Papers in Evolutionary Economic Geography

19.01

Technological Diversification of U.S. Cities during the Great Historical Crises

Mathieu P.A. Steijn and Pierre-Alexandre Balland and Ron Boschma and David L. Rigby

Utrecht University Human Geography and Planning

http://peeg.wordpress.com

Technological Diversification of U.S. Cities during the Great Historical Crises*

Mathieu P.A. Steijn[†‡] Pierre-Alexandre Balland[†§] Ron Boschma[†¶]
David L. Rigby[‖]

December 2018

Abstract — Regional resilience is high on the scientific and policy agenda. An essential feature of resilience is diversifying into new activities. But, little is known about whether major economic crises accelerate or decelerate regional diversification, and whether the impact differs between specialised and diverse regions. This paper offers systematic evidence on the effects of three of the largest crises in U.S. history (the Long Depression 1873-1879, the Great Depression 1929-1934, and the Oil Crisis 1973-1975) on the development of new technological capabilities within U.S. metropolitan areas. We find that crises reduce the pace of diversification in cities and that they narrow the scope of diversification to more closely related activities. We also find that more diverse cities outperform more specialised cities in diversifying during times of crisis but more diverse cities do not have a stronger focus on less related diversification during these unsettled times.

Keywords — Technological diversification, regional resilience, major historical crises, related diversification, U.S. cities, entry of technologies, patents

JEL codes — R11, D83, O33

*We thank Sergio Petralia, and participants at GEOINNO 2016 and DRUID 2016 for their valuable comments. This work has benefited from grant 438-13-406 from JPI Urban Europe (Steijn). (**contact:** m.p.a.steijn@uu.nl)
†Department of Human Geography and Planning, Utrecht University, Heidelberglaan 2, 3508TC Utrecht.
‡Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam.
§Collective Learning Group, The MIT Media Lab; Massachusetts Institute of Technology
¶UiS Business School, University of Stavanger.
‖Department of Geography and Department of Statistics, University of California, Los Angeles

1 Introduction

The financial crisis of 2007-2008 was seen as the largest global economic crisis since the Great Depression of the 1930s. Although the crisis was global, strong intra-country disparities in vulnerability to the shock have been noted (Martin, 2012). As a result, questions on how to prevent regions from entering crises and how to alleviate the impacts of crises on regions have once more returned to prominence on the research agenda. However, despite the wide interest, the literature on regional resilience is still largely considered as work in progress (Boschma, 2015).

A crucial component of regional resilience is the ability of regions to diversify into new activities (Pike et al., 2010; Boschma, 2015; Xiao et al., 2018). When regions are hit by a shock, it may be crucial to develop new growth industries to speed up the recovery process in regions during times of crisis. Several case-studies (Grabher, 1993; Glaeser, 2005) indeed suggest that diversifying into new activities aid overcoming downturns. However, little is actually known on how much diversification occurs in crises relative to periods of regular economic activity. Theories inspired by Schumpeter have expressed divergent views on this issue (Filippetti and Archibugi, 2011): some scholars claim major crises trigger technological breakthroughs (Schumpeter, 1939; Kleinknecht, 1987), while others suggest that dramatic drops in demand prevent the introduction of new (major) technologies during unsettled times (Schmookler, 1966; Scherer, 1982). Which of these theories prevails at the regional level remains unclear.

The literature on regional resilience has also claimed that specialised and diversified regions may respond differently to economic shocks. In particular, specialised regions are generally perceived to be unable to adapt to crises because of lock-in (Grabher, 1993; Boschma and Lambooy, 1999; Hassink, 2005; Essletzbichler, 2007). This might suggest that less specialised regions are more capable of diversifying into new activities than more specialised regions in times of major crises. Little systematic evidence is yet available.

Previously, empirical evidence on these questions relied primarily on case studies. Work of Hidalgo et al. (2007); Kogler et al. (2013); Boschma et al. (2015); Balland et al. (2015); Rigby (2015), among others, made it possible to quantify such a qualitative phenomenon as the relatedness between technologies, opening up the way for more systematic analyses. Advances in data availability complement this development. The recent HISTPAT U.S. patent data set (Petralia et al., 2016) reaching back to 1836, allows us to examine some of the deepest crises the United States has experienced. We focus on patterns of technological diversification within Metropolitan Statistical Areas (MSAs) during three of the most devastating economic shocks in U.S. history: the Long Depression, the Great Depression and the Oil Crisis [1].

The analysis yields four insights. First, we find that U.S. cities diversify less during major crises. Second, in periods of crisis, cities diversify more in stronger related activities than during periods of prosperity. Third, more diverse cities have a higher probability of diversifying during crises than do specialised cities. The advantage is twofold. In the first place, diverse cities have a larger technological portfolio and therefore have closer technological proximity to potential new technologies, enabling them to diversify more easily. Secondly, we show that even when controlling for this advantage, diverse cities outperform more specialised cities. Suggesting that local attitudes and/or vested interests differ in the regions, which induce more diverse cities to be more open to diversification. The fourth insight of the paper yields that even though diverse cities are more able to diversify into any new activity than their more specialised counterparts, they do not significantly differ in how much they focus on less related diversification.

The structure of the paper is as follows. In Section 2, we discuss recent theorizing on regional resilience and diversification, and how that is related to periods of crisis. Based on these theoretical considerations, we derive four hypotheses. In Section 3, we explain the data and the methodology used. In Section 4, we present the main empirical findings. Section 5 of the paper will conclude and discuss the findings in light of a future research agenda.

[1] The Financial Crisis (2007-2008) is too recent to be included in the analysis as we insist that for successful diversification, new technologies should persist in a region for a certain time period. As further explained in Section 3

2 Resilience of Regions and Diversification in Times of Crisis

In recent years, studies have investigated the ability of regions to bounce back after a crisis (Balland et al., 2015; Martin, 2012; Diodato and Weterings, 2015; Cuadrado-Roura et al., 2016; Crescenzi et al., 2016; Bristow and Healy, 2018; Fratesi and Perucca, 2018). The regional resilience literature is fundamentally interested in the capacity of regions to recover from a shock, and what processes drive that recovery. Many resilience studies follow an equilibrium approach, looking at the ability of regions to return to a pre-existing equilibrium state after a shock, or to move into a new equilibrium state (Fingleton et al., 2012). These studies tend to overlook the fact that a substantial part of the recovery process may depend on the ability of regions to develop new growing activities that offset processes of decline (Boschma, 2015; Balland et al., 2018). As such, to tackle the question of regional resilience requires understanding of how regions diversify into new activities.

In recent years, scholars have put a lot of effort in explaining why regions differ in their ability to develop new technologies, industries or jobs. This empirical literature suggests that regions do not start from scratch when diversifying: they tend to build on existing local capabilities, a process that has been labelled related diversification (Neffke et al., 2011; Boschma et al., 2015; Rigby, 2015). This is not to say that unrelated diversification (i.e. the successful development of new activities unrelated to local activities) does not occur in regions, but the evidence shows it is a rare phenomenon(Hidalgo et al., 2007; Neffke et al., 2018; Pinheiro et al., 2018).

Boschma (2015) has connected the growing literature on diversification to the field of regional resilience. Inspired by scholars who advocate an evolutionary approach to regional resilience (e.g. Christopherson et al., 2010; Pike et al., 2010; Simmie and Martin, 2010; Martin and Sunley, 2015; Webber et al., 2018; Cainelli et al., 2018), Boschma (2015) links resilience to the ability of regions to diversify and create new growth paths, to offset stagnation and decline during shocks. An idea that echos in certain case-studies (see: Grabher, 1993; Glaeser, 2005). This implies that instead of looking at the vulnerability of regions to a shock (conventionally measured as a decline in output levels) and the ability to recover from a shock (conventionally measured as a return to previous output levels, or to new equilibrium output levels), there is a need to examine to what extent shocks impact the ability of regions to diversify (Xiao et al., 2018).

