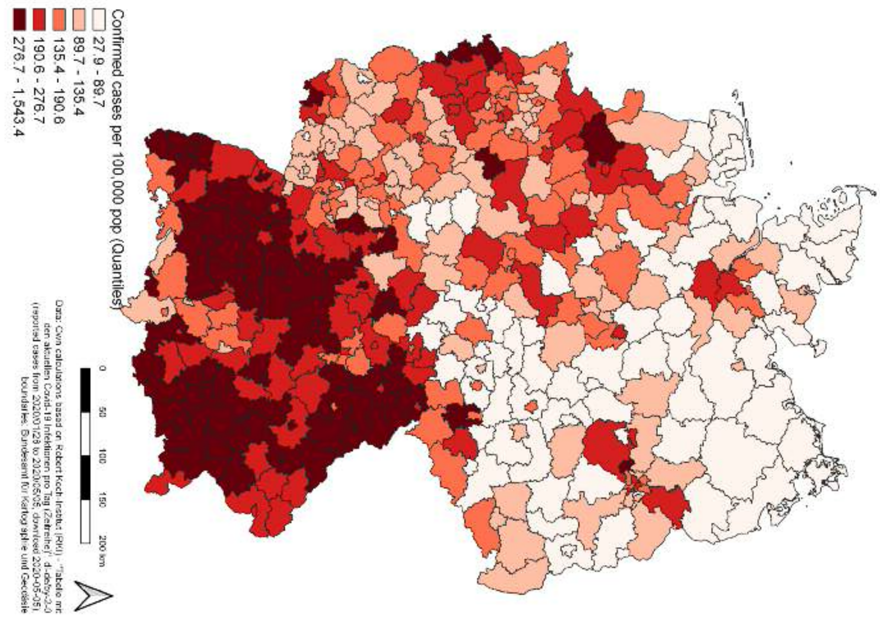
Figure 9: Growth rate and first day after inflection point vs. prevalence
Source: own illustration. Data source: own calculations based on RKI (2020b)

over time (see also Figures 5 and 8), more precisely, it declines in line with the time the disease is present in the regarded county (Pearson correlation coefficient of -0.47, $p < 0.001$). The longer the time between inflection point and now, the lower the growth speed, and vice versa. This process takes place over all German counties with a time delay depending on the first occurrence of the disease. As shown in Figure 9, there is also a negative correlation between regional prevalence and growth rate (Pearson correlation coefficient of -0.722, $p < 0.001$). These relationships, which are closely linked to the chronology of an infectious disease spread and the characteristics of the logistic growth model, respectively, are included into the regression models as control variables.

3.3 Regression models for intrinsic growth rates and mortality

The variables of most interest used within the models are mapped in Figures 10a (prevalence, *PRV*), 10b (mortality, *MRT*) and 11a (share of infected individuals of age ≥ 60, *POPS65*). Additionally, Figure 11b shows the current case fatality rate on the county level. Tables 7 and 8 show the estimation results for the OLS regression models explaining the intrinsic growth rates and the mortality, respectively, both transformed via natural logarithm. Table 9 displays the Moran's I-statistic for the dependent variables of the two models. Tables 10 and 11 show the estimation results for the spatial lag models.

In all of the OLS models, no variable exceeded the critical value of $VIF \geq 5$. For the prediction of the intrinsic growth rate, two model variants were estimated without and with the state dummy variables (Table 7). From the aspect of explained variance, the second OLS model provides a better fit ($R^2 = 0.731$ and Adj.$R^2 = 0.723$, respectively)



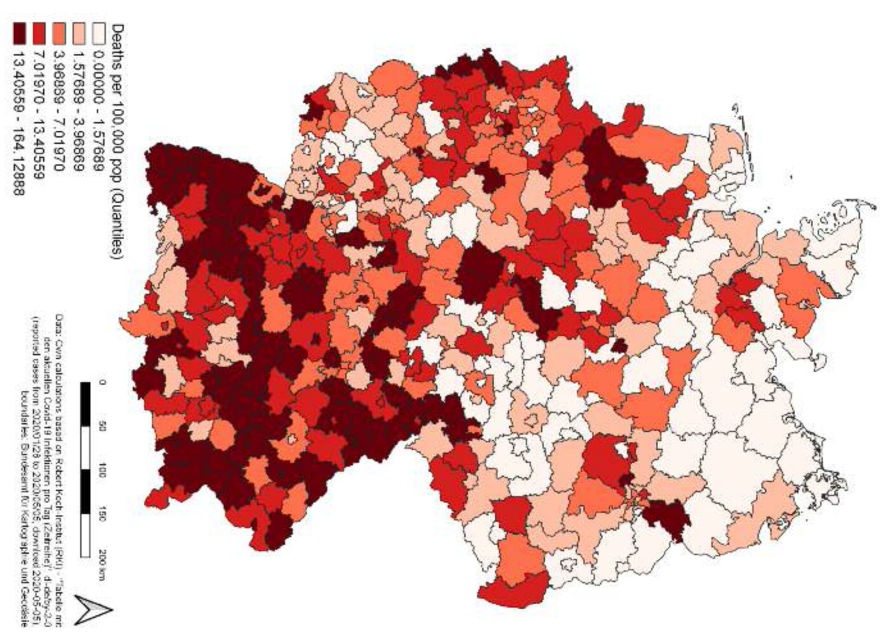


Figure 10: Prevalence by county (left) and Mortality by county (right)

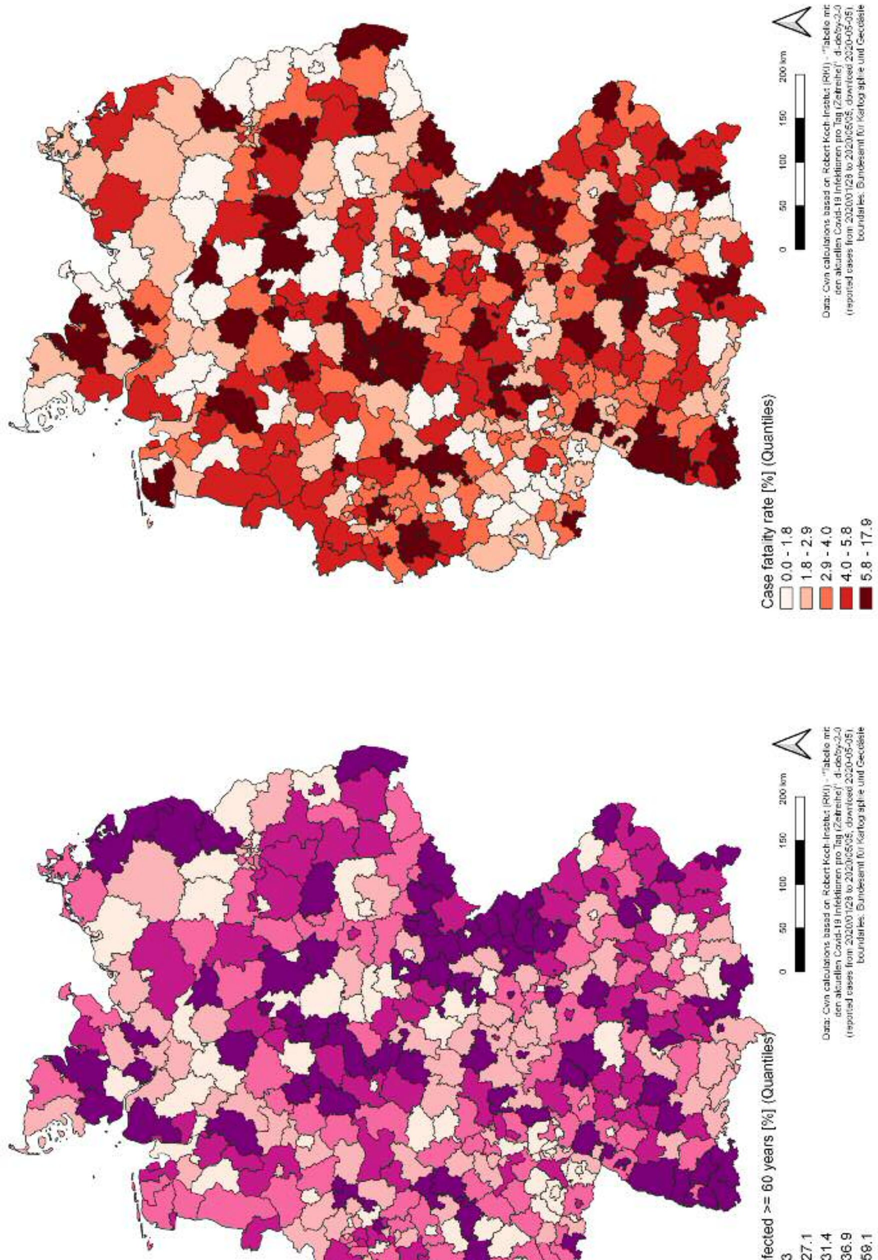


Figure 11: Share of reported infected individuals of age 60 and older by county (left) and CFR by county (right)

Table 7: Estimation results for the growth rate model (OLS)

	Dependent variable: ln(r)	
	(1)	(2)
ln (POPDENS)	-0.177^{***}	-0.102^{***}
	(0.028)	(0.027)
ln (POPS65)	0.830^{***}	1.165^{***}
	(0.284)	(0.269)
ln (CMI)	0.607^{***}	0.420^{***}
	(0.091)	(0.087)
ln (TOUR)	0.146^{***}	0.035
	(0.042)	(0.041)
EAST	−0.087	−0.045
	(0.091)	(0.089)
BV		0.473^{***}
		(0.091)
SL		0.118
		(0.231)
SX		-0.580^{***}
		(0.168)
NRW		-0.428^{***}
		(0.097)
BW		0.011
		(0.108)
ln (PRV)	-0.911^{***}	-1.039^{***}
	(0.049)	(0.056)
DAYS	-0.019^{***}	-0.014^{***}
	(0.003)	(0.003)
Constant	-3.660^{***}	-4.214^{***}
	(1.144)	(1.079)
Observations	412	412
R^2	0.678	0.731
Adjusted R^2	0.672	0.723
Residual Std. Error	0.587	0.540
Degrees of Freedom	404	399
F Statistic	121.278^{***}	90.307^{***}
Degrees of Freedom	7; 404	12; 399

Note: *p<0.1; **p<0.05; ***p<0.01

compared to the first ($R^2 = 0.678$ and Adj.$R^2 = 0.672$, respectively), thus, the second model is to be preferred for interpretation. Different than expected, population density (*POPDENS*) does not affect growth rate positively. A 1% increase of population density decreases the growth rate by 0.1%. Also the demographic indicator (*POPS65*) is correlated with growth contrary to expectations: An increase of 1% in the share of inhabitants of age ≥ 65 increases the growth rate by 1.2%. As expected, the intensity of commuting (*CMI*) has a significant positive effect on the intrinsic growth rate: An increase of commuting intensity by 1% increases the growth rate by 0.4%. Tourist density (*TOUR*) is only significant in the first model. On average, the intrinsic growth rate is significantly higher in Bavarian counties (dummy variable *BV*), and significantly lower in Saxony and North Rhine-Westphalian counties (*SL* and *SX*, respectively). The coefficients for Baden-Wuerttemberg (*BW*) and Saarland (*SL*) are not significant, which means that no significant deviation from the average growth rate is found for counties in these German states. The *EAST* dummy is insignificant in both models. Considering the necessary control variables, the current prevalence (*PRV*) and time (*DAYS*) decelerate the growth of infections significantly. The former has a nearly proportional impact: An increase of prevalence equal to 1% decreases the intrinsic growth rate by 1.04%. For each day

SARS-CoV-2/COVID-19 is present in the county, the growth speed declines on average by 1.4%. Here, one has to keep in mind that these relationships are reciprocal and represent the mandatory decline of susceptible individuals over time.

For the prediction of mortality (MRT), the growth rate (r) and the state dummy variables (BV, SL, SX and NRW) are entered into the model analyis successively, resulting in the three models shown in Table 8. When comparing models 1 and 2, adding the growth rate as independent variable increases the explained variance substantially (Adj.$R^2 = 0.219$ and 0.367, respectively). The third model provides the best fit, adjusted for the number of explanatory variables (Adj.$R^2 = 0.383$). No significant influence can be found for the spatial (*POPDENS*), demographic (*POPS65*), mobility (*PI* and *TOUR*), and air pollution variables (*PM10* and *NO2*). Life expectancy (*LEXP*) and the dummy for East German counties (*EAST*) are only significant in the first model. However, the share of infected people of age ≥ 60 (*INFS60*) significantly increases the regional mortality: An increase in the share of people of the "risk group" in all infected by 1% increases the mortality by approx. 0.5%. The only significant state-specific effect is found for Bavaria: The mortality in Bavarian counties is higher than in the counties of other states. Furthermore, a two-sample t-test reveals that Bavarian counties have a significantly higher share of infected belonging to the risk group ($\bar{x} = 31.11\%$) compared to the remaining states ($\bar{x} = 29.28\%$), with a difference of 1.83 percentage points ($p = 0.054$). The reciprocal relationship between mortality and growth rate is also significant.

As expected, spatial autocorrelation can be detected among the dependent variables. The Moran's I-statistic for both regional growth rate and mortality (0.49 and 0.16, respectively) is significant (see Table 9). Consequently, the OLS estimations are expected to be biased. Therefore, we apply a spatial lag model in the next step.

With respect to the spatial lag model for regional growth rates (Table 10), the ρ-parameter in both model variants is significant ($\rho = 0.158$ and $\rho = 0.095$, respectively), which indicates a significant spatial lag effect. However, when comparing the second spatial lag model (with $AIC = 674.96$, which is superior to model 1 with $AIC = 733.37$) to the second OLS model (see Table 7), there are only negligible differences in the parameter estimates and significance levels: The same independent variables are found to be significantly correlated with growth rates. They also have the same sign, which indicates the same direction of influence. Intrinsic growth rates on the county level are predicted by population density (approx. -0.9), population share of 65 and older (approx. 1.1), commuting intensity (approx. 0.4), and state-specific dummy variables (Bavaria: approx. 0.5, Saxony: approx. -0.5, North Rhine Westphalia: approx. -0.4) as well as the control variables (Prevalence: approx. -1.0 and time since first infection: approx. -0.01).

The same conclusion can be drawn with respect to the spatial lag models for mortality (Table 11), when comparing them to the OLS models (Table 8). The spatial lag effect is only significant in the first model ($\rho = 0.165$) but not in models 2 ($\rho = 0.042$) and 3 ($\rho = -0.044$). Regional mortality is significantly influenced by the share of infected people of age ≥ 60 (approx. 0.5) and the dummy variable for Bavarian counties (approx. 1.2) as well as correlated with regional growth rates (approx. -1.6). As spatial autocorrelation was also detected for the regional growth rate, which is an independent variable in the mortality models, a further robustness check of the estimations is necessary: Table 12 shows the results for the second and third mortality model in a spatial Durbin model, which incorporates a spatial lag effect for both the dependent variable (ρ) and the regional growth rate (*lag ln (r)*). The lag effect of $\ln(r)$ is statistically significant, but the other results remain qualitatively the same (share of infected people of age ≥ 60: approx. 0.5; dummy variable for Bavarian counties: approx. 1).

With respect to the regression models for regional growth and mortality, the results of the OLS estimations were confirmed by those from the models allowing for spatial autocorrelation. Although there is obvious spatial autocorrelation (which can be explained plausibly by interregional transmission of the infectious disease), both OLS and spatial regression models show nearly the same results with respect to strength and direction of correlations. From the spatial statistic point of view, this can be explained with the incorporated independent variables as regional differences in both growth rate and mortality are predicted entirely by the interregional variation in causal factors.

Table 8: Estimation results for the mortality model (OLS)

	Dependent variable: ln (MRT+0.0001) (1)	(2)	(3)
ln (POPDENS)	−0.060	−0.268**	−0.148
	(0.132)	(0.121)	(0.124)
ln (POPS65)	−0.738	0.883	1.730
	(1.316)	(1.196)	(1.230)
ln (CMI)	−0.627	0.397	0.025
	(0.412)	(0.385)	(0.396)
ln (TOUR)	−0.265	0.066	−0.081
	(0.188)	(0.173)	(0.178)
ln (LEXP)	54.205***	11.409	12.349
	(12.259)	(11.877)	(12.556)
ln (PM10)	0.053	−0.226	−0.200
	(0.716)	(0.645)	(0.653)
ln (NO2)	0.428	0.178	0.213
	(0.287)	(0.259)	(0.259)
ln (INFS60+0.0001)	0.618***	0.459***	0.462***
	(0.104)	(0.095)	(0.094)
EAST	−0.980**	−0.496	−0.148
	(0.395)	(0.359)	(0.385)
BV			1.110***
			(0.334)
SL			0.800
			(0.983)
SX			−0.909
			(0.738)
NRW			−0.144
			(0.433)
BW			0.193
			(0.465)
DAYS	0.017	−0.014	−0.011
	(0.013)	(0.012)	(0.012)
ln (r)		−1.570***	−1.535***
		(0.161)	(0.171)
Constant	−237.384***	−62.529	−69.684
	(55.406)	(53.006)	(55.923)
Observations	412	412	412
R^2	0.238	0.384	0.407
Adjusted R^2	0.219	0.367	0.383
Residual Std. Error	2.596	2.336	2.308
Degrees of Freedom	401	400	395
F Statistic	12.525***	22.679***	16.916***
Degrees of Freedom	10; 401	11; 400	16; 395

Note: *p<0.1; **p<0.05; ***p<0.01

Table 9: Moran's I-statistic for intrinsic growth rates and mortality (Weighting matrix: all adjacent counties)

Indicator	Moran's I	Expectation	Variance	Standard deviate
ln (r)	0.4856***	-0.0024	0.0011	14.738
ln (MRT+0.0001)	0.1635***	-0.0024	0.0011	5.0555

Note: *p<0.1; **p<0.05; ***p<0.01
Source: own calculation.

Table 10: Estimation results for the growth rate model (spatial lag model)

	Dependent variable: ln(r)	
	(1)	(2)
ln (POPDENS)	-0.151^{***}	-0.089^{***}
	(0.027)	(0.027)
ln (POPS65)	0.755^{***}	1.119^{***}
	(0.279)	(0.265)
ln (CMI)	0.561^{***}	0.402^{***}
	(0.092)	(0.087)
ln (TOUR)	0.121^{***}	0.024
	(0.041)	(0.041)
EAST	-0.148^{*}	−0.076
	(0.089)	(0.089)
BV		0.498^{***}
		(0.090)
SL		0.096
		(0.226)
SX		-0.538^{***}
		(0.166)
NRW		-0.368^{***}
		(0.100)
BW		0.073
		(0.110)
ln (PRV)	-0.827^{***}	-1.002^{***}
	(0.058)	(0.060)
DAYS	-0.018^{***}	-0.014^{***}
	(0.003)	(0.003)
Constant	-2.677^{**}	-3.562^{***}
	(1.159)	(1.110)
ρ	0.158^{***}	0.095^{*}
	(0.049)	(0.050)
Observations	412	412
Log Likelihood	−356.682	−322.481
σ^2	0.329	0.280
Akaike Inf. Crit.	733.365	674.962
LR Test (df = 1)	9.283^{***}	3.193^{*}
Wald Test (df = 1)	10.325^{***}	3.608^{*}

Note: *p<0.1; **p<0.05; ***p<0.01 *Source*: own calculation.

Table 11: Estimation results for the mortality model (spatial lag model)

	Dependent variable: ln (MRT+0.0001)		
	(1)	(2)	(3)
ln (POPDENS)	-0.085	-0.272^{**}	-0.138
	(0.129)	(0.119)	(0.121)
ln (POPS65)	-0.500	0.921	1.731
	(1.289)	(1.178)	(1.204)
ln (CMI)	-0.586	0.394	0.013
	(0.403)	(0.379)	(0.387)
ln (TOUR)	-0.250	0.066	-0.086
	(0.184)	(0.170)	(0.175)
ln (LEXP)	51.311^{***}	11.262	12.321
	(12.101)	(11.713)	(12.288)
ln (PM10)	0.197	-0.186	-0.224
	(0.700)	(0.637)	(0.639)
ln (NO2)	0.397	0.174	0.215
	(0.280)	(0.255)	(0.254)
ln (INFS60+0.0001)	0.614^{***}	0.460^{***}	0.461^{***}
	(0.102)	(0.094)	(0.092)
EAST	-0.764^{*}	-0.448	-0.180
	(0.399)	(0.365)	(0.382)
BV			1.185^{***}
			(0.349)
SL			0.800
			(0.962)
SX			-0.914
			(0.722)
NRW			-0.119
			(0.426)
BW			0.245
			(0.466)
DAYS	0.017	-0.013	-0.011
	(0.012)	(0.012)	(0.012)
ln (r)		-1.549^{***}	-1.550^{***}
		(0.162)	(0.168)
Constant	-225.876^{***}	-62.008	-69.584
	(54.613)	(52.268)	(54.731)
ρ	0.165^{**}	0.042	-0.044
	(0.064)	(0.062)	(0.065)
Observations	412	412	412
Log Likelihood	-969.231	-927.965	-920.308
σ^2	6.435	5.293	5.100
Akaike Inf. Crit.	1,964.461	1,883.930	1,878.615
Wald Test (df = 1)	6.714^{***}	0.449	0.454
LR Test (df = 1)	5.551^{**}	0.378	0.365

Note: *p<0.1; **p<0.05; ***p<0.01 *Source*: own calculation.

Table 12: Estimation results for the mortality model (spatial Durbin model)

	Dependent variable: ln (MRT+0.0001)	
	(1)	(2)
ln (POPDENS)	-0.328^{***}	-0.202
	(0.121)	(0.124)
ln (POPS65)	0.984	1.816
	(1.173)	(1.195)
ln(CMI)	0.476	0.092
	(0.379)	(0.386)
ln (TOUR)	0.146	-0.012
	(0.172)	(0.176)
ln (LEXP)	6.216	7.535
	(11.834)	(12.352)
ln (PM10)	-0.189	-0.358
	(0.633)	(0.636)
ln (NO2)	0.160	0.238
	(0.254)	(0.252)
ln (INFS60+0.0001)	0.467^{***}	0.465^{***}
	(0.093)	(0.092)
EAST	-0.236	0.022
	(0.373)	(0.388)
BV		0.980^{***}
		(0.357)
SL		0.829
		(0.955)
SX		-1.131
		(0.722)
NRW		-0.474
		(0.450)
BW		-0.139
		(0.489)
DAYS	-0.013	-0.009
	(0.012)	(0.011)
$\ln(r)$	-1.422^{***}	-1.456^{***}
	(0.172)	(0.172)
lag $\ln(r)$	-0.555^{**}	-0.634^{**}
	(0.254)	(0.278)
Constant	-43.482	-52.706
	(52.507)	(54.744)
ρ	-0.033	-0.116
	(0.070)	(0.072)
Observations	412	412
Log Likelihood	-925.583	-917.677
σ^2	5.233	5.025
Akaike Inf. Crit.	1,881.167	1,875.355
Wald Test (df = 1)	0.223	2.607
LR Test (df = 1)	0.190	2.145

Note: *p<0.1; **p<0.05; ***p<0.01 *Source*: own calculation.

4 Discussion

4.1 Curve flattening in the context of nonpharmaceutical interventions

Taking a look at the national level, the flattening of the epidemic curve in Germany occurred between three to six days before phase 3 of measures (as defined in this study) came into force. Due to this temporal mismatch, the decline of infections cannot be causally linked to the nationwide formal "lockdown" (including forced social distancing and ban of gatherings) of March 23. Note that the results for whole Germany, estimating the inflection point between March 17 and March 20, are rather conservative when compared with the RKI estimations. In the RKI nowcasting study, the peak of onset of symptoms (not infection time, which is not considered in the mentioned study) is found at March 18 and a stabilization of the reproduction number equal to $R = 1$ at March 22 (an der Heiden, Hamouda 2020). Subtracting an average incubation period of five days from these dates, the peak of infections occured around March 13 and the reproduction number stabilizes approximately at March 17. An earlier RKI study (an der Heiden, Buchholz 2020) estimated the peaks between June and July, depending on the parameters of the scenarios.

The main focus of this analysis is on the regional level, which reveals a more differentiated picture. In all German counties the curves of infections clearly flattened within a time period of about six weeks from the first to the last county. On average, it took one month from the first infection to the inflection point of the epidemic curve. However, the regional trend change in infections is not in line with the governmental nonpharmaceutical interventions to contain the virus spread. In nearly two thirds of the German counties which account for two thirds of the German population, the flattening of the infection curve occured before the "lockdown" (measures of phase 3) came into force (March 23). One in eight counties experienced a decline of infections even before the closures of schools, child day care facilities and retail facilities, which is attributed to phase 2 of interventions in this study. Consequently, in a majority of counties, the regional decline of infections cannot be attributed to the formal "lockdown". In a minority of counties, also closures of educational and retail facilities (measures of phase 2) cannot have caused the decline. Keeping in mind that SARS-CoV-2 emerged at different times across the counties, it is at least questionable whether these measures primarily caused the flattening of the infection curve in the other counties. Furthermore, in a minority of counties, the regional trend change occurred several weeks (up to about four weeks) after the nonpharmaceutical interventions came into force. One might argue that there could be a time lag between the date of official enforcement of the regulations and the time they became effective in practice. However, this could only be conceivable for the contact ban but not for the closures of schools and other services as these infrastructures are either closed or not and, thus, can be potential places of virus transmission or not. Moreover, it seems unlikely that an intervention like a contact ban becomes effective only after several weeks and regionally differentiated.

Bringing together these aspects, regional curve flattening seems to have occurred independently from the governmental measures of phase 2 and 3. Instead, regional pandemic growth appears as a function of time, reaching the peak of infection rates with a time lag depending on the date the virus emerged.

The results presented here tend to support the findings in the study by Ben-Israel (2020) that curve flattening in the SARS-CoV-2/COVID-19 pandemic occurs with or without a strict "lockdown". However, neither this study nor the study by Ben-Israel (2020) provides explicit epidemiological, virological or other kinds of clarifications for this phenomenon. The further interpretation must be limited to a collection of explanation attempts, which are non-mutually exclusive. Some reasons for the decline of infections relate to other types of interventions both voluntary and mandatory are:

- First of all, it must be pointed out that the focus of this study is on regional pandemic growth in the context of the nonpharmaceutical interventions of phase 2 and 3 starting in mid March 2020, especially the "lockdown" from March 23. One has to keep in mind that some interventions against virus spread were already established in the first half of March (phase 1), e.g. the cancellation of large events or "ghost games" in soccer (see Table 1). These early measures could have contributed substantially to curve flattening, as the cancellation of events might have prevented people from being infected in the context of so-called *super-spreading events*, which play an enormous role during infectious disease spreads (Al-Tawfiq, Rodriguez-Morales 2020, Stein 2011). Many infections and death cases attributed to COVID-19 in the early phase of the pandemic in Germany can be traced back to super-spreading events in February and early March 2020, such as in Heinsberg or Tirschenreuth county (Tagesspiegel 2020b). Also, the domestic quarantine of infected persons (which is the default procedure in the case of infectious diseases) might have reduced new infections. In Heinsberg county, about 1,000 people were in domestic quarantine at the end of February 2020 (Tagesschau.de 2020c), which could explain the early curve flattening in this Corona "hotspot".

- Also media reports from China or Italy as well as appeals and recommendations from the government could have influenced people's behavior on a *voluntary* basis already in the first half of March 2020 (or even earlier), e.g. with respect to physical distancing, thorough and frequent hand washing, coughing and sneezing in the arm fold, or reducing mobility in general. Unfortunately, there is no explicit indicator of changes in the individual behavior. However, some other findings give a hint towards voluntary behavioral changes: Several surveys show a high degree of public awareness in Germany (and other countries) towards the SARS-CoV-2/COVID-19 threat already in February and the first half of March 2020 (Ipsos 2020, YouGov 2020). This increasing awareness might be reflected by more caution in daily life: The RKI has documented an "abrupt" decline of *other* infectious respiratory diseases with shorter incubation periods (such as influenza) in Germany since the 10th calendar week (March 2 to 8, 2020). This decline is regarded as "extremely unusual" (Buchholz et al. 2020). This reduction might be attributed to voluntary cautious behavior in the context of the public discussion towards SARS-CoV-2, as this decrease started before any public health intervention came into force (except for the quarantines of SARS-CoV-2-infected persons, see Table 1). Furthermore, the analysis of mobility patterns shows a decline of mobility in Germany, starting already in the first half of March 2020. Additionally, a strong correlation between (aggregated) mobility and the acceptance of social restrictions (obtained by surveys) was found: The higher the agreement with the statement "I think the current measures are too strict", the higher the increase in mobility (Covid-19 Mobility Project 2020a,b). All these phenomenons suggest voluntary behavior changes within the (German) population, which reduce the transmission of infectious diseases and preceded the "lockdown" by several weeks. Another indicator for an increased awareness in the (German) population – although not intended or desired – is the enormous tendency of hoarding groceries, which started in the second half of February 2020 (Rheinische Post online 2020).

- Additionally, one has to keep in mind the seasonal cycle of respiratory viral diseases: Influenza viruses and most cold viruses (including those from the family of Coronaviridae) mainly occur during the winter months due to changes in environmental parameters (e.g., temperature and humidity) and human behavior (more or fewer activities outside, whilst the risk of infection is, all other things being equal, lower outside) (Moriyama et al. 2020). Several virologists expressed cautious optimism towards the sensitivity of the SARS-CoV-2 virus to increasing temperature and ultraviolet radiation (Focus 2020). Model-based analyses from biogeography show that temperate warm and cold climates facilitate the virus spread, while arid and tropical climates are less favorable (Araujo, Naimi 2020). By consequence, there might have also been a decline of SARS-CoV-2 infections due to weather changes in early spring (mid-March).

Furthermore, there is another possible reason for curve flattening with or without a "lockdown", which is of epidemiological nature and related to the transmission process of SARS-CoV-2/COVID-19 in the context of immunization. Note that "immunity" may have different causes (e.g., antibodies due to previous infections, vaccination, immunological memory) and does not necessarily prevent individuals from being infected (in terms of an invasion of an individual's body) but leads to an effective response of the immune system and prevents the emergence of (severe) symptoms (Mak, Saunders 2006). At the beginning of the pandemic, two assumptions towards the role of immunization were stated: 1) Nobody is immune, which means that all individuals of the population belong to the group of susceptibles, 2) Without any interventions (e.g., vaccine, nonpharmaceutical interventions), *herd immunity* – a share of a population is immune, which provides protection to those who are not immune, causing the pandemic to slow down and stop – is achieved when about 70% of the population was infected (D'Souza, Dowdy 2020). This percentage share is commonly known as *herd immunity threshold* (*HIT*) and is, in its basic form, calculated based on the infection's basic reproduction number, R_0: $HIT = 1 - 1/R_0$ (Fine et al. 2011). Early modeling studies focusing on the effect of nonpharmaceutical interventions are based on these (or similar) assumptions (an der Heiden, Buchholz 2020, Ferguson et al. 2020). However, there are some issues regarding the *HIT* for SARS-CoV-2/COVID-19 which need to be considered:

- In epidemiology, it is well known that disease transmission is mostly concentrated on a minor part of individuals causing a large majority of secondary infections: "In what became known as the 20/80 rule, a concept documented by observational and modeling studies and having profound implications for infection control, 20% of the individuals within any given population are thought to contribute at least 80% to the transmission potential of a pathogen, and many host-pathogen interactions were found to follow this empirical rule" (Stein 2011). Gomes et al. (2020) incorporate inter-individual variation in susceptibility and exposure to a SARS-CoV-2 infection into an epidemiological model (SEIR [susceptible-exposed-infectious-recovered] model). Depending on the assumptions on this *overdispersion*, the *HIT* of SARS-CoV-2 reduces to 10-20%. Thus, the achievement of herd immunity would require a considerably lower number of SARS-CoV-2 infections.

- Although the SARS-CoV-2 virus is highly infective, the "Heinsberg study" by Streeck et al. (2020) found a relatively low secondary infection risk (*secondary attack rate*, SAR). Infected persons did not even infect other household members in the majority of cases. The authors conjecture that this could be due to a present immunity (T helper cell immunity) not detected as positive in the test procedure. This kind of immunity is not to be confused with (temporal or everlasting) immunity due to antibodies against a specific virus but may be regarded as a functional immune memory. In a current virological study by Braun et al. (2020), 34% of test persons who have never been infected with SARS-CoV-2 had relevant T helper cells because of earlier infections with other harmless Coronaviruses causing common colds. In a study by Grifoni et al. (2020), SARS-CoV-2-reactive T cells were detected in even 40%–60% of unexposed individuals. If this explanation proves correct, the absolute number of susceptible individuals would have been substantially lower already at the beginning of the pandemy. Other Coronaviruses are responsible for about 10-15% of seasonal "common colds" (Padberg, Bauer 2006). Cross protection due to related virus strains has also been determined e.g. with respect to influenza viruses (Broberg et al. 2011).

- Considering the aforementioned aspects, we have to keep in mind that all data related to infections used here underestimate the real number of infected individuals in Germany as well as in nearly all countries where the Coronavirus emerged. Typically, at the beginning of the pandemic, only suspected cases with COVID-19 symptoms were tested, leading to a heavy underestimation of infected people without symptoms (see Section 2.2). Several recent studies have tried to estimate the real prevalence of the virus and/or the infection fatality rate (IFR), including all infected cases rather than the confirmed (see Table 13). Estimated rates of unreported cases (estimate PRV/reported PRV) lie between 5 (Gangelt, Germany) and 50-85 (Santa Clara County, USA). Obviously, when estimated CFR values exceed the estimated IFR values by ten times or more, there must be a large number of unreported cases and the total number of infected individuals must be

> considerably higher than reported, respectively. The logical consequence is that there is a hidden decrease of the absolute number of susceptible individuals because of many infected persons without symptoms not knowing that they have been infected (and probably immunized) in the past. These individuals were not tested for the *pathogen* (SARS-CoV-2 virus) because they did not suffer from the *disease* (COVID-19). Quantifying the "dark figure" of SARS-CoV-2 infections by using representative sample-based tests on current infection as well as seroprevalence will be a challenge in the near future.

Of course, the present empirical results cannot prove or disprove the presence or absence of (herd) immunity. However, the number of infected individuals is obviously higher than reported (see Table 13), whilst the number of susceptibles could have been considerably smaller than expected already at the beginning of the pandemic. If a SARS-CoV-2 infection leads to (lifelong or temporal) immunity (which is not yet clarified), the current level of immunity must be higher as well. Similar results were found in the UK: Stedman et al. (2020) also find decreasing infection rates (R_{ADIR}) related to (reported and unreported) prevalence on the regional level (Upper Tier Local Authority areas) and conclude that "the only factor that could be related to the R_{ADIR} in this analysis was the historic number of confirmed number infection/,000 population suggesting that some of the reduction in reported cases is due to the build-up of immunity due to larger numbers of historic cases in the population".

However, we have to keep in mind that even if herd immunity was achieved, this does not mean that *no* new infections occur. Furthermore, herd immunity implies a closed population, where there are too many immunized individuals to infect the remaining susceptibles. In reality, there are migratory and mobility flows between regions and nations (e.g., work-related commuting, tourism) and, thus, new infections may occur due to transmissions driven by spatial interactions. More precisely, a susceptible individual living in a given region with herd immunity might get infected when traveling to another region. Finally, there is a difference between *infection* and *disease*, whilst the infectiousness of asymptomatically infected individuals is not yet clarified, although they are regarded as much less likely to transmit the virus than infected with symptoms (World Health Organization 2020a).

4.2 Determinants of regional growth and mortality

Two regression models were estimated, with intrinsic growth rates (indicating the speed of pandemic growth) and mortality (indicating the severity of the disease) as dependent variables. For both variables, spatial autocorrelation was detected, which can be explained comprehensively by virus transmission across borders of nearby counties. However, both the OLS and the spatial regression models give qualitatively the same results.

With respect to the determinants of growth speed, two explanatory variables have a significant effect opposite to the expected: A slower disease spread is not due to lower population density. In contrast, intrinsic growth rates decrease with higher density values on the level of German counties. This could be explained with the validity of this indicator as it is questionable whether population density is a sufficient proxy for the amount and intensity of physical contacts between individuals. There is no empirical evidence that inhabitants living in larger and densely populated municipalities (such as large cities) have more social interactions than in people in rural areas (Mitterer 2013, Petermann 2001). There is also no dampening effect of virus spread by an older population on the county level. Instead, growth rates increase with the share of inhabitants of 65 years and older, although this age category covers the retired population. This result might be due to a bias in testing for SARS-CoV-2 infections: Most tests in the past were conducted on people with COVID-19 symptoms. As older people are more likely to have a severe course of the disease, these age groups are obviously overrepresented in the tested and confirmed cases (RKI 2020a). Differing from the expectations, no isolated effect of East German counties was found.