However, little is known on whether and how regions diversify during major crises. Do periods of deep economic distress accelerate or slow diversification? This question has not received a lot of attention in the regional resilience literature but a related debate has been taking place in the long wave literature for many years. Innovation theories, inspired on Schumpeter, that developed in the 1980s (Dosi et al., 1988) viewed radical innovations as clustering in waves rather than occurring randomly over time. Schumpeter referred to this as swarming of innovations which he believed happened during the downswing period of the long wave. In his work on basic innovation, Mensch (1975) developed the depression trigger hypothesis to explain the tendency for radical innovations to bunch during periods of crisis. This hypothesis was challenged by other scholars (Clark et al., 1981; Duijn, 1983) who argued that most innovations take place just after the crisis, during the upswing of a long wave. Kleinknecht (1981) reconciled both views, stating that "*the argument that depression is acting as a trigger for major innovations .... does not exclude the existence of a swarm of related innovations which accompany the diffusion of newly introduced products*" (p. 295).

Kleinknecht (1981, 1987) supported the depression trigger hypothesis, claiming that in periods of crisis, demand drops dramatically and returns on further improvements of mature products and technologies are low, and therefore the relative risk of introducing radical innovations for firms decreases. This incentive becomes even stronger when productive resources are set free during the downswing of the economy, leading to declining wages and lower capital costs, which makes it more attractive to invest (Krugman, 1993; Glaeser, 2005). Moreover, many innovative breakthroughs are technologically related to each other, showing interdependencies and complementarities (Rosenberg, 1982; Carlsson and Stankiewicz, 1991) which makes them cluster in time (Rosenberg and Frischtak, 1983; Boschma, 1999). And once radical innovations are introduced, they will attract new investments that will lead to a large stream of additional innovations, known as the bandwagon effect (Clark et al., 1981).

Diametrically opposing this depression trigger hypothesis, is the demand pull hypothesis suggesting that dramatic drops in demand during crises prevent the introduction of new (major) technologies (Schmookler, 1966; Freeman et al., 1982; Scherer, 1982). Freeman et al. (1982) argued that R&D activity is reduced considerably in long wave depressions. Instead, the rise in demand during the upswing provides more favourable conditions for firms to introduce breakthroughs and major innovations (Geroski and Walters, 1995). Schmookler (1966) claimed that upswings in inventive activity followed upswings in demand (Coombs et al., 1987). Moreover, depression phases are characterized by a mismatch between major technologies and institutions (Perez, 1983; Dosi, 1984): the successful introduction and diffusion of major breakthroughs in the economic system requires a set of new institutions that take a long time to develop (Freeman and Perez, 1988). The demand pull claims suggest that new major technologies are more likely to enter the economy in the growth phase of the long wave.

For similar reasons, but in a regional setting, agents faced by a drop in demand can opt to innovate in other technologies and possibly more specifically into technologies that are new to the region, or postpone diversification until demand rises again. Reformulating the neo-Schumpeterian ideas into the framework of the regional diversification literature, we could expect regions to introduce and develop new activities during downswings as much as during upswings.

Therefore, we develop a set of competing hypotheses on how regions adopt technologies new to them. Hypothesis 1a builds on the depression trigger hypothesis, stressing that diversification is more likely to occur during periods of crisis. As stated above, economic agents might be more willing to take risks and to try out something new when current products and technologies show decreasing returns. Institutional agents (like regional governments) may see major enduring crises as windows of opportunity and are therefore more prone to promote new ways of getting out of the crisis. By contrast, Hypothesis 1b builds on the demand-pull hypothesis and states that diversification is even more unlikely to take place in regions during periods of crisis. Inventions have to wait until upswings in demand arises. We therefore formulate the following two competing hypotheses:

Hypothesis 1 (a)**.** *cities diversify more during crises than non-crisis periods*

Hypothesis 1 (b)**.** *cities diversify less during crises than non-crisis periods*

Furthermore the contrasting Schumpeterian views on adopting new major technologies also yield different expectations on the level of unrelated diversification during crises. The depression trigger hypothesis suggests that unrelated diversification is more likely, while the demand pull hypothesis suggests that in case of diversification, this is more likely to be more closely related to other regional activities. Unrelated diversification would just add to the high uncertainty that is already inherent to a crisis period. When regions diversify during a major crisis, related diversification, though still risky, provides more certainty during highly crises.So, we formulate the following hypotheses:

Hypothesis 2 (a)**.** *cities diversify more in less related technologies during crises than non-crisis periods*

Hypothesis 2 (b)**.** *cities diversify more in related technologies during crises than non-crisis periods*

The regional resilience literature argues that variety is crucial for resilience because it can accommodate sector-specific shocks (Essletzbichler, 2007, 2015; Diodato and Weterings, 2015; Rocchetta and Mina, 2017). This is in line with numerous case studies on specialised regions that showed structural problems of adjustment (Boschma and Lambooy, 1999; Pike et al., 2010). Specialised regions may have a low capacity to diversify in new activities, because they are cognitively, socially and politically locked-in (Grabher, 1993; Hassink, 2005). For example, Grabher (1993) describes how the Ruhr-area in Germany was prosperous in the 1950s because of its strong specialisation in the steel industry. However, when adverse times hit in 1970s the region adapted poorly exactly because workers, firms, and instutions were strongly focused on a limited set of activities.

Variety in a region, instead, promotes regional diversification because a larger set of new activities are easier to undertake when a larger set of related capabilities is present (Boschma, 2015). Related variety has been mentioned in particular (Balland et al., 2015), as recombinations are more feasible and can be made more effective across activities that share similar knowledge and skills (Frenken et al., 2007). Xiao et al. (2018) showed that related variety indeed affected positively the resilience of European regions in terms of maintaining and increasing their ability to develop new industrial specializations after the 2008 economic shock. But also unrelated variety in regions facilitates diversification, especially in unrelated technologies (Castaldi et al., 2015). Moreover, in diverse regions, industries and vested interests are less likely to dominate the institutional and policy network that can block new key developments (Neffke et al., 2018; Boschma, 2015). Therefore, strongly specialised regions are expected to be less able to diversify.

Few studies have empirically tested whether more or less diverse regions are more capable of diversifying during major crises. The case-study by Grabher (1993) suggests that the mentioned weakness of strongly specialised regions in diversifying becomes even more apparent in times of crisis. When the main activity of a specialised region is prospering the lack in diversification potential would seem of little concern. However, when this industry faces a major downturn a specialised region will face difficulties in developing new activities to overcome decline. We argue that specialised regions are less able to diversify especially into more unrelated technologies than are more diversified regions. First of all because a larger technological portfolio means more technologies are easier to develop building on existing knowledge. But secondly, because we also expect this effect to hold when one controls for this fact, we expect regions that are more diverse to develop more unrelated activities in crises, because their diversity enables them to deviate more easily from the constraints of their past. This leads to the following hypotheses:

Hypothesis 3. *diverse cities diversify more than specialised cities during crises*

Hypothesis 4. *diverse cities diversify more in less related technologies than specialised cities during crises than non-crisis periods*

Where hypothesis 1 and hypothesis 2 explore the differences between periods of crisis and periods of prosperity, hypothesis 3 looks into the difference between diverse cities and specialised cities during crises, hence the sample for this hypothesis only consists of crisis periods. Hypothesis 4 compares the change in focus on unrelated technologies between diverse cities and specialised cities as they enter a crisis. In other words, hypothesis 4 evaluates hypothesis 2 for different levels of diversity. As such, the extent to which city-regions diversify more in less related technologies during crises can be compared across the level of diversity of these city-regions.