However, regional growth rates of infections are increased by inter-regional mobility, especially with respect to work-related commuting. This result is quite plausible and

Table 13: Studys on unreported cases and/or IFR of SARS-CoV-2/COVID-19

Study	Study area	n	Time of data collection	Est. PRV [%] (CI95)	Est.asymp-tomatic cases [%]	Est.PRV/ reported PRV	Est. IFR [%] (CI95)
Bendavid et al. (2020)	Santa Clara county (USA)	3,324	04/2020	2.8 (2.2, 3.4)	NA	50-85	NA
Bennett, Steyvers (2020) - re-analysis of Bendavid et al. (2020)	Santa Clara county (USA)	3,324	04/2020	0.27-3.21	NA	5-65	NA
Gudbjartsson et al. (2020)	Iceland						
Targeted testing 1		177	01-03/2020	9.2	13.6	NA	NA
Population screening 1		10,797	03/2020	0.8	41.4	NA	NA
Targeted testing 2		7,275	03/2020	14.4	5.7	NA	NA
Population screening 2 (Random sample)		2,283	04/2020	0.6	53.8	NA	NA
LAPH (2020)	Los Angeles county (USA)		04/2020	4.1 (2.8, 5.6)	NA	28-55	NA
Lavezzo et al. (2020)[1]	Vo						
First survey	(Italy)	2,812	02/2020	2.6[2] (2.1, 3.3)	41.1	NA	NA
Second survey		2,343	03/2020	1.2[2] (0.8, 11.8)	44.8	NA	NA
Nishiura et al. (2020)[1]	Wuhan (China)[3]	565	02/2020	1.4	62.5	5-20	0.3-0.6
Russell et al. (2020)[1]	Diamond Princess cruise ship	3,711	02/2020	17	51.4	NA	1.3 (0.38, 3.6)
Shakiba et al. (2020)	Guilan province (Iran)	551	04/2020	33[4] (28, 39)	18	NA	0.08-0.12
Streeck et al. (2020)	Gangelt (Germany)	919	03/2020	15.5 (12.3, 19)	22.2	5	0.36 (0.29, 0.45)

Source: own compilation. *Notes*: [1]full survey or nearly full survey, [2]only active infections (not including seroprevalence), [3]Japanese citizens evacuated from Wuhan (China), [4]adjusted for test performance.

supports previous findings in the literature on empirical and model-based approaches towards infectious disease spreads in the past (Charaudeau et al. 2014, Dalziel et al. 2014, Findlater, Bogoch 2018).

Taking a look at state-specific effects, the results show a significantly higher average growth rate in Bavarian counties and lower growth rates in Saxony and North Rhine Westphalia, while Saarland and Baden-Wuerttemberg counties did not deviate from the national average. Thus, no dampening effect on pandemic growth can be confirmed for German states with additional curfews for the containment of the virus spread (Bavaria, Saarland, Saxony). The growth rate in Bavaria is even considerably above the average, although the time since the virus emerged and the current prevalence were included into the models as control variables.

With respect to the determinants of mortality, few clear statements can be made. SARS-CoV-2/COVID-19 mortality on the county level is not significantly influenced by demographic, spatial or mobility factors. However, these variables explain the regional growth rate, which was added to the mortality model as a control variable (and is significant, as expected). There is no specific "East Germany effect" as well.

Considering previous studies on the influence of air pollution on COVID-19 severity (Ogen 2020, Wu et al. 2020), it was expected that particulate matter and NO_2 concentration would have an influence on mortality. Although it is plausible to assume that air pollution increases fatality rates of respiratory diseases, this hypothesis was not confirmed in the present study, which may result from an obvious data problem: The annual mean values of daily pollution was obtained on the level of monitoring stations, which are not evenly distributed across the counties and measure the air pollutant level at a specific point (e.g. traffic crossroad). It is unlikely that these obtained values are representative for the whole county. Thus, the validity of this indicator is questionable. Unfortunately, county-based data towards air pollution is not available nationwide.

Obviously, the variance of regional mortality reflects the regional variance of infected individuals belonging to the "risk group", defined as people of 60 years and above. Although no regional data is available for cases and deaths in retirement homes, a large share should be attributed to these facilities. Nationwide, people accommodated in facilities for the care of elderly make up at least 2,473 of 6,831 deaths (36.20%) as of May 5, 2020 (RKI 2020a). The share of residents of retirement homes in all COVID-19-related deaths is equal to 51% in France and 33% in Denmark, ranging internationally from 11% in Singapore to 62% in Canada (Comas-Herrera et al. 2020). The relevance of retirement homes in Germany can be underlined with examples based on information available in local media which depicts the regional situation:

- In the city of Wolfsburg (Lower Saxony), the current mortality (MRT) is equal to 41.08 deaths per 100,000 inhabitants, while the current case fatality rate (CFR) is the highest in all German counties (17.89%). Both values are calculated from the data used here (of date May 5, 2020). As of May 11, there have been 51 deaths attributed to COVID-19 in Wolfsburg, with 44 of these deaths (86.27%) stemming from residents of one retirement home (Wolfsburger Nachrichten 2020).
- In the Hessian Odenwaldkreis with MRT = 54.75 and CFR = 14.60%, 29 people who tested positive to SARS-CoV-2/COVID-19 died until April 14, 2020, 21 of them (72.41%) were living in retirements homes in this county (Echo online 2020).
- In the city of Würzburg (Bavaria) with MRT = 38.32 and CFR = 10.75%, there have been 44 COVID-19 positive deceased in two retirement homes until April 24, 2020, leading to investigations by the public prosecution authorities (BR24 2020). Up to April 23, 2020, in the whole administrative district Unterfranken, 64% of all people who died from or with Corona were residents of retirement homes for elderly people (Mainpost 2020).

With respect to state-specific effects, there is a clear significant impact regarding Bavaria: Although the measures in Bavaria, based on the Austrian model, were probably the strictest of all German states, both regional growth rates and mortality are significantly *higher* than in the other states. This effect is isolated, as other effects (time, population density etc.) were controlled. In addition, the share of individuals belonging to the "risk group" is slightly higher in Bavaria. In the other states with curfews, Saarland and Saxony, no significant impact of this additional intervention was found, especially with respect to mortality which does not differ significantly from other states.

5 Conclusions and limitations

In the present study, regional SARS-CoV-2/COVID-19 growth was analyzed as an empirical phenomenon from a spatiotemporal perspective. Using infection dates estimated from reported cases, logistic growth models were estimated for the disease spread at the level of German counties as well as at the national level. The resulting intrinsic growth rates vary across the 412 German counties. The inflection points of the epidemic curves were contrasted to the dates where nonpharmaceutical interventions against the disease spread came into force. As a result, Germany as a whole as well as the majority of German counties have experienced a decline of the infection rate – which means a flattening of the infection curve – before the main social-related measures (contact ban, ban of gatherings and closure of "nonessential" services) were established. In a minority of counties, curve flattening even occured before schools and child day care facilities were closed. In contrast, some regional trend changes took place several weeks after the measures came into force. *Due to this temporal mismatch, we have to conclude that the decline of infections cannot*

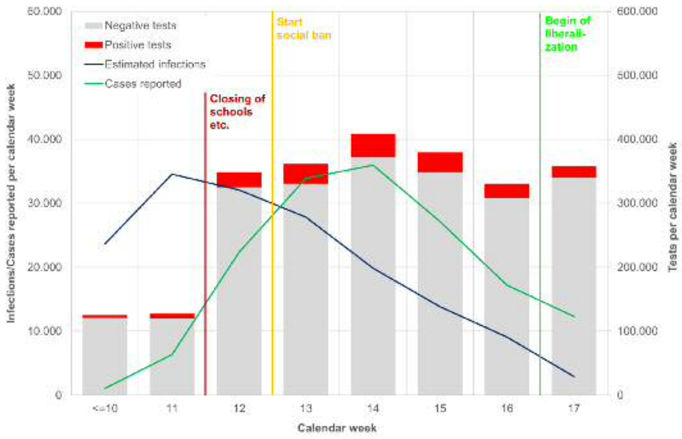


Figure 12: Estimated infections, reported cases and conducted tests by calendar week
Source: own illustration. Data source: own calculations based on RKI (2020b,c)

be causally linked to the "lockdown" of March 23. Moreover, also the impact of school and child care infrastructure closures on the pandemic spread remains questionable.

However, this does not mean that the disease spread slows down automatically. Four possible reasons have been identified for curve flattening independent from school closures and the "lockdown", with the first two relating to other types of (state-run and voluntary) measures which could reduce the transmission of an infectious disease: 1) Positive effects of *previous governmental nonpharmaceutical interventions* (especially the cancellation of large-scale events), 2) *voluntary behavior changes* (e.g., with respect to physical distancing and hygiene), 3) *seasonality* of the virus, and 4) a rising but undiscovered level of *immunity* within the population. However, whether these determinants may have contributed to the decline of infections, is outside the scope of the model-based analysis.

The determinants of regional intrinsic growth rates (as an indicator for the speed of pandemic spread) and mortality (as an indicator of the disease's severity) were explored using regression models. Among other things, regional pandemic growth is found to be driven by inter-regional mobility. Mortality on the county level obviously depends on the share of infected individuals belonging to the "risk group" (people of age 60 or older). This share is considerably influenced by SARS-CoV-2/COVID-19 outbreaks in retirement homes for the eldery, which have occurred in many German counties. Obviously, neither strict measures in Germany nor other countries were able to prevent these location-specific outbreaks. *By consequence, it must be concluded that the severity of SARS-CoV-2/COVID-19 depends on the local/regional ability to protect the "risk group", especially older people in care facilities. This is the more important as virus transmission in care homes is nearly independent from the nonpharmaceutical interventions concerning e.g. schools, commercial services, and private residences.* Three German states (Bavaria, Saarland, Saxony) established curfews additional to the nationwide interventions. *We must conclude that these state-specific curfews did not contribute to a more positive outcome with respect to growth speed and mortality.*

On the one hand, these findings pose the question as to whether contact bans and curfews are appropriate measures for containing the virus spread, especially when weighing the effects against the social and economic consequences as well as the curtailment of civil rights. On the other hand, when looking at regional mortality and case fatality rate, the protection of "risk groups", especially older people in retirement homes, is obviously of moderate success.

From the methodological point of view, two further conclusions must be stated: Nonpharmaceutical interventions aim at the reduction of new infections, thus, their impact must be assessed regarding temporal coincidences with new infections. *Regardless of the modeling approach used for the analysis of pandemic spread, any analysis concerning the effectiveness of nonpharmaceutical interventions must be based on realistic infection* *dates rather than reporting dates of infected persons. An over- or underestimation of the time between infection and report – in particular, the reporting delay – might lead to senseless conclusions towards the influence of specific measures.* Estimating the true infection dates from reported cases in official statistics is the biggest methodological challenge in this context. Moreover, the present study reveals the importance of a spatial perspective on pandemic spread: *Spatially varying growth rates and severity measures show that the spread of an infectious disease is to be regarded as a spatiotemporal phenomenon. Thus, further studies should address regional differences of epidemiological variables with respect to transmission.*

However, despite these conclusions, the study is faced with two important limitations:

- One has to keep in mind that previous model-based simulation studies which prove the effectiveness of nonpharmaceutical interventions already make *a priori* assumptions about the impact of these measures: In particular, the *input* parameters of the epidemiological models (such as the intensity of physical contacts between individuals) are set in a way that interventions (such as school closures or social distancing) reduce the transmission of the virus in any case and, thus, the simulation *output* shows a decline of infections subsequent to these interventions (an der Heiden, Buchholz 2020, Ferguson et al. 2020). This type of modeling approach (and the corresponding results) might be regarded either as a "causal model" or a tautology. In contrast, the modeling approach used here is of purely empirical nature, only incorporating time series of infections. By consequence, the results are not causal but correlative with respect to the presence or absence of temporal coincidences. It can be shown that curve flattening does not coincide with the focused interventions but occurred after previous interventions and might be due to several other causes. However, the actual reasons *why* the infections declined cannot be deduced from modeling results but must be explored based on interpretations of several empirical hints. Furthermore, as the focus is on inflection points and trend changes, respectively, the present empirical analysis cannot rule out *additional* impacts of the German "lockdown", e.g., in terms of a stabilization effect.

- Finally, it is necessary to take a look at the quality of the data on reported cases of SARS-CoV-2/COVID-19 used here. While several statistical uncertainties have been addressed by estimating the dates of infection in the present study, the method of data collection also raises concerns. The confirmed cases of infections reported by regional health departments to the RKI result from SARS-CoV-2 tests conducted in the case of specific symptoms. When aggregating the reported cases to time series and analyzing their temporal evolvement, it is implictly assumed that the testing strategy remains the same over time. However, the number of tests was increased enormously during the pandemic – which is to be welcomed from the point of view of public health. From a statistic perspective, it might cause a bias because an increase in testing *must* result in an increase of reported infections, as a larger share of infections is revealed, all other things being equal. In his statistical study, Kuhbandner (2020) argues that the detected SARS-CoV-2 pandemic growth is mainly due to increased testing, leading to the conclusion that "the scenario of a pandemic spread of the Coronavirus in based on a statistical fallacy". To confirm or deny this conclusion is not subject of the present study. However, taking a look at the conducted tests per calendar week (see Figure 12) reveals weekly differences. From calendar week 11 to 12, there has been an increase of conducted tests from 127,457 (5.9% positive) to 348,619 (6.8% positive), which means a raise by factor 2.7. The maximum of tests was conducted in CW 14 (408,348 with 9.0% positive results), decreasing beyond that time, rising again in CW 17. The absolute number of positive tests is reflected plausibly in the number of reported cases (green line), as the confirmed cases result from the tests. The most estimated infections occurred in CW 11 and 12, showing again the delay between infection and case confirmation. Apart from the fact that excessive testing is probably the best strategy to control the spread of a virus, the resulting statistical data may suffer from underestimation and overestimation, dependent on which time period is regarded.

References

Al-Tawfiq J, Rodriguez-Morales A (2020) Super-spreading events and contribution to transmission of MERS, SARS, and SARS-CoV-2 (COVID-19). *Journal of Hospital Infection* 105: 111–112. CrossRef.

an der Heiden M, Buchholz U (2020) Modellierung von Beispielszenarien der SARS-CoV-2-Epidemie 2020 in Deutschland. Report. https://doi.org/10.25646/6571.2 (accessed May 09, 2020)

an der Heiden M, Hamouda O (2020) Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting. *Epidemiologisches Bulletin* 17: 10–16. CrossRef.

Araujo MB, Naimi B (2020) Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate. *medRxiv*. CrossRef.

Backer JA, Klinkenberg D, Wallinga J (2020) Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance* 25. CrossRef.

Batista M (2020a) Estimation of the final size of the coronavirus epidemic by the logistic model (Update 4). Preprint. https://www.researchgate.net/publication/339240777_Estimation_of_the_final_size_of_coronavirus_epidemic_by_the_logistic_model

Batista M (2020b) Estimation of the final size of the coronavirus epidemic by the SIR model. Preprint. https://www.researchgate.net/profile/Milan_Batista/publication/339311383_Estimation_of_the_final_size_of_the_coronavirus_epidemic_by_the_SIR_model/links/5e767fca4585157b9a512f80/Estimation-of-the-final-size-of-the-coronavirus-epidemic-by-the-SIR-model.pdf

BBSR – Bundesinstitut für Bau, Stadt- und Raumforschung (2020) Wo in Deutschland die Lebenserwartung am höchsten ist. Dataset. https://www.bbsr.bund.de/BBSR/DE/Home/Topthemen/interaktive-karten/lebenserwartung/lebenserwartung.html (accessed May 01, 2020)

Ben-Israel I (2020) The end of exponential growth: The decline in the spread of coronavirus. The Times of Israel, 19 April 2020. https://www.timesofisrael.com/the-end-of-exponential-growth-the-decline-in-the-spread-of-coronavirus/ (accessed May 07, 2020)

Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, Lai C, Weissberg Z, Saavedra-Walker R, Tedrow J, Tversky D, Bogan A, Kupiec T, Eichner D, Gupta R, Ioannidis J, Bhattacharya J (2020) COVID-19 Antibody Seroprevalence in Santa Clara County, California. *medRxiv*. CrossRef.

Bennett ST, Steyvers M (2020) Estimating COVID-19 Antibody Seroprevalence in Santa Clara County, California. A re-analysis of Bendavid et al. *medRxiv*. CrossRef.

Bivand RS, Pebesma E, Gomez-Rubio V (2013) *Applied spatial data analysis with R* (second ed.). Springer, NY

BR24 – Bayerischer Rundfunk (2020) Corona: Vorermittlungen gegen weiteres Würzburger Seniorenheim. Online article of April 24, 2020. https://www.br.de/nachrichten/bayern/corona-vorermittlungen-gegen-weiteres-wuerzburger-seniorenheim,Rx4vSyc (accessed May 12, 2020)

Braun J, Loyal L, Frentsch M, Wendisch D, Georg P, Kurth F, Hippenstiel S, Dingeldey M, Kruse B, Fauchere F, Baysal E, Mangold M, Henze L, Lauster R, Mall M, Beyer K, Roehmel J, Schmitz J, Miltenyi S, Mueller MA, Witzenrath M, Suttorp N, Kern F, Reimer U, Wenschuh H, Drosten C, Corman VM, Giesecke-Thiel C, Sander LE, Thiel A (2020) Presence of SARS-CoV-2 reactive T cells in COVID-19 patients and healthy donors. *medRxiv*. CrossRef.

Broberg E, Nicoll A, Amato-Gauci A (2011) Seroprevalence to influenza A(H1N1) 2009 virus - where are we? *Clinical and vaccine immunology* 18: 1205–1212. CrossRef.

Buchholz U, Buda S, Prahm K (2020) Abrupter Rückgang der Raten an Atemwegserkrankungen in der deutschen Bevölkerung. *Epidemiologisches Bulletin* 16: 7–9. CrossRef.

Capital (2020) Thomas Straubhaar: Die öffentliche Meinung wird kippen. Online article of March 21, 2020. https://www.capital.de/wirtschaft-politik/thomas-straubhaar-die-oeffentliche-meinung-wird-kippen (accessed March 23, 2020)

Carvalho Aguiar Melo M, de Sousa Soares D (2020) Impact of social distancing on mental health during the COVID-19 pandemic: An urgent discussion. *International Journal of Social Psychiatry* 66: 625–626. CrossRef.

Charaudeau S, Pakdaman K, Boëlle PY (2014) Commuter Mobility and the Spread of Infectious Diseases: Application to Influenza in France. *PLOS ONE* 9: 1–9. CrossRef.

Chi G, Zhu J (2008) Spatial Regression Models for Demographic Analysis. *Population Research and Policy Review* 27: 17–42. CrossRef.

Chowell G, Simonsen L, Viboud C, Yang K (2014) Is West Africa Approaching a Catastrophic Phase or is the 2014 Ebola Epidemic Slowing Down? Different Models Yield Different Answers for Liberia. *PLoS currents* 6. CrossRef.

Chowell G, Viboud C, Hyman J, Simonsen L (2015) The Western Africa Ebola Virus Disease Epidemic Exhibits Both Global Exponential and Local Polynomial Growth Rates. *PLOS Currents Outbreaks*. CrossRef.

CNN – Cable News Network (2020) China goes into emergency mode as number of confirmed Wuhan coronavirus cases reaches 2,700. Online article of January 27, 2020. https://edition.cnn.com/2020/01/26/asia/wuhan-coronavirus-update-intl-hnk/index.html (accessed May 8, 2020)

Comas-Herrera A, Zalakaín J, Litwin C, Hsu AT, Lane N, Fernández JL (2020) Mortality associated with COVID-19 outbreaks in care homes: early international evidence (Last updated 3 May 2020). Report, International Long-term Care Policy Network. https://ltccovid.org/wp-content/uploads/2020/05/Mortality-associated-with-COVID-3-May-final-6.pdf (accessed May 12, 2020)

Covid-19 Mobility Project (2020a) First Report: Mobility in Germany and Social Distancing (April 5, 2020). Website. https://www.covid-19-mobility.org/reports/first-report-general-mobility/ (accessed August 13, 2020).

Covid-19 Mobility Project (2020b) Second Report: Mobility on the rise (April 15, 2020). Website. https://www.covid-19-mobility.org/reports/second-report/ (accessed August 13, 2020).

Dalziel BD, Pourbohloul B, Ellner SP (2014) Human mobility patterns predict divergent epidemic dynamics among cities. *Proceedings. Biological sciences* 280: 20130763. CrossRef.

Destatis (2020a) Beherbergungsbetriebe, Schlafgelegenheiten, Gästeankünfte, Gästeübernachtungen – Jahressumme – regionale Tiefe: Kreise und krfr. Städte (ab 2018) (Tab. 45412-01-03-4). Dataset. https://www.regionalstatistik.de/genesis/online/ (accessed August 13, 2020), dl-de/by-2-0

Destatis (2020b) Bevölkerung: Kreise, Stichtag, Altersgruppen – Fortschreibung des Bevölkerungsstandes (Tab. 12411-0017). Dataset. https://www.regionalstatistik.de/genesis/online/ (accessed May 6, 2020), dl-de/by-2-0

Destatis (2020c) Sozialversicherungspflichtig Beschäftigte am Arbeits- und Wohnort, Ein- und Auspendler über Kreisgrenzen nach Geschlecht – Stichtag 30.06. – regionale

Ebenen (Tab. 13111-09-01-4-B). Dataset. https://www.regionalstatistik.de/genesis/online/ (accessed August 14, 2020), dl-de/by-2-0

Deutsche Welle (2020a) Coronavirus: What are the lockdown measures across Europe? Online article of April 14, 2020. https://www.dw.com/en/coronavirus-what-are-the-lockdown-measures-across-europe/a-52905137 (accessed May 8, 2020)

Deutsche Welle (2020b) What are Germany's new coronavirus social distancing rules? Online article of March 22, 2020. https://www.dw.com/en/what-are-germanys-new-coronavirus-social-distancing-rules/a-52881742 (accessed May 8, 2020)

D'Souza G, Dowdy D (2020) What is Herd Immunity and How Can We Achieve It With COVID-19? Online article of April 10, 2020. https://www.jhsph.edu/covid-19/articles/achieving-herd-immunity-with-covid19.html (accessed May 10, 2020)

Echo online (2020) Coronavirus: Altenheime im Odenwaldkreis beklagen 21 Tote. Online article of April 14, 2020. https://www.echo-online.de/lokales/odenwaldkreis/odenwaldkreis/coronavirus-altenheime-im-odenwaldkreis-beklagen-21-tote_21547352 (accessed May 12, 2020)

Engel J (2010) Parameterschätzen in logistischen Wachstumsmodellen. *Stochastik in der Schule* 1: 13–18. CrossRef.

Engina AB, Engin ED, Engin A (2020) Two important controversial risk factors in SARS-CoV-2 infection: Obesity and smoking. *Environmental toxicology and pharmacology* 78: 103411. CrossRef.

European Centre for Disease Prevention and Control (2020) COVID-19 situation update worldwide, as of 10 May 2020 . Website. https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases (accessed May 11, 2020)

Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunubá Z, Cuomo-Dannenburg G, Dighe A, Dorigatti I, Fu H, Gaythorpe K, Green W, Hamlet A, Hinsley W, Okell LC, van Elsland S, Thompson H, Verity R, Volz E, Wang H, Wang Y, Walker PG, Walters C, Winskill P, Whittaker C, Donnelly CA, Riley S, Ghani AC (2020) Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand (16 March 2020). Report, Imperial College COVID-19 Response Team. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf (accessed May 12, 2020)

Findlater A, Bogoch II (2018) Human Mobility and the Global Spread of Infectious Diseases: A Focus on Air Travel. *Trends in Parasitology* 34: 772–783. CrossRef.

Fine P, Eames K, Heymann DL (2011) "Herd Immunity": A Rough Guide. *Clinical Infectious Diseases* 52: 911–916. CrossRef.

Focus (2020) Bei 9 Grad fühlt sich Corona am wohlsten: Verschwindet das Virus im Sommer? Online article of April 14, 2020. https://www.focus.de/gesundheit/ratgeber/erkaeltung/neue-modellrechnung-bei-9-grad-fuehlt-sich-corona-am-wohlsten-verschwindet-das-virus-im-sommer_id_11885151.html (accessed May 10, 2020)

Fox J, Weisberg S (2019) *An R Companion to Applied Regression* (3rd ed.). Sage, Thousand Oaks CA

Gomes MGM, Corder RM, King JG, Langwig KE, Souto-Maior C, Carneiro J, Goncalves G, Penha-Goncalves C, Ferreira MU, Aguas R (2020) Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *medRxiv*. CrossRef.

Greene WH (2012) *Econometric Analysis. Seventh Edition.* Pearson

Griffith D (2009) Spatial Autocorrelation. In: Kitchin R, Thrift N (eds), *International Encyclopedia of Human Geography*. 308–316

Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, Rawlings SA, Sutherland A, Premkumar L, Jadi RS, Marrama D, de Silva AM, Frazier A, Carlin AF, Greenbaum JA, Peters B, Krammer F, Smith DM, Crotty S, Sette A (2020) Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell*. CrossRef.

Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, Saemundsdottir J, Sigurdsson A, Sulem P, Agustsdottir AB, Eiriksdottir B, Fridriksdottir R, Gardarsdottir EE, Georgsson G, Gretarsdottir OS, Gudmundsson KR, Gunnarsdottir TR, Gylfason A, Holm H, Jensson BO, Jonasdottir A, Jonsson F, Josefsdottir KS, Kristjansson T, Magnusdottir DN, le Roux L, Sigmundsdottir G, Sveinbjornsson G, Sveinsdottir KE, Sveinsdottir M, Thorarensen EA, Thorbjornsson B, Löve A, Masson G, Jonsdottir I, Möller AD, Gudnason T, Kristinsson KG, Thorsteinsdottir U, Stefansson K (2020) Spread of SARS-CoV-2 in the Icelandic Population. *New England Journal of Medicine*. CrossRef.

Guth D, Holz-Rau C, Maciolek M (2010) Indikatoren für Berufspendelanalysen. Datengrundlagen und Anwendungsbeispiele. Raum und Mobilität. Arbeitspapiere des Fachgebiets Verkehrswesen und Verkehrsplanung der TU Dortmund, 18. http://www.vpl.tu-dortmund.de/cms/Medienpool/PDF_Dokomunte/Arbeitspapiere/AP18_von_Dennis_Guth_und_Christian_Holz-Rau_und_Markus_Maciolek.pdf (accessed August 15, 2020)

Hlavac M (2018) *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). R package version 5.2.2

IBBS – Informationsstelle des Bundes für Biologische Gefahren und Spezielle Pathogene (2020) 2019-nCoV: Verdachtsabklärung und Maßnahmen. Technical report. CrossRef.

Ipsos (2020) Internationale Umfrage: Mehrheit empfindet Coronavirus als große Bedrohung. Press release from February 12, 2020. https://www.ipsos.com/sites/default/files/ct/news/documents/2020-02/ipsos-pi_coronavirus_feb2020.pdf (accessed May 11, 2020)

Johns Hopkins University (2020) New Cases of COVID-19 In World Countries. Website. https://coronavirus.jhu.edu/data/new-cases (accessed May 11, 2020)

Kaw AK, Kalu EE, Nguyen D (2011) *Numerical Methods with Applications*. http://nm.mathforcollege.com/topics/textbook_index.html (accessed May 08, 2020)

Kuhbandner C (2020) The Scenario of a Pandemic Spread of the Coronavirus SARS-CoV-2 is Based on a Statistical Fallacy. Preprint. CrossRef.

Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR (2020) Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents* 55: 105924. CrossRef.

LAPH – Los Angeles County Department of Public Health (2020) USC-LA County Study: Early Results of Antibody Testing Suggest Number of COVID-19 Infections Far Exceeds Number of Confirmed Cases in Los Angeles County. Press release from April 20, 2020. http://publichealth.lacounty.gov/phcommon/public/media/mediapubhpdetail.cfm?prid=2328 (accessed May 09, 2020

Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J (2020, 03) The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*. CrossRef.

Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, Rossi L, Manganelli R, Loregian A, Navarin N, Abate D, Sciro M, Merigliano S, Decanale E, Vanuzzo MC, Saluzzo F, Onelia F, Pacenti M, Parisi S, Carretta G, Donato D, Flor L, Cocchio S, Masi G, Sperduti A, Cattarino L, Salvador R, Gaythorpe KA, , Brazzale AR, Toppo S, Trevisan M, Baldo V, Donnelly CA, Ferguson NM, Dorigatti I, Crisanti A (2020) Suppression of COVID-19 outbreak in the municipality of Vo, Italy. *medRxiv*. CrossRef.

Leung C (2020) The difference in the incubation period of 2019 novel coronavirus (SARS-CoV-2) infection between travelers to Hubei and nontravelers: The need for a longer quarantine period. *Infection Control & Hospital Epidemiology* 41: 594–596. CrossRef.

Li MY (2018) *An Introduction to Mathematical Modeling of Infectious Diseases*. Springer International Publishing

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TT, Wu JT, Gao GF, Cowling BJ, Yang B, Leung GM, Feng Z (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine* 382: 1199–1207. CrossRef.

Linton N, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov A, Jung SM, Yuan B, Kinoshita R, Nishiura H (2020) Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: A statistical analysis of publicly available case data. *Journal of Clinical Medicine* 9: 538. CrossRef.

Ma J (2020) Estimating epidemic exponential growth rate and basic reproduction number. *Infectious Disease Modelling* 5: 129–141. CrossRef.

Mainpost (2020) Hat Unterfranken das Coronavirus bald besiegt? Online article of May 8, 2020. https://www.mainpost.de/regional/wuerzburg/hat-unterfranken-das-coronavirus-bald-besiegt;art735,10443791 (accessed May 12, 2020)

Mak TW, Saunders ME (2006) Immunity to Pathogens. *The Immune Response*: 641–694. CrossRef.

Mitterer B (2013) *Raumbezogene Lebensstile und Konsummuster - eine Annäherung an die Dorfbewohner des 21. Jahrhunderts. Untersucht an zwei Gemeinden im Einzugsgebiet von Großstädten*. PhD thesis, Universität Würzburg. https://opus.bibliothek.uni-wuerzburg.de/files/9811/Dissertation_Mitterer_Brigitte_Raumbezogene_Lebensstile.pdf (accessed August 21, 2020).

Moriyama M, Hugentobler WJ, Iwasaki A (2020) Seasonality of Respiratory Viral Infections. *Annual Review of Virology* 7: 83–101. CrossRef.

Mucci F, Mucci N, Diolaiuti F (2020) Super-spreading events and contribution to transmission of MERS, SARS, and SARS-CoV-2 (COVID-19). *Clinical Neuropsychiatry* 17: 63–64. CrossRef.

New York Times (2020) A New Covid-19 Crisis: Domestic Abuse Rises Worldwide. Online article of April 6, 2020. https://www.nytimes.com/2020/04/06/world/coronavirus-domestic-violence.html (accessed May 10, 2020)

Nishiura H, Kobayashi T, Yang Y, Hayashi K, Miyama T, Kinoshita R, Linton NM, Jung SM, Yuan B, Suzuki A, Akhmetzhanov AR (2020) The Rate of Underascertainment of Novel Coronavirus (2019-nCoV) Infection: Estimation Using Japanese Passengers Data on Evacuation Flights. *Journal of clinical medicine* 9: 419. CrossRef.

Ogen Y (2020) Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality. *Science of The Total Environment* 726: 138605. CrossRef.

Padberg J, Bauer T (2006) Erkältungskrankheiten. *Deutsche Medizinische Wochenschrift* 131. CrossRef.

Pell B, Kuang Y, Viboud C, Chowell G (2018) Using phenomenological models for forecasting the 2015 ebola challenge. *Epidemics* 22: 62 – 70. CrossRef.

Petermann S (2001) Soziale Vernetzung städtischer und ländlicher Bevölkerungen am Beispiel der Stadt Halle. Abschlussbericht und Codebuch. Report. https://www2.soziologie.uni-halle.de/publikationen/pdf/0102.pdf (accessed August 21, 2020).

Porta M (2008) *A Dictionary of Epidemiology*. Oxford University Press

QGIS Development Team (2019) *QGIS Geographic Information System*. Open Source Geospatial Foundation. http://qgis.osgeo.org

R Core Team (2019) R: A language and environment for statistical computing. Software:, Vienna, Austria. https://www.R-project.org/

Rheinische Post online (2020) Hamsterkäufe und leere Regale. Online article of February 28, 2020. https://rp-online.de/wirtschaft/hamsterkaeufe-und-leere-regale-bei-aldi-lidl-und-co_aid-49285165 (accessed May 10, 2020)

Ritz C, Streibig JC (2008) *Nonlinear Regression with R*. Springer

RKI – Robert Koch Institut (2020a) Coronavirus Disease 2019 (COVID-19) – Daily Situation Report of the Robert Koch Institute 05/05/2020. Report. https://www.rki.de/DE/-Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-05-05-en.pdf (accessed May 07, 2020)

RKI – Robert Koch Institut (2020b) Tabelle mit den aktuellen Covid-19 Infektionen pro Tag (Zeitreihe). Dataset. https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/-dd4580c810204019a7b8eb3e0b329dd6_0/data (accessed May 05, 2020), dl-de/by-2-0

RKI – Robert Koch Institut (2020c) Täglicher Lagebericht des RKI zur Coronavirus-Krankheit-2019 (COVID-19) 06.05.2020. Report. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-05-06-de.pdf (accessed May 11, 2020)

Rusche K (2008) Quality of Life in the Regions: An Exploratory Spatial Data Analysis for West German Labor Markets. CAWM Discussion Paper, 10. https://www.econstor.eu/obitstream/10419/51262/1/671597892.pdf (accessed August 15, 2020

Russell TW, Hellewell J, Jarvis CI, van Zandvoort K, Abbott S, Ratnayake R, working group CC, Flasche S, Eggo RM, Edmunds WJ, Kucharski AJ (2020) Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. *Eurosurveillance* 25. CrossRef.

Süddeutsche Zeitung (2020a) Um jeden Preis? (guest commentary by René Schlott). Online article of March 17, 2020. https://www.sueddeutsche.de/leben/corona-rene-schlott-gastbeitrag-depression-soziale-folgen-1.4846867 (accessed March 19, 2020)

Süddeutsche Zeitung (2020b) Wenn das Kind verborgen bleibt. Online article of May 6, 2020. https://www.sueddeutsche.de/politik/coronavirus-haeusliche-gewalt-jugendaemter-1.4899381?print=true (accessed May 11, 2020)

Selvan ME (2020) Risk factors for death from COVID-19. *Nature Reviews Immunology* 20: 407. CrossRef.

Shakiba M, Hashemi Nazari SS, Mehrabian F, Rezvani SM, Ghasempour Z, Heidarzadeh A (2020) Seroprevalence of COVID-19 virus infection in Guilan province, Iran. *medRxiv*. CrossRef.

Stedman M, Davies M, Lunt M, Verma A, Anderson SG, Heald AH (2020) A phased approach to unlocking during the COVID-19 pandemic - Lessons from trend analysis. *International Journal of Clinical Practice* 74: e13528. CrossRef.

Stein RA (2011) Super-spreaders in infectious diseases. *International Journal of Infectious Diseases* 15: e510–e513. CrossRef.

Streeck H, Schulte B, Kuemmerer B, Richter E, Hoeller T, Fuhrmann C, Bartok E, Dolscheid R, Berger M, Wessendorf L, Eschbach-Bludau M, Kellings A, Schwaiger A, Coenen M, Hoffmann P, Noethen M, Eis-Huebinger AM, Exner M, Schmithausen R, Schmid M, Kuemmerer B (2020) Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. *medRxiv*. CrossRef.

Stuttgarter Zeitung (2020) Gewaltambulanz verzeichnet deutlich mehr Kindesmisshandlungen. Online article of May 5, 2020. https://www.stuttgarter-zeitung.de/inhalt.-coronavirus-in-baden-wuerttemberg-gewaltambulanz-verzeichnet-deutlich-mehr-kindesmisshandlungen.c73a0841-0538-4c80-9035-1e81436cfd47.html (accessed May 10, 2020)

Sun K, Chen J, Viboud C (2020) Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *The Lancet Digital Health* 2: e201–e208. CrossRef.

Tagesschau.de (2020a) 16 Wege aus der Corona-Krise. Online article of May 8, 2020. https://www.tagesschau.de/inland/corona-lockerung-bundeslaender-103.html (accessed May 8, 2020)

Tagesschau.de (2020b) Obergrenze mehrfach überschritten. Online article of May 10, 2020. https://www.tagesschau.de/inland/corona-landkreise-105.html (accessed May 10, 2020)

Tagesschau.de (2020c) Quarantäne im Kreis Heinsberg teils beendet. Online article of March 1, 2020. https://www.tagesschau.de/inland/coronavirus-heinsberg-101.html (accessed May 10, 2020)

Tagesspiegel (2020a) Gibt keinen Grund, das ganze Land in häusliche Quarantäne zu schicken. Online article of March 24, 2020. https://www.tagesspiegel.de/politik/epidemiologe-warnt-vor-noch-schaerferen-massnahmen-gibt-keinen-grund-das-ganze-land-in-haeusliche-quarantaene-zu-schicken/25672822.html (accessed March 27, 2020)

Tagesspiegel (2020b) In diesen Regionen wütet das Coronavirus besonders stark weiter. Online article of April 21, 2020. https://www.tagesspiegel.de/gesellschaft/tirschenreuth-heinsberg-potsdam-in-diesen-regionen-wuetet-das-coronavirus-besonders-stark-weiter/25755328.html (accessed March 27, 2020)

The Guardian (2020) 'Great Lockdown' to rival Great Depression with 3% hit to global economy, says IMF. Online article of April 14, 2020. https://www.theguardian.com/-business/2020/apr/14/great-lockdown-coronavirus-to-rival-great-depression-with-3-hit-to-global-economy-says-imf (accessed May 10, 2020)

Tsoularis A (2002) Analysis of logistic growth models. *Mathematical Biosciences* 179: 21 – 55. CrossRef.