3 Data and Methodology

The hypotheses outlined above are tested with a unique dataset of U.S. patents covering the period 1836 – 2002. This long time span allows us to test the hypotheses across three major crises in U.S. history: the Long Depression, the Great Depression and the Oil Crisis. Although we are aware of limitations of patent records (Griliches, 1981), patent records hold a wealth of information regarding the process of invention and the nature of additions to the expanding stock of knowledge. The patent data originates from the efforts of Petralia et al. (2016) geographical locations for all patents over the period 1836 – 1974 from Google scans of historical U.S. patents. Patents since 1974 are available from the NBER patent data (Hall et al., 2001).

Diversification in a Metropolitan Statistical Area (MSA) is captured by the development of a comparative advantage in a new technology within that MSA[1]. Technologies are represented by the different primary classes of the United States Patent and Trademark Office (USPTO) patent classification system[2]. When diversification occurs, it is possible to calculate the relatedness of the new technology to the technologies present in the MSA in the previous time period.

We restrict our sample to MSAs within the contiguous U.S.A. We also impose a minimum of 10 patents per year for a time period of a MSA to be taken into account and a minimum of 0.5 patents[3] per year in a certain primary technology class. As a result, data is drawn from a sample of 274 MSAs and 2,171 MSA-time periods. Below, we introduce our definitions and measurements of crises, diversification, relatedness and diversity.

3.1 Crises

Like Balland et al. (2015), we build on trends in patenting per region to indicate when regions are in crisis, as patent counts are highly correlated with economic performance (Glaeser and Ponzetto, 2007; Rothwell et al., 2013). To ascertain this link with economic performance, we focus on the great historical crises of the United States, identified independently of the patent data, while using patent counts to indicate the breadth and depth of these crises per MSA.

[1] We note that if a region diversifies in activities where patenting is uncommon this will not be captured by our methodology.

[2] Primary technology class are comparable over time as the USPTO reclassifies all patents when new class definitions are introduced.

[3] Patents that are assigned to inventors in multiple MSAs, only count as 1 divided by the number of MSAs on that patent for each of the MSAs.

Each nationwide crisis is regarded as a shock at the regional level. A metropolitan area can then either enter into a crisis or not. At the regional level, the emergence and the duration of crises are identified from patent records using an adapted version of the business cycle algorithm of Harding and Pagan (2002), after Balland et al. (2015). We follow the definition of technological crises by Balland et al. (2015) as sustained periods of negative growth in patent activity: "*more formally, a time series recording yearly patenting activity can be defined as a continuum of local maxima (peaks) and minima (troughs) that divide the series into periods of technological growth from trough to peak and technological crisis from peak to trough.*" (p. 6).

The algorithm to detect business cycles "*identifies potential turning points as the local minima (trough) and maxima (peak) in the series. Let P_t be a patent count yearly series. A trough is identified as $(p_{(t-j)}, , p_{(t-1)}) > p_t^{trough} < (p_{(t+j)}, , p_{(t+1)})$ while a peak follows the condition that $(p_{(t-j)}, , p_{(t-1)}) < p_t^{peak} > (p_{(t+j)}, , p_{(t+1)})$.*" (Balland et al., 2015, p. 172). To prevent noise due to years of random growth or decline, two extra conditions are imposed: "*The phases (technological growth or technological crisis) should be at least 2 years long, while complete cycles (period between 2 peaks or between 2 troughs) should be at least 5 years long.*"

As a result of this procedure, time periods are defined separately for each MSA. Therefore, the length of time periods varies and does not necessarily match those of other MSAs. For each of the U.S. metropolitan areas that we examine, all periods of technological crisis and growth are identified between 1836 and 2002. If a period of regional technological crisis overlaps with one of the three great U.S. crises, it is retained in our model. Crises that do not overlap with one of the three great U.S. crises or experience a decrease in patenting activity of less than 35 percent[4] during the crisis are ignored.

The decision to ignore regional downturns in patenting that do not occur during a nation-wide shock decreases the risk of including local crises that are unrelated to major economic downturns or spurious decreases in patent counts. Figure A1, in Appendix A, depicts the number of MSAs entering a period of growth in green, respectively a period of crisis in red per year, during the time periods associated with each of the great historical crises. The impact of the crises on

[4] Lowering this threshold leads to greater standard errors in the results, indicating that diversification behaviour during small (less significant) crises does not differ much from diversification behaviour during periods of growth, which is to be expected.

patenting activity is clearly visible. One can also note a small time lag between the actual start of the great historical crises and MSAs entering a period of downturn in patenting for the first two major crises. The effect of the Oil crisis is immediately visible[5]. Because of the time lag in the reaction of patenting activity, we retain the regional crises that start in years when more MSAs enter a crisis period than MSAs enter a period of growth. For the Long depression this is 1876 to 1878, for the Great depression 1932 to 1938, and for the Oil crisis 1972 to 1976. All other crisis periods are dropped from the sample. Regional periods of growth are kept regardless of when they occur.

Table 1 shows the strong impact of the great historical crises on the patent production at the regional level. Affected MSAs, in the second column, indicates the number of MSAs that enter a crisis that meets the aforementioned requirements and respective time period. Unaffected MSAs are MSAs that were in a growth phase before the start of the crisis and remain so over the course of the crisis. #MSAs gives the total number of MSAs that meets the requirement of producing on average ten patents per years in that time period. Note that the number of unaffected MSAs and affected MSAs do not sum to the total number of MSAs. MSAs could already be in crisis upon entering the respective time periods or could enter in a crisis in which less than 35% of patenting activity is lost. The last two columns respectively give the average time duration of the crises, and the average percentage of patent activity lost at the trough compared to the peak for the affected MSAs. The Great Depression stands out as having impacted the most regions.

Table 1 – The Regional Impact of the Great Historical Crises

Crisis	Aff. MSAs	Unaff. MSAs	#MSAs	Avg. Crisis Duration	Avg. Crisis Depth
Long depression	30	19	101	∼ 3.7 years	∼ -53.6%
Great depression	139	2	205	∼ 6.4 years	∼ -74.3%
Oil crisis	128	44	252	∼ 6.7 years	∼ -59.4%

[5] Note that the years indicate the end year of the previous cycle period and the start year of the next cycle period. E.g. a period of crisis starting in 1972 indicates that the peak was in 1972 and the first year of downturn is 1973.

3.2 Diversification

We use the notion of Revealed Comparative Advantage (RCA) (see Hidalgo et al., 2007) to identify in which technologies each MSA is specialised across the time periods examined. In equation 1, x represents the number of patents, c denotes the city-region (MSA), i is the primary technology class, and t indicates the time period. RCA values are bounded on the left by zero. A RCA value of 1 indicates that an MSA has the same share of patenting activity in a particular technology class as the national average. RCA values of 1 or greater indicate regional specialization in a technology. A technology enters the technological portfolio of an MSA when an MSA develops a specialization in a technology class that it did not have in the previous time period we refer. An entry is considered a diversification of the region. To account for spurious entries of technologies we add the condition that an entering technology has to remain present in the portfolio of an MSA (with $RCA => 1$) for at least two time periods.

$$RCA_{cit} = \frac{\frac{x_{cit}}{\sum_{i=1}^{I} x_{cit}}}{\frac{\sum_{c=1}^{C} x_{cit}}{\sum_{c=1}^{C}\sum_{i=1}^{I} x_{cit}}}, \tag{1}$$

3.3 Relatedness

Technologies that are not in the technological portfolio of an MSA in time period $t-1$ (those for which the RCA value is below one) enter or do not enter in time period t. An important predictor of the entry of a technology within an MSA is how closely related it is to technologies that are already present in the region (Boschma et al., 2015; Balland et al., 2018). This notion of relatedness is essential for hypotheses 2, and 4, where we focus on less related diversification. The co-occurrence of technology classes on patents is used to measure the relatedness between *technologies*. Technology classes are more related to one another as they co-occur with a frequency that is greater than that which would be predicted based on the overall counts of classes in the population of patents of a given time period. The formula for relatedness, outlined by van Eck and Waltman (2009) and improved by Steijn (2018), is reported in equation 2. Where C_{ijt} is the number of co-occurrences between technology i and technology j in time period t. S_{it} and S_{jt} is the number of co-occurrences involving respectively technology i and technology j in time period t, N is the total number of technologies, and m is the total number of co-occurrences.

$$TR_{ijt} = \frac{C_{ijt}}{(\frac{S_{it}}{\sum_{n=1}^{N} S_n} \frac{S_{jt}}{\sum_{n=1}^{N} S_n - S_{it}} + \frac{S_{jt}}{\sum_{n=1}^{N} S_n} \frac{S_{it}}{\sum_{n=1}^{N} S_n - S_{jt}})m}, \tag{2}$$

Building on relatedness, relatedness density (see Hidalgo et al., 2007) gives the relatedness of a *region* to a *technology* that is not yet present in its technological portfolio. Relatedness density is equal to the sum of relatedness values of the technologies in the region to the possibly entering technology divided over the sum of relatedness values of all technologies to this technology, as can be seen in equation 3.

$$Rel.density_{cit} = \frac{\sum_{j \in c, j \neq i} TR_{ijt}}{\sum_{j \neq i} TR_{ijt}}, \tag{3}$$

3.4 Diversity

For hypotheses 3 and 4, a measure of the extent of specialization in a region is required. Here we follow Duranton and Puga (2000) who propose a simple diversity index, known as the Relative Diversity Index (RDI). The intuition is that if the relative distribution of patenting activity over technology classes in an MSA resembles the national distribution, then the city is relatively diverse. However, when the patents of an MSA cluster strongly above the national average in a few classes then it is seen as specialised. Duranton and Puga (2000) use the inverse of the formula[6] in equation 4, like before x stands for the number of patents, c indicates the MSA, i the respective technology, and t the respective time period. A value close to zero denotes a diverse city, whereas the larger the value the more specialised a city is.

$$RDI_{ct} = \sum_{i=1}^{I} \left| \frac{x_{cit}}{x_{ct}} - \frac{x_{it}}{x_t} \right|, \tag{4}$$

[6] We choose not to use the inverse as this transformation skews the variable, which makes outliers more influential in the estimation of its coefficient.

3.5 *Control variables*

Other factors that are correlated with our variables of interest may influence the development of specialisation in a new technology by an MSA. Having MSAs nearby that have an RCA in a technology can be expected to positively influence the entry of that technology to the technological portfolio of a city (see Rigby, 2015). Therefore, we develop a spatial weight matrix using the inverse distance for the presence of technology in neighboring MSAs . We also include the average population of MSAs in the time periods. Boschma et al. (2015) employ several other control variables using information on the inventors. This type of data is unavailable in HISTPAT or other sources. However, this shortfall can largely be mitigated by the inclusion of fixed effects. Table 2, gives the descriptive statistics of our variables.

TABLE 2 – DESCRIPTIVE STATISTICS

Statistic	N	Mean	St. Dev.	Min	Max
Entry	724,752	0.031	0.174	0	1
Crisis	724,752	0.141	0.348	0	1
Relatedness density	724,752	0.089	0.126	0.000	1.000
Population	724,752	416,093	961,537	20,402	17,019,060
Present*$\mathbf{W}$	724,752	0.00003	0.00003	0.00000	0.0003
RDI	724,752	1.086	0.303	0.255	1.833

3.6 *Empirics*

Entry models are a common tool in the literature that yield insights on the role of relatedness in diversification (e.g. Boschma et al., 2015; Balland et al., 2018). In spite of the popularity, some underestimation of risks exists in relation to two particular traits of the econometric specification. First, the dependent variable entry is binary and is strongly right skewed: there are very little incidences of successful entries (values of 1) compared to technologies that do not enter (values of 0). Second, the independent variable relatedness density is truncated and strongly right skewed.

The first has already been noticed by Boschma et al. (2015), referring to work by King and Zeng (2001). They argue that the coefficient estimates of nonlinear models might not be consistent when there are too many zeros in the dependent variable. They therefore use an OLS to estimate the entry model. The use of such a Linear Probability Model (LPM) has certain risks that can be considered to be outweighed by the benefits of easier interpretation (see Hellevik, 2009). However, when the probability of success of the dependent variable is on the extreme ends of the distribution[7], the slope of a logit or probit is not well approximated by the slope of a linear regression and the flaw of the LPM in predicting probabilities outside the possible range of 0 to 1 generally becomes apparent.

In the case of the entry model, probabilities can be considered to be on the lower end of the probability distribution. Therefore, a logit model seems more appropriate. Boschma et al. (2015) run a logit model as robustness check, confirming their results. As said, this is not without a risk as King and Zeng (2001) warn for inconsistent estimates when probabilities are extremely low. King and Zeng (2001) also provide guidelines when this risk is more likely to exist. They show in a simulation that the inconsistency tends to zero as the sample size tends to infinity and/or the percentage of ones tend to 50%. In our data, there are 724,752 observations and an average probability of entry of 3.1%. Following guidelines and simulation results of King and Zeng (2001), the risk can be assumed to be negligible[8].

A logistic regression would seem the way to proceed in estimating the entry model. However, the skewness of the distribution of relatedness density[9] values poses a serious threat to the effective estimation of its effect on entry. Although OLS and logit models alike do not require variables to be normally distributed, unlike the error terms, a variable that is highly skewed has the risk of outliers exerting a strong influence on the results. Regressing relatedness density on entry, gives the mean effect of relatedness density on entry, but as the mean is far to the right of the median, this coefficient is unreliable. Figure 1 gives the histogram of relatedness density.

[7]Von Hippel (2015) suggests that probabilities of success should be in the range of 20% to 80% for logit and linear models to be used interchangeably.

[8]The rules of thumb of Allison (2012) also suggest that the risks are negligible.

[9]The skewness was also noticed by Uhlbach et al. (2017)