Tveite H (2019) *NNJoin: Nearest neighbour join. Join vector layers based on nearest neighbour relationships*. http://arken.nmbu.no/ havatv/gis/qgisplugins/NNJoin/

UBA – Umweltbundesamt (2020a) Aktuelle Luftdaten – Jahresbilanzen. Website. https:-//www.umweltbundesamt.de/daten/luft/luftdaten/jahresbilanzen/ (accessed August 13, 2020).

UBA – Umweltbundesamt (2020b) Stationsdatenbank – Downloads. Website. https://-www.env-it.de/stationen/public/downloadRequest.do (accessed August 13, 2020).

Vasconcelos GL, Macêdo AMS, Ospina R, Almeida FAG, Duarte-Filho GC, Souza ICL (2020) Modelling fatality curves of covid-19 and the effectiveness of intervention strategies. *medRxiv*. CrossRef.

Welt online (2020a) In Hamburg ist niemand ohne Vorerkrankung an Corona gestorben. Online article of April 8, 2020. https://www.welt.de/regionales/hamburg/article-207086675/Rechtsmediziner-Pueschel-In-Hamburg-ist-niemand-ohne-Vorerkrankung-an-Corona-gestorben.html (accessed April 8, 2020)

Welt online (2020b) Warum Ostdeutschland weniger von Corona betroffen ist. Online article of May 4, 2020. https://www.welt.de/wissenschaft/article207718789/Epidemiologe-zu-Corona-Darum-kommt-der-Osten-so-gut-durch-die-Pandemie.html (accessed May 8, 2020)

Wieland T (2019) REAT: A Regional Economic Analysis Toolbox for R. *REGION* 6: R1–R57. CrossRef.

Williams SN, Armitage CJ, Tampe T, Dienes K (2020) Public perceptions and experiences of social distancing and social isolation during the COVID-19 pandemic: A UK-based focus group study. *medRxiv*. CrossRef.

Wolfsburger Nachrichten (2020) Corona in Wolfsburg: Die Fakten auf einen Blick. Online article of May 11, 2020. https://www.wolfsburger-nachrichten.de/wolfsburg/article228716311/corona-wolfsburg-infizierte-geschaefte-bus-bahn-infos-informationen-arzt.html (accessed May 12, 2020)

World Health Organization (2020a) Transmission of COVID-19 by asymptomatic cases. Website. http://www.emro.who.int/health-topics/corona-virus/transmission-of-covid-19-by-asymptomatic-cases.html (accessed August 20, 2020)

World Health Organization (2020b) WHO announces COVID-19 outbreak a pandemic. Press release from March 12, 2020. http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic (accessed May 10, 2020)

Wu K, Darcet D, Wang Q, Sornette D (2020) Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. Technical report. Preprint: https://arxiv.org/ftp/arxiv/papers/2003/2003.05681.pdf (accessed April 24, 2020)

Wu X, Nethery RC, Sabath BM, Braun D, Dominici F (2020) Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. *medRxiv*. CrossRef.

Xia W, Liao J, Li C, Li Y, Qian X, Sun X, Xu H, Mahai G, Zhao X, Shi L, Liu J, Yu L, Wang M, Wang Q, Namat A, Li Y, Qu J, Liu Q, Lin X, Cao S, Huan S, Xiao J, Ruan F, Wang H, Xu Q, Ding X, Fang X, Qiu F, Ma J, Zhang Y, Wang A, Xing Y, Xu S (2020) Transmission of corona virus disease 2019 during the incubation period may lead to a quarantine loophole. *medRxiv*. CrossRef.

YouGov (2020) Der Corona-Tracker: Angst der Bevölkerung vor Virus wächst. Press release from March 17, 2020. https://yougov.de/news/2020/03/17/der-corona-tracker-angst-der-bevolkerung-vor-virus/ (accessed May 11, 2020)

Zhou X, Ma X, Hong N, Su L, Ma Y, He J, Jiang H, Liu C, Shan G, Zhu W, Zhang S, Long Y (2020) Forecasting the Worldwide Spread of COVID-19 based on Logistic Model and SEIR Model. *medRxiv*. CrossRef.

REGION

The Journal of ERSA
Powered by WU

Volume 7, Number 2, 2020, 85–106
DOI: 10.18335/region.v7i2.308

journal homepage: region.ersa.org

Does the crisis change the nature of agglomeration economies in Indonesia? A productivity analysis of pre-post 1997-1998 financial crisis

Khoirunurrofik Khoirunurrofik[1]

[1] Universitas Indonesia, Depok, Indonesia

Received: 15 February 2020/Accepted: 5 December 2020

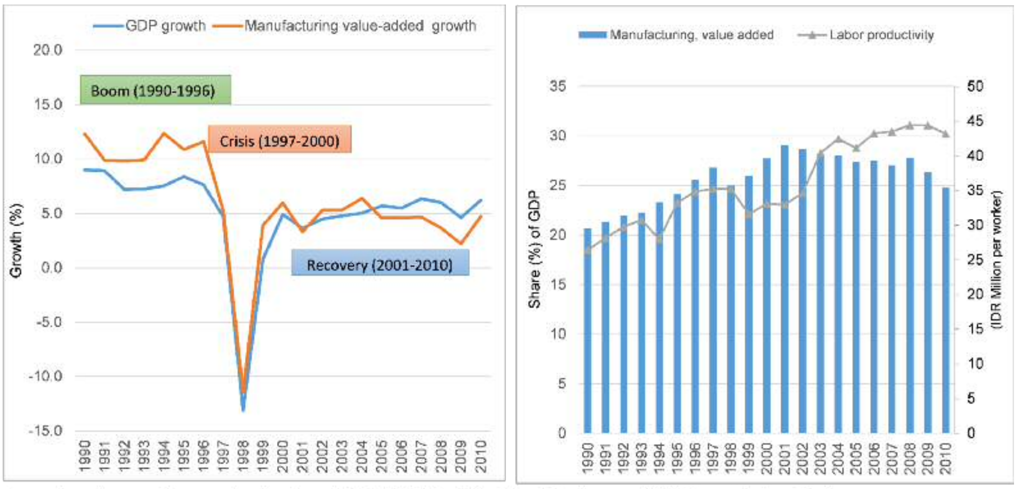
Abstract. By examining their source and magnitude, this paper looked at the changes in the nature of Indonesia's agglomeration economies over two distinct successive periods: the pre-crisis (1990–1998) and the post-crisis (1999-2010). We found that economies of localisation and urbanisation are both present in the post-crisis period, with the former generating a more substantial effect than the latter, though there has been a growing role of urbanisation in recent years. There is relatively strong evidence of plant size and sectoral heterogeneity with respect to types of agglomeration externalities. These factors shed some light on the nature of agglomeration economies and how the agglomeration sources after the crisis period have been visibly in favour of localisation economies. It confirms that the plants have been improved to benefit from the external environment given the policy in place. The results also demonstrate the firm and sectoral life cycle at work, as evidenced by the changing industry structure during the post-crisis period.

JEL classification: R11, R12, R30, L25, L60

Key words: Industry, agglomeration economies, economic cycles, financial crisis, Indonesia

1 Introduction

Since industrialisation began its heyday in Indonesia's economy in the mid-1960s, the manufacturing industry has now become one of its leading sectors. In the boom years of the early 1990s, the sector's role in the growth and vitality of the economy was especially remarkable. Nevertheless, this trend was overturned by the financial crisis of 1997–1998 as the industry saw decreased growth despite the introduction of certain deregulations, for example, in foreign ownership and tariff reduction, to stimulate investment in the sector (Aswicahyono et al. 2010). Eventually, Indonesia's manufacturing-driven economic growth was effectively paralysed (Poczter et al. 2013). Nevertheless, the sector is still a dominant force in the economy and contributes an average of 27.5% of gross domestic product (GDP) despite only registering about 4% growth performance from 2001 to 2010, compared to 10% from 1990 to 1996. Also interesting is the fact that the slower growth of the manufacturing sector is not linked to the increase in labour productivity, which



Source: Annual survey of large and medium firms 1990-2010, BPS and World Bank Development (WDI) data, author's calculation.

Figure 1: Indonesia's Economic and Manufacturing Sector Growth: The Manufacturing Sector's Contribution and Its Productivity

suggests that external factors like higher wage rates in agglomerated regions might have affected productivity (see Figure 1).

This paper is an attempt to investigate how external economies of scale, or agglomeration economies, affect productivity throughout the economic cycle, which in this case is represented by the period before, during, and after the Asian financial crisis of 1997–1998. Therefore, it is also an attempt to close the academic gap by studying the effects of external agglomeration economies – i.e., localisation and urbanisation economies – on productivity over the pre-crisis and post-crisis recovery periods. According to the literature, localisation economies capture the economic activities of similar firm clusters or industry concentration (Rosenthal, Strange 2004), where firms gain benefits from input sharing, spatial concentration, intra-industry knowledge sharing, labour pooling, and innovative competition (Gill, Goh 2010). Meanwhile, urbanisation economies focus on the diversity of a region's industrial structure (Rosenthal, Strange 2004), in which firms are more productive in a diversified region and enjoy the variation of business services, inter-industry exchange of ideas, enlarged market size, and more innovations (Gill, Goh 2010).

Even though a great deal of literature has explained the connection between productivity and agglomeration economies, there is arguably no study that compares the sources and magnitude of agglomeration economies along the financial crisis periods. De Groot et al. (2016) argued that using micro-panel data can enhance our understanding of how agglomeration externalities have played a key role in impacting productivity and find systemic variations among the element of agglomeration. Few authors, including Martin et al. (2011) and Henderson (2003), have utilised micro-panel data to examine the properties and sources of agglomeration economies. Nevertheless, none of them discussed how changing economic circumstances might have influenced the sources of agglomeration economies, especially surrounding the 1997-1998 financial crisis. The effect of the crisis on the nature of agglomeration economies might have been overlooked since these studies have either only utilised data collected over a short period or demonstrated little knowledge of the crisis. Conversely, several studies (for instance, Aswicahyono et al. 2010, Narjoko, Hill 2007, Poczter et al. 2013) have chosen to examine the financial crisis' effect on firm productivity, but their approaches were preoccupied with the effects related to firm's internal economies of scale, such as export performance, labour cost, and ownership.

Notwithstanding a few previous studies that have investigated similar issues (Henderson 2003, Martin et al. 2011, Day, Ellis 2013, Graham 2009, Kuncoro 2009), the novelty of this paper is that it examines the effects of localisation and urbanisation and how they differ across plants of different sizes, industries, and periods. The linking of the issue of agglomeration economies and productivity to external shocks and the cyclical behaviour of the economy is worthy of examination. Nonetheless there is available literature that has raised the issue in a rigorous style, such as explaining that growth and productivity might be linked to various types of structural or more short-term cycles. However, there is nothing obviously evident regarding the direction in which causality might run and whether the connection between agglomeration economies and productivity with external shocks result in the form of positive or negative externalities. This paper provides new empirical evidence about the impact of agglomeration economies on plant-level productivity by accounting for distinct economic stages; in doing so this paper aims to broaden the empirical literature on agglomeration economies studies.

In this introduction, we have presented a brief overview of the importance and novelty of this study. In Section 2 that follows, we survey the existing body of literature surveys on related subjects. The empirical modelling of this study is discussed in Section 3, while Section 4 reports the data and variable construction, and Section 5 discusses the analysis and results. We offer our conclusion for this study in Section 6.

2 Related Literature

The academic discussion over whether scale externalities are related to localisation or urbanisation economies have led to the question of the validity of the sources of agglomeration economies. Many authors have pointed to extensive literature which attempts to address the question by providing empirical evidence and guidelines on formulating better estimations and identifications (Beaudry, Schiffauerov 2009, Rosenthal, Strange 2004). Rosenthal, Strange (2004) stressed the importance of temporal, industrial, and geographic scopes in examining the sources and nature of agglomeration economies. Gill, Goh (2010) laid out the differences between urbanisation and localisation economies while highlighting the inter- or intra-industry exchange of ideas and technology to attain productivity growth. Moreover, Beaudry, Schiffauerov (2009) discovered the methodologies and measurements to determine the supported types of externalities. The debate has also been the subject of other studies which focus on empirical estimation. The seminal paper by Glaeser et al. (1992) inspired many empirical works attempting to explain the relationship between local industrial structure – namely, specialisation, competition, and diversity – and growth patterns in cities. Glaeser et al. (1992) explained how urban areas and local economies develop over time through the contributions of three types of externalities: intra-industry knowledge spillover, inter-industry knowledge spillover, and local competition.

According to Gill, Goh (2010), firms and workers within a particular industry located near each other can enjoy knowledge spillover from similar or different technologies, access a pooled market of labour and employment skill, and benefit from intermediate input sharing, all of which enhances firm productivity. In a dynamic context, Marshal, Arrow, and Romer (MAR) externalities explain the existence of external scale economies, or intra-industry knowledge spillover effects (Glaeser et al. 1992). The importance of knowledge as a source of both firm dynamics and local growth calls into debate which type of economic activity facilitates knowledge spillover (De Groot et al. 2016). The spillover of knowledge is believed to improve technological change, subsequently increasing economic growth.

On the other hand, Gill, Goh (2010) defines inter-industry exchanges of ideas and technology among different kinds of industries that could create more variety in business services, enlarge market size on the supply and demand sides, and facilitate more product innovation and firm growth. In a dynamic context, these effects are known as Jacobs externalities (Glaeser et al. 1992). Jacobs' notion of externalities is believed to have an impact on employment growth when the diversity is positively related and specialisation is negatively related to growth (Glaeser et al. 1992, Frenken et al. 2007). Finally, the third type of externality known as Porter externalities stems from the recognition that local competition also plays a role in firms' development. Local competition is a main source of pressure on firms to create innovative products and adopt new technologies (Glaeser et al. 1992).

Related to this study, by employing plant-level data, several empirical studies have produced the evidence suggesting localisation economies as the source of agglomeration economies; but this result might be different across various industries (Graham 2009, Henderson 2003, Martin et al. 2011). Moreover, variation in magnitude and source of agglomeration economies might also be the result of different data aggregation levels and estimation techniques, as argued by Melo et al. (2009) in their meta-study. Henderson (2003), who analysed the firm-level data of high-tech and machinery industries in the United States, argued that agglomeration is the result of localisation economies and found no evidence that pointed toward urbanisation economies. Similarly, Graham (2009) used the findings from the service and manufacturing industries in the United Kingdom to suggest that urbanisation and localisation economies exist. Still, only localisation economies returned significant positive effects on productivity. Martin et al. (2011) in their study supported the above findings by concluding that, while limited evidence of urbanisation economies was found, localisation economies were proven to enhance plant-level productivity in France. As these cases from the developed world have shown, localisation economies have been considered more dominant than urbanisation economies.

The prevalence of localisation economies in the discussion of agglomeration sources in Indonesia is also supported by Kuncoro (2009), who examined four selected industries and concluded that localisation as a form of agglomeration generates stronger benefits than those produced by urbanisation. More recently, Day, Ellis (2013) in their study of Indonesia's agglomeration economies, argued that identified benefit comes from localisation economies contributing to manufacturing growth, rather than from urbanisation. Similarly, we applied the classification of agglomeration effects into urbanisation and localisation economies as utilised in Kuncoro (2009) and Day, Ellis (2013). Nevertheless, this paper differs from the preceding works through its utilisation of a unique, long-panel data set at the plant level which enabled us to observe plant behaviours spanning years and over an economic cycle. As supported by Rosenthal, Strange (2004), using microdata in agglomeration studies improves estimation results reliability, as it positions the econometric model to address the endogeneity concerns. Recently, the survey by De Groot et al. (2016) emphasised the importance of micro-data in conjunction with applying spatial analysis to examine the impact of agglomeration.

3 Data and Variables

This study utilises time-series data from 1990 to 2010, which contains two different economic periods in Indonesia pertaining to the 1997-1998 Asian financial crisis: pre-crisis (1990–1998) and post-crisis (1999–2010). This study uses the Central Bureau of Statistics (BPS)'s unpublished electronic dataset of their annual medium and large firms survey (*Statistik Industri*). The publication surveyed companies with 20 or more employees, including new industrial companies with newly established commercial production. In this study, an establishment or a plant represents each individual unit of observation since the information did not differentiate a firm with many establishments from a standalone establishment. We used the terms 'firm' and 'establishment' interchangeably in the analysis, but the latter should be considered the primary concept. It includes data for 459,677 plants in the period 1990–2010. After data cleaning and adjustments, we were able to construct an unbalanced panel of cleaned observations consisting of 441,187 unique observations, representing 95.98% of the original sample (see Appendix for the cleaning process). The cross-tabulation report of entry and exit rates between an observational year and group classes is provided in Table A.2 of the Appendix. The average exit and entry rates were about 9 to 11 %, whereas the small size classes reported the highest rates. The average exit and entry rates were around 12 to 14 %, 6 to 9%, and 4 to 5% for small, medium, and large size classes, respectively.

In the *Statistik Industri*, each plant is identified using either a Plant Identity Code (PSID) or *Nomor Kode Induk Perusahaan* (NKIP) across different periods of the survey publication. To bridge the different coding, we developed a concordance table with data series for some years that are available in both codes. Moreover, we also classified a plant according to the Indonesian Field Business Classification (KLUI), which is published by

BPS following the International Standard Industrial Classification (ISIC). Multiple ISIC versions are present in the dataset. Data from 1990 until 1997 are classified based on ISIC Revision 2. Meanwhile, ISIC Revision 3 (ISICrev3) is used for 1998–2009. The 2010 data has followed the United Nation's standards of the updated ISIC Revision 4 (ISICrev4). We were able to construct a complete time series dataset for the whole data periods using the BPS-provided bridge table of the five-digit ISIC.

We expressed all values in a given year in constant 2000 prices and we used wholesale price indices (WPIs) which are published monthly in BPS's bulletin, Statistik Bulanan Indikator Ekonomi. The data were collected from the CEIC database and the Statistik Indonesia annual publication. Using the manufacturing WPI in five-digit ISIC, output, value-added, intermediate input, and materials were deflated. Furthermore, we deflated wage and energy and electricity using a GDP deflator and a weighted price of oil for the industry sector, accordingly.

Perpetual investment method (PIM) was applied to estimate firms' capital stock in Indonesia (Arnold, Javorcik 2009, Jacob, Meister 2005). Investment values of a plant is understood as the sum of land, building, machinery, vehicles, and other types of investment. Building investment was deflated using a residential and non-residential WPI according to Jacob, Meister (2005). Meanwhile, machinery and vehicles were deflated using the imported machinery and imported transport equipment WPI. Using the construction WPI, other investments were changed into real values.

Regional district data were utilised to specify natural endowments and regional characteristics. Road-length data were collected from BPS, and data for land-area were retrieved from the Ministry of Home Affairs. We employed the BPS' Village Potential Survey (PODES) to obtain data of share of coastal areas and share of households with electricity. Due to changing number of districts over the years – especially since 2001 reflecting the start of regional autonomy in Indonesia – we maintained the number of districts at 284 (the number in 1990) by reintegrating new districts into their original districts. By this approach, cross-district comparison from 1990 to 2010 was achievable. Since 2001, regional autonomy and fiscal decentralisation have stimulated the splitting of the country into new regions. In 2016 there were 34 provinces and 508 districts and cities in Indonesia.

4 Empirical Estimation

A two-step empirical approach was employed in the agglomeration economies modelling. The first step was running the semiparametric estimation of total-factor productivity (TFP) for each three-digit ISIC, as suggested by Levinsohn, Petrin (2003). The semiparametric estimators were used to control the endogeneity problem when firm- or plant-level data were used. LP methods propose a control function approach using a proxy variable to estimate the production function. This proxy variable should not be correlated at all with the unobserved productivity shock that is represented by a firm's investment decision or capital stock (Van Beveren 2012). We applied the LP method because of inadequate reliable investment data in the manufacturing data from Indonesia. As is common in the data from developing countries, there is a significant number of zero investments reported. Fortunately, that is not the case when using intermediate inputs such as materials, energy, or electricity consumption as a proxy variable for capital stock because such information is available from Indonesian manufacturing data (Vial 2006). To calculate plant-level production function, we ran the 'levpet' command on Stata that was developed by Petrin et al. (2004).

The second step involved running a fixed-effect panel data analysis to investigate how plant-level TFP is influenced by agglomeration economies while controlling for regional and plant characteristics. We referred to the application of an augmented standard production function model for the general framework of the agglomeration economies modelling by Rosenthal, Strange (2004).

$$y_i = g(A_i) f(\boldsymbol{x}_i) \tag{1}$$

and hence

$$TFP_{it} = g(A_{it}) \tag{2}$$

where y_i is the value-added from the plant, $g(A_I)$ indicates the production function shift from external economies, and x_i is a vector of the plant's levels of traditional inputs such as capital and labour. Through this framework, neutrality of productivity or a status of balance between labour and capital is assumed and agglomeration economies through $g(A_i)$ in which $g^{'}(.) \geq 0$ may be estimated. Based on Equation (2), it is now possible to set the econometric parameters into a model for testing the effects of urbanisation and localisation on plant-level productivity.

We computed localisation and urbanisation variables at the three-digit industrial classification in order to accurately separate agglomeration externalities, as advised by Beaudry, Schiffauerov (2009). Adhering to Henderson (2003), we estimated the static localisation through the decomposition of the local industry's employment into that of the local industry plant as well as employment average of other local plants. Additionally, population density was used to quantify static urbanisation, as this was more robust when concerning district areas and precisely exposed potential congestion costs or productivity benefits arising from urbanisation economies in a region (Melo et al. 2009). In addition to population density in reflecting urbanisation economies, we measured a diversity index to scrutinise whether an area is relatively more diverse when compared to other regions in Indonesia, following Khoirunurrofik (2018) and Marrocu et al. (2013). Log form was used for all non-dummies and share variables.

We specified the fixed-effect model at the plant level and incorporated industry-year (two-digit ISIC) fixed effects.

$$\begin{aligned} \ln TFP_{irt} \quad = \quad & \alpha_0 \beta_1 \ln Age_{irt} + \beta_2 \ln Size_{irt} + \beta_3 DFDI_{irt} + \beta_4 DGov_{irt} + \\ & \beta_5 DExport_{irt} + \beta_6 Coastal_{rt} + \beta_7 Electricity_{rt} + \beta_8 Roaddens_{rt} + \\ & \beta_9 \ln Distport_{rt} + \beta_{10} \ln Avrindregemp_{jrt} + \beta_{11} \ln Locplant_{jrt} + \\ & \beta_{12} \ln Popdens_{rt} + \beta_{13} \ln Diversity_{rt} + \epsilon_{irt} \end{aligned} \tag{3}$$

TFP_{irt} indicates the TFP of plant i in region r in year t. *Avrindregemp* calculates average number of employees in the same industry j but excludes the number of employees of their own plant i. *Locplant* signifies number of plants in industry j in region r at time t. *Popdens* denotes population density in region r at time t, not the total number of people in a region. *Diversity* measures the variety level challenged by a plant in a specific industry j in region r at time t.

Age and *Size* were defined respectively as the plant's age and plant's number of employees. Meanwhile, *DFDI*, *DGov*, and *DExport* respectively act as dummy variables for foreign ownership, government ownership, and export activity. If at least 10% of the plant was subject to foreign ownership, then *DFDI* was valued at 1, while if the share of the government in the plant was greater than 50%, then *DGov* was equal to 1. Nevertheless, although we named these variables as Dummies, they represent the conditions in a particular year and may change ownership from 1990 to 2010. Additionally, *DExport* was equal to 1 if the plants exported in the respective year. *DExport* may also change across years.

Furthermore, *Coastal* denotes the percentage of villages within a coastal area. Since there were some proliferation of villages, the number of villages with coastal areas changes over time, and thus also changes the value of coastal areas. It should be noted that this coastal area is not a time-invariant variable since number of villages in a district can increase. Finally, the percentage of households with access to electricity was represented by *Electricity*, while *Roaddens* indicates the ratio of the total length of three road types – national, provincial, and district roads – to provincial roads. We provide a complete description of variable definition and data sources in the Appendix (Table A.1).

Plant fixed effects were applied in order to control for the unobservable characteristics of the plant and location selection bias (Henderson 2003). These treatments were intended to avoid the plant behaviour bias, which could possibly lead to plants in the most agglomerate and productive regions being located. Nevertheless, endogeneity bias or

Table 1: Descriptive Statistics of Variables

Variable	Mean	SD	Min	Max
Dependent Variable (441,187 observations)				
ln (*TFP*)	4.172	1.432	-8.039	11.874
Plant Characteristics (441,187 observations)				
ln(Size)	2.396	0.914	0.000	4.700
ln(Age)	4.193	1.179	2.996	10.661
DFDI	0.068	0.252	0.000	1.000
DGov	0.069	0.253	0.000	1.000
Dexp	0.149	0.356	0.000	1.000
Regional Characteristics (283 observations)				
Coastal (%)	12.615	17.366	0.000	89.000
Electricity (%)	89.659	11.483	16.613	99.920
ln(Roaddens)	-0.902	0.920	-3.311	2.327
ln(Distport)	6.812	0.470	6.159	8.408
Agglomeration Economies (4,864 observations)				
ln(Locplant)	2.399	1.166	0.000	5.832
ln(Avrindregemp)	2.340	2.058	0.000	8.022
ln(Popdens)[a]	5.680	1.889	0.902	9.840
ln(Diversity)	-2.286	1.654	-10.216	1.136

Note: SD = Standard deviation. [a] Number of observations = 283.

possible reverse causality still might have existed owing to unobservable industries and regions characteristics which might influence plant productivity. To account for this, following Henderson (2003) and Maré, Graham (2013) we then split the error term into two – industry-time-period fixed effects and residual white noise error – so that the remaining shocks not absorbed by plant fixed effects would be absorbed by industry-year fixed effects. Additionally, to anticipate the possibility of correlation among plants in the same industry sector in a given region but not across industries which may create errors by in-cluster correlation, we allowed the clustering of standard errors by industry district to avoid underestimated errors and the more likely possibility that the null hypothesis would be rejected Cameron et al. (2011), Nichols, Schaffer (2007).

5 Results and Analysis

We used the TFP level that was calculated from the plant-level production function estimation as the dependent variable. The estimation results of the plant-level production function for each three-digit SIC are reported in the Appendix (Table A.3). It indicates that in 66% of the sectors, constant returns to scale could not be rejected. We then categorised independent variables into three groups: plant characteristics, regional characteristics, and agglomeration economies. At first glance, there is considerable heterogeneity concerning plant size and age as depicted in Table 1. Likewise, a high variation of road density implies an imbalance in the amount of transport infrastructure across regions. The table also demonstrates that urbanisation economies' measurement is slightly more dispersed than that of localisation economies. Nevertheless, the diversity index, with an average of negative values, indicates that the industry in most of regions are less diverse, implying that those regions tend to specialise in a specific industry.

5.1 Aggregate Estimate

For the dependent variable, we employed the TFP level estimated from the plant-level production function for each three-digit SIC. The empirical model estimation returned the main results depicted in Table 2. Columns (1) and (2) provide the estimation results for the baseline model, which investigates the existence of agglomeration externalities in Indonesia. As the estimated coefficients from the OLS estimation results show in

Column (1), we found significant effects of urbanisation and localisation economies. The pairwise correlation coefficients between the level and first difference of lnlocplant and lnavrindregemp were respectively 0.254 and -0.009. Therefore, no multicollinearity was found between variables representing localisation. However, the true values might have been overestimated in the results due to the likelihood of reversed causality between agglomeration variables and productivity.

Upon the application of industry-year dummies and fixed-effects methods, we observed that localisation economies greatly determined productivity as indicated in the Column (2) results and a significant coefficient value at 0.066, which means a 1% increase in the number of plants of industry for each district would lead to 0.066 improvements in plant productivity. Contrary to Kuncoro (2009), who observed significant coefficients for all specifications between 0.13 and 0.24, our estimate of the localisation economies was below half. The discrepancies are the results of our improvements in the estimation method by eliminating possible input endogeneity and plant self-selection biases and accommodating the unobserved plant fixed effects. Nevertheless, when compared to localisation economies in other countries' cases, our result apparently exhibited a similar magnitude. These include the 0.02 to 0.08 for the United States manufacturing (Henderson 2003), 0.03 for British manufacturing (Graham 2009), 0.032–0.063 for Korean manufacturing (Lee et al. 2010), and 0.05–0.06 for French manufacturing (Martin et al. 2011). This finding contradicts the survey of De Groot et al. (2016) that conclude it would be a less likely insignificant effect of specialisation when using micro-panel data as it might be less important at the firm level. Although between localisation and specialisation, some time is interchangeable, it shows a different impact of static and dynamic agglomeration measurement on productivity direction when we utilise the micro-data level. Additionally, we found the insignificantly negative findings for diversity to be consistent with the conclusion of De Groot et al. (2016), but it does not support Jacobs externalities theory, which states that it has to be positive. Another fascinating result is that the estimation result of the agglomerated regions of Java and two megapolitan areas, namely Greater Jakarta and Greater Surabaya, shows that localisation economies are consistent, revealed in columns 3 to 6, which is even higher than the basic estimation (column 2).

We discovered that all plant characteristics of the control variables significantly determine productivity, with the exception of the export dummy. Plant's age returned significant with a positive coefficient, which demonstrates internalisation of the accumulated knowledge of the plants over a period of improved productivity. We also found company size to be positive and statistically significant, which means larger plant sizes generate higher productivity. Government plants and foreign direct investment (FDI) returned statistically significant and had a positive effect, indicating that higher productivity plants possibly have better access to overseas markets and capital sources (Narjoko, Hill 2007). To further explore of the role of FDI in industrial development in Indonesia, we then looked at different effects of agglomeration between Java and Outside Java as well as in the area where the Multi-National Enterprises (MNEs) are mostly located (Greater Jakarta and Greater Surabaya). We have provided sub-sample estimations to examine whether the MNEs have acquired agglomeration benefits, particularly in Java and Major cities. These findings indicate that the manufacturing sector is saturated, concentrated in Java. Since the beginning of Indonesia's industrial development, the government has put in a great deal of investment to build a large number of industrial zones in Java, particularly around the capital city, Jakarta, contributing to the global economy (Firman et al. 2007). However, the negative effect of urbanisation in Greater Jakarta is due to the fact that urban areas contain many types of industries and tend to develop rapidly over time, thus facing problems of congestion and over-utilisation of infrastructure.

Furthermore, we obtained a confirmation about the significance of network externalities as depicted by road density. Between 0.065 and 1.687, the estimated coefficients were fairly robust, which implies that road infrastructure improvement across cities or districts in a province not only creates network connectivity between plants and jobs and their equivalents in other regions, but also raises productivity. Meanwhile, electricity and coastal location showed positive effects on productivity, which suggests that regional competitiveness is a necessary factor in improving productivity at the plant level. We

Table 2: Agglomeration Externalities: Main Results

Dependent Variable:	Total Factor Productivity (LnTFP)					
Empirical Method:	OLS	FE	FE	FE	FE	FE
Sample	All	All	Java	Non Java	Grt Jakarta	Grt Surabaya
	(1)	(2)	(3)	(4)	(5)	(6)
Age (Ln)	-0.067***	0.109***	0.100***	0.133***	0.152***	0.079***
	[0.009]	[0.009]	[0.010]	[0.023]	[0.017]	[0.022]
Size (Ln)	0.287***	0.060***	0.075***	-0.007	0.078***	0.080***
	[0.011]	[0.012]	[0.013]	[0.024]	[0.015]	[0.016]
DFDI (1=Foreign)	0.330***	0.119***	0.131***	0.085**	0.136***	0.054
	[0.036]	[0.018]	[0.019]	[0.039]	[0.026]	[0.057]
DGov (1=Gov)	0.402***	0.238***	0.273***	0.110***	0.138***	0.280***
	[0.036]	[0.026]	[0.030]	[0.026]	[0.033]	[0.033]
Dexp (1=Exp)	0.004	-0.005	-0.009	0.020	0.028*	0.020
	[0.021]	[0.009]	[0.010]	[0.016]	[0.014]	[0.018]
Coastal (%)	0.001	0.005***	0.008**	0.004	-0.002	-0.048
	[0.001]	[0.002]	[0.004]	[0.003]	[0.004]	[0.031]
Electricity (%)	-0.004**	0.002**	0.001	0.004***	-0.001	0.002
	[0.002]	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]
Roaddens (Ln)	0.000	0.065***	0.094***	0.032	0.021	1.687**
	[0.020]	[0.023]	[0.027]	[0.033]	[0.036]	[0.664]
Distport (Ln)	0.526***	-1.152**	-3.048**	-0.512	0.024	-2.440
	[0.069]	[0.464]	[1.462]	[0.476]	[1.206]	[1.523]
Avregindemp (Ln)	0.114***	0.005	0.001	0.009	0.005	-0.026**
	[0.010]	[0.004]	[0.005]	[0.006]	[0.009]	[0.011]
Locplant (Ln)	-0.013	0.066***	0.083***	-0.002	0.117***	0.116**
	[0.023]	[0.018]	[0.022]	[0.027]	[0.028]	[0.047]
Popdens (Ln)	0.041***	0.015	0.013	0.017	-0.117**	0.024
	[0.014]	[0.013]	[0.021]	[0.016]	[0.051]	[0.016]
Diversity (Ln)	-0.008	-0.001	0.005	-0.010	-0.015	0.013
	[0.010]	[0.007]	[0.010]	[0.008]	[0.013]	[0.023]
Constant	-1.958***	10.665***	22.390**	7.281**	4.175	19.440*
	[0.624]	[3.034]	[9.399]	[3.341]	[7.843]	[9.871]
Industry-Year Dummies	Y	Y	Y	Y	Y	Y
Plant Fixed Effects	N	Y	N	N	Y	N
N x T	441,187	441,187	360,163	81,024	116,012	66,699
R^2	0.376	0.054	0.073	0.083	0.086	0.124

Notes: Robust standard errors for correcting at the industry-district level are reported in brackets. Significance levels: * p <0.10, ** p <0.05, *** p <0.01..

also found the GIS-Euclidean distance that measured the distance from district capital to the seaport as a qualified approximation of transport costs and travel time, despite not considering the quality and availability of the network. The statistically significant and consistently negative estimated coefficients demonstrate that, as the distance to the international seaports increase, so do the travel time and costs.

5.2 Agglomeration Externalities over Economic Stages

In this section, we present our analysis of the disaggregated data to investigate the effects of agglomeration on plant productivity across economic stages. The plants were categorised as 'small' (20–49 employees), 'medium' (50–249), and 'large' (250+). The classification of plants according to size is important for knowing which group is more adapted to economic change (crisis) as part of the firm and industrial life cycles in acquiring agglomeration externalities (Neffke et al. 2011). Since small firms may spend a lower sunken cost of investment, they may be more flexible to enter and exit from the market, and usually, their ages are typically young. On the contrary, large firms are more established and are classified as mature industries with large sunken costs. In terms of plant size, a specific pattern was observable among the small, medium, and large plants as presented in Table 3. Considerable differences in agglomeration effects were present related to plant-size heterogeneity.

As shown in the table, small plants were flexible – they exhibited the capacity of adjustment and dynamic behaviour as a reaction to the economic situations. During pre-crisis, all plants seem to not have benefited from agglomeration economies except the large plants that received a small benefit from diversity. At later stages, however, the agglomeration sources of localisation economies have appeared in post-crisis for all plant sizes, in addition to urbanisation effects for small and medium plants.

Evidence of changing industrial structure was also found, where small and medium plants have gathered urbanisation externalities in the post-crisis period. It shows that small and medium plants have an advantage in the diversity of the environment across industries in the whole region and are more likely to have much stronger productive advantages in large cities. Meanwhile, large manufacturing plants are inclined to gather small external economies from localisation, benefiting from Marshallian externalities such as labour pooling, knowledge spillover, and input sharing in that period. Strong agglomeration effects on small plants' productivity after the crisis are echoed by Aswicahyono et al. (2010), who found that small plants were the single contributor to employment growth while recording a robust growth at around 8.8% between 1996 and 2006.

Interestingly, at pre-crisis, the urbanisation economies have had different effects on medium and large plants. Table 3 also depicts a significant negative impact of urbanisation economies against medium plants, which signifies that the de-clustering for medium plants might be the results of labour cost, congestion, or institutional costs in large areas. It may be due to the relatively lower scale of economies of medium plants against the cost of urbanisation. However, it does not make a case for a large plant that acquired the benefit of diversity from other sectors in a region.

The agglomeration effects on productivity were examined further through industry grouping. The industries from the three-digit SIC were divided into six groups of industry following Henderson et al. (2001): (a) traditional, (b) heavy, (c) transportation equipment, (d) machinery and electronics, (e) high-technology, and (f) other industries. Externalities of labour pooling were assumed to occur between plants sharing the same two-digit SIC and region. Detailed information about the industrial grouping of industries is provided in the Appendix (Table A.3).

Table 4 shows the different effects of economic stages across industry types where traditional industries, machinery, and electronics were more productive in a localised area and absorbed external benefits from localisation economies. In the post-crisis period, localisation economies persistently benefited traditional industries (such as food and beverage, wood and furniture, and tobacco). It is evident that a specialised environment might be preferable to these labour-intensive, resource-based, and typical industries (as they are described in the OECD classification; OECD 1987). In addition, traditional industries have also acquired benefits from urbanisation economies.