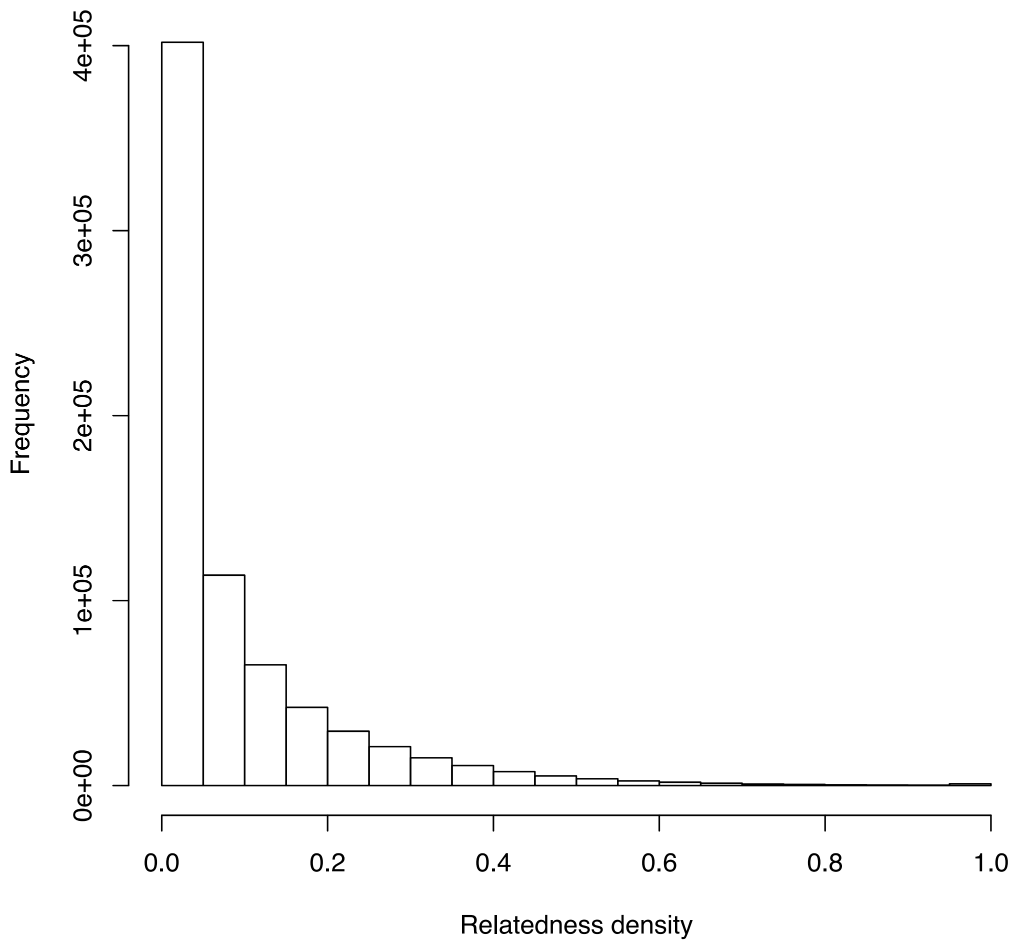


Figure 1 – Histogram of relatedness density

A standard approach in econometrics is to apply some type of monotonic transformation to the variable to preserve order but alter the distribution. The most common option, a log transformation, cannot be applied due to the fact that there are zeros among the values. Other common transformation methods in dealing with right skewed data (like box-cox, taking the square, or an inverse hyperbolic sine) are able to deal with zeros but due to the large number of zeros (12.9% of observations involve zero relatedness density) fail to give a distribution in which a strong influence of outliers can be ruled out. As a result, we will estimate our econometric specifications for quantile groups of relatedness density[10] and compare how the coefficients of interest change between the groups. Next, we introduce the econometric specification.

[10]Note that this is different from quantile regressions where the *dependent* variable is divided over quantiles

3.7 Econometric specification

Equation 5 gives a regression formula, which is in line with previous work of amongst others Boschma et al. (2015). If a technology i enters the technological portfolio of city c in time period t, the value of the dependent variable is 1. If it was not in the portfolio of city c and it did not enter, its value is 0. The dependent variable is regressed on the relatedness density of the technology class to the portfolio of each city in the previous time period, on a dummy variable which indicates if a city is experiencing a crisis or not, on the population in the MSA in time period t, and on the presence of technology i in the technological portfolio of neighboring MSAs multiplied by a spatial weight matrix $\mathbf{W}$, a city-fixed effect (ϕ), a technology-fixed effect (θ) and a time-fixed effect (τ). Population and Present*$\mathbf{W}$ are scaled to have 0 mean and a standard deviation of 1, as the interpretation of these coefficients is not so straightforward in its units of measurement.

$$Entry_{cit} = \beta_1 Rel.density_{cit-1} + \beta_2 Crisis_{ct} + \beta_3 Population_{ct} + \beta_4 Present_{it} * \mathbf{W} + \phi_c + \theta_i + \tau_t + \epsilon_{cit}, \tag{5}$$

For the main specification of Hypothesis 1, and 2 relatedness density is left out of the specification but the data is split according to quintiles of relatedness density, due to the mentioned skewness of this variable. When using quintiles we have to exclude fixed effects, because proper estimation of fixed effects requires observations with "successful" entry and "unsuccessful" entry of each city, technology, and time period. However, we add them in a robustness check where data is split according to terciles.

For Hypothesis 3 the RDI variable is added to specification 5, and the crisis variable is left out as only data from cities in times of crisis is used. Relatedness density is left in the specification for Hypothesis 3. Here it is an important control variable to assure that more effective diversification of diverse regions is not driven by their, on average, higher relatedness density values, due to having more technologies in their portfolio. This advantage even holds within quantiles of relatedness density values, in the sense that very diverse cities (low RDI), have on average values that are closer to the upper bound of a quantile compared to more specialised cities (high RDI).

To investigate risks of imposing a functional form on the relation between RDI and entry, we also show results by assigning dummy variables to different ranges of RDI values[11] next to the results using the RDI values proper. However, for this hypothesis we only find minor differences in interpretation.

Hypothesis 4 is tested in a similar fashion to analysing Hypothesis 1 and 2, hence using the full dataset, but including the RDI variable. The RDI values are added as the aforementioned dummy variables, as here there are major differences in interpretation of the results. The results suggest that an increase or decrease in RDI value leads to different changes in outcome depending on the initial RDI value. As such, adding RDI values linearly imposes an unfit functional form.

We also did robustness checks with various different compositions of relatedness values and RDI values, seperately standardizing the relatedness values in crisis periods and non-crisis periods such that the mean is equal in both types of periods, and we tried out penalized log-likihood methods (see Firth, 1993; Kosmidis and Firth, 2009). These robustness checks yield similar results as presented below.

[11] e.g. a dummy variable called RDI 0.3-0.8 is one for all RDI values ranging from 0.3 to 0.8 and zero for all other RDI values.

4 Results

4.1 Diversification in times of crisis

Table 3 gives the results for specification 5, as mentioned without fixed effects, for five subsets of data, ranging from the quintile with relatedness density values among the 20% lowest to the quintile with the 20% highest values. As expected, the coefficients on the control variables are positive and significant. This indicates that both a larger population as having more neighbouring MSAs specialised in the technology increases the probability of entry ceteris paribus. The coefficient on crisis is negative and significant across the quintiles. This indicates that the probability of an MSA entering a new technological specialization is lower during a crisis regardless of relatedness density. This confirms Hypothesis 1b and rejects the depression trigger Hypothesis 1a: cities diversify less during crises.

Table 3 – Regression results (Hypothesis 1 and 2)

	Dependent variable:				
	entry				
	0-20%	20-40%	40-60%	60-80%	80-100%
Crisis	-1.232^{***}	-0.747^{***}	-0.734^{***}	-0.514^{***}	-0.295^{***}
	(0.184)	(0.102)	(0.071)	(0.045)	(0.028)
Population	0.400^{***}	0.257^{***}	0.148^{***}	0.083^{***}	0.073^{***}
	(0.032)	(0.027)	(0.018)	(0.011)	(0.004)
Present*$\mathbf{W}$	0.573^{***}	0.395^{***}	0.432^{***}	0.381^{***}	0.288^{***}
	(0.032)	(0.020)	(0.013)	(0.010)	(0.007)
Constant	-4.706^{***}	-4.484^{***}	-4.004^{***}	-3.377^{***}	-2.510^{***}
	(0.035)	(0.028)	(0.022)	(0.017)	(0.011)
Time Fixed Effects	No	No	No	No	No
Technology Fixed Effects	No	No	No	No	No
MSA Fixed Effects	No	No	No	No	No
Observations	144,950	144,950	144,951	144,950	144,951
Log Likelihood	−5,304.951	−8,149.655	−13,199.480	−22,153.370	−40,856.230
Akaike Inf. Crit.	10,617.900	16,307.310	26,406.960	44,314.740	81,720.460
Note:					*p<0.1; **p<0.05; ***p<0.01

To facilitate interpretation, the average probability of entry per quintile group outside of crisis is given in blue in Figure 2. This average probability is equal to the intercept converted to probabilities of each quintile group as population and present*W have been scaled. The red line in Figure 2 then gives the change from the no crisis base scenario when a crisis occurs. As can be deduced from the coefficient on crisis. The shaded area gives the 95% confidence interval. As has been noticed earlier in the literature, the probability of entry increases when the entering technology is more strongly related to local technologies.

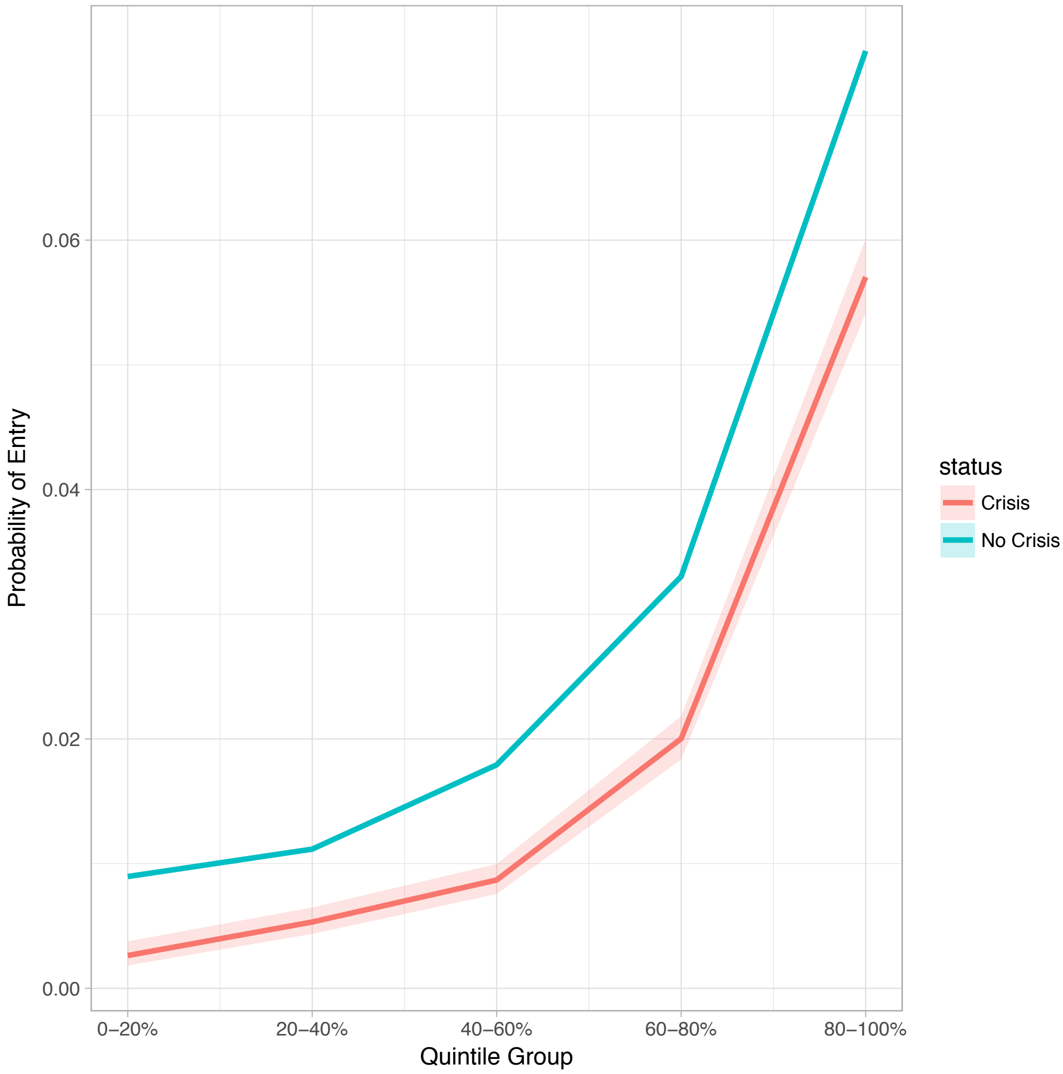


Figure 2 – Probability of entry according to crisis status

From figure 2, we can deduce how the crisis affects the development of new technological specialization (entry) by MSAs across relatedness density groups. Figure 3 depicts the difference in percentages between the entry probabilities during a crisis and outside a crisis across quintiles. Technologies with low relatedness values, those in the lowest quintile, are 70.6% less likely to be added to the technological portfolio of cities during crises than outside of crises, whereas for technologies with relatedness density values in the highest quintile, the entry probability is only about 24.1% smaller during crises. This confirms Hypothesis 2b: cities diversify more in related technologies during crises. Apparently, in times of high uncertainty, diversification is more likely in technology classes that are more closely related to the knowledge core of the region. This likely reflects the uncertainty of economic agents in terms of future technological development during the highly turbulent phases of major crises.

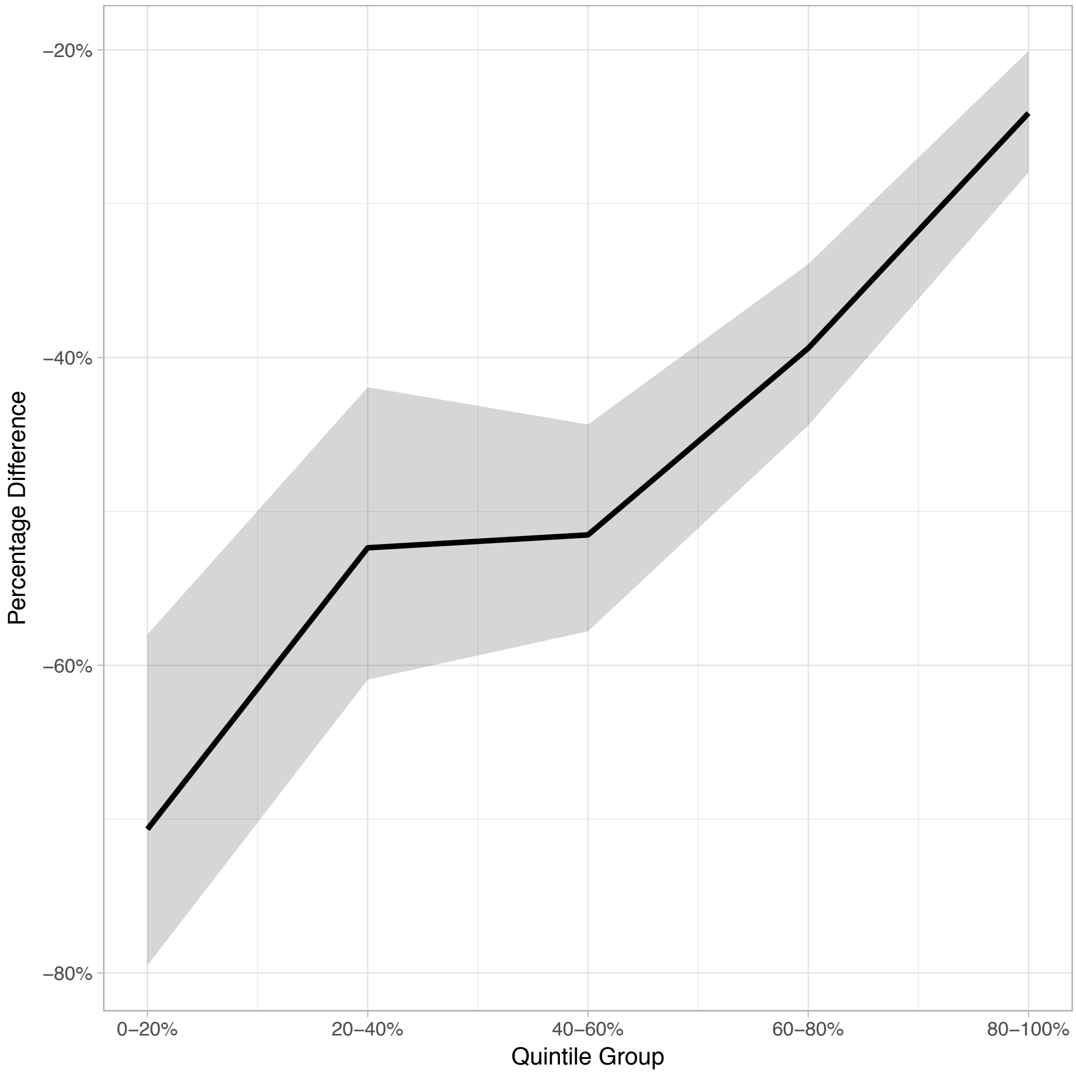


Figure 3 – Percentage difference in probability of entry between crisis and no crisis across quintile groups

However, unobserved time-invariant characteristics at the MSA-level, technological class-level, and/or time period level may be correlated with our variable of interest. Therefore, we re-run the regressions while including the fixed effects mentioned in the main specification. For fixed effects to be correctly estimated, at least one incident of successful and non-successful entry is required for all values associated to a particular dummy. To maintain a sufficient amount of observations, we run the analysis using terciles instead of quintiles. The resulting figures are based on a respective MSA, Technology and Time Period with the respective coefficient closed to the median across all regressions. As shown in Table A1, Figure A2 and Figure A3 in Appendix A, the results are confirmed, albeit it less large.