The fact that the location of mature firms is attractive to new plants further supports this finding, as it provides information about the most suitable area compared to others with comparable situations (Henderson, Kuncoro 1996). Infrastructure improvements and localisation economies also strongly inform firms' decisions in Indonesia over plant location and other activities, as pointed out by Deichmann et al. (2005). In a similar vein, Amiti, Cameron (2007) observed that the firms benefited from at least two of the three agglomeration sources, i.e., labour-market pooling and input sharing. Their findings imply that Indonesia's localisation economies come to light by observing the inter-firm interaction in supply and demand relations.

Additionally, transport equipment was found as the only industry that obtained negative externalities from urbanisation in period preceding the crisis – that said, external benefits emerging out of this industry's agglomeration economies were weakened by the crisis. Through this finding, it was possible to explain the properties of the industries receiving more external costs from large areas and diversified environments. The failure of transport industries to absorb external benefits from any agglomeration economies source is contrary to the conclusions from Lee et al. (2010) and Henderson et al. (2001). They studied the same industry in Korea and found that it sourced external benefits from localisation. The Korean case shows that the transport industry is composed of businesses operating in specialised and concentrated areas. Also important to note is the finding

Table 3: Agglomeration Externalities by Plant Size over Economic Cycles

Dependent Variable:	Total Factor Productivity (Ln TFP)	
Economic Cycles	Pre-Crisis (1990–98)	Post-Crisis (1999–2010)
Small Firm (20–49 Workers)		
N x T	89,938	146,997
Locplant (Ln)	0.030	0.101***
	[0.025]	[0.030]
Popdens (Ln)	0.006	0.056***
	[0.036]	[0.018]
Diversity (Ln)	0.001	0.004
	[0.009]	[0.013]
Medium Firm (50–249 Workers)		
N x T	54,028	84,070
Locplant (Ln)	0.010	0.106***
	[0.031]	[0.035]
Popdens (Ln)	-0.090*	0.045**
	[0.052]	[0.018]
Diversity (Ln)	-0.016	-0.014
	[0.013]	[0.012]
Large Firm (≥ 250 Workers)		
N x T	26,119	40,035
Locplant (Ln)	0.050	0.082*
	[0.041]	[0.044]
Popdens (Ln)	0.030	-0.027
	[0.073]	[0.031]
Diversity (Ln)	0.031*	-0.011
	[0.017]	[0.016]
All Firm		
N x T	170,085	271,102
Locplant (Ln)	0.046**	0.101***
	[0.020]	[0.024]
Popdens (Ln)	0.026	0.033**
	[0.029]	[0.013]
Diversity (Ln)	0.003	0.002
	[0.008]	[0.010]

Notes: Estimations include fixed effects at the plant level and industry-year dummies. Each regression includes control for the plant's characteristics of age, size, dummies of ownership (DFDI, DGov), and export activity, and regional characteristics of coastal area, access to electricity, road density, and distance to the closest international port. Robust standard errors for correcting at the industry-district level are reported in brackets. Significance levels: * p <0.10, ** p <0.05, *** p <0.01.

Table 4: Agglomeration Externalities by Industry over Economic Cycles

Dependent Variable:	Total Factor Productivity (Ln TFP)	
Economic Cycles	Pre-Crisis (1990–98)	Post-Crisis (1999–2010)
Traditional Industries		
N x T	106,848	178,582
Locplant (Ln)	0.041*	0.112***
	[0.025]	[0.028]
Popdens (Ln)	0.015	0.023*
	[0.037]	[0.013]
Diversity (Ln)	0.002	0.014
	[0.010]	[0.009]
Heavy Industries		
N x T	42,724	64,891
Locplant (Ln)	0.037	0.058
	[0.040]	[0.055]
Popdens (Ln)	0.031	0.072***
	[0.060]	[0.027]
Diversity (Ln)	0.016	-0.028
	[0.014]	[0.029]
Transport Industries		
N x T	4,634	7,090
Locplant (Ln)	0.109	-0.137
	[0.094]	[0.131]
Popdens (Ln)	0.089	-0.143*
	[0.173]	[0.076]
Diversity (Ln)	-0.008	-0.113
	[0.043]	[0.076]
Machinery and Electronic Industries		
N x T	9,056	8,952
Locplant (Ln)	0.077	0.206*
	[0.067]	[0.124]
Popdens (Ln)	0.155*	0.047
	[0.082]	[0.099]
Diversity (Ln)	-0.042	0.003
	[0.028]	[0.037]
High-Technology Industries		
N x T	1,486	3,156
Locplant (Ln)	-0.452**	0.222
	[0.198]	[0.306]
Popdens (Ln)	0.497	-0.066
	[0.326]	[0.077]
Diversity (Ln)	0.042	-0.031
	[0.089]	[0.104]
Other Industries		
N x T	5,337	8,431
Locplant (Ln)	0.103	0.278***
	[0.072]	[0.085]
Popdens (Ln)	-0.272**	0.061
	[0.118]	[0.094]
Diversity (Ln)	-0.039	0.047
	[0.036]	[0.034]

Notes: Estimations include fixed effects at the plant level and industry-year dummies. Each regression includes control for the plant's characteristics of age, size, dummies of ownership (DFDI, DGov), and export activity, and regional characteristics of coastal area, access to electricity, road density, and distance to the closest international port. Robust standard errors for correcting at the industry-district level are reported in brackets. Significance levels: * p <0.10, ** p <0.05, *** p <0.01.

that the high-technology industry gained negative externalities or faced deagglomeration economies before the crisis; the sector started to benefit from agglomeration economies, though it is insignificant. Accordingly, we found the different result to that of Henderson (2003), who observed positive effects and significance in the industry in the United States concerning productivity.

Especially during the post-crisis period, the existence of localisation economies generally dominated that of urbanisation economies. The varied effects of agglomeration economies across industrial groups in relation to the shifting economic circumstances should be a revelation for policymakers to help them in formulating the applicable policies. Localisation economies positively and strongly affected productivity in resource-based industries such as traditional industries (e.g., apparel, food and beverage, paper, tobacco, textile, leather, furniture, and wood) and other industries (e.g., waste and recycling), and moderately improved productivity in machinery industries (such as machinery, electrical motor, wire, cable, battery) during the post-crisis period. In contrast, we found heavy industries to be more productive in a diversified environment created by urbanisation economies. Additionally, the productivity of traditional industries is also considerably affected by urbanisation economies, showing the importance of the size of the population for this sector as a market target.

For heavy and traditional industries, we looked at the agglomeration economies by plant size across an economic cycle in order to learn the changing industrial structure of small-sized plants. Both sectors make up 89.1% of the total number of observations in our study period and thus are the two biggest industry groups. As shown in Table 5 below, small plants in both industries were behind the industry structure change from no benefits of agglomeration to capture both localisation and urbanisation economies during post-crisis. The results imply that the larger the plant size, the smaller the magnitude of agglomeration benefits from both sources. In particular, for large plants in traditional industries, benefits came only from localization economies. Furthermore, the larger the size of plants in heavy industries, the effect of urbanisation become insignificant.

The ability to capture agglomeration sources of small plants in heavy and traditional industries was due to a number of developments. Firstly, after the credit rationing in the recovery periods, small plants had limited access to finance (Aswicahyono et al. 2010), which might have led small plants to switch their strategy from relying on diverse industries in a region to taking advantage of a specialised environment created by similar industries. To reduce production costs, smaller plants tapped into input sharing, knowledge transfer, and labour pooling in an industry. Aswicahyono et al. (2010) added that the crisis might have increased an entry barrier, which resulted in a higher exit rate compared to the entry rate. As a consequence, a reduced entry rate made urbanisation economies appear on a smaller scale. At the same time, the surviving plants had matured during the post-crisis phase and mostly moved to specialised areas while receiving the perks of localisation economies.

Second, these results are consistent with the findings of Khoirunurrofik (2018) that younger and smaller industries within Indonesian manufacturing can grow faster in diversified cities due to the competition that creates pressure for firms to innovate to survive. As Duranton, Puga (2001) describe it in the industry life cycle theory, urbanisation economies usually suit new-entry and small plants that considerably rely on their external environments during their early days; this is referred to as a 'nursery city', where an urban area provides a diversified environment for productivity growth.

Additionally, to some classes of firms our finding explains the importance of the diversity that major agglomerations afford, particularly for traditional industries, and would perhaps constitute a natural part of Jacobs externalities alongside the traditional pairing of localisation and urbanisation economies. It can be argued that a diversified economy may facilitate an inter-industry knowledge environment that supports the sustainability of firms by diversifying their products. As the economic stage enters the recovery phase, the industrial diversity would mean that firms could operate in more stable demand conditions with a wide choice of inputs that reduce the revenue risk and operational cost due to external shocks and price fluctuations (Neffke et al. 2011, Potter, Watts 2011). Likewise, our finding is in line with those of Brown, Greenbaum

Table 5: Agglomeration Externalities by Plant Size over Economic Cycles for Traditional and Heavy Industries

Dependent Variable:	Total Factor Productivity (Ln TFP)			
Industry Groups	Traditional Industries		Heavy Industries	
Economic Cycles	Pre-Crisis (1990–98)	Post-Crisis (1999–2010)	Pre-Crisis (1990–98)	Post-Crisis (1999–2010)
Small Firm (20–49 Workers)				
N x T	57,510	101,745	21,935	32,674
Locplant (Ln)	0.007	0.105***	0.028	0.096
	[0.032]	[0.034]	[0.050]	[0.069]
Popdens (Ln)	0.016	0.051***	0.000	0.099**
	[0.046]	[0.019]	[0.059]	[0.042]
Diversity (Ln)	-0.001	0.024**	0.015	-0.033
	[0.012]	[0.012]	[0.015]	[0.032]
Medium Firm (50–249 Workers)				
N x T	31,713	50,520	15,091	23,254
Locplant (Ln)	0.007	0.120***	0.003	0.020
	[0.040]	[0.043]	[0.065]	[0.075]
Popdens (Ln)	-0.099	0.042**	-0.084	0.065*
	[0.066]	[0.020]	[0.104]	[0.033]
Diversity (Ln)	-0.020	-0.013	0.016	-0.038
	[0.016]	[0.013]	[0.028]	[0.034]
Large Firm ($\geq$ 250 Workers)				
N x T	17,625	26,317	5,698	8,963
Locplant (Ln)	0.054	0.124**	0.078	-0.050
	[0.050]	[0.048]	[0.082]	[0.114]
Popdens (Ln)	-0.037	-0.052	0.244	0.027
	[0.080]	[0.035]	[0.183]	[0.084]
Diversity (Ln)	0.032	-0.005	0.016	-0.009
	[0.021]	[0.017]	[0.027]	[0.040]

Notes: Estimations include fixed effects at the plant level and industry-year dummies. Each regression includes control for the plant's characteristics of age, size, dummies of ownership (DFDI, DGov), and export activity, and regional characteristics of coastal area, access to electricity, road density, and distance to the closest international port. Robust standard errors for correcting at the industry-district level are reported in brackets. Significance levels: * p <0.10, ** p <0.05, *** p <0.01.

(2017) that during periods of economic downturn, counties in United States with a more diverse industry structure performed better compared to more concentrated counties performing well in recovery phases or 'periods of prosperity.' Finally, the power of small plants in adapting to economic conditions indicated that the life cycles of Indonesian industries occurred, at least for small and medium sized plants along with various types of dynamic agglomeration economies, precisely diversity, in accordance with localisation and urbanisation.

6 Conclusion

The central finding of this study is that both agglomeration sources – localisation and urbanisation economies – co-exist, which both have a positive effect on plant-level productivity. At large, the localisation effects are stronger than urbanisation effects. It is supported by the fact that, even after more than 20 years since the financial crisis hit the nation in 1998, there are small significant changes in the concentration of economic activity across the country's main islands.

Looking at the effects of agglomeration economies on plants of different sizes and plants in different industries, the present study discovered that for small and medium plants, the effects of localisation are stronger than those of urbanisation in post-crisis. This finding sheds some light on the nature of agglomeration economies and suggests that the sources of agglomeration experienced a shift in accordance with changing economic stages. We also demonstrated that the positive externalities of agglomeration economies on productivity are the response of a plant to benefit more from within an industry and a region.

In terms of industrial groups and plant-size heterogeneity, productivity is improved by localization economies for all plant sizes, in addition to traditional, machinery, electronics, and "other" manufacturing sectors. Meanwhile, productivity in a heavy industry is enhanced by urbanisation economies. Furthermore, the productivity of small and medium-sized plants is demonstrably enhanced by both agglomeration sources, but not for large-sized plants. These differential effects are compelling, and the results differ considerably depending on which of the various subsamples are used, although it is quite difficult to discern any clear patterns in the differential effects. Moreover, the breakdown estimation across economic cycles suggests an adjustment and change of agglomeration magnitudes and sources. After the financial crisis of 1997–1998, the agglomeration externalities have demonstrated themselves to be in favour of localisation economies for productivity and advise that the industry has, to a certain extent, undergone a structural change for seizing benefits from external economies.

It is conceivable that some limitations might have influenced the results obtained. As we identified a spatial autocorrelation of productivity in cross-border region beyond the administrative boundary[1], further studies on spatial scope of externalities are important to investigate the impacts of agglomeration with the attention of spatial and temporal variation as suggested by De Groot et al. (2016). In addition, as the data is a survey and not census, we do not know the exact number of entry and exit firms. Therefore, we were not able to measure the precise level of competition for each industry which is enormously important for agglomeration economies.

Acknowledgments

The author is grateful for the 2020 PUTI Q2 Grant-Universitas Indonesia's support and to the editor and anonymous reviewer(s) for their valuable suggestions.

References

Amiti M, Cameron L (2007) Economic geography and wages. *The Review of Economics and Statistics* 89: 15–29. CrossRef.

Arnold JM, Javorcik BS (2009) Gifted kids or pushy parents? foreign direct investment and plant productivity in Indonesia. *Journal of International Economics* 79: 42–53. CrossRef.

Aswicahyono H, Hill H, Narjoko DA (2010) Industrialisation after a deep economic crisis, Indonesia. *Journal of Development Studies* 46: 1084–1108. CrossRef.

Beaudry C, Schiffauerov A (2009) Who's right, marshall or jacobs? the localisation versus urbanization debate. *Research Policy* 38: 318–337. CrossRef.

Brown L, Greenbaum RT (2017) The role of industrial diversity in economic resilience: An empirical examination across 35 years. *Urban Studies* 54: 1347–1366. CrossRef.

Cameron AC, Gelbach JB, Miller DL (2011) Robust inference with multi-way clustering. *Journal of Business & Economic Statistics* 29: 238–249

Day J, Ellis P (2013) Growth in Indonesia's manufacturing sectors: Urban and localisation contributions. *Regional Science Policy and Practice* 5: 343–368. CrossRef.

De Groot HL, Poot J, Smit MJ (2016) Which agglomeration externalities matter most and why? *Journal of Economic Surveys* 30: 756–782

[1] Morans' I and LM diagnostic for residual spatial autocorrelation test on the relationship between lntfp and agglomeration variables is statistically significant.

Deichmann U, Kaiser K, Lall SV, Shalizi Z (2005) Agglomeration, transport and regional development in Indonesia. Word bank policy research working paper 3477. CrossRef.

Duranton G, Puga D (2001) Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review* 91: 1454–1477. CrossRef.

Firman T, Kombaitan B, Pradono P (2007) The dynamics of Indonesia's urbanisation, 1980-2006. *Urban Policy and Research* 25: 433–454. CrossRef.

Frenken K, van Oort F, Verburg T (2007) Related variety, unrelated variety and regional economic growth. *Regional studies* 41: 685–697. CrossRef.

Gill IS, Goh C (2010) Scale economies and cities. *World Bank Research Observer* 25: 235–262. CrossRef.

Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. *Journal of Political Economy* 100: 1126–52. CrossRef.

Graham DJ (2009) Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science* 88: 63–84. CrossRef.

Henderson JV (2003) Marshall's scale economies. *Journal of Urban Economics* 53: 1–28. CrossRef.

Henderson JV, Kuncoro A (1996) Industrial centralization in Indonesia. *The World Bank Economic Review* 10: 513–540. CrossRef.

Henderson JV, Lee T, Lee YJ (2001) Scale externalities in korea. *Journal of Urban Economics* 49: 479–504. CrossRef.

Jacob J, Meister C (2005) Productivity gains, technology spillovers and trade: Indonesian manufacturing, 1980-96. *Bulletin of Indonesia Economic Studies* 41: 37–56. CrossRef.

Khoirunurrofik (2018) Local economic structure, productivity growth, and industry life cycle: Evidence from Indonesia. *Asia Pacific Journal of Regional Science* 2: 453–475. CrossRef.

Kuncoro A (2009) Spatial agglomeration, firm productivity and government policies in Indonesia: Concentration and deconcentration in the manufacturing sector. *In H. Yukon & A. M. Bocchi (Eds.), Reshaping Economic Geography in East Asia, A Companion to the World Development Report 2009*

Lee BS, Jang S, Hong SH (2010) Marshall's scale economies and jacob's externality in korea: The role of age, size and the legal form of organization of establishments. *Urban Studies* 47: 3131–3156. CrossRef.

Levinsohn J, Petrin A (2003) Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70: 317–341. CrossRef.

Maré DC, Graham DJ (2013) Agglomeration elasticities and firm heterogeneity. *Journal of Urban Economics* 75: 44–56. CrossRef.

Marrocu E, Paci R, Usai S (2013) Productivity growth in the old and new Europe: The role of agglomeration externalities. *Journal of Regional Science* (3): 418–442. CrossRef.

Martin P, Mayer T, Mayneris F (2011) Spatial concentration and plant-level productivity in france. *Journal of Urban Economics* 69: 182–195. CrossRef.

Melo PC, Graham DJ, Noland RB (2009) A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics* 39: 332–342. CrossRef.

Narjoko D, Hill H (2007) Winners and losers during a deep economic crisis: Firm-level evidence from Indonesian manufacturing. *Asian Economic Journal* 21: 343–368. CrossRef.

Neffke F, Henning M, Boschma R, Lundquist KJ, Olander LO (2011) The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies* 45: 49–65. CrossRef.

Nichols A, Schaffer ME (2007) Clustered standard errors in stata. Paper presented at united kingdom stata users' group meeting

OECD (1987) Structural adjustment and economic performance. Paris: Oecd

Petrin A, Poi BP, Levinsohn J (2004) Production function estimation in stata using inputs to control for unobservables. *Stata Journal* 4: 113–123

Poczter S, Gertler P, Rothenberg AD (2013) Financial crisis and productivity evolution: Evidence from Indonesia. *World Economy* 37: 705–731. CrossRef.

Potter A, Watts HD (2011) Evolutionary agglomeration theory: Increasing returns, diminishing returns, and the industry life cycle. *Journal of Economic Geography* 11: 417–455. CrossRef.

Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson JV, Thisse JF (eds), *Handbook of Regional and Urban Economics (vol.* Elsevier, 4: 2119-2171). Amsterdam. CrossRef.

Van Beveren I (2012) Total factor productivity estimation: A practical review. *Journal of economic surveys* 26: 98–128. CrossRef.

Vial V (2006) New estimates of total factor productivity growth in Indonesian manufacturing. *Bulletin of Indonesian Economic Studies* 42: 357–369

A Appendix

A.1 Data cleaning process

1. Possible mistakes in data keypunching:
    - The constructed panel was adjusted for possible mistakes in data keypunching and inconsistencies in the input across firms or plants such as starting year of operation, different ISIC used, and the sum of the percentage of ownership.
    - By spotting the firm identifier, we examined the consistencies of imputing the information of similar firms. If we found inaccurate information, we made an adjustment to retain correct and consistent information.
    - Generating variables such as output, value-added, intermediate input, materials, and so on, we resorted to manual accounting for calculating those variables instead of using reported variables that may have contained mistakes due to typing errors
2. Missing observation and non-reporting items
    - These may be because some firms opted out of the survey or they exited the market because they downsized to less than 20 employees, and the firm no longer met the definition of a medium or large manufacturer.
    - To solve these problems, we estimated the cell value by conducting linear interpolation or an average of the value within a window of two consecutive years for certain variables.
    - This approach does not apply for missing observations in the beginning or end period of series since we do not know whether the firm still exists.
3. Duplicate observations
    - We found that a few observations had similar numbers for the main part of the variable set such as the number of employees, output, value-added, etc.
    - We suspected confidently that these double observations were due to the plants that belonged to a similar firm. The manufacturing survey asked for plant-level information. Therefore, for a multi-plant firm, the headquarters may have completed the questionnaire with the consolidated value of all the plants owned.
    - To account for this, we selected only one observation for these duplicate observations.

Finally, to generate a panel series with unique observations, we resorted to the following steps:

- Excluding East Timor as part of Indonesia.
- Removing the observations if it has zero values of a key variable such as input, output, value-added, and labour.
- Removing observation with repeated values of the key variables or similar PSID.
- Removing outlier observations that have productivity values of ratio between output to labour and value-added to labour were below the lowest (1 percentile) and higher than the highest (99 percentile).
- Removing observation for which capital stock cannot be estimated.

A.2 Additional Information

Table A.1: Variable Definition and Data Source

Variable	Label	Definition	Source
Dependent Variable			
Total factor productivity	TFP	Total factor of productivity using the Letvin-Petrin control function approach	Estimated from SI 1990–2010, BPS
Plant Characteristics			
Age	Age	Age of plant as a difference between the year production started and year of survey	SI 1990–2010, BPS
Size	Size	Number of workers	SI 1990–2010, BPS
Foreign ownership	DFDI	= 1 if foreign has at least 10% share of ownership	Constructed
Government ownership	DGov	= 1 if central or local government has at least 50% share of ownership	Constructed
Exporter	DEexp	= 1 if plant exports	Constructed
Regional Characteristics			
Coastal	Coastal	Percentage of villages located offshore in a district/city	PODES 1990–2011
Electricity	Electricity	Percentage of households that has access to electricity in a district/city	PODES 1990–2011
Road density	Roaddens	Length of road infrastructure per square kilometres in a province	BPS and Ministry of Home Affairs
Distance to intl. seaport	Distport	GIS distance from capital of district/city to capital of city where the closest international port is located	Constructed
Agglomeration Economies			
Localisation (plants)	Locplant	Own industry plant in the district/city (plants)	Calculated from SI 1990–2010, BPS
Average industry-region employment	Avrindregemp	Average industry employment in the district/city minus own plant (person)	Calculated from SI 1990–2010, BPS
Urbanisation (population)	Popdens	Employment density in the district/city	BPS
Diversity	Diversity	The diversity of the various industry in the district/city	Calculated from SI 1990–2010, BPS

Notes: BPS is the Indonesian Central Bureau of Statistics. SI is the Annual Survey of Large and Medium Firms. PODES is the Village Potential Survey.

Table A.2: Plants' Observation and Exit-Entry rate

Year	All N	All Entry	All Exit	All Entry rate (%)	All Exit rate (%)	Small N	Small Entry	Small Exit	Small Entry rate (%)	Small Exit rate (%)	Medium N	Medium Entry	Medium Exit	Medium Entry rate (%)	Medium Exit rate (%)	Large N	Large Entry	Large Exit	Large Entry rate (%)	Large Exit rate (%)
1990	15,562					8,651					4,832					2,079				
1991	15,885	2,910	2,117	18.32	13.33	8,383	1,639	1,485	19.55	17.71	5,071	909	501	17.93	9.88	2,431	362	131	14.89	5.39
1992	17,074	2,597	1,383	15.21	8.10	8,941	1,534	913	17.16	10.21	5,434	819	373	15.07	6.86	2,699	244	97	9.04	3.59
1993	17,543	1,940	1,401	11.06	7.99	8,821	1,161	952	13.16	10.79	5,825	622	335	10.68	5.75	2,897	157	114	5.42	3.94
1994	18,389	2,119	1,217	11.52	6.62	9,115	1,359	833	14.91	9.14	6,202	630	288	10.16	4.64	3,072	130	96	4.23	3.13
1995	20,853	3,587	1,118	17.20	5.36	11,076	2,684	766	24.23	6.92	6,585	780	267	11.85	4.05	3,192	123	85	3.85	2.66
1996	22,297	3,234	1,804	14.50	8.09	12,185	2,386	1,273	19.58	10.45	6,833	709	411	10.38	6.01	3,279	139	120	4.24	3.66
1997	21,718	1,733	2,294	7.98	10.56	11,632	1,075	1,701	9.24	14.62	6,825	523	435	7.66	6.37	3,261	135	158	4.14	4.85
1998	20,764	1,551	2,319	7.47	11.17	11,134	955	1,740	8.58	15.63	6,421	475	478	7.40	7.44	3,209	121	101	3.77	3.15
1999	21,410	1,400	979	6.54	4.57	11,378	982	813	8.63	7.15	6,686	334	138	5.00	2.06	3,346	84	28	2.51	0.84
2000	21,502	1,091	1,083	5.07	5.04	11,307	723	801	6.39	7.08	6,809	283	203	4.16	2.98	3,386	85	79	2.51	2.33
2001	20,724	3,698	4,568	17.84	22.04	10,712	2,448	3,181	22.85	29.70	6,635	911	994	13.73	14.98	3,377	339	393	10.04	11.64
2002	20,491	1,000	1,094	4.88	5.34	10,542	569	764	5.40	7.25	6,599	327	216	4.96	3.27	3,350	104	114	3.10	3.40
2003	19,716	950	1,657	4.82	8.40	9,992	622	1,145	6.22	11.46	6,425	263	361	4.09	5.62	3,299	65	151	1.97	4.58
2004	20,071	1,847	1,487	9.20	7.41	10,290	1,125	906	10.93	8.80	6,466	522	379	8.07	5.86	3,315	200	202	6.03	6.09
2005	20,057	1,485	1,360	7.40	6.78	10,354	946	899	9.14	8.68	6,482	433	320	6.68	4.94	3,221	106	141	3.29	4.38
2006	28,525	11,019	2,463	38.63	8.63	16,686	8,124	1,728	48.69	10.36	8,389	2,423	495	28.88	5.90	3,450	472	240	13.68	6.96
2007	27,205	1,243	2,663	4.57	9.79	15,832	896	1,978	5.66	12.49	7,958	261	522	3.28	6.56	3,415	86	163	2.52	4.77
2008	24,967	818	3,241	3.28	12.98	14,253	576	2,396	4.04	16.81	7,421	185	654	2.49	8.81	3,293	57	191	1.73	5.80
2009	23,781	491	1,881	2.06	7.91	13,336	328	1,449	2.46	10.87	7,183	118	339	1.64	4.72	3,262	45	93	1.38	2.85
2010	22,653	834	2,327	3.68	10.27	12,315	525	1,793	4.26	14.56	7,017	210	413	2.99	5.89	3,321	99	121	2.98	3.64
(1990-10)	441,187					236,935					236,935					66,154				
(1991-09)	402,972	44,713	38,456	10.56	9.02	215,969	30,132	26,031	13.95	12.05	126,249	11,527	7,621	9.13	6.04	60,754	3,054	2,687	5.03	4.42

Notes: Years of The Census of Manufacturing: 1996 and 2006

Table A.3: Group and ISIC3-Industry

Group	3 Digits-ISIC	Industry
Traditional	151	Meat, fish, fruit, vegetables, oils
	152	Dairy products
	153	Grain mill products, animal feeds
	154	Other foods
	155	Beverages
	160	Tobacco products
	171	Spinning, weaving & textile finish
	172	Other textiles
	173 & 174	Knitted, crocheted fab., articles, and Kapok
	181 & 182	Apparel and fur
	191	Leather tanning and products
	192	Footwear
	201	Wood saw milling and planning
	202	Wood product
	210	Paper and products
	221 & 222	Publishing and printing
	223	Media recording reproduction
	361	Furniture
	369	Jewelry, sports goods, games
	231 & 232	Coke oven and refined petroleum products
	241	Basic chemicals
	242	Industries other chemical products
	243	Manmade fibers
Heavy	251	Rubber products
	252	Plastic products
	261	Glass products
	262	Porcelain products
	263	Clay products
	264	Cement and lime products
	265	Marble and granite product
	266	Asbestos products
	269	Other nonmetallic products
	271	Basic iron and steel
	272	Basic precious, nonferrous
	273	Iron and steel smelting product
	289	Other fabricated metal products
	281	Structural metal products
Transportation	341	Motor vehicle assembly
	342	Motor vehicle bodies
	343	Motor vehicle components
	351	Building and repairing ships and boats
	352 & 353	Manufacture of railway and aircraft
	359	Motorcycle, bicycle, other
Machinery and Electronic	291	General purpose machinery
	292	Special purpose machinery
	293	Domestic appliances n.e.c.
	311	Electrical motors, generators, etc.
	312	Electrical distribution equipment
	313	Insulated wire, cable
	314	Batteries and cells
	315	Lamps and equipment
	319	Other electrical equipment n.e.c.
High-technology	300 & 321	Office, acc., computing machinery & electronic components
	322 & 323	TV and radio transmitters, and TV, radio, video equipment
	331	Medical, measuring equipment
	332& 333	Optical, photographic equipment, watches, and clocks
Other	371	Metal waste and scrap recycling
	372	Non-metal waste and scrap recycling

Table A.4: Plant-Level Production Function Estimation

3 Digits ISIC	Industry	OLS (Factor share) α	β	Levin Petrin Production Function α	β	$\alpha+\beta$	Wald test
151	Meat, fish, fruit, vegetables, oils	0.296	0.704	0.086	0.666	0.752	44.3***
152	Dairy products	0.105	0.895	0.169	1.074	1.243	3.04*
153	Grain mill products, animal feeds	0.252	0.748	0.247	0.600	0.847	7.64***
154	Other foods	0.195	0.805	0.135	0.850	0.985	0.3
155	Beverages	0.154	0.846	0.213	0.876	1.089	2.1
160	Tobacco products	0.171	0.829	0.146	0.848	0.994	0.0
171	Spinning, weaving & textile finish	0.200	0.800	0.131	0.620	0.751	49.04***
172	Other textiles	0.166	0.834	0.250	0.747	0.997	0.0
173&174	Knitted, crocheted fab., articles, and Kapok	0.162	0.838	0.172	0.750	0.922	4.57**
181&182	Apparel and fur	0.145	0.855	0.188	0.783	0.970	1.7
191	Leather tanning and products	0.173	0.827	0.245	0.859	1.105	1.8
192	Footwear	0.076	0.924	0.021	0.876	0.897	2.4
201	Wood saw milling and planning	0.170	0.830	0.104	0.720	0.823	22.66***
202	Wood product	0.190	0.810	0.119	0.817	0.936	4.38**
210	Paper and products	0.177	0.823	0.224	0.836	1.060	1.3
221&222	Publishing and printing	0.087	0.913	0.151	0.800	0.951	2.0
223	Media recording reproduction	0.112	0.888	0.604	0.733	1.337	1.1
231&232	Coke oven and refined petroleum products	0.036	0.964	0.273	0.793	1.067	0.1
241	Basic chemicals	0.203	0.797	0.137	0.788	0.925	1.3
242	Industries other chemical products	0.230	0.770	0.170	0.672	0.843	15.66***
243	Manmade fibers	0.129	0.871	0.494	1.111	1.605	4.00**
251	Rubber products	0.251	0.749	0.191	0.573	0.764	32.99***
252	Plastic products	0.198	0.802	0.222	0.739	0.961	3.08*
261	Glass products	0.086	0.914	0.595	0.818	1.413	7.8***
262	Porcelain products	0.269	0.731	0.323	0.605	0.928	0.3
263	Clay products	0.162	0.838	0.278	0.774	1.051	0.2
264	Cement and lime products	0.193	0.807	0.261	0.838	1.100	1.8
265	Marble and granite product	0.167	0.833	0.246	0.817	1.063	0.8
266	Asbestos products	0.077	0.923	0.186	1.032	1.218	0.4
269	Other nonmetallic products	0.196	0.804	0.313	0.808	1.121	0.2
271	Basic iron and steel	0.167	0.833	0.011	0.835	0.845	0.8
272	Basic precious, nonferrous	0.311	0.689	0.039	0.479	0.519	4.81**
273	Iron and steel smelting product	0.248	0.752	0.264	0.525	0.789	1.5
281	Structural metal products	0.126	0.874	0.116	0.978	1.094	2.6
289	Other fabricated metal products	0.202	0.798	0.247	0.671	0.918	6.81***
291	General purpose machinery	0.139	0.861	0.235	0.845	1.080	1.3
292	Special purpose machinery	0.221	0.779	0.175	0.686	0.861	6.11**
293	Domestic appliances n.e.c.	0.027	0.973	0.044	0.798	0.842	1.8
311	Electrical motors, generators, etc.	0.194	0.806	0.333	0.741	1.075	0.2
312	Electrical distribution equipment	0.175	0.825	0.344	0.809	1.153	1.0
313	Insulated wire, cable	0.030	0.970	0.077	0.872	0.949	0.1
314	Batteries and cells	0.113	0.887	0.267	1.008	1.276	3.65*
315	Lamps and equipment	0.124	0.876	0.158	0.616	0.774	0.8
319	Other electrical equipment n.e.c.	0.039	0.961	0.508	0.715	1.224	0.3
300&321	Office, acc., computing machinery & electronic components	0.190	0.810	0.122	0.595	0.718	2.83*
322&323	TV and radio transmitters, and TV, radio, video equipment	0.083	0.917	0.060	0.809	0.869	1.0
331	Medical, measuring equipment	0.113	0.887	0.245	0.841	1.086	1.7
332&333	Optical, photographic equipment, watches, and clocks	0.105	0.895	0.237	0.829	1.066	0.1
341	Motor vehicle assembly	0.035	0.965	0.935	0.600	1.535	0.2
342	Motor vehicle bodies	0.134	0.866	0.033	0.789	0.822	1.8
343	Motor vehicle components	0.214	0.786	0.291	0.740	1.030	0.0
351	Building & repairing ships & boats	0.225	0.775	0.213	0.854	1.067	1.3
352&353	Manufacture of railway & aircraft	0.540	0.460	0.826	0.673	1.498	0.3
359	Motorcycle, bicycle, other	0.126	0.874	0.200	0.797	0.997	0.0
361	Furniture	0.138	0.862	0.072	0.783	0.855	28.54***
369	Jewelry, sports goods, games	0.131	0.869	0.127	0.784	0.911	8.26***
371	Metal waste and scrap recycling	0.127	0.873	0.485	1.458	1.942	1.7
372	Non-metal waste & scrap recycling	0.080	0.920	0.326	0.808	1.134	0.4

Notes: α is the capital coefficient and β is the labor coefficient. Wald test of constant returns to scale is a test where the sum of the coefficients equals 1. Significance levels: * p <0.10, ** p <0.05, *** p <0.01.

Resources

REGION
The Journal of ERSA Powered by WU

Volume 7, Number 2, 2020, R1–R13
DOI: 10.18335/region.v7i2.326

journal homepage: region.ersa.org

A classification for English primary schools using open data

Stephen Clark, Nik Lomax, Mark Birkin[1]

[1] University of Leeds, Leeds, United Kingdom

Received: May 28 2020/Accepted: September 10 2020

Abstract. England has statutory regulations in place that ensure state funded schools deliver broadly the same curriculum. However, there still exists a wide range of contexts in which this education takes place, including: the management of schools; how schools choose to spend their budgets; individual policies in regards to staffing; behaviour and attendance; and perhaps most importantly, the composition of the pupil population. Given these factors, one outcome of interest is the attainment profile of schools, and it is important that this performance is judged in context, for the benefits of pupils, parents and schools. To this end, this study develops a classification using contemporary data for English primary schools. The open data used captures aspects of the gender, ethnic, language, staffing and affluence makeup of each school. The nature of these derived groupings is described and made available as a mapping resource. These groupings allow the identification of "families of schools" to act as a resource for fostering better collaboration between schools and more nuanced benchmarking.

1 Introduction

The learning that takes place in a child's early years is often cited as one of the most critical phases in their education (Bruce 2012, Nores, Barnett 2010, Sammons 2011). Therefore, parents understandably want to ensure that their child receives a good education, particularly at the start of their education experience. In England, parents are able to rank their choices of schools, not being limited to the closest (Burgess et al. 2006, Harris, Johnston 2008). However, selecting a school does not necessarily mean that their child will be allocated a place there, especially if the school is oversubscribed. However, in the primary phase of English education, covering ages up to 11, parents are often able to send their children to local schools (Burgess et al. 2011).

Not all primary schools are the same. They are shaped by the composition of their pupil intake (e.g. gender, ethnicity or deprivation) (Harris 2010) and the ethos of the school (Day et al. 2016). These characteristics can have an important impact on the performance of the pupils and the school. Thus, many authorities and parents are keen to benchmark schools, in particular in regards to their academic performance. The question then arises as to which schools to benchmark against. Commonly, the options are benchmarked against a pool of schools within the same administrative area, or all schools nationally. However, given the heterogeneity of schools, this comparison can be unfair or meaningless.

Therefore, this study aims to capture this diversity in the characteristics of mainstream primary schools in England and establish a grouping of such schools. This in turn allows for benchmarking against schools in the same group, or those in the same group but in close geographical proximity. This categorisation allows for a fairer assessment of schools' performance against their natural peers, which is critical if we are to ensure that the funding system does not favour schools purely on headline comparisons, penalising the ones that are performing better than headline attainment and progress statistics may suggest.

The mapping of these groupings of schools is available via this interactive map resource https://qgiscloud.com/tra6sdc/Map_QGISCloud/ with an accompanying guide in the Appendix. The map shows the neighbourhood contexts of schools in terms of the percentage of the non-White British or Irish living in the area and the rank of the degree of income deprivation affecting children (where 1 is the most deprived).