4.2 Diversity and diversification in times of crisis

In Hypothesis 3, we posit that more diverse cities tend to diversify more than specialised cities during crises, even when controlling for relatedness density. The sample here only consists of data from cities in crisis. RDI captures the diversity effect. Regression results are given in Table 4, while Table 4 gives the marginal effects of the coefficients keeping all else at its average.

Across all levels of relatedness density, including the least related, the effect of RDI is negative and significant, indicating that when cities are more specialised (RDI is higher) the probability of entry decreases. As we control for relatedness density, this increase in the probability of entry when cities are more diverse is not due to the fact that having a larger technological portfolio means having more relatedness to other technologies. This corroborates the existing literature reviewed in Section 2, which claims that there is more to diverse cities that allows them to diversify into new activities, than just having increased technological proximity. Table A3 confirm that the advantage of diverse cities holds respectively when introducing fixed effects. Table A4 gives the results when using dummy variables for ranges of RDI values. These suggest that overall the relation between RDI and entry is reasonably well approximated[12] by the functional form assumed in Table 4.

We note that contrary to Table 3, Population has an insignificant effect in some specifications. As we find that diversity has a positive effect on entry, this suggests that the positive coefficient on population in Table 3 was more likely a “diversity-effect” rather than a “size-effect”, as population is correlated with diversity. As expected, relatedness density and the presence of technology i in nearby MSAs have a positive effect on the possible entry of technology i.

[12] Only for one out of four samples (the one with the least related relatedness density values) the functional form is less clear, but that is the focus of Hypothesis 4.

Table 4 – Regression results (Hypothesis 3)

	Dependent variable:			
	entry			
	0-25%	25-50%	50-75%	75-100%
RDI	-1.384^{**}	-0.984^{***}	-0.577^{***}	-1.529^{***}
	(0.655)	(0.317)	(0.200)	(0.118)
Relatedness dens.	15.676	10.486	10.716^{***}	2.404^{***}
	(64.786)	(8.291)	(1.990)	(0.166)
Population	-0.407	0.120^{**}	-0.001	-0.023^{**}
	(0.751)	(0.053)	(0.041)	(0.012)
Present***W**	0.658^{***}	0.615^{***}	0.465^{***}	0.419^{***}
	(0.119)	(0.045)	(0.030)	(0.018)
Constant	-4.197^{***}	-4.157^{***}	-4.433^{***}	-2.332^{***}
	(0.828)	(0.446)	(0.292)	(0.126)
Time Fixed Effects	No	No	No	No
Technology Fixed Effects	No	No	No	No
MSA Fixed Effects	No	No	No	No
Observations	19,951	25,164	28,182	28,966
Log Likelihood	-329.487	$-1{,}011.821$	$-2{,}402.095$	$-6{,}187.553$
Akaike Inf. Crit.	668.973	2,033.642	4,814.191	12,385.110
Note:				*p<0.1; **p<0.05; ***p<0.01

As diverse regions outperform specialised regions in diversifying across all levels of relatedness in times of crisis, a relevant question becomes how diverse regions change their diversification behaviour in relative terms compared to more specialised regions. Do diverse cities switch more strongly to less related technologies during crises than specialised cities? As noted in hypothesis 4. For this we reproduce figure 3 but by estimating the effect for different groups according to RDI[13]. The resulting Figure 4, based on Table A5 is shown below.

When entering a crisis, the most diverse regions, with RDI values 0.3-0.8, lose over 75% of the diversification in the least related technologies. While the most specialised regions, with RDI values 1.4-1.8, only lose under 50%, and the intermediate group of regions with RDI values 1.2-1.4 does not even diversify less in the least related technologies during crises than outside of crises. At the same time, there is a negligible difference between the groups when it comes

[13]Note that for sake of legibility the two intermediate groups with RDI values between 0.8-1 and 1-1.2 are removed.

to the strongest related technologies. Suggesting that in relative terms it are actually more specialised regions who diversify more strongly into less related technologies. With the addition that extreme specialisation (RDI 1.4-1.8) decreases the focus on less related diversification compared to slightly lesser specialised regions. However, these results are not robust as adding fixed effects leads to the disappearance of differences between the RDI groups, as can be seen in Figure A4, based on Table A6. This suggests that unobserved factors related to the specific cities, technologies, or time periods associated with each group explain the difference between the RDI groups observed in figure 4 rather than the differences in diversity. Therefore hypothesis 4 cannot be accepted, nor is the opposite true.