2 Capturing school heterogeneity

The issue of school effectiveness has received much academic attention. Whilst issues around leadership and teaching should not be neglected (Sammons et al. 2011, 2014), a common finding is that the socio-demographic and socio-economic compositions of a school's pupil population can have a big influence on its effectiveness and academic performance (Ainscow et al. 2016, Dustmann et al. 2010, Strand 2010, 2014).

2.1 Geodemographics

The method used here to develop a typology of English primary schools is a classification based on the characteristics of the schools. Such approaches are widely deployed in the field of geodemographics (Singleton, Spielman 2014), which attempts to classify neighbourhoods based on the characteristics of the people who live in the area (Gale et al. 2016) or work there (Cockings et al. 2015). However, such techniques are not limited to geographic areas; they can also be applied to other typologies such as individuals (Burns et al. 2017) or organisations (Phillip, Iyer 1975).

2.2 Groupings of schools

An early article by Bennett (1975) provides an introduction to an approach for categorisation, outlining many of the concepts needed to ensure a meaningful outcome. Dorabawila et al. (2002) classified schools in the Galle district of Sri Lanka into six groups, using information on school facilities and pupil performance. The authors commend the utility of their classification since it enables a fair distribution and targeting of funds to schools. A classification of French middle schools by Thaurel-Richard, Thomas (2006) used information on family socio-economic status, foreign national pupils, progress, attainment and the nature of the school to derive five classes of schools: urban privileged; under-privileged urban; small; under-privileged socially mixed and privileged socially mixed. Johnston et al. (2005) defined a grouping of English Secondary schools based on their ethnic composition and determined the membership into five groups as a function of: (1) the percentage of the White pupil population; and (2) the dominance of a non-White group. This approach placed schools into a grid based on the mono- or multi-ethnic nature of their pupil population. In perhaps the closest study to the work here, Gibbs et al. (2011) used ethnicity and deprivation data from London primary schools to identify 14 classes, which were then allocated into four groups defined by their relative position on a deprivation (well-off vs in-need) and ethnicity (White vs non-White) scale.

This examination into the literature has highlighted that there exists no up to date grouping of primary schools in England. What is available however is a range of databases that attempt to identify the closest "statistical neighbours" for schools based on their characteristics and performance (Education Endowment Foundation 2018, SchoolDash 2018). Such studies allow schools to benchmark against similar schools. However, it is left for schools to decide how extensive this search for statistical neighbours should be and when the comparisons become less valid. The approach proposed in this study allows schools to select statistical neighbours from a defined set of schools that share the same characteristics, recognising that "... geodemographic typologies are structured methods for making sense of the spatial and socioeconomic patterns [in schools].” (Harris et al. 2007, p. 556).

3 Data and Methods

The data used in this study are obtained from the Department for Education (2018a) and relates to the academic year from September 2018 to August 2019. The data are derived from a number of sources, primarily the annual census of schools and pupils that takes place in the Spring term (Department for Education 2018b).

The composition of the pupils attending the school forms a vital component of the data. These include demographic information about the number of pupils, their gender and ethnic backgrounds. Further information is also available, including the number of pupils eligible for free school meals; number with a statement of Special Educational Needs (SEN); the rate of authorised and unauthorised absences; and the number of pupils whose first language is not English. Information on the staff composition of schools is also available. Finally, there is a measure of deprivation of the schools’ catchments.

The data contain 24,952 schools. Not all these schools are appropriate for analysis, with an initial sub-set of 20,472 consisting of those that are designated as primary schools. Of these, 18,683 were open during the whole of the academic year 2018-2019. Some primary schools do not have cohorts covering the required ages, and restricting our sample to schools with starting ages of 2, 3, 4 or 5 and a highest age of 11, gives us 14,091 schools. Further elimination of Special, Independent and unknown school types leaves 13,443 mainstream primary schools for consideration in this study.

Table 1 lists the variables used to define the groups of primary schools. They fall into six sets: the ethnic composition of the schools’ pupils; the degree of classroom over-crowding; the staffing structure; the demand for Special Educational Needs SEN provision; the degree of absences; and finally, the deprivation. Most of these variables are expressed as a percentage of the pupil or staff population, whilst two are direct measures: the pupil teacher ratio and a population weighted child-centred deprivation measure taken from Ministry of Housing Communities and Local Government (2019).

For some schools these data items are based on small numbers suppressed in the supplied data tables for confidentiality reasons. The complete case analysis therefore involves 13,363 primary schools. These data have the advantage of being provided by a trusted source, and are all openly available.

The method used to establish the groupings of schools is the widely applied k-means approach (Everitt et al. 2001). This method attempts to form a given number of clusters of schools based on their similarity. This similarity is measured in how close schools are in the “data space”, i.e. the variables described above, from a cluster mean. A school is always allocated to the cluster whose mean is closest. However, the process of forming these clusters is iterative, with schools moving between clusters and cluster centres being updated until all schools are stable in their cluster. The quality of the final solution can be measured using a within-group-sum-of-squares, with better solutions having lower values.

This categorisation approach works best when the variables are uncorrelated, not skewed and are measured on a similar scale. The two variables, ‘percentage of White British or Irish’ and ‘percentage with English as first language’, have an absolute correlation above 0.75. Using the criteria that the variables that are, on average, highly correlated with the remaining variables should be removed, both the ‘percentage White British or Irish’ and the ‘percentage with English as first language’ are discarded as categorisation variables. Similarly, the percentage of pupils eligible for free school meals and the deprivation ranking are also correlated, and the free school meal measure is not included. The percentages of pupils who are boys and who are girls are also highly correlated and the percentage of female pupils is removed. In regards to skewness, Tukey’s ladder of powers transformation (Mosteller, Tukey Mosteller, Tukey) is used to correct for a positive skew in these data. For standardisation, a range standardisation is applied.

Similar to the approach used to derive 2001 and 2011 Output Area Classifications from the UK census data, a hierarchical approach to categorisation is adopted (Gale et al.

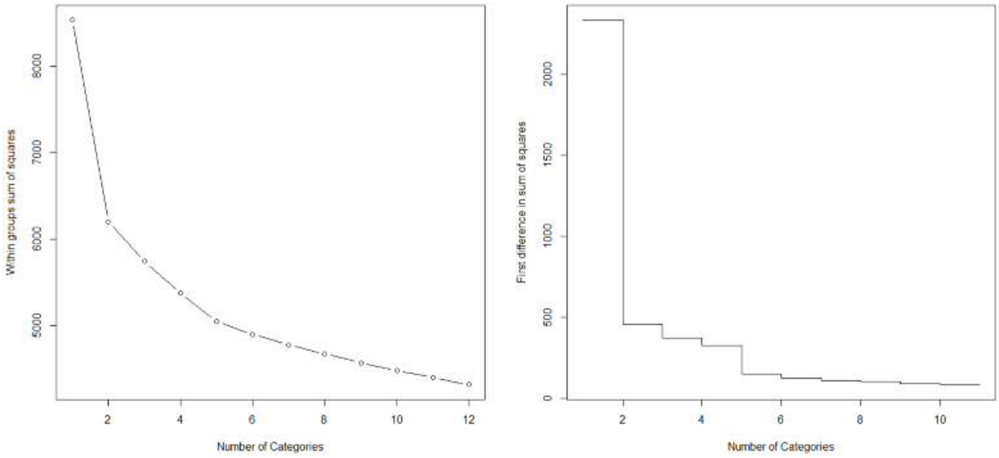


Figure 1: Scree plots of number of groupings and (a) within group sum of squares, (b) first difference in within group sum of squares

2016). Firstly, a number of Groups are formed, and then a further categorisation of the schools within each Group is undertaken, after re-transformation and re-standardisation, to define a series of Sub-groups. To gain an understanding of the nature of each Group and Sub-group, reference is made to the mean centres of each grouping, taken as averages of the variables for all schools in that group. These are presented on the raw scale, and on a scale that standardises each group centre relative to both all schools and schools within the same Group.

4 Results

Choosing the number of groupings using k-means is not an exact science, however, there is a range of methods which can support the decision-making process. The scree plots of the within-group-sum-of-squares for a given value of k and its first difference identify the point at which additional clusters do not materially reduce this measure of fit. In both these plots we are looking for an 'elbow' where the change in trajectory reduces or levels off. For the Group level of categorisation, the within-group-sum-of-squares are calculated using the k-means function in R (R Core Team 2017) with 100 random starting points, and plotted in Figure 1.

The scree plot of weighted sum of squares (Figure 1a) suggests that there are five groupings at this Group level, and this is confirmed by looking at the first differences (Figure 1b), where they level-off after moving beyond five groupings.

4.1 Category Groups

The raw group centres and the standardised versions of these centres are provided in Table 1 and Table 2 (scree and radial plots of this information are also provided in the supplementary material file SupplimentalScreeRadial.pdf). What is of interest for interpretation purposes are those variables where the group centre is particularly different from either all schools or schools in other groupings.

In these tables, it is clear that the nature of these groupings is most distinctly defined by the ethnic composition and the deprivation of the schools' pupils or catchments, and can be described as:

A: Multi-ethnic and Affluent : these schools have pupils coming from a range of ethnic backgrounds, but with the White British or Irish group still being dominant at nearly 70%. This is an affluent Group, having one of the largest rankings for

Table 1: Group centres on the raw scale

Variable / Group	A	B	C	D	E
Boys (%)	49.05	49.01	49.10	49.12	49.25
Girls (%) [1]	50.95	50.99	50.90	50.88	50.75
White British or Irish (%) [1]	68.41	91.01	92.09	83.91	30.41
White Other (%)	8.24	2.68	2.36	5.94	12.09
Traveller (%)	0.29	0.33	0.54	0.72	0.76
Mixed (%)	7.74	3.23	2.90	3.94	9.91
Indian (%)	4.15	0.39	0.19	0.48	5.38
Pakistani or Bangladeshi (%)	2.76	0.29	0.10	0.74	15.33
Other Asian (%)	1.86	0.27	0.17	0.60	3.75
Black (%)	3.26	0.40	0.26	1.78	15.45
Other (%)	2.14	0.48	0.31	1.13	5.67
Ethnicity unclassified (%)	1.14	0.91	1.07	0.74	1.26
First language is English (%) [1]	83.43	97.07	97.79	91.02	52.40
First language is not English (%)	16.37	2.83	2.09	8.89	47.33
First language is unclassified (%)	0.20	0.10	0.13	0.09	0.27
Pupils in classes of 31 to 35 with one teacher (%)	11.73	31.29	0.12	6.53	6.73
Pupils in classes of 36 or more with one teacher (%)	0.81	0.82	0.32	0.91	0.86
Pupil-Teacher ratio	21.80	22.19	18.82	20.64	20.28
Teaching staff (%)	46.74	46.22	48.77	42.79	43.71
Teaching Assistant (%)	33.21	33.71	31.11	37.31	35.48
Non-class based (%)	11.30	11.04	11.72	11.13	12.46
Auxiliary (%)	8.79	9.09	8.51	8.83	8.39
SEN pupils (%)	1.54	1.40	1.57	1.68	1.78
Authorised absence (%)	2.87	2.91	3.15	3.21	2.87
Unauthorised absence (%)	0.69	0.65	0.69	1.25	1.26
FSM pupils (%) [1]	13.59	13.15	13.90	36.06	33.79
Catchment IMD	21987	20884	21750	7384	8338

Notes: [1] Percentage of girls; percentage White British or Irish; percentage with English as first language; and percentage eligible for free school meals are not used in the categorisation but are reported here.

affluence. There are a large number of pupils in over-sized classes, suggesting these schools are popular with parents.

B: White British or Irish and Popular : this is the first of three Groups with a dominant White British or Irish ethnic grouping. These schools are very popular with parents, meaning that nearly a third of pupils are in classes with more than 30 pupils. Whilst not as affluent as some Groups, these schools are located in comfortable neighbourhoods.

C: White British or Irish and Affluent : This is another Group dominated by pupils of a White British or Irish ethnicity, but in contrast to the previous Group, there is no evidence of oversubscription from larger class sizes. These schools are also located in affluent neighbourhoods.

D: White and Deprived : in this Group the White British or Irish in combination with the White groups of other ethnic backgrounds form a large proportion, at nearly 90%. This Group is further differentiated by the level of deprivation, which is high, both from the perspective of the percentage of pupils that are eligible for free school meals and also the deprivation of the schools' neighbourhoods.

Table 2: Group centres on the standardised scale

Variable / Group	A	B	C	D	E
Boys (%)	0.9987	0.9980	0.9998	1.0002	1.0029
Girls (%) [2]	1.0013	1.0020	1.0002	0.9998	0.9972
White British or Irish (%) [2]	0.9488	1.2622	1.2772	1.1638	0.4217
White Other (%)	1.2981	0.4225	0.3713	0.9357	1.9041
Traveller (%)	0.5337	0.6117	1.0053	1.3453	1.4154
Mixed (%)	1.3776	0.5754	0.5160	0.7021	1.7631
Indian (%)	1.9093	0.1782	0.0895	0.2225	2.4738
Pakistani or Bangladeshi (%)	0.6580	0.0703	0.0246	0.1763	3.6547
Other Asian (%)	1.3435	0.1964	0.1239	0.4349	2.7012
Black (%)	0.7155	0.0880	0.0566	0.3908	3.3901
Other (%)	1.0490	0.2350	0.1520	0.5548	2.7774
Ethnicity unclassified (%)	1.1065	0.8778	1.0364	0.7134	1.2179
First language is English (%) [2]	0.9991	1.1624	1.1710	1.0899	0.6274
First language is not English (%)	1.0022	0.1733	0.1279	0.5444	2.8984
First language is unclassified (%)	1.2465	0.5991	0.7801	0.5641	1.7086
Pupils in classes of 31 to 35 with one teacher (%)	1.0731	2.8621	0.0109	0.5969	0.6158
Pupils in classes of 36 or more with one teacher (%)	1.1021	1.1141	0.4315	1.2304	1.1731
Pupil-Teacher ratio	1.0544	1.0733	0.9100	0.9982	0.9810
Teaching staff (%)	1.0236	1.0124	1.0681	0.9372	0.9573
Teaching Assistant (%)	0.9733	0.9878	0.9116	1.0933	1.0396
Non-class based (%)	0.9770	0.9538	1.0126	0.9616	1.0768
Auxiliary (%)	1.0104	1.0443	0.9773	1.0143	0.9638
SEN pupils (%)	0.9643	0.8733	0.9797	1.0485	1.1164
Authorised absence (%)	0.9581	0.9711	1.0487	1.0691	0.9557
Unauthorised absence (%)	0.7524	0.7169	0.7503	1.3646	1.3782
FSM pupils (%) [2]	0.6095	0.5897	0.6235	1.6170	1.5153
Catchment IMD	1.3801	1.3108	1.3652	0.4635	0.5234

Notes: [2] Percentage of girls; percentage of White British or Irish; percentage with English as first language; and percentage eligible for free school meals are not used in the categorisation but are reported here.

E: Multi-ethnic and Deprived : this final Group is the most multi-ethnic, with all ethnicities being present in large numbers, and the White British or Irish ethnic group comprising less than a third of pupils at the schools. There is also substantial deprivation associated with these schools, but less so than in Group D.

4.2 Category Sub-groups

The Groups presented above provide useful summary measures, but there is also some variation at a Sub-group level. The Sub-groups are constructed by applying k-means to just those schools in each group, following re-transformation and re-standardisation and using the methods outlined in Section 3. The scree plots, first difference in scree plots and tables of group centres are provided in the supplementary materials. Table 3 provides the number of schools in each sub-group along with information on how the Sub-groups differ within their Groups.

Table 3: Names for the Sub-groups

Code	Description	# schools
A	**Multi-ethnic Affluent**	**2402**
A.1	Comfortable	425
A.2	White Other	385
A.3	Affluent	448
A.4	Oversubscribed	442
A.5	Unclassified	364
A.6	Traveller	338
B	**White British/Irish & Over subscribed**	**2591**
B.1	White Other	414
B.2	Deprived	511
B.3	Affluent	624
B.4	Very oversubscribed	123
B.5	High absences	332
B.6	Comfortable	408
B.7	Unclassified	179
C	**White British/Irish**	**2873**
C.1	Girls	287
C.2	Unclassified	158
C.3	Traveller	267
C.4	Oversubscribed	227
C.5	Very White British/Irish	626
C.6	Low teachers	241
C.7	White Other	433
C.8	Deprived	634
D	**White British/Irish & Deprived**	**2461**
D.1	Comfortable	847
D.2	Unclassified	205
D.3	Very deprived	732
D.4	Oversubscribed	677
E	**Multi-ethnic Deprived**	**3036**
E.1	Black	558
E.2	Deprived	507
E.3	Indian	328
E.4	Oversubscribed	596
E.5	Very oversubscribed	207
E.6	Oversubscribed	492
E.7	Pakistan/Bangladesh	348

5 Geographic distribution

The regional distribution for each group is shown in Figure 2. This map reveals some significant spatial variations. There are few multi-ethnic schools in the North East and South West of England, whilst there are a sizeable number of such schools in the West Midlands and the South East. London stands out as particularly different to all the other regions. The two multi-ethnic Groups (A: Multi-ethnic Affluent and E: Multi-ethnic Deprived) dominate London schools, and the remaining mono-ethnic white groupings are very uncommon in London. A closer look at the distribution by London Boroughs in Figure 3 also reveals differences within London. Multi-ethnic schools dominate in the inner Boroughs of Tower Hamlets and Newham, whilst they are far less common in some outer Boroughs (Bromley and Barnet) where some white groupings are represented.

To furher illustrate the utility of this categorisation, an example map in Figure 4 is provided for the city of Derby. Each primary school is displayed by its Group and labeled with its Sub-group. The background maps show the Index of Multiple Deprivation (IMD) (the higher the rank, the higher the deprivation) and the percentage of the White British or Irish for the neigbourhood from the 2011 Census. This map shows that the

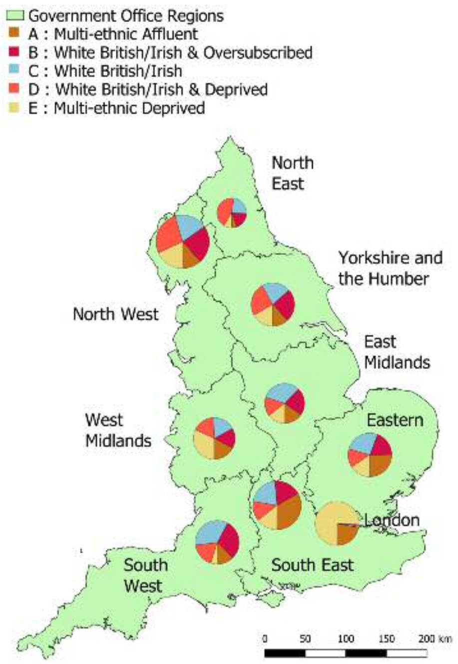


Figure 2: Geographic distribution of groupings by English region

more prosperous schools, in Groups A and B are located in areas with low deprivation ranks and that the multi-ethnic schools, Groups A and E are located in areas with lower percentages of the White British or Irish populations.

6 Discussion

This study has demonstrated that ethnicity is an important characteristic that differentiates primary schools in England. Within England, there are large areas within towns and cities with concentrations of particular ethnic groups, e.g. White British in rural towns, South Asians in ex-industrial northern towns, and Black populations in London. Given that the pupil catchments of primary schools are concentrated around their location, it is inevitable that such schools will have an intake of the dominant ethnic groups in their vicinity (this is especially the case in London, where Gibbs et al. (2011) note that "… *the vast majority of schools reflect the ethnic mix of their immediate neighbourhood.*". page 37-38). The affluence or deprivation of the school's pupil population is also similar, for concentrations of these measures in the locality of the school will dominate in the character of that school. These two aspects of schools are largely constrained by geography (Ainscow et al. 2016), however, the school can influence some of the other characteristics. A school may choose to employ more classroom teachers in preference over teaching assistants or non-classroom based staff – at a cost. This will then influence the composition of its staff and also whether pupils will be taught in large classes of 30 or more . However, some schools could struggle to manage an ideal staffing structure with overcrowding resulting from either the school being popular and having to take more pupils than its capacity allows or from the school being unable to attract enough teachers to provide a full complement of staff.

Another area over which a school has some control is the absence profile of pupils (Taylor 2012). Education welfare officers can be employed by schools to work with families whose children are failing to attend regularly. Schools can also fine parents for days that they take their children out of school. In reality, the percentage of authorised absences is similar between all the Groups and Sub-groups, but the percentage of un-authorised absences is larger in the groupings that are defined as challenged.

The relationship between this national categorisation of primary schools and the

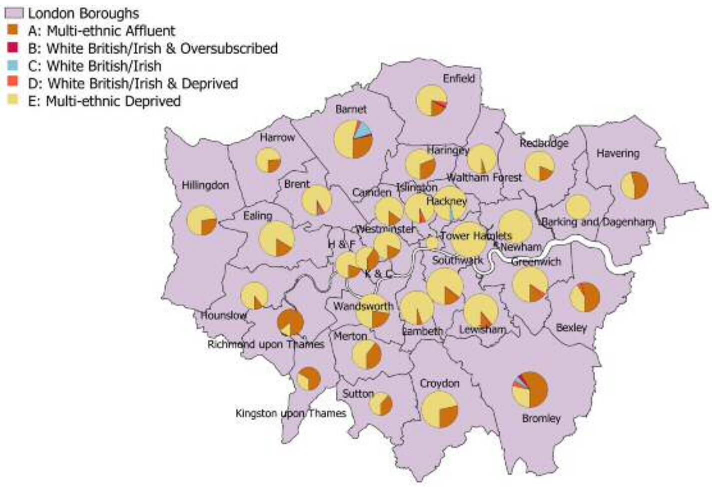


Figure 3: Geographic distribution of groupings by London Borough (H & F : Hammersmith and Fulham and K & C : Kensington and Chelsea)

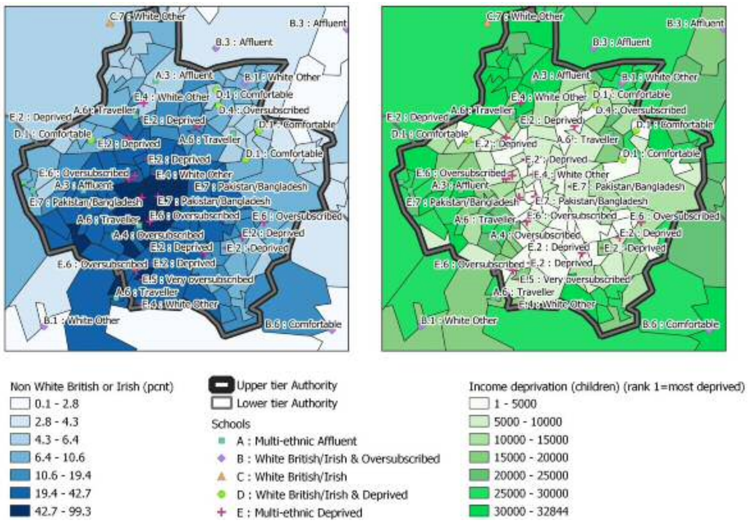


Figure 4: Categorisation of schools in Derby

London categorisation in Gibbs et al. (2011) is similar to that between the national OAC (Gale et al. 2016) and the London specific OAC (Greater London Authority 2017), the motivation for the latter being a desire to provide a categorisation that reflects the unique nature of London.

In this study a number of assumptions have been made. Amongst these is the use of the k-means classification technique in preference to other available methods. K-means is undoubtedly the most widely applied classification technique with applications in a wide number of domains. Another limitation has been our reliance on open data. This is however seen as a desirable feature for a number of reasons. Firstly, it allows the data to be shared with no restrictions, so that the work can be reproduced or extended. Secondly, it ensures that no information about any individual is disclosive. The decision not to include the performance of a school as a categorisation measure was also made. This assumption allows such measures to be treated as an output measure from the school, and there would undoubtedly be interest in how this independent measure varies amongst schools that otherwise look similar.

The circumstances of a school are not fixed over time. Transformations can take place in schools, for example, the ethnic or socio-economic compositions of a school population can change, being driven by how a school's catchment evolves as a result of population migration or as other neighbouring schools expand or contract their intake. A new leadership team or governance arrangement may also take the school in a different direction. In light of this, it will be useful to re-visit these groupings over time. Luckily, the information used in this study, derived mainly from the annual school census, will keep being generated, allowing for this resource to be maintained.

References

Ainscow M, Dyson A, Hopwood L, Thomson S (2016) Primary schools responding to diversity: Barriers and possibilities. Cambridge primary review trust, York

Bennett N (1975) Cluster analysis in education research: A non-statistical introduction. *Research Intelligence* 1: 64–70

Bruce T (2012) Early childhood education: Hachette UK. Technical report

Burgess S, Briggs A, McConnell B, Slater H (2006) School choice in England: Background facts. DP, 6, 159

Burgess S, Greaves E, Vignoles A, Wilson D (2011) Parental choice of primary school in England: What types of school do different types of family really have available to them? *Policy Studies* 32: 531–547

Burns L, See L, Heppenstall A, Birkin M (2017) Developing an individual-level geodemographic classification. *Applied Spatial Analysis and Policy* 11: 417–437. CrossRef.

Cockings S, Martin D, Harfoot A (2015) A classification of workplace zones for England and Wales (COWZ-EW). Southampton: University of Southampton

Day C, Gu Q, Sammons P (2016) The impact of leadership on student outcomes. *Educational Administration Quarterly* 52: 221–258. CrossRef.

Department for Education (2018a) Get information about schools. Retrieved from https://www.get-information-schools.service.gov.uk/

Department for Education (2018b) School census 2016 to 2017: Guide for schools and local authorities. Retrieved from https://www.gov.uk/government/publications/school-census-2016-to-2017-guide-for-schools-and-las

Dorabawila S, Samita S, Thattil R (2002) Classification of state schools in Sri Lanka: A multivariate approach. *Ceylon Journal of Science: Physical Sciences* 9: 102–109

Dustmann C, Machin S, Schönberg U (2010) Ethnicity and educational achievement in compulsory schooling. *The Economic Journal* 120: F272–F297. CrossRef.

Education Endowment Foundation (2018) Families of schools database. Retrieved from https://educationendowmentfoundation.org.uk/tools/families-of-schools-database/

Everitt B, Landau S, Leese M (2001) *Cluster analysis* (4th ed.). Arnold, London

Gale C, Singleton A, Bates A, Longley P (2016) Creating the 2011 area classification for output areas (2011 OAC). *Journal of Spatial Information Science* (12): 1–27. CrossRef.

Gibbs A, Stillwell J, See L (2011) Understanding ethnic sorting and school performance through a classification of primary schools in London. *Education Research Journal* 1: 31–45

Greater London Authority (2017) London output area classification. Retrieved from https://data.london.gov.uk/dataset/london-area-classification

Harris R (2010) Segregation by choice?: The debate so far. Centre for market and public organisation, University of Bristol

Harris R, Johnston R (2008) Primary schools, markets and choice: Studying polarization and the core catchment areas of schools. *Applied Spatial Analysis and Policy* 1: 59–84

Harris R, Johnston R, Burgess S (2007) Neighborhoods, ethnicity and school choice: Developing a statistical framework for geodemographic analysis. *Population Research and Policy Review* 26: 553–579. CrossRef.

Johnston R, Wilson D, Burgess S (2005) England's multiethnic educational system? a classification of secondary schools. *Environment and Planning A* 37: 45–62

Ministry of Housing Communities and Local Government (2019) The English indices of deprivation 2019. Technical report. Retrieved from https://www.gov.uk/government/-publications/english-indices-of-deprivation-2019-research-report

Mosteller F, Tukey J *Data analysis and regression: A second course in statistics*. Addison-Wesley, Reading, Mass.

Nores M, Barnett W (2010) Benefits of early childhood interventions across the world: (under) investing in the very young. *Economics of education review* 29: 271–282. CrossRef.

Phillip P, Iyer R (1975) Classification of community hospitals. *Health Services Research* 10: 349–368. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1241783

R Core Team (2017) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria

Sammons P (2011) Do the benefits of pre-school last? Investigating pupil outcomes to the end of key stage 2 (aged 11). In: Sylva K, Melhuish E, Sammons P, Siraj-Blatchford I, Taggart B (eds), *Early Childhood Matters: Evidence from the Effective Pre-school and Primary Education project*. Routledge, London and New York, 114–148

Sammons P, Davis S, Day C, Gu Q (2014) Using mixed methods to investigate school improvement and the role of leadership: An example of a longitudinal study in England. *Journal of Educational Administration* 52: 565–589

Sammons P, Gu Q, Day C, Ko J (2011) Exploring the impact of school leadership on pupil outcomes: Results from a study of academically improved and effective schools in England. *International Journal of Educational Management* 25: 83–101

SchoolDash (2018) Understand education. Retrieved from https://www.schooldash.com/-about.html#dashboards

Singleton A, Spielman S (2014) The past, present and future of geodemographic research in the United States and United Kingdom. *Professional Geographer* 66: 558–567. CrossRef.

Strand S (2010) Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement* 21: 289–314

Strand S (2014) School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11. *Oxford Review of Education* 40: 223–245

Taylor C (2012) Improving attendance at school. Retrieved from https://www.gov.uk/-government/publications/improving-attendance-at-school

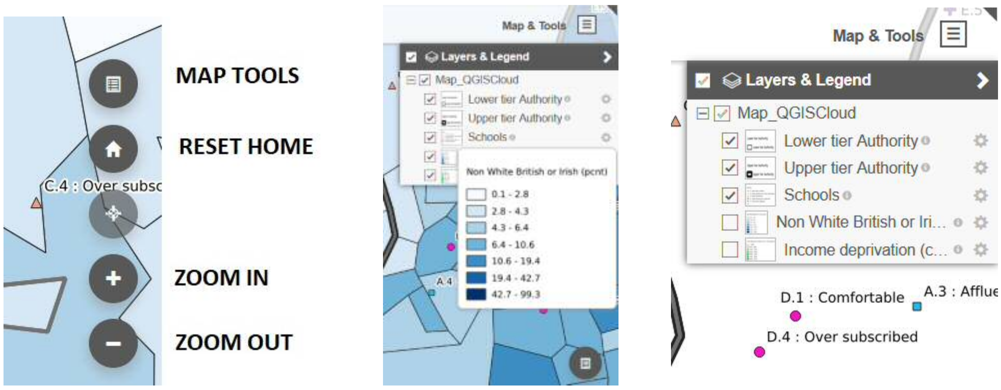
Thaurel-Richard M, Thomas F (2006) Typologies des colleges publics. *Donnees sociales - La societe francaise*: 147–156. Retrieved from https://www.insee.fr/fr/statistiques/-fichier/1372011/donsoc06p.pdf

A Appendix A: VIEWING OUR DATASET IN QGIS CLOUD

This short appendix introduces the functionality of QGIS Cloud in order to provide a guide to viewing our dataset.

A.1 *FUNCTION BUTTONS*

To the lower right hand side of the map there are five function buttons in a column (Figure A.1a).



(a) Function Buttons (b) Map Legends (c) Managing Layers

Figure A.1: Function buttons and Map tools

Clicking on the top button opens up the MAP TOOLS (more on these later). The next button down returns the map to the HOME view, this is the view displayed when the map was first loaded. The bottom two buttons zoom in and out of the map. It is also possible to pan around the map using the left mouse button; click on the map and, whilst keeping the mouse button pressed down, move the mouse to change the area of the map in view.

A.2 *MAP TOOLS*

The MAP TOOLS allow the user to see the legends that correspond to the displays on the map and to switch layers on and off.

A.2.1 Opening MAP TOOLS

To see the MAP TOOLS click on the MAP TOOLS button.

A.2.2 Map legends

To see the legend associated with each layer hover the mouse above the miniature image of the legend. This will temporarily display a larger version of the legend. As an example, Figure A.1b shows the legend for the percent White British or Irish layer of our resource. To remove the legend, move the mouse off the miniature image of the legend.

A.2.3 Managing layers

To the left of the name of the layer in the MAP TOOLS display there is a tick box. If this box contains a tick then the layer is displayed on the map. If the box is empty then the layer is not displayed. To toggle the display of the layer, click in this box. If the layer is on (shown with a tick) and this box is clicked then the layer will be turned off. If the layer if off (shown with an empty box) and this box is clicked then the layer will be turned on. Figure A.1c shows the map with the layers for the Percentage White British or Irish and Income Deprivation Rank turned off.

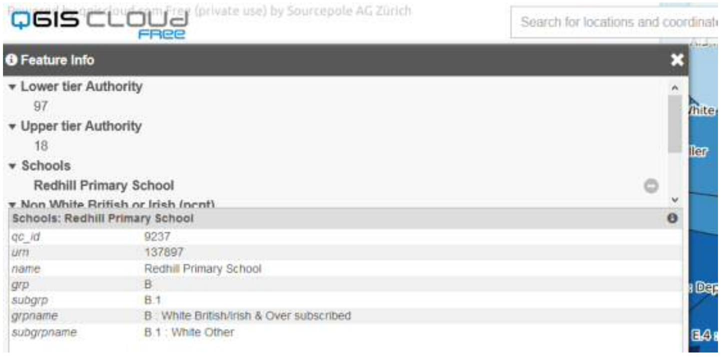
A.2.4 Closing MAP TOOLS

To close the MAP TOOLS click on the MAP TOOLS button.

A.3 IDENTIFYING FEATURES

To identify a feature in the map, click on (or very close to) the feature or within the area of the map of interest. To the left of the map a display will show information on the objects selected in the five layers.

To see more detail on an object, click on the object text. For example, clicking on the text "Redhill Primary School" provides the classification details on Redhill School.



Clicking on the index number 22501 for the Non White British or Irish (pcnt) shows that the percentage of non White British or Irish population in the lower level super output area that contains Redhill School is 4.35%.

A.3.1 Close the Information Screen

To close the information screen, click on the white cross in the top right hand corner.

REGION
The Journal of ERSA
Powered by WU

Volume 7, Number 2, 2020, R15–R46
DOI: 10.18335/region.v7i2.295

journal homepage: region.ersa.org

A reproducible notebook to acquire, process and analyse satellite imagery: Exploring long-term urban changes

Meixu Chen, Dominik Fahrner, Daniel Arribas-Bel, Francisco Rowe[1]

[1] University of Liverpool, Liverpool, UK

Received: 22 December 2019/Accepted: 1 December 2020

Abstract. Satellite imagery is often used to study and monitor changes in natural environments and the Earth surface. The open availability and extensive temporal coverage of Landsat imagery has enabled to monitor changes in temperature, wind, vegetation and ice melting speed for a period of up to 46 years. Yet, the use of satellite imagery to study cities has remained underutilised in Regional Science, partly due to the lack of a practical methodological approach to capture data, extract relevant features and monitor changes in the urban environment. This notebook offers a framework to demonstrate how to batch-download high-resolution satellite imagery; and enable the extraction, analysis and visualisation of features of the built environment to capture long-term urban changes.

Key words: satellite imagery, image segmentation, urbanisation, cities, urban change, computational notebooks

1 Introduction

Sustainable urban habitats are a key component of many global challenges. Efficient management and planning of cities are pivotal to all 17 UN Sustainable Development Goals (SDGs). Over 90% of the projected urban population growth by 2050 will occur in less developed countries (United Nations 2019). Concentrated in cities, this growth offers an opportunity for social progress and economic development but it also imposes major challenges for urban planning. Prior work on urbanisation has identified the benefits of agglomeration and improvements in health and education, which tend to outweigh the costs of congestion, pollution and poverty (Glaeser, Henderson 2017). Yet research has remained largely focused on Western cities (e.g. Burchfield et al. 2006), developing a good understanding of urban areas in high-income, developed countries (Glaeser, Henderson 2017). Much less is known about the long-term evolution of urban habitats in less developed countries. Analysis of historical census data exist exploring changes at discrete points over time such as slum detection (e.g. Giada et al. 2003, Kit, Lüdeke 2013, Kohli et al. 2016). Less applications can be identified tracking changes in urban settings over a continuous temporal scale (Ibrahim et al. 2020). This gap is partly due to the lack of comprehensive and consistent data sources capturing the long-term dynamics of urban structures in less developed countries.

Cities in Asia provide a unique setting to explore the challenges triggered by rapid urbanisation. The share of urban population in Asia is currently at a turning point transitioning to exceed the share of rural population. Currently Asia is home to over 53% of the urban population globally and the share of urban population is projected to increase to 66% by 2050 (United Nations 2019). Developing tools to monitor and understand the past and current urbanisation process is key to guide appropriate urban planning and policy strategies.

Recent technological developments can help overcome the paucity in spatially-detailed urban data in less developed countries. The combination of geospatial technology, cheap computing and new machine learning algorithms has ushered in an age of new forms of data, producing brand new data sets and repurposing existing sources. Satellite imagery represents a key source of information. Photographs from the sky have existed for decades, but their use in the context of socioeconomic urban research has been limited. Image data has been hard to process and understand for social scientists. Yet recent developments in machine learning and artificial intelligence have made images computable and turned these data into brand new information to be explored by quantitative urban researchers. Further, satellite data has become more abundant and openly accessible in the past decade, and offers new possibilities for data exploration through increasing spatial and temporal resolution. This, together with more computational power being available, allows to process these data in an efficient and meaningful way.

This notebook illustrates an easy-to-use analytical framework based on Python tools which enables batch download, image feature extraction, analysis and visualisation of high-resolution satellite imagery to capture long-term urban changes. Our purpose is to fill in the absence of a systematic and reproducible framework to acquire, process and analyse satellite imagery in urban built environment related to the field of Regional Science. The source of satellite data and administrative boundaries data are from NASA's Landsat satellite programme and ArcGIS Online. The Python libraries used in this notebook are the following:

- Landsat images in Google Cloud Storage: The Google Cloud Storage is accessed using an API to download Landsat imagery (version used: 0.4.9)
- Matplotolib:A Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.
- Numpy: Adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions
- Pandas: Provides high-performance, easy-to-use data structures and data analysis tools
- GeoPandas: Python library that simplifies working with geospatial data (version used: 0.6.2)
- Folium: Python library that enables plotting interactive maps using leaflet (version used: 0.10.0)
- Glob: Unix style pathname pattern expansion
- GDAL: Library for geospatial data processing (version used: 2.4.4)
- Landsat578: Simple Landsat imagery download tool
- L8qa: Landsat processing toolbox (version used: 0.1.1)
- Rasterio: Library for raster data processing (version used: 1.1.3)
- Scikit-image: Collection of algorithms for image processing
- Wget: Pure python download utility (version used: 3.2)
- OpenCV: Library for image processing
- scikit-learn: Machine learning in Python. Simple and efficient tools for data mining and data analysis.

We can import them all as follows:

[1]:
```
%matplotlib inline

#load external libraries
import matplotlib.pyplot as plt
from matplotlib import colors
import pandas as pd
import numpy as np
import geopandas as gpd
import folium
import os, shutil
import glob
import gdal
import wget
from landsat import google_download
from google_download import GoogleDownload
from l8qa.qa import write_cloud_mask
import rasterio
import rasterio as rio
from rasterio import merge
from rasterio.plot import show
from rasterio.mask import mask
from skimage import io,exposure, transform,data
from skimage.color import rgb2hsv, rgb2gray
from skimage.feature import local_binary_pattern
from sklearn.cluster import KMeans
import matplotlib.cm as cm
from sklearn import preprocessing
from rasterio.enums import Resampling
import seaborn as sns
import itertools

wdir= os.getcwd()
```

The remainder of this paper is structured as follows. The next section introduces the Landsat satellite imagery, study area Shanghai, and process on how to batch download and pre-process satellite data. Section 3 proposes our methods to extract different features including colour, texture, vegetation and built-up from imagery. Section 4 performs a clustering method on the extracted features, and section 5 interprets the results and gain insights from them. Finally, section 6 concludes by providing a summary of our work and avenues for further research using our proposed framework.

2 Data and Study Area

2.1 Landsat Imagery

We draw data from the NASA's Landsat satellite programme. It is the longest standing programme for Earth observation (EO) imagery (NASA 2019). Landsat satellites have been orbiting the Earth for 46 years providing increasingly higher resolution imagery. Landsat Missions 1-3 offer coarse imagery of 80m covering the period from 1972 to 1983. Landsat Missions 4-5 provides images of 30m resolution covering the period from 1983 to 2013 and Landsat Missions 7-8 are currently collecting enhanced images at 15m capturing Cirrus and Panchromatic bands, in addition to the traditional RGB, Near-, Shortwave-Infrared, and Thermal bands. The Landsat 6 mission was unsuccessful due to the transporting rocket not reaching orbit. Landsat imagery is openly available and offers extensive temporal coverage stretching for 46 years. Table 1 provides a summary overview of the operation, revisit time and image resolution for the Landsat programme, with other Earth observation satellite missions being shown in Table 2.

Additional Earth observation programmes exist. These programmes also offer freely accessible imagery at a higher resolution.

2.2 Study Area

In this analysis, we examine urban changes in Shanghai, China. Shanghai has experienced rapid population growth. Between 2000 and 2010, Shanghai's population rose by 7.4

Table 1: Overview of Landsat missions, their revisit time and spatial resolution

Mission	Operational time	Revisit time	Resolution
Landsat 1	1972-1978	18 d	80 m
Landsat 2	1975-1982	18 d	80 m
Landsat 3	1978-1983	18 d	80 m
Landsat 4	1983-1993	16 d	30 m
Landsat 5	1984-2013	16 d	30 m
Landsat 7	1999-present	16 d	15 m
Landsat 8	2013-present	16 d	15 m

Table 2: Overview of other Earth observation satellites, their revisit time and spatial resolution

Provider	Programme	Operational time	Revisit time	Resolution
European Space Agency	Sentinel	2015-present	5 d	10m
Planet Labs	Rapideye PlanetscopeSkysat	2009-present	4/5 d to daily	up to 0.8 m
NASA	Orbview 3	2003-2007	<3 d	1-4 m
NASA	EO-1	2003 -2017	–	10-30 m

million from 16.4 million to 23.8 million. It has an annual growth rate of 3.8 percent over 10 years. While the pace of population expansion has been less acute, Shanghai's population has continued to grow. In 2018, an estimated 24.24 million people were living in Shanghai experiencing a population expansion of approximately 8 million since 2010. The city is therefore a well suited example to explore long-term changes in urbanisation.

To extract satellite imagery, a first step is to identify the shape of the geographical area of interest. To this end, we use a polygon shapefile (https://www.arcgis.com/home/item.html?id=105f92bd1fe54d428bea35eade65691b). These polygons represent the Shanghai metropolitan area, so they include the city centre and surrounding areas. These polygons will be used as a bounding box to identify and extract relevant satellite images. We need to ensure the shapefile is in the same coordinate reference system (CRS) as the satellite imagery (WGS84 or EPSG:4326).

[2]:
```
# Specify the path to your shapefile
directory = os.path.dirname(wdir)
shp = 'shang_dis_merged/shang_dis_merged.shp'
```

[3]:
```
# Certify that the shapefile is in the right coordinate system, otherwise reproject
# it into the right CRS
def shapefile_crs_check(file):
    global bbox
    bbox = gpd.read_file(file)
    crs = bbox.crs
    data = crs.get("init", "")
    if 'epsg:4326' in data:
        print('Shapefile in right CRS')
    else:
        bbox = bbox.to_crs({'init':'epsg:4326'})
    f,ax = plt.subplots(figsize=(5,5))
    plt.title('Fig.1: Shapefile of Shanghai urban area',y= -0.2)
    bbox.plot(ax=ax)
```

[4]:
```
shapefile_crs_check(shp)
```

[4]:
```
Shapefile in right CRS
image/png<Figure size 360x360 with 1 Axes>
```

The world reference system (WRS) from NASA is a system to identify individual satellite imagery scenes using path-row tuples instead of absolute latitude/longitude

Figure 1: Shapefile of Shanghai urban area

coordinates. The latitudinal centre of the image corresponds to the row, the longitudinal centre to the path. This system allows to uniformly catalogue satellite data across multiple missions and provides an easy to use reference system for the end user. It is necessary to note that the WRS was changed between Landsat missions, due to a difference in swath patterns of the more recent Landsat satellites (NASA 2019). The WRS1 is used for Landsat missions 1-3 and the WRS2 for Landsat missions 4,5,7,8. In order to obtain path-row tuples of relevant satellite images for an area of interest (AOI), it is necessary to intersect the WRS shapefile (either WRS1 or WRS2, depending on the Landsat satellite you would like to obtain data from) with the AOI shapefile. The resulting path-row tuples will later be used to locate and download the corresponding satellite images from the Google Cloud Storage. The output of the intersection between WRS and AOI files can be visualised using an interactive widget. The map below shows our area of interest in purple and the footprints of the relevant Landsat images on top of an OpenStreetMap basemap.

[5]:
```
# Donwload the WRS 2 file to later intersect the shapefile with the WRS path/row
# tuples to identify relevant Landsat scenes
#
def sat_path():

    url = 'https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/
    ...s3fs-public/atoms/files/WRS2_descending_0.zip'
    # Create folder for WRS2 file
    if os.path.exists(os.path.join('Landsat_images','wrs2')):
        print('folder exists')
    else:
        os.makedirs(os.path.join('Landsat_images','wrs2'))

    WRS_PATH = os.path.join('Landsat_images','WRS2_descending_0.zip')
    LANDSAT_PATH = os.path.dirname(WRS_PATH)

    # The WRS file is only needed once thus we add this loop
    if os.path.exists(WRS_PATH):
        print('File already exists')
    # Downloads the WRS file from the URL given and unzips it
    else:
        wget.download(url, out = LANDSAT_PATH)
        shutil.unpack_archive(WRS_PATH, os.path.join(LANDSAT_PATH, 'wrs2'))
```

[6]:
```
%%time
# WARNING: this will take time the first time it's executed
# depending on your connection
sat_path()
```

[6]:
```
folder exists
File already exists
Wall time: 1e+03 mu s
```

[7]:
```
# Intersect the shapefile with the WRS2 shapefile to determine relevant path/row tuples
def get_pathrow():
    global paths,rows,path,row, wrs_intersection

    wrs=gpd.GeoDataFrame.from_file(os.path.join('Landsat_images','wrs2',
            'WRS2_descending.shp'))
    wrs_intersection=wrs[wrs.intersects(bbox.geometry[0])]
    paths,rows=wrs_intersection['PATH'].values, wrs_intersection['ROW'].values

    for i, (path,row) in enumerate(zip(paths,rows)):
        print('Image', i+1, ' -path:', path, 'row:', row)
```

[8]:
```
get_pathrow()
```

[8]:
```
Image 1  -path: 118 row: 38
Image 2  -path: 119 row: 38
```

[9]:
```
# Visualise the output of the intersection with the shapefile using Folium

# Get the center of the map
xy = np.asarray(bbox.centroid[0].xy).squeeze()
center = list(xy[::-1])

# Select a zoom
zoom = 8

# Create the most basic OSM folium map
m = folium.Map(location = center, zoom_start = zoom, control_scale=True)

# Add the bounding box (bbox) GeoDataFrame in red using a lambda function
m.add_child(folium.GeoJson(bbox.__geo_interface__, name = 'Area of Interest',
                      style_function = lambda x: {'color': 'purple', 'alpha': 0}))

loc = 'Fig 2.: Landsat satellite tiles that  cover the Area of Interest'
title_html = '''
             <figcaption align="center" style="font-size:12px"><b>{}</b></figcaption>
             '''.format(loc)
m.get_root().html.add_child(folium.Element(title_html))

# Iterate through each polygon of paths and rows intersecting the area
for i, row in wrs_intersection.iterrows():
    # Create a string for the name containing the path and row of this Polygon
    name = 'path: %03d, row: %03d' % (row.PATH, row.ROW)
    # Create the folium geometry of this Polygon
    g = folium.GeoJson(row.geometry.__geo_interface__, name=name)
    # Add a folium Popup object with the name string
    g.add_child(folium.Popup(name))
    # Add the object to the map
    g.add_to(m)
m
```

[9]:
```
text/html<folium.folium.Map at 0x1f0ea0d7dd8>
```

[10]:
```
+fvtextcolorcomment_color# Display number of images and Path/Row of the image
for i, (path,row) in enumerate(zip(paths,rows)):
    print('Image', i+1, ' -path:', path, 'row:', row)
```

[10]:
```
Image 1  -path: 118 row: 38
Image 2  -path: 119 row: 38
```

Note that here you have two options: 1) continuing and executing the code reported in the next two sections on data donwload and image cropping, or 2) skipping these sections and proceeding to the image mosaicing sections. We recommend 2) as the processing of unzipping every folder may take long causing the JupyterLab instance to crash.

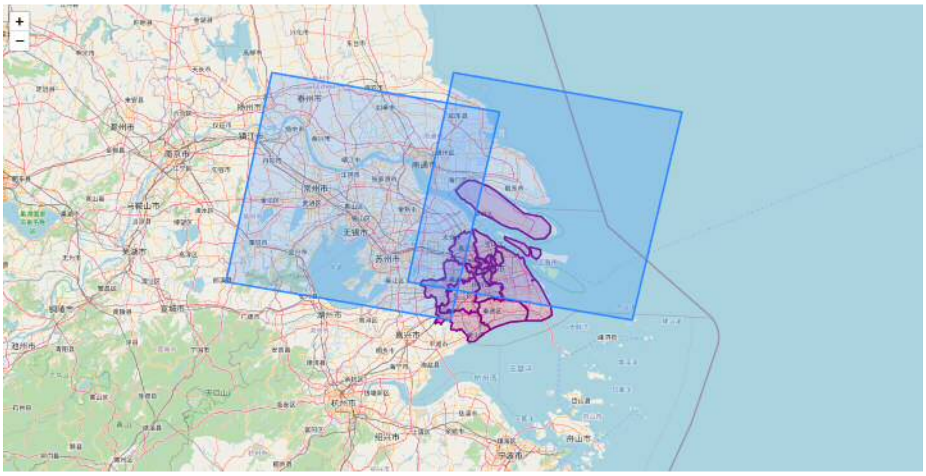


Figure 2: Landsat satellite tiles that cover the Area of Interest

2.3 Data download and pre-processing

We now have relevant path and row tuples for our area of analysis. So we can proceed to download satellite images, which are stored on the Google Cloud. To download images, we specify certain parameters: time frame, cloudcover in percentage (0-100 %) and satellite mission (1-5,7,8). The here used Landsat578 API automatically searches the Google Cloud for scenes with the specified parameters and downloads matching images. In order to search the Google Cloud for relevant images, a list of available needs to be downloaded when the code is run for the first time. The list provides basic information of the satellite images and since Landsat data acquisition is ongoing, is updated continuously. Thus, if data from the latest acquisition date is required, it is recommended to re-download the file list before running the code.

We use satellite imagery from a Landsat 5 scene taken in 1984 and a Landsat 8 taken in 2019 to determine neighbourhood changes over time. Landsat 5 scenes can be obtained from two different sensors, the Multispectral Scanner System and the Thematic Mapper, which provide 4 and 7 bands, respectively. The Multispectral Scanner System (MSS) is used in Landsat 1-3 and was superseded by the Thematic Mapper (TM). The MSS provides a green and red band (Band numbers: 1,2) and two infrared bands (Band numbers: 3,4), while the TM provides bands covering red, blue and green (Band numbers: 1,2,3), near-infrared (Band numbers: 4), short-wave infrared (Band numbers: 5,7) and thermal infrared (6). Each downloaded scene contains all bands with one image per band. The different bands can then be stacked in order to highlight various Earth surface processes. In this exercise, scenes from the MSS and TM are downloaded, but only data from the TM is used for analysis.

The Operational Land Imager (OLI) aboard Landsat 8 provides multispectral bands (bands 1-7 and 9) with a resolution of 30 metres and a panchromatic band (band 8) with a resolution of 15 metres (Barsi et al. 2014a). The Thermal Infrared Sensor (TIRS) provides thermal infrared images (bands 10 and 11) with a resolution of 100 meters (Barsi et al. 2014b). The Landsat 8 satellite has a swath width of 185 km for the OLI and TIRS instruments, so one scene usually captures the extent of a city. In other cases, the geographical area of interest may extend beyond one image so that multiple images may be needed (Barsi et al. 2014b, Knight, Kvaran 2014). Given the revisit time of 16 days, usually cloud free images can be retrieved for most cities on a bi-weekly or monthly basis (Roy et al. 2014). The folder and filename of each scene provides information about the satellite, instrument, path/row tuple and date.

Table 3 and Table 4 show which general information of the downloaded scenes can be inferred from the folder and file names of each individual scene:

Table 3: Overview of folder naming convention for Landsat images

Parameter	Meaning
L	Landsat
X	Sensor (“C”=OLI/TIRS combined, “O”=OLI-only, “T”=TIRS-only, “E”=ETM+, “T”=TM, “M”=MSS)
PPP	WRS path
RRR	WRS row
YYYY	Year
DDD	Julian day of year
GSI	Ground station identifier
VV	Archive version number

Note: Folder names are structured as `LXPPPRRRYYYYDDDGSIVV`

Table 4: Overview of file naming convention for Landsat images

Parameter	Meaning
L	Landsat
X	Sensor (“C”=OLI/TIRS combined, “O”=OLI-only, “T”=TIRS-only, “E”=ETM+, “T”=TM, “M”=MSS)
SS	Satellite (“0””=Landsat 7, “08”=Landsat 8)
LLL	Processing correction level (L1TP/L1GT/L1GS)
PPP	WRS path
RRR	WRS row
YYYYMMDD	Acquisition year, month, day
yyyymmdd	Processing year, month, day
CC	Collection number (01, 02, ...)
TX	Collection category (“RT”=Real-Time, “T1”=Tier 1, “T2”=Tier 2)

Note: File names are structured as `LXSS_LLLL_PPPRRR_YYYYMMDD_yyyymmdd_CC_TX`

2.3.1 Landsat imagery download

We will now download two Landsat satellite images, one from 1984 and one from 2019. The starting year was chosen due to the increase in spatial resolution to 30 metres with Landsat 4, whereas the end year was chosen at random. The specific dates were selected as the cloud cover was below 5%, ensuring an unobstructed view of the urban area.

[11]:
```
# Download Tile list from Google - only needs to be done when first running the code
# NOTE this cell is using the ! magic, which runs command line processes from a Jupyter
# notebook. Make sure the ‘landsat‘ tool, from the ‘landsat578‘ package is installed
# and available

# Path to index file
Index_PATH = os.path.join(directory  ’/index.csv.gz’)
if os.path.exists(Index_PATH):
    print(’File already exists’)
else:
    !landsat --update-scenes yes
```

[12]:
```
# Define Download function to acquire scenes from the Google API
def landsat_download(start_date, end_date, sat,path,row,cloud,output):
    g=GoogleDownload(start=start_date, end=end_date, satellite=sat, path=path,
    ...row=row, max_cloud_percent=cloud, output_path=output)
    g.download()
```

[13]:
```
# Specify start/end date (in YYYY-MM-DD format), the cloud coverage of the image (in %)
# and the satellite you would like to acquire images from (1-5,7,8). In this case we
# acquire a recent scene from Landsat 8 with a cloud coverage of 5 %.

start_date = ’2019-01-01’
end_date = ’2019-02-20’
```

```
cloud = 5
satellites = [8]
output = os.path.join(directory  '/Lansat_images/')
```

[14]:
```
# Loop through the specified satellites for each path and row tuple
for sat in satellites:
    for i, (path,row) in enumerate(zip(paths,rows)):
        print('Image', i+1, ' -path:', path, 'row:', row)
        landsat_download(start_date, end_date,sat,path,row,cloud,output)
```

[15]:
```
# The above step is repeated to acquire a Landsat 5 scene from 1984 with 5 % cloud
# coverage.
start_date = '1984-04-22'
end_date = '1984-04-24'
cloud = 5
satellites = [5]
output = os.path.join(directory  '/Lansat_images/')
```

[16]:
```
# Loop through the specified satellites for each path and row tuple
for sat in satellites:
    for i, (path,row) in enumerate(zip(paths,rows)):
        print('Image', i+1, ' -path:', path, 'row:', row)
        landsat_download(start_date, end_date,sat,path,row,cloud,output)
```

[17]:
```
# Delete Scenes that were acquired using the MSS:
outdir = os.listdir(output)
for i in outdir:
    if 'LM' in os.path.basename(i):
        try:
            shutil.rmtree(os.path.abspath(os.path.join(output,os.path.basename(i))))
        except OSError as e:
            print ("Error: %s - %s." % (e.filename, e.strerror))
```

2.3.2 Image Cropping

Satellite imagery is large. The size per image can easily equate to 1 GB. It often makes the data processing and analysis computationally expensive. Cropping the obtained scenes to the relevant region of the image enables faster processing and analysing by significantly reducing the size of the input.

[18]:
```
# Define cropping function using command line gdalwarp.
## Note: The BQA band is the quality assessment band, which has a different no data
## value (1) than the other bands (0), which makes it necessary to us a different
## croping function.
def crop(inraster,outraster,shape):
        !gdalwarp -cutline {shape} -srcnodata 0 -crop_to_cutline {inraster} {outraster}
def crop_bqa(inraster,outraster,shape):
        !gdalwarp -cutline {shape} -srcnodata 1 -crop_to_cutline {inraster} {outraster}
```

[19]:
```
# Loop through every folder and a create an image cropped to the extent of the shapefile
# save it with the original name and the extension _Cropped
for t in range(0,12):
    for filename in glob.glob((output/**/*_B{}.tif').format(t), recursive=True):
        inraster = filename
        outraster = filename[:-4]  '_Cropped.tif'
        crop(inraster, outraster, shp)
for filename in glob.glob(output/**/*.tif'):
    if 'BQA.TIF' in i:
        inraster = i
        outraster = i[:-4]  '_Cropped.tif'
        crop_bqa(inraster,outraster,shp)
```

2.3.3 Image mosaic

As indicated above, a single Landsat scene may not cover the full extent of a city due to the satellite's flight path as can be observed from the interactive map. Creating a mosaic of two or more images is thus often needed to produce a single image that covers the entirety of the area under analysis.

[20]:
```
# Read in the relevant Landsat 8 files
output = 'Landsat_images/'
images = sorted(os.listdir(output))
dirpath1 = os.path.join(output, images[0])
dirpath2 = os.path.join(output, images[1])
mosaic_n = os.path.join(output,'Mosaic/')
search = 'L*_Cropped.tif'
query1 = os.path.join(dirpath1,search)
query2 = os.path.join(dirpath2,search)
files1 =  glob.glob(query1)
files2 =  glob.glob(query2)
files1.sort()
files2.sort()
if os.path.exists(mosaic_n):
    print('Output Folder exists')
else:
    os.makedirs(mosaic_n)
```

[21]:
```
# Match bands together and create a mosaic. Since the BQA band and the cloudmask have
# different denominations than the other bands, these images have to be merged
# together separately.
def mosaic_new(scene1,scene2):
    src_mosaic =[]
    string_list=[]
    for i,j in zip(scene1,scene2):
        for k in range(1,12):
            string_list.append('B{}_Cropped'.format(k))
        for l in range(0,11):
            if string_list[l] in os.path.basename(i) and os.path.basename(j):
                src1 = rasterio.open(i)
                src2 = rasterio.open(j)
                src_mosaic = [src1,src2]
                mosaic,out_trans = rasterio.merge.merge(src_mosaic)
                out_meta = src1.meta.copy()
                out_meta.update({"driver": "GTiff", 'height':mosaic.shape[1],
                                'width':mosaic.shape[2],'transform':out_trans})
                outdata = os.path.join(mosaic_n,'B{}_mosaic.tif'.format(l))
                with rasterio.open(outdata,'w',**out_meta) as dest:
                    dest.write(mosaic)
        # Mosaic Quality Assessment Band
        if 'BQA_Cropped' in os.path.basename(i) and os.path.basename(j):
            bqa1 = rasterio.open(i)
            bqa2 = rasterio.open(j)
            bqa_mosaic = [bqa1,bqa2]
            mosaic_,out_trans = rasterio.merge.merge(bqa_mosaic,nodata=1)
            out_meta = bqa1.meta.copy()
            out_meta.update({"driver": "GTiff", 'height':mosaic_.shape[1],
                            'width':mosaic_.shape[2],'transform':out_trans})
            outdata = os.path.join(mosaic_n,'BQA_mosaic.tif')
            with rasterio.open(outdata,'w',**out_meta) as dest:
                dest.write(mosaic_)
            # Mosaic of Cloudmask
            search = 'cloudmask.tif'
            query3 = os.path.join(dirpath1,search)
            query4 = os.path.join(dirpath2,search)
            files3 =  glob.glob(query3)
            files4 =  glob.glob(query4)
            for i,j in zip(files3,files4):
                if 'cloudmask' in os.path.basename(i)and os.path.basename(j):
                    cloudmask1 = rasterio.open(i)
                    cloudmask2 = rasterio.open(j)
                    cloud_mosaic = [cloudmask1,cloudmask2]
                    mosaic_c,out_trans = rasterio.merge.merge(cloud_mosaic,nodata=1)
                    out_meta = cloudmask1.meta.copy()
                    out_meta.update({"driver": "GTiff", 'height':mosaic_c.shape[1],
                                    'width':mosaic_c.shape[2],'transform':out_trans})
                    outdata = os.path.join(mosaic_n,'Cloudmask_mosaic.tif')
                    with rasterio.open(outdata,'w',**out_meta) as dest:
                        dest.write(mosaic_c)
```

[22]:
```
mosaic_new(files1,files2)
```

[23]:
```
# Read in the relevant files for the Landsat 5 scenes
images = sorted(os.listdir(output))
dirpath_o1 = os.path.join(output, images[2])
dirpath_o2 = os.path.join(output, images[3])
mosaic_o = os.path.join(output,'Mosaic_old/')
query_o1 = os.path.join(dirpath_o1,search)
query_o2 = os.path.join(dirpath_o2,search)
files_o1 =  glob.glob(query_o1)
files_o2 =  glob.glob(query_o2)
files_o1.sort()
files_o2.sort()
if os.path.exists(mosaic_o):
    print('Output Folder exists')
else:
    os.makedirs(mosaic_o)
```

[24]:
```
# Match bands together and create a mosaic. Since the BQA band and the cloudmask have
# different denominations than the other bands, these images have to be merged together
# separately.
def mosaic_old(scene_o1,scene_o2):
    src_mosaic =[]
    string_list=[]
    for i,j in zip (scene_o1,scene_o2):

        for k in range(1,8):
            string_list.append('B{}_Cropped'.format(k))
        for l in range(0,7):
            if string_list[l] in os.path.basename(i) and os.path.basename(j):
                src1 = rasterio.open(i)
                src2 = rasterio.open(j)
                src_mosaic = [src1,src2]
                mosaic,out_trans= rasterio.merge.merge(src_mosaic)
                out_meta = src1.meta.copy()
                out_meta.update({"driver": "GTiff", 'height':mosaic.shape[1],
                                'width':mosaic.shape[2],'transform':out_trans})
                outdata = os.path.join(mosaic_o,'B{}_mosaic.tif'.format(l))
                with rasterio.open(outdata,'w',**out_meta) as dest:
                    dest.write(mosaic)
        # Mosaic Quality Assessment Band
        if 'BQA_Cropped' in os.path.basename(i) and os.path.basename(j):
            bqa1 = rasterio.open(i)
            bqa2 = rasterio.open(j)
            bqa_mosaic = [bqa1,bqa2]
            mosaic_,out_trans= rasterio.merge.merge(bqa_mosaic,nodata=1)
            out_meta = bqa1.meta.copy()
            out_meta.update({"driver": "GTiff", 'height':mosaic_.shape[1],
                            'width':mosaic_.shape[2],'transform':out_trans})
            outdata = os.path.join(mosaic_o,'BQA_mosaic.tif')
            with rasterio.open(outdata,'w',**out_meta) as dest:
                dest.write(mosaic_)

            # Mosaic of Cloudmask
            search = 'cloudmask.tif'
            query_o3= os.path.join(dirpath_o1,search)
            query_o4 = os.path.join(dirpath_o2,search)
            files_o3 =  glob.glob(query_o3)
            files_o4 =  glob.glob(query_o4)
            for i,j in zip(files_o3,files_o4):
                if 'cloudmask' in os.path.basename(i)and os.path.basename(j):
                    cloudmask1 = rasterio.open(i)
                    cloudmask2 = rasterio.open(j)
                    cloud_mosaic = [cloudmask1,cloudmask2]
                    mosaic_c,out_trans= rasterio.merge.merge(cloud_mosaic,nodata=1)
                    out_meta = cloudmask1.meta.copy()
                    out_meta.update({"driver": "GTiff", 'height':mosaic_c.shape[1],
                                    'width':mosaic_c.shape[2],'transform':out_trans})
                    outdata = os.path.join(mosaic_o,'Cloudmask_mosaic.tif')
                    with rasterio.open(outdata,'w',**out_meta) as dest:
                        dest.write(mosaic_c)
```

[25]:
```
mosaic_old(files_o1,files_o2)
```

2.3.4 Natural-colour (True-colour) composition

Our downloaded data from Landsat 8 and Landsat 5 have different band designations. Combining different satellite bands are useful to identify features of the urban environment: vegetation, built-up areas, ice and water. We create a standard natural-colour composition image using Red, Green and Blue satellite bands. This colour composition best reflects the natural environment. For instance, trees are green; snow and clouds are white; and water is blue. Landsat 8 has 11 bands with bands 4, 3 and 2 corresponding to Red, Green and Blue respectively. Landsat 5 has 7 bands with bands 3, 2 and 1, corresponding to Red, Green and Blue. We perform layer stacking to produce a true colour image composition to gain understanding of the local area before extracting and analysing features of the urban environment.

[26]:
```
# Normalise the bands to so that they can be combined to a single image
def normalize(array):
    """Normalizes numpy arrays into scale 0.0 - 1.0"""
    array_min, array_max = array.min(), array.max()
    return ((array - array_min)/(array_max - array_min))
```

[27]:
```
# Adjust the intensity of each band for visualisation.
# This is a way of rescaling each band by clipping the pixels that are outside the
# specified range to the range we defined. By adjusting the gamma, we change the
# brightness of the image with gamma >1 resulting in a brighter image. However
# there are more complex methods such as top of the atmosphere corrections, which
# subtracts any atmospheric interference from the image.
# For the purpose of this notebook, this way is sufficient.
def rescale_intensity(image):
    p2, p98 = np.percentile(image, (0.2, 98))
    img_exp = exposure.rescale_intensity(image, in_range=(p2, p98))
    img_gamma = exposure.adjust_gamma(img_exp, gamma=2.5,gain=1)
    return(img_gamma)
```

[28]:
```
# Downsample image resolution with factor 0.5 for displaying purposes.
def downsample(file):
    downscale_factor=0.5
    data = file.read(1,
        out_shape=(
            file.count,
            int(file.height * downscale_factor),
            int(file.width * downscale_factor)
        ),
        resampling=Resampling.bilinear
    )
    # scale image transform
    transform = file.transform * file.transform.scale(
        (file.width / data.shape[-1]),
        (file.height / data.shape[-2])
    )
    return data
```

[29]:
```
# Use rasterio to open the Red, Blue and Green bands of the mosaic image from 1984
# to create an RGB image
# **NOTE**: The Mosaic names do not correspond to the actual band designations as
# python starts counting at 0!
with rasterio.open('Landsat_images/Mosaic_old/B0_mosaic.tif') as band1_old:
    b1_old=downsample(band1_old)
with rasterio.open('Landsat_images/Mosaic_old/B1_mosaic.tif') as band2_old:
    b2_old=downsample(band2_old)
with rasterio.open('Landsat_images/Mosaic_old/B2_mosaic.tif') as band3_old:
    b3_old=downsample(band3_old)
```

[30]:
```
# Normalise the bands so that they can be combined to a single image
red_old_n = normalize(b3_old)
green_old_n = normalize(b2_old)
blue_old_n = normalize(b1_old)

# Apply the function defined before to make more natural-looking image
red_adj = rescale_intensity(red_old_n)
green_adj = rescale_intensity(green_old_n)
blue_adj = rescale_intensity(blue_old_n)
```

Figure 3: True colour Landsat image of the Shanghai urban area from 1984

```
# Stack the three different bands together
rgb_2 = np.dstack((red_adj,green_adj,blue_adj))

# Visualise the true color image
fig,ax = plt.subplots(figsize=(10,10))
ax.imshow(rgb_2)
plt.title('Fig.3: True color Landsat image of the Shanghai urban area from 1984',
          y=-0.1, fontsize=12)
plt.show()
plt.close()
del rgb_2,b1_old,b2_old,b3_old,red_adj,green_adj,blue_adj
```

[30]:
```
image/png<Figure size 720x720 with 1 Axes>
```

[31]:
```
# Use rasterio to open the Red, Blue and Green bands of the mosaic image from 2019
# to create an RGB image
# **NOTE**: The Mosaic names do not correspond to the actual band designations as
# python starts counting at 0!!
with rasterio.open('Landsat_images/Mosaic/B1_mosaic.tif') as band2_new:
    b2_new = downsample(band2_new)
with rasterio.open('Landsat_images/Mosaic/B2_mosaic.tif') as band3_new:
    b3_new = downsample(band3_new)
with rasterio.open('Landsat_images/Mosaic/B3_mosaic.tif') as band4_new:
    b4_new = downsample(band4_new)
```

[32]:
```
# Normalise the bands so that they can be combined to a single image
red_new_n = normalize(b4_new)
green_new_n = normalize(b3_new)
blue_new_n = normalize(b2_new)

# Apply the function defined before to make more natural-looking image
red_rescale = rescale_intensity(red_new_n)
green_rescale = rescale_intensity(green_new_n)
blue_rescale = rescale_intensity(blue_new_n)

# Stack the three different bands together
rgb = np.dstack((red_rescale, green_rescale, blue_rescale))

# Here we adjust the gamma (brightness) for the stacked image to achieve a more
# natural looking image.
rgb_adjust = exposure.adjust_gamma(rgb, gamma = 1.5, gain=1)

# Visualise the true color image
fig,ax = plt.subplots(figsize=(10,10))
```

Figure 4: True colour Landsat image of the Shanghai urban area from 2019

```
ax.imshow(rgb_adjust)
plt.title('Fig.4: True color Landsat image of the Shanghai urban area from 2019',
           y=-0.1, fontsize=12)
plt.show()
plt.close()
del rgb,red_new_n,green_new_n,blue_new_n,red_rescale,green_rescale,blue_rescale,
          rgb_adjust
```

[32]:
```
image/png<Figure size 720x720 with 1 Axes>
```

When comparing the true colour Landsat satellite images in Figures 3 and 4, the urbanisation of Shanghai between 1984 and 2019 is apparent. In the following steps, we will analyse and quantify these urban changes.

3 Feature extraction

Since the above two maps show that urban neighbourhoods of Shanghai have undergone dramatic changes over time in colour, texture, greenery, buildings, etc., the next stage is to gain valuable information out of satellite images and interpret these changes. Since the images we have downloaded are on a city-wide scale, which covers more than a thousand kilometre spatial resolution and less detailed. Therefore, feature extraction is performed to get a reduced representation of the initial image but informative and sufficiently accurate for subsequent analysis and interpretation.

We examine four sets of features based on the above two true colour maps and the scale, where the colour, texture, greenery, and buildings changed a lot during the past 25 years in Shanghai. Specifically, colour and texture features extracted from true colour imagery (i.e. RGB bands composition represented by bands 1-3 and bands 2-4 in 1984 and 2019), and vegetation features and built-up features extracted from Red, near infrared (NIR) and shortwave infrared (SWIR) bands, represented by bands 3-5 and bands 4-6 in 1984 and 2019. More detailed information about the meaning of each band can be found at https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research?qt-news_science_products=0#qt-news_science_products. In this analysis, colour features measure the colour moments of true colour imagery to interpret colour distribution; texture features apply LBP (Local binary patterns) texture spectrum model to show spatial distribution of intensity values in an image; vegetation features calculate the NDVI (Normalised difference vegetation index) to capture the amount of vegetation, and built-up features calculate NDBI (Normalised difference built-up index) to highlight artificially constructed areas.

Figure 5: Spatial distribution of all administrative divisions of Shanghai

The administrative divisions of Shanghai have experienced tremendous changes in the last tens of years (Ministry of Civial Affairs of the People's Republic of China 2018), thus, we will conduct feature extraction of imagery on the current administrative boundaries to explore if satellite imagery can be used to reflect and interpret urban changes. The figure below shows the spatial distribution of each administrative area with relative labels in Shanghai.

[33]:
```
# read administritive boundary shapefile of Shanghai
poly = gpd.read_file(shp)


f, ax = plt.subplots(1, figsize = (9,9))
poly.plot(ax = ax)
# create a new colomn, in order to plot polygon labels (i.e. name) in the map
poly['coords']=poly['geometry'].apply(lambda x:x.representative_point().coords[:])
poly['coords']=[coords[0] for coords in poly['coords']]
for idx, row in poly.iterrows():
    ax.annotate(text=row['Name'],xy=row['coords'],va='center',ha='center',alpha = 0.8,
                fontsize = 8)
plt.axis('equal')
plt.axis('off')
f.suptitle('Fig.5: Spatial distribution of all administrative divisions of Shanghai',
                y=-0.1,fontsize = 12)
```

[33]:
```
Text(0.5, -0.1, 'Fig.5: Spatial distribution of all administrative divisions of
Shanghai')
image/png<Figure size 648x648 with 1 Axes>
```

Figure 5 shows that administrative divisions of 'Chongming' in the north appear three geometries. Therefore, it is necessary to check if they belong to a single administrative unit.

[34]:
```
poly.loc[poly['Name']== 'Chongming','Name']
```

[34]:
```
0    Chongming
3    Chongming
5    Chongming
Name: Name, dtype: object
```

Chongming administrative division consist of three separate geometries, which may confuse our further analysis. As a result, we dissolved these geometries into a single geometric feature and take a look at the new dataset. The below table shows that the Chongming administrative division now consists of multipolygons which includes all polygons as a whole.

[35]:
```
# Dissolve geometries with the identical names together
poly = poly.dissolve(by = 'Name').reset_index()
# Have a look at the name of all administritive unit and we can see that chongming
# districts have been dissolved into a single administritive unit
poly['Name'].values
```

[35]:
```
array(['Baoshan', 'Changning', 'Chongming', 'Fengxian', 'Hongkou',
       'Huangpu', 'Jiading', 'Jinshan', 'Minhang', 'Pudong New', 'Putuo',
       'Qingpu', 'Songjiang', 'Xuhui', 'Yangpu', 'Zhabei'], dtype=object)
```

3.1 Image processing

Further pre-processing of satellite imagery is needed before feature extraction. This pre-processing involves three steps:

1. Masking (cropping) of raster files (i.e., Blue, Green, Red, Nir and SWIR bands) into each administrative district polygon;
2. Image enhancement to improve the quality and content of the original image; and,
3. Band stacking based on each neighbourhood unit.

[36]:
```
# open raster files
file_list_old = sorted(glob.glob('Landsat_images/Mosaic_old'  "/*.tif",recursive = True))
files_old = [rio.open(filename) for filename in file_list_old]
```

[37]:
```
file_list = sorted(glob.glob('Landsat_images/Mosaic'  "/*.tif"))
files = [rio.open(filename) for filename in file_list]
```

Before cropping all raster files into each polygon in the vector file (i.e. Shanghai administrative area shapefile), we have to ensure they have the same coordinate reference system (CRS). Once matched, the cropping process is prepared to go.

[38]:
```
poly.crs
```

[38]:
```
{'init': 'epsg:4326'}
```

[39]:
```
# check the crs of one band of satellite imagery
files[0].crs
```

[39]:
```
CRS.from_epsg(32651)
```

[40]:
```
# reproject the vector file to make it consistent with raster files
poly = poly.to_crs('EPSG:32651')
```

[41]:
```
# get each neighbourhood geographic boundary based on administritive area data
geo = [poly.__geo_interface__['features'][i]['geometry']
            for i in range(len(poly))]
```

[42]:
```
# clip R,G,B bands separately by each poly, so get pixel values in each poly and save
# them into a list
out_image = [[] for i in range(5)]
img_old = [[] for i in range(5)]

# x: Blue,Green,Red,NIR and SWIR bands, y: 16 polygons from vertor file
for x,y in itertools.product(range(5),range(len(geo))):
    # out_image[0] means masked Blue band polygon
    out_image[x].append(mask(files_old[0:5][x], [geo[y]], crop=True))
    # image enhancement: normalisation and Histogram Equalization
    img_old[x].append(exposure.equalize_hist(normalize(out_image[x][y][0][0])))
del out_image,files_old
```

[43]:
```
# clip R,G,B bands separately by each poly, so get pixel values in each poly and save
# them into a list
out_image = [[] for i in range(5)]
img_new = [[] for i in range(5)]

# x: Blue,Green,Red,NIR and SWIR bands, y: 16 polygons from vertor file
```

```
for x,y in itertools.product(range(5),range(len(geo))):
    # out_image[0] means masked Blue band polygon
    out_image[x].append(mask(files[0:5][x], [geo[y]], crop=True))
    # image enhancement: normalisation and Histogram Equalization
    img_new[x].append(exposure.equalize_hist(normalize(out_image[x][y][0][0])))
del out_image,files
```

[44]:
```
# have a look at the pixel values of one geographic area in blue band
img_new[0][0]
```

[44]:
```
array([[0.48515378, 0.48515378, 0.48515378, ..., 0.48515378, 0.48515378,
        0.48515378],
       [0.48515378, 0.48515378, 0.48515378, ..., 0.48515378, 0.48515378,
        0.48515378],
       [0.48515378, 0.48515378, 0.48515378, ..., 0.48515378, 0.48515378,
        0.48515378],
       ...,
       [0.48515378, 0.48515378, 0.48515378, ..., 0.48515378, 0.48515378,
        0.48515378],
       [0.48515378, 0.48515378, 0.48515378, ..., 0.48515378, 0.48515378,
        0.48515378],
       [0.48515378, 0.48515378, 0.48515378, ..., 0.48515378, 0.48515378,
        0.48515378]])
```

[45]:
```
# stack R,G,B bands together for later feature extraction
bb = [img_old[0][x].astype(np.float) for x in range(len(geo))]
bg = [img_old[1][x].astype(np.float) for x in range(len(geo))]
br = [img_old[2][x].astype(np.float) for x in range(len(geo))]
```

[46]:
```
rgb_old = [np.dstack((br[x],bg[x],bb[x])) for x in range(len(geo))]
```

[47]:
```
bb = [img_new[0][x].astype(np.float) for x in range(len(geo))]
bg = [img_new[1][x].astype(np.float) for x in range(len(geo))]
br = [img_new[2][x].astype(np.float) for x in range(len(geo))]
```

[48]:
```
rgb_new = [np.dstack((br[x],bg[x],bb[x])) for x in range(len(geo))]
```

3.2 Colour features

Colour features are used to extract the characteristics of colours from satellite imagery. A commonly used method to extract colour features is to compute colour moments of an image. Colour moments provide a measurement of colour similarity between images (Keen 2005). Basically, colour probability distributions of an image are characterised by a range of unique moments. The mean, standard deviation and skewness these three central moments are generally used to identify colour distribution. Here we extract colour features on HSV (Hue, Saturation and Value) colour space because it corresponds to human vision and has been widely used in computer vision. HSV colour space can be converted from RGB colour channels, Hue represents the colour portion, saturation represents the amount of grey in a particular colour (0 is grey), and Value represents the brightness of the colour (0 is black). Therefore, the true-colour imagery is characterised by a total of nine moments - three moments for each HSV channel in the same units.

[49]:
```
# interpret the color probability distribution by computing low order color
# moments(1,2,3)
def color_moments(img):
    if img is None:
        return
    # Convert RGB to HSV colour space
    img_hsv = rgb2hsv(img)
    # Split the channels - h,s,v
    h, s, v = [img_hsv[:,:,i] for i in [0,1,2]]
    # Initialize the colour feature
    color_feature = []
    # N = h.shape[0] * h.shape[1]
    # The first central moment - average
    h_mean = np.mean(h)  # np.sum(h)/float(N)
    s_mean = np.mean(s)  # np.sum(s)/float(N)
    v_mean = np.mean(v)  # np.sum(v)/float(N)
```

Table 5: Partial colour features identified in 1984

Name	h_mean	s_mean	v_mean	h_std	s_std	v_std	h_skew	s_skew	v_skew
Baoshan	0.27216	0.05208	0.64415	0.32709	0.07246	0.18531	0.35671	0.09006	0.19945
Changning	0.22141	0.05156	0.65989	0.28837	0.07518	0.17446	0.33080	0.09250	0.18763
Chongming	0.15381	0.01739	0.74231	0.27216	0.03563	0.10235	0.33292	0.05118	0.12060
Fengxian	0.33961	0.11292	0.60576	0.32194	0.12281	0.24362	0.34723	0.14439	0.25767
Hongkou	0.24951	0.06370	0.65073	0.30983	0.08744	0.18781	0.34797	0.10619	0.20013

```
    color_feature.extend([h_mean, s_mean, v_mean])
    # The second central moment - standard deviation
    h_std = np.std(h)  # np.sqrt(np.mean(abs(h - h.mean())**2))
    s_std = np.std(s)  # np.sqrt(np.mean(abs(s - s.mean())**2))
    v_std = np.std(v)  # np.sqrt(np.mean(abs(v - v.mean())**2))
    color_feature.extend([h_std, s_std, v_std])
    # The third central moment - the third root of the skewness
    h_skewness = np.mean(abs(h - h.mean())**3)
    s_skewness = np.mean(abs(s - s.mean())**3)
    v_skewness = np.mean(abs(v - v.mean())**3)
    h_thirdMoment = h_skewness**(1./3)
    s_thirdMoment = s_skewness**(1./3)
    v_thirdMoment = v_skewness**(1./3)
    color_feature.extend([h_thirdMoment, s_thirdMoment, v_thirdMoment])

    return color_feature
```

[50]:
```
# create and initialize a data table to store colour feastures
color_mom_old = pd.DataFrame(color_moments(rgb_old[0]))
# add the rest columns by assigning 9 color moments in each poly
for i in range(1,len(rgb_old)):
    color_mom_old[i]= color_moments(rgb_old[i])
    i = i+1
```

[51]:
```
# create and initialize a data table
color_mom_new = pd.DataFrame(color_moments(rgb_new[0]))
# add the rest columns by assigning 9 color moments in each poly
for i in range(1,len(rgb_new)):
    color_mom_new[i]= color_moments(rgb_new[i])
    i = i+1
```

[52]:
```
# Data manipulation
color_old_var = color_mom_old.T
# assign column names
color_old_var.columns =
   ['h_mean','s_mean','v_mean','h_std','s_std','v_std','h_skew','s_skew','v_skew']
# set geographic name as index
color_old_var= color_old_var.set_index(poly.Name)
```

[53]:
```
color_new_var = color_mom_new.T
color_new_var.columns =
   ['h_mean','s_mean','v_mean','h_std','s_std','v_std','h_skew','s_skew','v_skew']
color_new_var= color_new_var.set_index(poly.Name)
```

As we have created two new tables for colour features in the year 1984 and 2019, it would be helpful to have a view of the tables and see how they look like. Table 5 and Table 6 show nine variables (column) representing colour features within five administrative division of Shanghai (row).

[54]:
```
# check the information of colour feature
color_old_var.head().style.set_caption('Table 5: Partial colour features
... identified in 1984')
```

[54]:
```
text/html<pandas.io.formats.style.Styler at 0x1f0ed9c4be0>
```

[55]:
```
color_new_var.head().style.set_caption('Table 6: Partial colour features
... identified in 2019')
```

[55]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081f31518>
```

Table 6: Partial colour features identified in 2019

Name	h_mean	s_mean	v_mean	h_std	s_std	v_std	h_skew	s_skew	v_skew
Baoshan	0.23107	0.03587	0.63894	0.29785	0.05205	0.18019	0.33678	0.06787	0.19559
Changning	0.23185	0.03129	0.64924	0.30469	0.04847	0.16700	0.34439	0.06297	0.18248
Chongming	0.15731	0.01647	0.74240	0.28250	0.03184	0.10177	0.34529	0.04336	0.11980
Fengxian	0.29554	0.08620	0.60543	0.30273	0.09770	0.24360	0.32939	0.11570	0.25723
Hongkou	0.23994	0.03758	0.63861	0.30383	0.05505	0.18294	0.33962	0.07055	0.19762

3.3 Texture features

To extract texture features, we use a Local Binary Pattern (LBP) approach. LBP searches for pixels adjacent to a central point and tests whether these surrounding pixels are greater or less than the central pixel and generate a binary classification (Pedregosa et al. 2011) (https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_local_binary_pattern.html)). In theory, eight adjacent neighbour pixels in greyscale are set to compare with one central pixel value by 3 * 3 neighbourhood threshold, and consider the result as 1 or 0 (Ojala et al. 1996). Thus, these eight surrounding binary numbers correspond to LBP code for the central pixel value, determining the texture pattern of that threshold. Texture features are then the distribution of a collection of LBPs over an image.

[56]:
```
# convert a RGB image into Grayscale,which takes less space for analysis
gray_images_old = [rgb2gray(rgb_old[i]) for i in range(len(rgb_old))]
gray_images_new = [rgb2gray(rgb_new[i]) for i in range(len(rgb_new))]
```

[57]:
```
# settings for LBP
radius = 1 # radius = 1 refers to a 3*3 patch/window scale
n_points = 8 * radius #  the number of circularly symmetric neighbour set points
method = 'uniform' # finer quantization of the angular space which is gray scale and
                   # rotation invariant

lbps_old = [local_binary_pattern(gray_images_old[i],n_points,radius,method)
           ... for i in range(len(rgb_old))]
lbps_new = [local_binary_pattern(gray_images_new[i],n_points,radius,method)
           ... for i in range(len(rgb_new))]
```

[58]:
```
# n_bins are the same in each neighbourhood
n_bins = int(lbps_old[0].max()+1)
# define a function to count the number of points in a given bin of LBP distribution
# histogram
def count_hist(x):
    return np.histogram(lbps_old[x].ravel(),density=True, bins=n_bins,range=(0, n_bins))
# Assign counts to a new list, return the higtogram vector features in this cell(polygon)
hist_features_old = [count_hist(i)[0] for i in range(len(rgb_old))]
```

[59]:
```
# Extract texture features of another year based on same method
n_bins = int(lbps_new[0].max()+1)

def count_hist(x):
    return np.histogram(lbps_new[x].ravel(),density=True, bins=n_bins,range=(0, n_bins))

# Assign counts to a new list, return the higtogram vector features in this cell(polygon)
hist_features_new = [count_hist(i)[0] for i in range(len(rgb_new))]
```

Same with operations on colour features, this time we build two new tables (Table 7 and 8) for texture features, with each row present administrative division and each column represent texture feature.

[60]:
```
# The histogram features are the texture features
texture_old_var = pd.DataFrame([hist_features_old[a] for a in range(len(rgb_old))])
texture_old_var.columns = ['LBP'+ str(i) for i in range(n_bins)]
texture_old_var = texture_old_var.set_index(poly.Name)
# Have a look at the table with texture features of administrive division of
# Shanghai in 1984
texture_old_var.head().style.set_caption('Table 7: Partial texture features
  ... identified in 1984')
```

Table 7: Partial texture features identified in 1984

Name	LBP0	LBP1	LBP2	LBP3	LBP4	LBP5	LBP6	LBP7	LBP8	LBP9
Baoshan	0.03509	0.04196	0.04071	0.06839	0.07839	0.06748	0.04034	0.04110	0.52005	0.06648
Changning	0.03609	0.04608	0.04196	0.05979	0.06004	0.06442	0.03754	0.04339	0.53942	0.07129
Chongming	0.02582	0.02995	0.02176	0.03406	0.03958	0.03679	0.02404	0.02916	0.70944	0.04941
Fengxian	0.05551	0.06647	0.05123	0.07200	0.07377	0.07393	0.05291	0.06510	0.38211	0.10698
Hongkou	0.04202	0.05056	0.04354	0.05933	0.05676	0.07072	0.03969	0.04649	0.51043	0.08047

Table 8: Partial texture features identified in 2019

Name	LBP0	LBP1	LBP2	LBP3	LBP4	LBP5	LBP6	LBP7	LBP8	LBP9
Baoshan	0.04306	0.04774	0.04077	0.05862	0.06808	0.05681	0.03741	0.04557	0.52419	0.07777
Changning	0.04264	0.05012	0.03767	0.05137	0.05842	0.06153	0.03524	0.04708	0.53945	0.07648
Chongming	0.02547	0.02964	0.02333	0.03522	0.04762	0.03641	0.02359	0.02865	0.70442	0.04565
Fengxian	0.05121	0.06105	0.05288	0.08141	0.09716	0.07923	0.05152	0.06044	0.36993	0.09517
Hongkou	0.04703	0.05417	0.04294	0.05439	0.05501	0.06805	0.03879	0.04894	0.50732	0.08335

[60]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081ebe630>
```

[61]:
```
# The histogram features are the texture features
texture_new_var = pd.DataFrame([hist_features_new[a] for a in range(len(rgb_new))])
texture_new_var.columns = ['LBP' str(i) for i in range(n_bins)]
texture_new_var = texture_new_var.set_index(poly.Name)
# Have a look at the table with texture features of administrive division of
# Shanghai in 2019
texture_new_var.head().style.set_caption('Table 8: Partial texture features
  ... identified in 2019')
```

[61]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081ebeac8>
```

3.4 Vegetation and built-up features

Vegetation features and built-up features can be measured by calculating fundamental NDVI and NDBI indices in each administrative area respectively. The Normalized Difference Vegetation Index (NDVI) is a normalized index, using Red and NIR bands to display the amount of vegetation (NASA 2000). The use of NDVI maximizes the reflectance properties of vegetation by minimizing NIR and maximizing the reflectance in the red wavelength. The measure is used to distinguish vegetation in regions, as more vegetation will affect the ratio of visible light absorbed and near-infrared light reflected. The formula is as follows:

$$\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{Red})/(\mathrm{NIR} + \mathrm{Red})$$

The output value of this index is between -1.0 and 1.0. Close to 0 represents no vegetation, close to 1 indicates the highest possible density of green leaves, and close to -1 indicates water bodies.

The Normalized Difference Built-up Index (NDBI) uses the NIR and SWIR bands to highlight artificially constructed areas (built-up areas) where there is a typically a higher reflectance in the shortwave infrared region than the near infrared region (Zha et al. 2003). The index is a ratio type that reduces the effects of differences in terrain illumination and atmospheric effects. The formula is as follows:

$$\mathrm{NDBI} = (\mathrm{SWIR} - \mathrm{NIR})/(\mathrm{SWIR} + \mathrm{NIR})$$

Also, the output value of this index is between -1 to 1. Higher values represent built-up areas whereas negative values represent water bodies.

After calculating these two indices, vegetation features and built-up features can be measured by calculating average values of index values within each administrative area.

Table 9: Partial vegetation and built-up features identified in 1984

Name	veg_mean	builtup_mean
Baoshan	-0.002218	0.000611
Changning	-0.002147	0.000582
Chongming	-0.000805	0.000190
Fengxian	-0.007201	0.001499
Hongkou	-0.004648	-0.000313

3.4.1 Vegetation features

[62]:
```
# identify red and NIR band to each neighbourhood unit in 1984
red_old, nir_old = img_old[2],img_old[3]
# Calculate ndvi, assign 0 to nodata pixels
ndvi_old = [np.where((nir_old[i]  red_old[i])==0, 0,
                    (nir_old[i]-red_old[i])/(nir_old[i]  red_old[i]))
            for i in range(len(poly))]
```

[63]:
```
# identify red and NIR band to each neighbourhood unit in 1984
red_new, nir_new = img_new[2],img_new[3]
# Calculate ndvi, assign 0 to nodata pixels
ndvi_new = list(map(lambda i: np.where((nir_new[i]  red_new[i])==0, 0,
                    (nir_new[i]-red_new[i])/(nir_new[i]  red_new[i])),
            list(range(len(poly)))
    ))
```

[64]:
```
veg_old_var = pd.DataFrame([np.mean(ndvi_old[i]) for i in range(len(poly))],
                       index = poly.Name, columns = ['veg_mean'])
```

[65]:
```
veg_new_var = pd.DataFrame([np.mean(ndvi_new[i]) for i in range(len(poly))],
                       index = poly.Name, columns = ['veg_mean'])
```

3.4.2 Built-up features

[66]:
```
# identify red and NIR band to each neighbourhood unit in 1984
nir_old, swir_old = img_old[3],img_old[4]
# Calculate ndvi, assign 0 to nodata pixels
ndbi_old = [np.where((nir_old[i]  swir_old[i])==0., 0,
                    (swir_old[i] - nir_old[i])/(nir_old[i]  swir_old[i]))
            for i in range(len(poly))]
```

[67]:
```
# identify red and NIR band to each neighbourhood unit in 1984
nir_new, swir_new = img_new[3],img_new[4]
# Calculate ndvi, assign 0 to nodata pixels
ndbi_new = list(map(lambda i: np.where((nir_new[i]  swir_new[i])==0., 0,
                    (swir_new[i] - nir_new[i])/(nir_new[i]  swir_new[i])),
            list(range(len(poly)))
                   ))
```

[68]:
```
builtup_old_var = pd.DataFrame([np.mean(ndbi_old[i]) for i in range(len(poly))],
                       index = poly.Name, columns = ['builtup_mean'])
```

[69]:
```
builtup_new_var = pd.DataFrame([np.mean(ndbi_new[i]) for i in range(len(poly))],
                       index = poly.Name, columns = ['builtup_mean'])
```

Table 9 and Table 10 created as shown below contain both vegetation features (NDVI) and builtup features (NDBI), with the mean value of vegetation features and built-up features (two columns) calculated at each administrative division (row).

[70]:
```
veg_built_old = pd.concat([veg_old_var,builtup_old_var], axis = 1)
veg_built_old.head().style.set_caption('Table 9: Partial vegetation and built-up
 ... features identified in 1984')
```

[70]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081e1b1d0>
```

Table 10: Partial vegetation and built-up features identified in 2019

Name	veg_mean	builtup_mean
Baoshan	-0.001801	0.001938
Changning	-0.001515	0.000774
Chongming	-0.000705	0.000318
Fengxian	-0.008185	-0.000408
Hongkou	-0.002057	-0.000277

[71]:
```
veg_built_new = pd.concat([veg_new_var,builtup_new_var], axis = 1)
veg_built_new.head().style.set_caption('Table 10: Partial vegetation and built-up
 ... features identified in 2019')
```

[71]:
```
text/html<pandas.io.formats.style.Styler at 0x1f0ed9c4828>
```

4 Feature clustering

Now we have four types of features: colour, texture, vegetation and built-up area for Shanghai in 1984 and 2019. These features are the embodiment of urban changes and vary greatly due to rapid urbanisation and development. Therefore, the subsequent task is to identify systematic patterns from these integrated features for analysis of urban changes, such as whether several administrative areas share similar patterns. A clustering method is required within this context to group these geographical divisions that are similar within each other but different between them. Considering the ease of computation and fast implementation, we use generalised and the most popular k-means clustering to identify representative types of neighbourhoods based on multiple features. K-means clustering partitions the data by creating k groups of equal variance, minimising the within-cluster sum of squares (Pedregosa et al. 2011). We can perform K-means using the package scikit-learn, which is a powerful machine learning package for Python.

[72]:
```
# merge all features together
features_old_var = pd.concat([color_old_var,texture_old_var,veg_old_var,
   builtup_old_var], axis = 1)
features_old_var.head().style.set_caption('Table 11: Four types of features
 ... (21 in total) identified in 1984')
```

[72]:
```
text/html<pandas.io.formats.style.Styler at 0x1f0ed9c4908>
```

[73]:
```
# merge all features together
features_new_var = pd.concat([color_new_var,texture_new_var,veg_new_var,
   builtup_new_var], axis=1)
features_new_var.head().style.set_caption('Table 12: Four types of features
 ... (21 in total) identified in 2019')
```

[73]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081f31438>
```

Table 11 and Table 12 reveal the integrated 21 features across our four sets of image features and their differences at geographical division in magnitude between 1984 and 2019. Since k-means clustering is one of the machine learning algorithms, which generally expect data transformation for preprocessing before fitting the algorithm. We therefore use one of the most popular rescale methods to standardise these features to lie between 0 and 1 based on MinMaxScaler() function in scikit-learn package. The motivation of this method relies on the robustness to very small standard deviation. This preprocess ensures individual features of dataset have the same scale that standard normally distributed.

[74]:
```
# Last preprocessing step before machine learning: data rescaling
min_max_scaler = preprocessing.MinMaxScaler()
np_scaled = min_max_scaler.fit_transform(features_old_var)
oldvar_scale = pd.DataFrame(np_scaled)
oldvar_scale.columns = features_old_var.columns
```

Table 11: Four types of features (21 in total) identified in 1984

Name	h_mean	s_mean	v_mean	h_std	s_std	v_std	h_skew
Baoshan	0.272161	0.052081	0.644148	0.327094	0.072457	0.185309	0.356713
Changning	0.221412	0.051564	0.659894	0.288368	0.075177	0.174455	0.330803
Chongming	0.153807	0.017394	0.742309	0.272162	0.035627	0.102347	0.332916
Fengxian	0.339613	0.112915	0.605758	0.321941	0.122805	0.243621	0.347226
Hongkou	0.249526	0.063704	0.650725	0.309825	0.087439	0.187805	0.347968

Name	s_skew	v_skew	LBP0	LBP1	LBP2	LBP3	LBP4
Baoshan	0.090057	0.199446	0.035093	0.041960	0.040705	0.068394	0.078389
Changning	0.092504	0.187627	0.036086	0.046078	0.041956	0.059792	0.060040
Chongming	0.051184	0.120603	0.025822	0.029946	0.021757	0.034058	0.039580
Fengxian	0.144392	0.257670	0.055508	0.066468	0.051230	0.072002	0.073767
Hongkou	0.106194	0.200131	0.042018	0.050562	0.043542	0.059326	0.056759

Name	LBP5	LBP6	LBP7	LBP8	LBP9	veg_mean	builtup_mean
Baoshan	0.067483	0.040339	0.041101	0.520053	0.066483	-0.002218	0.000611
Changning	0.064422	0.037538	0.043385	0.539416	0.071285	-0.002147	0.000582
Chongming	0.036787	0.024035	0.029158	0.709444	0.049413	-0.000805	0.000190
Fengxian	0.073928	0.052907	0.065099	0.382110	0.106981	-0.007201	0.001499
Hongkou	0.070718	0.039691	0.046490	0.510429	0.080465	-0.004648	-0.000313

[75]:
```
min_max_scaler = preprocessing.MinMaxScaler()
np_scaled = min_max_scaler.fit_transform(features_new_var)
newvar_scale = pd.DataFrame(np_scaled)
newvar_scale.columns = features_new_var.columns
```

Above two steps are the results of data transformation in 1984 and 2019. To identify robust and consistent clustering results, we merge them into a single one based on their common geographical units (see Table 13). The column names ended with '_x' and '_y' represent features extracted in 1984 and 2019, respectively. This table is the one prepared for the final k-mean clustering analysis. The dominant parameter in k-means clustering is the number of clusters (i.e., k), determining the optimal numbers of clusters is therefore a fundamental issue. We select a direct and popular elbow method as an example to assess the resulting partitions, testing nine different solutions varying k from 2 to 10. Basically, the idea of elbow method is to define clusters to minimise the total intra-cluster variation or total within-cluster sum of square (WSS). The optimal number can be determined by plotting the curve of WSS according to different k clusters and the location of a bend is considered as an indicator of the appropriate number for k.

[76]:
```
merged_var = pd.merge(oldvar_scale, newvar_scale, left_index = True, right_index = True)
merged_var.head().style.set_caption('Table 13: Integrated preprocessed features
 ... identified in 1984 and 2019 seperately')
```

[76]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081e1b6d8>
```

[77]:
```
# elbow analysis
cluster_range = range( 2, 11 )
cluster_errors = []

for num_clusters in cluster_range:
        clusters = KMeans( num_clusters )
        clusters.fit( merged_var)
        cluster_errors.append( clusters.inertia_ )
clusters_df = pd.DataFrame( { "num_clusters":cluster_range,
                              "cluster_errors": cluster_errors } )
plt.figure(figsize=(12,6))
plt.title('Fig.6: Elbow method to determine the optimal k for k-mean clustering',y=-0.2)
plt.plot( clusters_df.num_clusters, clusters_df.cluster_errors, marker = "o" )
```

[77]:
```
[<matplotlib.lines.Line2D at 0x1f0817c2550>]image/png<Figure size 864x432 with 1 Axes>
```

Table 12: Four types of features (21 in total) identified in 2019

Name	h_mean	s_mean	v_mean	h_std	s_std	v_std	h_skew
Baoshan	0.231070	0.035865	0.638941	0.297847	0.052048	0.180189	0.336779
Changning	0.231849	0.031294	0.649237	0.304689	0.048471	0.167002	0.344394
Chongming	0.157306	0.016473	0.742402	0.282495	0.031843	0.101771	0.345289
Fengxian	0.295539	0.086197	0.605431	0.302731	0.097695	0.243596	0.329388
Hongkou	0.239944	0.037582	0.638613	0.303830	0.055047	0.182938	0.339615

Name	s_skew	v_skew	LBP0	LBP1	LBP2	LBP3	LBP4
Baoshan	0.067866	0.195591	0.043059	0.047740	0.040768	0.058617	0.068075
Changning	0.062974	0.182483	0.042641	0.050118	0.037668	0.051370	0.058422
Chongming	0.043360	0.119800	0.025468	0.029637	0.023332	0.035217	0.047621
Fengxian	0.115702	0.257234	0.051206	0.061050	0.052882	0.081410	0.097157
Hongkou	0.070552	0.197624	0.047032	0.054172	0.042940	0.054392	0.055014

Name	LBP5	LBP6	LBP7	LBP8	LBP9	veg_mean	builtup_mean
Baoshan	0.056809	0.037410	0.045565	0.524189	0.077767	-0.001801	0.001938
Changning	0.061528	0.035235	0.047082	0.539452	0.076482	-0.001515	0.000774
Chongming	0.036412	0.023590	0.028650	0.704424	0.045649	-0.000705	0.000318
Fengxian	0.079233	0.051521	0.060437	0.369931	0.095174	-0.008185	-0.000408
Hongkou	0.068051	0.038789	0.048937	0.507320	0.083353	-0.002057	-0.000277

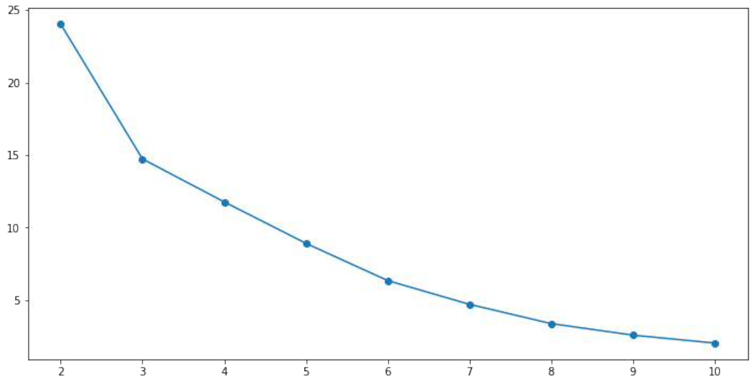


Figure 6: Elbow method to determine the optimal k for k-mean clustering

Figure 6 indicates that 2 and 6 (i.e. knee in the plot) can be the optimal numbers of k clusters for the features extracted from both years of satellite imagery. Considering the context of the paper, the number of 6 is finally assigned to k to fit the kmeans clustering model, varying labels are subsequently matched to features dataset.

[78]:
```
np.random.seed(0)
k = 6
cls = pd.Series(KMeans(n_clusters=k, max_iter = 1000, n_init = 1000,
                random_state = 24).fit_predict(merged_var))
```

After implementing k-means clustering on our constructed dataset, the label of each cluster is assigned to the last columns of data for further interpretation (as shown in Table 14).

[79]:
```
# Assign the each cluster number to the merged data
merged_var = merged_var.assign(lbls=cls)
merged_var.index = features_old_var.index
# last columns represent class labels
merged_var.head().style.set_caption('Table 14: Assign cluster number to each
 ... administritive area')
```

[79]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081e58550>
```

Table 13: Integrated preprocessed features identified in 1984 and 2019 seperately

	h_mean_x	s_mean_x	v_mean_x	h_std_x	s_std_x	v_std_x	h_skew_x	s_skew_x	v_skew_x	LBP0_x	LBP1_x	LBP2_x	LBP3_x	LBP4_x
0	0.636975	0.359694	0.281144	1.000000	0.422465	0.580121	1.000000	0.417053	0.568198	0.286043	0.328943	0.588494	0.717523	0.837853
1	0.363843	0.354334	0.396450	0.295018	0.453664	0.504220	0.000000	0.443305	0.483028	0.316677	0.441716	0.627346	0.537772	0.441710
2	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.081547	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	1.000000	0.990530	0.000000	0.906183	1.000000	0.987872	0.633860	1.000000	0.987806	0.915923	1.000000	0.915397	0.792921	0.738052
4	0.515153	0.480226	0.329302	0.685631	0.594329	0.597572	0.662480	0.590183	0.573137	0.499704	0.564467	0.676601	0.528034	0.370871

	LBP5_x	LBP6_x	LBP7_x	LBP8_x	LBP9_x	veg_mean_x	builtup_mean_x	h_mean_y	s_mean_y	v_mean_y	h_std_y	s_std_y	v_std_y	h_skew_y
0	0.613055	0.564716	0.332298	0.714937	0.032829	0.647652	0.766995	0.442195	0.278123	0.301061	0.463123	0.308221	0.547531	0.504942
1	0.551933	0.467707	0.395850	0.744082	0.042063	0.654899	0.764480	0.446860	0.212560	0.370616	0.659603	0.254002	0.455458	0.773771
2	0.000000	0.000000	0.000000	1.000000	0.000000	0.790941	0.730085	0.000000	0.000000	1.000000	0.022307	0.002006	0.000000	0.805338
3	0.741784	1.000000	1.000000	0.507310	0.110712	0.142243	0.844884	0.828667	1.000000	0.074677	0.603370	1.000000	0.990252	0.244056
4	0.677669	0.542270	0.482237	0.700451	0.059718	0.401200	0.686044	0.495392	0.302755	0.298842	0.634913	0.353666	0.566728	0.605083

	s_skew_y	v_skew_y	LBP0_y	LBP1_y	LBP2_y	LBP3_y	LBP4_y	LBP5_y	LBP6_y	LBP7_y	LBP8_y	LBP9_y	veg_mean_y	builtup_mean_y
0	0.338747	0.541837	0.526221	0.472274	0.548142	0.506561	0.410501	0.406220	0.494801	0.499487	0.726866	0.060848	0.856451	1.000000
1	0.271126	0.448129	0.513894	0.534312	0.450686	0.349686	0.216767	0.500209	0.416926	0.544291	0.749995	0.058415	0.893850	0.699181
2	0.000000	0.000000	0.007423	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	1.000000	0.581435
3	1.000000	0.982533	0.766492	0.819500	0.928936	1.000000	0.994141	0.852811	1.000000	0.938629	0.493097	0.093827	0.020274	0.393903
4	0.375880	0.556374	0.643383	0.640062	0.616412	0.415111	0.148374	0.630101	0.544153	0.599054	0.701302	0.071432	0.822844	0.427538

5 Interpretation

To understand the analysis result, the mean of each feature across each cluster can be calculated to uncover the feature differences among clusters. A categorical bar-plot shown below presents how the average of all features changed between 1984 and 2019. Besides, a choropleth map is created to visualise the spatial distribution of categories/clusters by varying colours.

[80]:
```
# calculate the mean of features for each class
k6_mean = merged_var.groupby('lbls').mean()
k6_mean.style.set_caption('Table 15: Mean values of each feature at each cluster for
... different years')
```

[80]:
```
text/html<pandas.io.formats.style.Styler at 0x1f081654b00>
```

Table 15 displays the mean values of all features in two years at varying groups. For more interpretability, a few data munging steps are required to generate visual representations.

[81]:
```
# Rearrange our data in a way that every row is one feature in a class
k6_mean = k6_mean.stack()
k6_mean.head()
```

[81]:
```
lbls
0    h_mean_x     0.803863
     s_mean_x     0.749195
     v_mean_x     0.146109
     h_std_x      0.895068
     s_std_x      0.749907
dtype: float64
```

[82]:
```
# convert multi-indices into single index

k6_mean = k6_mean.reset_index()
# renmae the columns
k6_mean = k6_mean.rename(columns = {'lbls': 'Class','level_1': 'Features', 0: 'Values'})
# rename feature names in Feature column
old  = k6_mean.loc[k6_mean['Features'].str.contains('x') ==  True, :]
new  = k6_mean.loc[k6_mean['Features'].str.contains('y') ==  True, :]
# add a new column to represent time
old = old.assign(Time = 1984)
new = new.assign(Time = 2019)
# remove '_x' and '_y' in the table to make feature names for both years are the same
old['Features'] = old['Features'].str.replace('_x', '')
new['Features'] = new['Features'].str.replace('_y', '')
```

[83]:
```
# create a new dataframe to store the mean of each feature each cluster with time

data = pd.concat([old,new])
data.head().style.set_caption('Table 16: Tidy table represents mean values of features
 ... for each cluster at different years')
```

[83]:
```
text/html<pandas.io.formats.style.Styler at 0x1f08cee1d68>
```

Table 16 reveals different categorical information, with each row represents the number of class, the feature name, the mean value of the feature and the year when the feature is extracted. We can then visualise this table in the bar-plot in Figure 7 to understand the pattern from image features.

[84]:
```
# visualise the distribution of mean values by features, class and time

g = sns.catplot( data = data, x = 'Features', y = 'Values',row = 'Class',
            hue = 'Time',kind = 'bar', aspect = 5, height = 3, palette = 'Accent')
g.fig.suptitle('Fig.7: Visual representation of patterns extracted from k-mean
 ... clustering', y = -0.1, fontsize = 18)
```

[84]:
```
Text(0.5, -0.1, 'Fig.7: Visual representation of patterns extracted from k-mean
 clustering')
image/png<Figure size 1141.5x1296 with 6 Axes>
```

Table 14: Assign cluster number to each administrative area

Name	h_mean_x	s_mean_x	v_mean_x	h_std_x	s_std_x	v_std_x	h_skew_x	s_skew_x	v_skew_x	LBP0_x	LBP1_x	LBP2_x	LBP3_x	LBP4_x
Baoshan	0.636975	0.359694	0.281144	1.000000	0.422465	0.580121	1.000000	0.417053	0.568198	0.286043	0.328943	0.588494	0.717523	0.837853
Changning	0.363843	0.354334	0.396450	0.295018	0.453664	0.504220	0.000000	0.443305	0.483028	0.316677	0.441716	0.627346	0.537772	0.441710
Chongming	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.081547	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Fengxian	1.000000	0.990530	0.000000	0.906183	1.000000	0.987872	0.633860	1.000000	0.987806	0.915923	1.000000	0.915397	0.792921	0.738052
Hongkou	0.515153	0.480226	0.329302	0.685631	0.594329	0.597572	0.662480	0.590183	0.573137	0.499704	0.564467	0.676601	0.528034	0.370871

Name	LBP5_x	LBP6_x	LBP7_x	LBP8_x	LBP9_x	veg_mean_x	builtup_mean_x	h_mean_y	s_mean_y	v_mean_y	h_std_y	s_std_y	v_std_y	h_skew_y
Baoshan	0.613055	0.564716	0.332298	0.714937	0.032829	0.647652	0.766995	0.442195	0.278123	0.301061	0.463123	0.308221	0.547531	0.504942
Changning	0.551933	0.467707	0.395850	0.744082	0.042063	0.654899	0.764480	0.446860	0.212560	0.370616	0.659603	0.254002	0.455458	0.773771
Chongming	0.000000	0.000000	0.000000	1.000000	0.000000	0.790941	0.730085	0.000000	0.000000	1.000000	0.022307	0.002006	0.000000	0.805338
Fengxian	0.741784	1.000000	1.000000	0.507310	0.110712	0.142243	0.844884	0.828667	1.000000	0.074677	0.603370	1.000000	0.990252	0.244056
Hongkou	0.677669	0.542270	0.482237	0.700451	0.059718	0.401200	0.686044	0.495392	0.302755	0.298842	0.634913	0.353666	0.566728	0.605083

Name	s_skew_y	v_skew_y	LBP0_y	LBP1_y	LBP2_y	LBP3_y	LBP4_y	LBP5_y	LBP6_y	LBP7_y	LBP8_y	LBP9_y	veg_mean_y	builtup_mean_y	lbls
Baoshan	0.338747	0.541837	0.526221	0.472274	0.548142	0.506561	0.410501	0.406220	0.494801	0.499487	0.726866	0.060848	0.856451	1.000000	1
Changning	0.271126	0.448129	0.513894	0.534312	0.450686	0.349686	0.216767	0.500209	0.416926	0.544291	0.749995	0.058415	0.893850	0.699181	1
Chongming	0.000000	0.000000	0.007423	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	1.000000	0.581435	2
Fengxian	1.000000	0.982533	0.766492	0.819500	0.928936	1.000000	0.994141	0.852811	1.000000	0.938629	0.493097	0.093827	0.020274	0.393903	3
Hongkou	0.375880	0.556374	0.643383	0.640062	0.616412	0.415111	0.148374	0.630101	0.544153	0.599054	0.701302	0.071432	0.822844	0.427538	1

Table 15: Mean values of each feature at each cluster for different years

lbls	h_mean_x	s_mean_x	v_mean_x	h_std_x	s_std_x	v_std_x	h_skew_x	s_skew_x	v_skew_x	LBP0_x	LBP1_x	LBP2_x	LBP3_x	LBP4_x
0	0.803863	0.749195	0.146109	0.895068	0.749907	0.834330	0.633322	0.720249	0.823005	0.454389	0.565946	0.821528	0.973220	0.949141
1	0.426195	0.381664	0.384955	0.543494	0.481686	0.523486	0.450601	0.478586	0.504149	0.382849	0.444543	0.574545	0.510795	0.408613
2	0.113097	0.088470	0.837305	0.210939	0.118478	0.127151	0.316506	0.107534	0.124469	0.070301	0.089850	0.162624	0.139945	0.089317
3	0.949856	0.995265	0.004559	0.876415	0.991048	0.993936	0.676387	0.975688	0.993903	0.882717	0.957562	0.957699	0.873867	0.798240
4	0.721478	0.979475	0.114017	0.567528	0.998762	0.933619	0.244557	0.971796	0.925791	1.000000	0.990273	0.845726	0.668671	0.407605
5	0.471719	0.427622	0.373358	0.634773	0.554207	0.559062	0.554391	0.560426	0.534555	0.426212	0.502654	0.647752	0.516126	0.345374

lbls	LBP5_x	LBP6_x	LBP7_x	LBP8_x	LBP9_x	veg_mean_x	builtup_mean_x	h_mean_y	s_mean_y	v_mean_y	h_std_y	s_std_y	v_std_y	h_skew_y
0	0.714188	0.767025	0.567415	0.603671	0.062757	0.336476	0.964021	0.621423	0.582989	0.125378	0.621145	0.621750	0.781063	0.466975
1	0.491368	0.480169	0.415924	0.747533	0.045784	0.579325	0.782732	0.400459	0.233562	0.370746	0.449325	0.279210	0.479726	0.520610
2	0.114726	0.140034	0.083018	0.941070	0.009137	0.739326	0.768474	0.081734	0.042801	0.835523	0.144022	0.062326	0.114256	0.757332
3	0.870892	0.996516	0.943707	0.495601	0.106946	0.071121	0.808680	0.914333	0.849361	0.087764	0.539909	0.845232	0.995126	0.122028
4	0.847049	0.819912	0.961713	0.538442	0.114071	1.000000	0.000000	0.883471	0.529215	0.000000	1.000000	0.546832	0.871779	0.698488
5	0.299768	0.528165	0.568006	0.000000	1.000000	0.549489	0.733977	0.428816	0.207932	0.329358	0.435917	0.246778	0.505767	0.409755

lbls	s_skew_y	v_skew_y	LBP0_y	LBP1_y	LBP2_y	LBP3_y	LBP4_y	LBP5_y	LBP6_y	LBP7_y	LBP8_y	LBP9_y	veg_mean_y	builtup_mean_y
0	0.629142	0.768836	0.559714	0.585206	0.760955	0.841492	0.838157	0.722830	0.794093	0.652172	0.600350	0.070401	0.714016	0.835832
1	0.300671	0.470766	0.465958	0.451277	0.510348	0.446913	0.309060	0.470514	0.466278	0.466282	0.748782	0.053443	0.852035	0.622854
2	0.077196	0.114218	0.050600	0.054194	0.140384	0.172060	0.158486	0.147885	0.115339	0.054215	0.939642	0.006619	0.959421	0.572280
3	0.870699	0.991266	0.883246	0.909750	0.964468	0.940470	0.808656	0.863404	0.987573	0.969315	0.488089	0.104292	0.010137	0.196952
4	0.575903	0.866917	0.981285	0.905575	0.907557	0.660840	0.487015	1.000000	0.950033	0.913027	0.521429	0.107485	0.713309	0.986139
5	0.280930	0.499205	0.621626	0.525659	0.570189	0.416572	0.116072	0.221324	0.445842	0.718360	0.000000	1.000000	0.895112	0.568988

Figure 7: Visual representation of patterns extracted from k-mean clustering

[85]:

```
# plot clustering results for two different years
f, ax = plt.subplots(1, figsize=(10, 12))
# plot cluster results
poly = poly.drop('coords', axis = 1)
poly.assign(lbls=cls)\
     .plot(column='lbls', categorical=True, linewidth=1, alpha=0.5, ax=ax,
           legend = True,cmap = 'Accent', edgecolor = 'black')
# add labels for geographical units
poly['coords']=poly['geometry'].apply(lambda x:x.representative_point().coords[:])
poly['coords']=[coords[0] for coords in poly['coords']]
for idx, row in poly.iterrows():
    ax.annotate(text=row['Name'],xy=row['coords'],va='center',ha='center',
                alpha = 0.8, fontsize = 10)
plt.title('Fig.8: Spatial distribution of classification results', y=-0.01)
# remove axes and set aspect ratio so that the data units are the same in every direction
ax.axis('off')
ax.axis('equal')
```

[85]:

```
(290053.0696196473, 407301.6741094636, 3389866.639388826, 3533566.430983904)
image/png<Figure size 720x864 with 1 Axes>
```

Table 16: Tidy table represents mean values of features for each cluster at different years

	Class	Features	Values	Time
0	0	h_mean	0.803863	1984
1	0	s_mean	0.749195	1984
2	0	v_mean	0.146109	1984
3	0	h_std	0.895068	1984
4	0	s_std	0.749907	1984

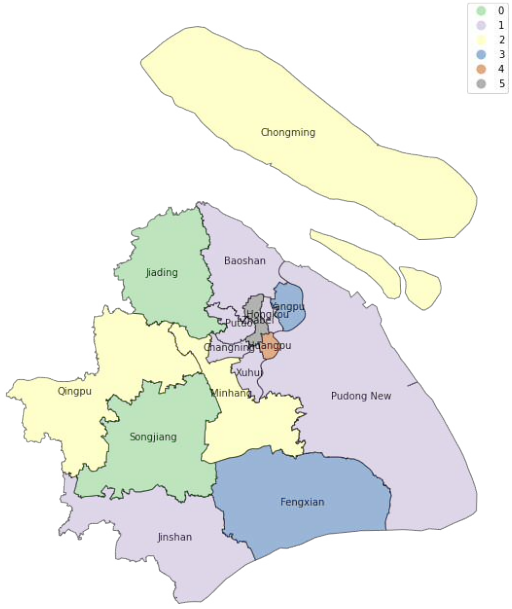


Figure 8: Spatial distribution of classification results

From Figures 7 and 8 we can see a few striking differences across clusters, or classes. For class 4, only one administrative area (i.e. Huangpu area) is grouped, displayed in the middle of north-east areas. The mean values for this class are mostly high in both years except a couple of features such as v_mean, LBP4 and LBP9 features. The brightness (v_mean) for this area is highly low and it became completely black over time. H_mean value is high in both years, demonstrating that the dominating colour is blue, which represent water. This corresponds to the famous area of The Bund, with its river skyline, which is part of this polygon. The vegetation built-up features indicate that this area has experienced a remarkable change, from more vegetation and few buildings to less vegetation and completely constructed/urbanisation.

Class 0 and Class 1 are relatively consistent compared to other classes, implying that the urban areas in purple and green colours almost remained unchanged during the past 35 years. Besides, these two classes have similar transformation such as more vegetation coverage and less buildings for the current year of 2019. However, Class 0 has more brightness and more green colour based on v_mean, h_mean and veg_mean features, and Class 1 has higher h_mean, h_std, h_skew and built-up_mean, implying these two areas have water covered and were highly constructed.

Class 2 distributed at north and middle-west areas in the map, which is extremely diverse and unique among all categories. It has the highest brightness features and LBP8 texture features, while the rest mean values of colour and texture features are highly low, especially for LBP9 where almost zero values in both years. The values for h_mean, s_mean and v_mean display that the primary colour for these areas is red with little grey and much brightness, representing that these areas include more bare ground or soil and thus probably rural areas. Adversely, Class 5 has zero values for LBP8 but highest values for LBP9 in both years. It contains only one administrative area (i.e. Zhabei area), surrounded by Class 4 and Class 0. Similarly, the area in Class 5 has more vegetation but slightly less built-up areas over the past years. Class 3 contains two areas distributed at the south and surrounded by Class 1 from the map. The feature values in Class 3 are mostly extremely high, while the veg_mean and built-up_mean for current year are the least, thus indicating that these areas have more water over the time.

6 Conclusion

Urbanisation has significantly changed the interaction between humans and the surrounding environment, which poses new challenges in a multitude of fields including construction and city planning, hazard mitigation or disease control. It is essential to quantify and assess urbanisation over time to enable policy makers and planners to make informed decisions about future urban changes. The sustainability of urban spaces will become particularly important in the light of future climate change. Satellite imagery could play a vital role in assessing cities for their livability by i.e. quantifying the greenspace to built environment ratio. This notebook shows the potential of open source satellite imagery to exploring urban changes and proposes a simple method framework for automatic data collection and features extraction to determine urbanisation over time using Python as a tool.

References

Barsi JA, Lee K, Kvaran G, Markham BL, Pedelty JA (2014a) The spectral response of the Landsat-8 operational land imager. *Remote Sensing* 6: 10232–10251. CrossRef.

Barsi JA, Schott JR, Hook SJ, Raqueno NG, Markham BL, Radocinski RG (2014b) Landsat-8 thermal infrared sensor (TIRS) vicarious radiometric calibration. *Remote Sensing* 6: 11607–11626. CrossRef.

Burchfield M, Overman HG, Puga D, Turner MA (2006) Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics* 121: 587–633. CrossRef.

Giada S, De Groeve T, Ehrlich D, Soille P (2003) Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania. *International Journal of Remote Sensing* 24: 4251–4266. CrossRef.

Glaeser E, Henderson JV (2017) Urban economics for the developing world: An introduction. *Journal of Urban Economics* 98: 1–5. CrossRef.

Ibrahim MR, Haworth J, Cheng T (2020) Understanding cities with machine eyes: A review of deep computer vision in urban analytics. *Cities* 96. CrossRef.

Keen N (2005) Color moments. School of informatics, University of Edinburgh

Kit O, Lüdeke M (2013) Automated detection of slum area change in Hyderabad, India using mulittemporal satellite imagery. *Journal of Photogrammetry and Remote Sensing* 83: 130–137. CrossRef.

Knight EJ, Kvaran G (2014) Landsat-8 operational land imager design, characterization and performance. *Remote Sensing* 6: 10286–10305. CrossRef.

Kohli D, Sliuzas R, Stein A (2016) Urban slum detection using texture and spatial metrics derived from satellite imagery. *Journal of Spatial Science* 61: 405–426. CrossRef.

Ministry of Civial Affairs of the People's Republic of China (2018) Change of administrative divisions at or above the county level. Available at: http://202.108.98.30/description?dcpid=1 [Accessed 10 Oct. 2019]

NASA (2000) Normalized difference vegetation index (NDVI). Available at: https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php [Accessed 30 Oct. 2019]

NASA (2019) Landsat science. Available at: https://landsat.gsfc.nasa.gov/ [Accessed 10 Sep. 2019]

Ojala T, Pietikäinen M, Harwood D (1996) A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* 19: 51–59. CrossRef.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J (2011) Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12: 2825–2830

Roy DP, Wulder MA, Loveland TR, Woodcock CE, Allen RG, Anderson MC, Helder D, Irons JR, Johnson DM, Kennedy R, Scambos TA, Schaaf CB, Schott JR, Sheng Y, Vermote EF, Belward AS, Bindschadler R, Cohen WB, Gao F, Hipple JD, Hostert P, Huntington J, Justice CO, Kilic A, Kovalskyy V, Lee ZP, Lymburner L, Masek JG, McCorkel J, Shuai Y, Trezza R, J. Vogelmann J, R.H. Wynne RH, Zhu Z (2014) Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment* 145: 154–172. CrossRef.

United Nations (2019) World urbanization prospects 2018: Highlights. United Nations, Department of Economic and Social Affairs, Population Division (ST/ESA/SER.A/421)

Zha Y, Gao J, Ni S (2003) Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing* 24: 583–594. CrossRef.

Letters

REGION
The Journal of ERSA
Powered by WU

Volume 7, Number 2, 2020, L1–L8
DOI: 10.18335/region.v7i2.314

journal homepage: region.ersa.org

Macroeconomic determinants of Port and Douro wine exports: An econometric approach

Anthony Macedo, Sofia Gouveia, João Rebelo[1]

[1] University of Trás-os-Montes and Alto Douro, Vila Real, Portugal

Received: 2 April 2020/Accepted: 2 November 2020

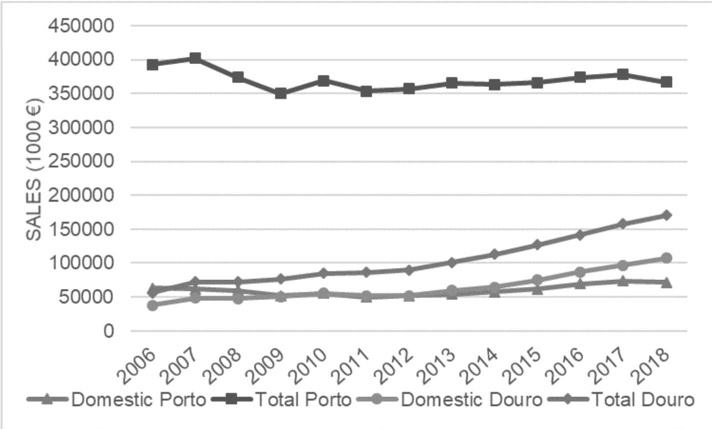
Abstract. This study analyses recent trends (2006-2018) in Port and Douro wine exports and estimates their macroeconomic determinants. The results of the gravity equation reveal that the gross domestic product in importing countries is the most important export determinant of both wines, and Douro wine exports are negatively affected by the distance to the destination country, but positively influenced by sharing a common language and the level of wine production in importing countries. Therefore, in order to increase exports, the industry's strategic decision-makers should pay special attention to the markets in wealthier countries or with a high potential for economic growth, taking into account issues such as market access, adaptation of the market to wine consumption, and regulation.

Key words: International trade, Gravity model, Poisson pseudo-maximum likelihood, Wine, Demarcated Douro region

1 Introduction

During the last 50 years, the international wine market has dramatically changed. New competitors appeared on the international scene and joined the top wine producers and exporters. Initially, the production and exports were concentrated in a small number of European countries, the Old World, namely France, Italy, Spain, Portugal and Germany. Since the 1950s, a new set of wine producing and exporting countries emerged, the New World (South Africa, Argentina, Australia, United States of America, and New Zealand). In recent years there appeared a second new group of wine producers, the New New World, including China, India, Brazil, and the countries around the Black Sea (Cardebat 2019). It is expected that the relatively new hierarchy of wine-producing countries could remain in place for the ten next years, but a process of convergence between Old and New World seems to be under way (Holmes, Anderson 2017, Morrison, Rabellotti 2017, Bargain et al. 2018), with consequences for wine output and also in the market of inputs, including direct investment and labour market.

Portugal is a historical wine-producing country, with a rich history, where wine is inseparable from the culture, heritage, terroirs and economy of the country. Representative of this history is the wine from the Demarcated Douro Region (DDR), demarcated since 1756. This region fits the typical terroir model and it is the largest and the most heterogeneous wine mountain region in the world, with high production costs and low productivity. Essentially the two main types of wines produced in this region are Port



Source: Authors' computation from IVDP data.

Figure 1: The evolution of total Port and Douro wine sales and domestic market, 2006-2018

Table 1: Main destination countries of Port wine exports, in real value (million €), 2006 and 2018

R	Country	2006	%	Country	2018	%
1	France	92	28	France	75	25
2	United Kingdom	52	16	United Kingdom	44	15
3	Netherlands	49	15	Netherlands	38	13
4	Belgium	42	13	USA	36	12
5	USA	35	11	Belgium	31	11
6	Canada	17	5	Denmark	15	5
7	Germany	10	3	Germany	12	4
8	Denmark	7	2	Canada	11	4
9	Spain	6	2	Spain	4	1
10	Switzerland	4	1	Switzerland	4	1
	Top 10	313	95	Top 10	240	91
	World	330	100	World	295	100

Source: Authors' computation from IVDP data.
Notes: R = ranking; % = share of total Port wine exports; real values computed using export deflator of goods from Banco de Portugal; USA = United States of America.

(fortified) and Douro (still) wines, each one presenting a completely different history and market position. With a total sale of 6,389 million euro in the period 2006-2018, Port wine represented 75% and Douro wine 21%, these two wines are the core of the region.

Port wine has a successful presence in the international markets for more than three centuries, currently exporting almost 80% of its total sales, although suffering declining sales in the last decade. Unlike the Port wine, the commercial history of Douro is much shorter - just thirty years, it is a new entrant in the mature international market, and its exports represent almost 40% of the total sales. Both wines are sold (on average) at a lower price than the average price of wine in the EU (Hogg, Rebelo 2018).

As shown in Figure 1, the recent trend of Port wine sales reveals a slight decrease of 7% in total sales from 2006 to 2018.

The share of domestic sales for Port wine increased from 16% to 20% between 2006 and 2018, becoming in the last year the most important market for the first time in many years, replacing France as the main market. This was probably encouraged by the boom in tourism which has occurred in Portugal, especially in the Porto-Gaia region. The growth in domestic sales is also explained intrinsically by a worse performance in foreign markets (Table 1). Comparing 2018 to 2006, exports decreased from 330 to 295 million euro (in real terms).

Table 2: Main destination countries of Douro wine exports, in real value (million €), 2006 and 2018

R	Country	2006	%	Country	2018	%
1	Canada	3,1	17	Canada	12,1	19
2	USA	2,1	12	United Kingdom	8,2	13
3	Brazil	1,9	11	Brazil	6,5	10
4	France	1,5	8	Switzerland	5,8	9
5	Germany	1,5	8	Germany	4,8	8
6	Switzerland	1,4	8	USA	4,5	7
7	Norway	1,1	6	Angola	2,9	5
8	United Kingdom	1,1	6	France	2,5	4
9	Angola	0,8	4	Belgium	2,5	4
10	Belgium	0,7	4	China	2,3	4
	Top 10	15	84	Top 10	52	82
	World	18	100	World	64	100

Source: Authors' computation from IVDP data.
Notes: R = ranking; % = share of total Douro wine exports; real values computed using export deflator of goods from Banco de Portugal; USA = United States of America.

Compared with Port wine, Douro wine sales presented a different trend (Figure 1). Between 2006 and 2018, total sales increased 205% and domestic sales 183%. As it can be observed in Table 2, during the period of study, exports of Douro wine increased from 18 million euro (in real terms) to 64 million euro.

The comparison of Tables 1 and 2 highlights the high concentration of Port wine exports to European countries (USA and Canada are the only non-European countries in the top 10), while Douro wine exports are more dispersed, with the particularity of Portuguese-speaking countries being important destinations.

Since the cost of grape production are high in the DDR, due to high labour costs and low yields, and adding the increased competition in the international wine market, the DDR wine industry faces a double challenge. On the one hand, to break the negative trend of Port wine sales and, on the other hand, to achieve a higher valuation of both wines (Hogg, Rebelo 2018). In this way, a better knowledge of the macroeconomic determinants of international demand for Port and Douro wine may contribute to the achievement of these goals.

Thus, based on data provide by IVDP, a panel econometric model is applied to a gravity equation, considering the period 2006-2018 and 80 main destination countries of both wines. This work also contributes to the enrichment of the wine economics literature on international trade by comparing the exports determinants of two wines produced in the same region, through the use of two estimation approaches.

After this introduction, the next section is dedicated to the theoretical framework. In Section 3 data and results are presented. Finally, section 4 concludes the study.

2 Theoretical framework: The gravity equation

As countries or sectors have turned toward export-led growth strategies, the estimation of export equations has become more relevant to support policy decisions. Bayar (2018) grouped export equations into four levels of analysis, from macro to micro: aggregated level exports, country-level, sector-level and firm-level analysis or combinations of them, each one requiring the availability of specific data. The research included in this work is a mix between country-level and sector-level analysis. On the one hand, as in most country-level analysis, the gravity equation is formulated with explanatory variables representing trade frictions or trade facilitators between countries. On the other hand, it is focused on a single product, examining sector-specific characteristics, as in sector-level analysis.

Gravity equations started to be developed in the 1960's, however it took some time for the link between the model and trade theory to be demonstrated by several authors (Anderson 1979, Anderson, van Wincoop 2003, Bergstrand 1989, Eaton, Kortum 2002, Helpman 1987, Helpman, Krugman 1985). Starting from a model explaining that trade flows are negatively affected by bilateral distance between two trade partners whilst being positively affected by the respective economic masses, researchers then developed the gravity model to address more trade frictions/facilitators.

The literature analysing world wine trade with the gravity model (Castillo et al. 2016, Dal Bianco et al. 2016, Dascal et al. 2002, Gouveia et al. 2018, Macedo et al. 2019, 2020) is the base for the following export function of Port and Douro wines:

$$exports_{w,it} = f(gdp_{it}, er_{it}, eu_{it}, ave_{w,it}, prod_{it}, dist_i, land_i, lang_i) \tag{1}$$

Where the dependent variable *exports* represents the exports of wine w (Port or Douro) to country i in year t. Among the explanatory variables there are typical variables of gravity equations, such as the GDP of country i in year t (gdp), the average annual exchange rate of country i's currency in relation to the euro in year t (er), a dummy variable equal to 1 if country i is member of the European Union (EU) in year t or 0 otherwise (eu), the geographical distance between Lisbon and the capital city of country i ($dist$), a dummy variable equal to 1 if country i is landlocked ($land$), and a dummy variable equal to 1 if country i has a common official language with Portugal ($lang$). Because the research is focused in each wine, considered equivalent to a single sector, two explanatory variables were specifically included to represent sectoral characteristics (Bayar 2018), which are the ad valorem equivalent tariff applied by country i to wine w (Port or Douro) imports in year t (ave) and the litres of wine produced by the importing country in year t ($prod$).

In order to overcome problems of specification of the gravity equation, multilateral resistance terms should be considered (Anderson, van Wincoop 2003) and, following Head, Mayer (2014), two different approaches are used to deal with this issue. The first one is derived from "remoteness indexes" (Baldwin, Harrigan 2011, Wei 1996) and calculated through weighting bilateral distance with the degree of trade openness of the importing country ($distw$), so it is expected that the costs associated to transport decreases as openness increases, due to economies of scale in transport.

The alternative solution to proxy multilateral resistance terms is to consider countries fixed effects (Dal Bianco et al. 2016, Dal Bianco et al. 2017, Macedo et al. 2020, Santeramo et al. 2019), not requiring strong structural assumptions on the underlying model (Head, Mayer 2014). However, its main drawback is not allowing estimation of the coefficients for time-invariant variables (such as distance, landlockedness, and common language).

Another well-known issue in gravity equations is the "zero problem". For a long time, gravity equations were estimated with linear methods using the logarithm of exports as a dependent variable, resulting in an undefined logarithm of zero when there was no trade between two countries. To overcome this drawback, several approaches were put forward, the seminal one being proposed by Silva, Tenreyro (2006). They avoid the log-form using a multiplicative form and, additionally, recommend a non-linear estimation method, preferably Poisson pseudo-maximum likelihood (PPML).

Using the PPML method, the following gravity equations for Port and Douro wine exports are estimated:

$$\begin{aligned} exports_{w,it} &= \exp\left[\beta_1 \ln gdp_{it} + \beta_2 \ln er_{it} + \beta_3 \ln(ave_{w,it}+1) + \beta_4 eu_{it} + \right. \\ &\quad \left. \beta_5 \ln prod_{it} + \beta_6 land_i + \beta_7 lang_i + \beta_8 \ln distw_{it} + \varphi_t + u_{it}\right] \qquad (2) \\ exports_{w,it} &= \exp\left[\beta_1 \ln gdp_{it} + \beta_2 \ln er_{it} + \beta_3 \ln(ave_{w,it}+1) + \beta_4 eu_{it} + \right. \\ &\quad \left. \beta_5 \ln prod_{it} + \omega_i + \varphi_t + u_{it}\right] \qquad (3) \end{aligned}$$

Equation (2) stands for the method considering time-invariant variables and equation (3) includes importing countries fixed effects (ω_i). Time fixed effects (φ_t) are included, and the statistical error (u_{it}) is assumed to be identically and independently distributed. Explanatory variables are in the log-form, dummy variables being the exception.

Table 3: Data sources and descriptive statistics

Variables	Data Source	Unit	N	Mean	P50	SD	Min	Max
exports_{Port}	IVDP	10^3€	1032	3891	118	12218	0	87688
exports_{Douro}	IVDP	10^3€	1040	492	20	1247	0	11238
gdp	WDI	10^9€	1036	647	150	1779	0.1	17405
er	WDI	LCU per €	1040	565	6	3136	0.3	39857
ave_{Port}	Macmap	%	974	12.9	1.4	23.3	0	150
ave_{Douro}	Macmap	%	974	12.8	1.0	23.2	0	150
eu	EU	Binary	1040	0.3	0.0		0	1
prod	OIV	10^3hl	1040	3222	65	8935	0	54800
distw	WDI/CEPII	Weighted km	1036	15092	9394	15507	969	78030
lang	CEPII	Binary	1040	0.1	0		0	1
land	CEPII	Binary	1040	0.1	0		0	1

Source: Authors' computation.
Notes: IVDP = Instituto de Vinhos do Douro e do Porto (https://www.ivdp.pt/); WDI = World Development Indicators (https://databank.worldbank.org/source/world-development-indicators); Macmap = Market Access Map (https://www.macmap.org/en/query/customs-duties); OIV = International Organisation of Vine and Wine (http://www.oiv.int/en/statistiques/recherche); CEPII = (http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=6); hl = hectolitres; LCU = local currency unit; km = Kilometres N = number of observations; P50 = median; SD = standard deviation; Min = minimum value; Max = maximum value; For the binary variables the mean represents the percentage of observations equal to one.

3 Data and results

Table 3 introduces the data sources and the main descriptive statistics. The data for Port and Douro wine exports were extracted from the website of IVDP and they cover a sample of 80 importing countries from 2006 to 2018 (99% of total exports during the period of study).

Table 4 presents the results of the estimations of equation (2) in columns (I) and (III), and the results of the estimations of equation (3) in columns (II) and (IV). Time effects are statistically significant, which are considered using yearly dummy variables (omitted due to space considerations). Standard errors account for intra-group correlation. The Wald test was applied to compare the estimated coefficients for Port and Douro wines and the results ($\chi^2_{21} = 7412.70$ comparing columns I and III and $\chi^2_{17} = 7768.04$ comparing columns II and IV) suggest they are statistically different. Therefore, export determinants of Douro and Port wines are fairly different.

Starting from the analysis of gravity variables, GDP presents in the four models a positive and statistically significant coefficient and the magnitude of the coefficients is considerably higher in columns (I) and (III), which highlights the importance of including countries fixed effects to not overstate the effect of GDP. The positive effect of GDP is in line with the results of previous research, suggesting that wine exports increases with the size of the destination market.[1]

The distance, present in columns (I) and (III), is only statistically significant for Douro wine and has the expected negative effect. Ceteris paribus, this means that, on average, relatively more distant importing countries import less Douro wine due to higher transport costs. For Port wine, the estimated coefficient is not significant, converging with the results of Gouveia et al. (2018), which can be explained by the observation that the main importers of Port wine are European Union countries. Therefore, distance plays a lesser role for Port wine exports than for Douro wine, whose exports are more spread around the world.

Regarding the other time-invariant variables, the importing country being landlocked is not a statistically significant determinant for the exports of any of the wines. Sharing a common official language has a significant impact only for Douro wine, which can be supported by the good commercial relationship with former colonies, such as Angola

[1] This result is different from the one presented by Gouveia et al. (2018), in which GDP has not a statistically significant effect on Port wine exports, but this may be due to potential high correlation with another significant and positive regressor considered by the authors, the per capita GDP.

Table 4: Estimation of a gravity model for Port and Douro wine exports (in €), 2006-2018

Variables	(I) Port PPML	(II) Port PPML-FE	(III) Douro PPML	(IV) Douro PPML-FE
$\ln gdp_{it}$	0.903***	0.825***	1.022***	0.545**
	(0.235)	(0.183)	(0.351)	(0.271)
$\ln er_{it}$	0.055	0.042	0.026	-0.264
	(0.196)	(0.194)	(0.162)	(0.241)
$\ln(ave_{w,it}+1)$	0.051	0.051	-0.099	-0.095
	(0.102)	(0.105)	(0.113)	(0.111)
eu_{it}	0.519	0.533	0.044	0.100
	(0.386)	(0.390)	(0.428)	(0.442)
$\ln prod_{it}$	-0.019	-0.023	0.183*	0.130
	(0.056)	(0.056)	(0.094)	(0.102)
$\ln distw_{it}$	-0.173		-0.771***	
	(0.246)		(0.192)	
$land_i$	-0.085		-0.044	
	(0.555)		(1.231)	
$lang_i$	0.370		5.549***	
	(1.214)		(1.251)	
Constant	-8.004		-8.554	
	(5.208)		(7.976)	
Observations	963	963	970	970
Time effects'	151.15***	133.30***	410.50***	373.60***
significance	[0.000]	[0.000]	[0.000]	[0.000]

Source: Authors' computation.
Notes: Robust standard errors in ( ); *** p<0.01, ** p<0.05, * p<0.1; p values in [ ].

and Brazil, that allows Portuguese still wines to capture important market share. This result was also pointed to by Macedo et al. (2019). However, for historical reasons, these countries have never been important destination markets of Port wine, thus the estimated coefficient is not significant.

Concerning the results associated with wine production in the importing country, the estimated coefficients are almost all nonsignificant, with the exception of the barely statistically significant coefficient in column (III). For Port wine, following Macedo et al. (2020), this effect may be explained by the phenomenon that the top wine-producing countries do not find a valid substitute for Port wine in their domestic markets. This is also a possible explanation for Douro wine, but it is more likely that Douro wine arouses greater interest in more mature wine markets. As Macedo et al. (2019) suggest, taste for Douro wine is not homogeneous across the world, with some countries valuing quality more than others when they choose among distinct categories of Douro wines.

Regarding ad valorem equivalent tariffs, they do not have a statistically significant effect. This result was also highlighted by Gouveia et al. (2018) and Macedo et al. (2020), suggesting that the differentiating characteristics of these two wines more than compensate for customs duties. However, another explanation may lie in the fact that changes in tariffs have been scarce in recent years and fixed effects capture their impact. Additionally, neither the exchange rate nor the dummy variable for EU membership present statistically significant coefficients.

4 Conclusion

The results point to some discrepancy in export determinants between Douro and Port wines. The GDP of importing countries is a crucial determinant for both wines, with an increase in GDP provoking a positive variation of exports. Additionally, some inference can be made to point to specific determinants. For example, Douro wine is negatively influenced by the distance of the importing country and its marketability may increase in more mature wine markets or countries speaking Portuguese. Besides, exports of Port and Douro wine are not significantly influenced by the wine production in the importing country, exchange rates, tariffs, the importer being member of the EU, or the importer not having direct access to the sea.

From the results it can be concluded that decision-makers from the DDR wine industry can increase exports aimed at markets with high GDP or potential for economic growth. However, increasing international competitiveness can only be achieved via supporting market access, wine quality adaptation to the market, and appropriate regulation. As stated by Hogg, Rebelo (2018), to increase exports it is necessary to gain better knowledge of both markets and consumers, which can then be used to attract new consumers, and promote online sales through digital marketing tools. Further, it is necessary to create better and more effective regulation that will be able to address the market challenges.

References

Anderson JE (1979) A theoretical foundation for the gravity equation. *The American Economic Review* 69: 106–116. CrossRef.

Anderson JE, van Wincoop E (2003) Gravity with gravitas: A solution to the border puzzle. *The American Economic Review* 93: 170–192. CrossRef.

Baldwin R, Harrigan J (2011) Zeros, quality, and space: trade theory and trade evidence. *American Economic Journal: Microeconomics* 3: 60–88. CrossRef.

Bargain O, Cardebat JM, Chiappini R, Laffitte C (2018) Long-term prospects of the international wine trade. *Journal of Wine Economics* 13: 461–468. CrossRef.

Bayar G (2018) Estimating export equations: A survey of the literature. *Empirical Economics* 54: 629–672. CrossRef.

Bergstrand JH (1989) The generalized gravity equation, monopolistic competition, and factor proportions theory in international trade. *The Review of Economics and Statistics* 71: 143–153. CrossRef.

Cardebat JM (2019) Introduction: Does a national model exist which favors trade performance? In: Ugaglia A, Cardebat JM, Corsi A (eds), *The Palgrave Handbook of Wine Industry Economics*. Palgrave Macmillan, 455–461

Castillo JS, Villanueva EC, García-Cortijo MC (2016) The international wine trade and its new export dynamics (1988-2012): A gravity model approach. *Agribusiness* 32: 466–481. CrossRef.

Dal Bianco A, Boatto VL, Caracciolo F, Santeramo FG (2016) Tariffs and non-tariff frictions in the world wine trade. *European Review of Agricultural Economics* 43: 31–57. CrossRef.

Dal Bianco A, Estrella-Orrego MJ, Boatto VL, Gennari AJ (2017) Is Mercosur promoting trade? Insights from Argentinean wine exports. *Spanish Journal of Agricultural Research* 15: 1–10. CrossRef.

Dascal D, Mattas K, Tzouvelekas V (2002) An analysis of EU wine trade: A gravity model approach. *International Advances in Economic Research* 8: 135–147. CrossRef.

Eaton J, Kortum S (2002) Technology, geography, and trade. *Econometrica* 70: 1741–1779. CrossRef.

Gouveia S, Rebelo J, Lourenço-Gomes L (2018) Port wine exports: A gravity model approach. *International Journal of Wine Business Research* 30: 218–242. CrossRef.

Head K, Mayer T (2014) Gravity equations: Workhorse, toolkit, and cookbook. In: Gopinath G, Helpman E, Rogoff K (eds), *Handbook of International Economics*, Volume 4. Elsevier, 131–195

Helpman E (1987) Imperfect competition and international trade: Evidence from fourteen industrial countries. *Journal of the Japanese and International Economies* 1: 62–81. CrossRef.

Helpman E, Krugman P (1985) *Market structure and foreign trade: Increasing returns, imperfect competition and the international economy.* MIT Press, Cambridge

Hogg T, Rebelo J (2018) Rumo estratégico para o setor dos vinhos do Porto e Douro. IVDP/UTAD report, Porto. http://www.ivdp.pt/pt/docs/SINTESE.pdf

Holmes JA, Anderson K (2017) Convergence in national alcohol consumption patterns: New global indicators. *Journal of Wine Economics* 12: 117–148. CrossRef.

Macedo A, Gouveia S, Rebelo J (2019) Does wine quality have a bearing on exports? *Agris on-line Papers in Economics and Informatics* 11: 49–59. CrossRef.

Macedo A, Gouveia S, Rebelo J (2020) Horizontal differentiation and determinants of wine exports: Evidence from Portugal. *Journal of Wine Economics* 15: 163–180. CrossRef.

Morrison A, Rabellotti R (2017) Gradual catch up and enduring leadership in the global wine industry. *Research Policy* 46: 417–430. CrossRef.

Santeramo FG, Lamonaca E, Nardone G, Seccia A (2019) The benefits of country-specific non-tariff measures in world wine trade. *Wine Economics and Policy* 8: 28–37. CrossRef.

Silva JS, Tenreyro S (2006) The log of gravity. *The Review of Economics and Statistics* 88: 641–658. CrossRef.

Wei SJ (1996) Intra-national versus international trade: How stubborn are nations in global integration? National Bureau of Economic Research Working Paper 5531, Cambridge. https://ssrn.com/abstract=4196

Acknowledgement

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by the project VINCI – Wine, Innovation and International Competitiveness, under the operation number SOE3/P2/F0917, FEDER – Interreg SUDOE, and national funds, through the FCT – Portuguese Foundation for Science and Technology under the project UIDB/SOC/04011/2020.

Funded by

ersa

WU WIRTSCHAFTSUNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

FWF