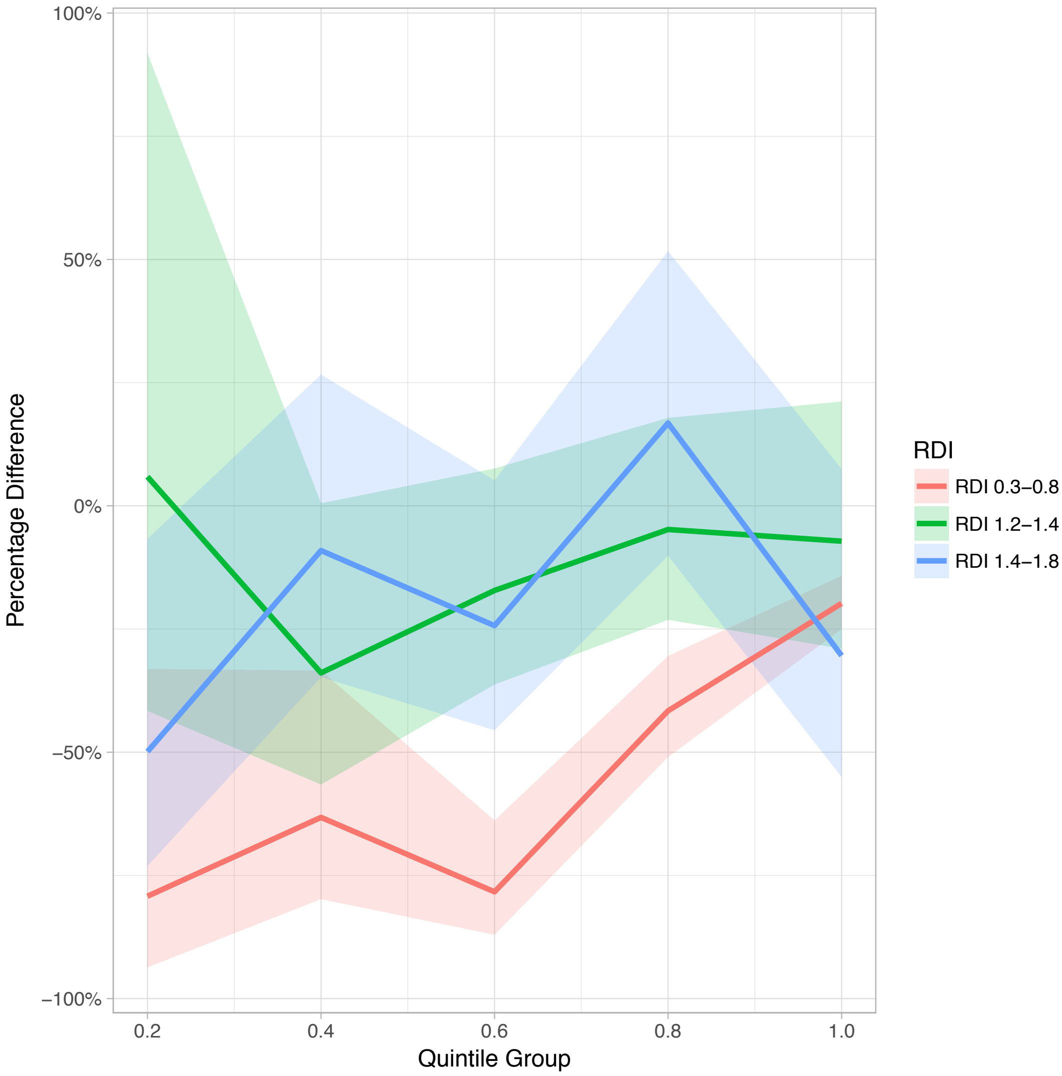


Figure 4 – Percentage difference in probability of entry between crisis and no crisis across quintile groups and across RDI groups

5 Conclusion

In this paper, we provide systematic evidence on the diversification behaviour of regions in times of major crisis. Diversification is considered to be a crucial part of regional resilience, as developing new capabilities may allow regions to come out of crises. For a long time, questions like the ones asked here relied on case studies, which although insightful were difficult to generalize. Combining developments in data availability and in methods to quantify relatedness, we were able to examine technological diversification of MSAs in the U.S. during the Long Depression, the Great Depression, and the Oil Crisis.

We found that crises have a strong dampening effect on diversification, and that especially diversification in less related technologies is reduced compared to more prosperous times, which is in line with the demand pull hypothesis (Schmookler, 1966; Freeman et al., 1982; Scherer, 1982). We also show that more diverse cities manage to diversify more than their more specialised counterparts during crises. When it comes to diversification, diverse cities profit from their diversity in two ways. First, a larger technological portfolio means having more capabilities that are related to regionally new capabilities, increasing their probability of entry. Second, we show that when controlling for this advantage, diverse regions still outperform more specialised cities. Suggesting that more diverse cities are less likely to have interests to block new developments making them generally more open to diversification. The results show that more diverse regions also outperform their counterparts when it comes to diversification in less related activities during crises. However, it is not true that diverse regions focus more strongly on unrelated diversification than more specialised regions when entering a crisis. In relative terms, the two types of regions do not significantly differ in their focus on less related technologies during crises.

These results give a detailed description on how regions diversify during major crises. However, the study remains largely descriptive, causal mechanisms can be suggested from theory but are not tested directly. In particular, future research could explore which features of diverse regions allow them to diversify more strongly than their relatedness density would suggest.

Furthermore, this paper describes how regions diversify during times of crisis but not how this impacts the depth and duration of crises. Do regions that diversify more strongly or more into less related activities experience less damage from crises, and under which circumstances?

Related diversification is suggested to be more sustainable in the long run in a city because it can build on local capabilities (Balland et al., 2018). However, a recent study (Pinheiro et al., 2018) suggests that unrelated diversification brings additional economic growth in countries in the long run but it is unsure if this also holds for cities (Coniglio et al., 2018).

The framework also allows to retrace previous case-studies in the data and compare them with a large sample of other cities. This would for example allow us to examine if the story of "Reinventing Boston" (Glaeser, 2005) is a story of unrelated diversification against the odds or a more common case of related diversification, and whether major crises like the ones we examined had a major impact on the diversification pattern in the Boston region. The described diversity of economic activities in Boston and the associated diversification through economic downturns is in line with the results here that more diverse regions outperform more specialised regions in diversification during crises. In this sense, the results also shed light on how cities like New York remain among the top largest cities through economic cycles.

This study is limited by its focus on technological diversification based on patent data. Consequently, it picks up only that part of new knowledge that is embodied in patents. In order to get a more comprehensive picture of resilience of cities, it is important to account for other forms of knowledge that may provide opportunities for cities to diversify. This would include other forms of new activities like new products, industries or new jobs in which cities can diversify, which are not captured by patent data, like in most tertiary activities (Xiao et al., 2018).

Finally, a possible improvement for future research would be to include the role of institutions in regional resilience research (Boschma, 2015). Recent research has shown that regional institutions like bridging social capital matter for the ability of regions to diversify (Cortinovis et al., 2017). This might be especially relevant in times of crises when high demands are put on institutional agents to renew their economies, adapt their institutions, and enable the development of new growth paths (Freeman and Perez, 1988; Amable, 2000; Hall and Soskice, 2001). To our knowledge, the effect of regional institutions on regional resilience, and whether institutional agents like policy makers can make the difference during major crises (Bristow and Healy, 2014; Dawley, 2014; Sotarauta et al., 2017), has not yet been systematically tested.

References

Allison, P. (2012). Logistic Regression for Rare Events [Blog post].

Amable, B. (2000). Institutional complementarity and diversity of social systems of innovation and production. *Review of International Political Economy*, 7(4):645–687.

Balland, P.-A., Boschma, R., Crespo, J., and Rigby, D. L. (2018). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, pages 1–17.

Balland, P.-A., Rigby, D. L., and Boschma, R. (2015). The technological resilience of US cities. *Cambridge Journal of Regions, Economy and Society*, 8(2):167–184.

Boschma, R. (1999). The rise of clusters of innovative industries in Belgium during the industrial epoch. *Research Policy*, 28(8):853–871.

Boschma, R. (2015). Towards an Evolutionary Perspective on Regional Resilience. *Regional Studies*, 49(5):733–751.

Boschma, R., Balland, P.-A., and Kogler, D. F. (2015). Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and Corporate Change*, 24(1):223–250.

Boschma, R. and Lambooy, J. (1999). The prospects of an adjustment policy based on collective learning in old industrial regions. *GeoJournal*, 49(4):391–399.

Bristow, G. and Healy, A. (2014). Regional Resilience: An Agency Perspective. *Regional Studies*, 48(5):923–935.

Bristow, G. and Healy, A. (2018). Innovation and regional economic resilience: an exploratory analysis. *The Annals of Regional Science*, 60(2):265–284.

Cainelli, G., Ganau, R., and Modica, M. (2018). Industrial relatedness and regional resilience in the European Union. *Papers in Regional Science*.

Carlsson, B. and Stankiewicz, R. (1991). On the nature, function and composition of technological systems. *Journal of Evolutionary Economics*, 1(2):93–118.

Castaldi, C., Frenken, K., and Los, B. (2015). Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting. *Regional Studies*, 49(5):767–781.

Christopherson, S., Michie, J., and Tyler, P. (2010). Regional resilience: theoretical and empirical perspectives. *Cambridge Journal of Regions, Economy and Society*, 3(1):3–10.

Clark, J., Freeman, C., and Soete, L. (1981). Long waves, inventions, and innovations. *Futures*, 13(4):308–322.

Coniglio, N. D., Lagravinese, R., Vurchio, D., and Armenise, M. (2018). The pattern of structural change: testing the product space framework. *Industrial and Corporate Change*, 27(4):763–785.

Coombs, R., Saviotti, P., and Walsh, V. (1987). *Economics and technological change*. Rowman & Littlefield, Totowa.

Cortinovis, N., Xiao, J., Boschma, R., and van Oort, F. G. (2017). Quality of government and social capital as drivers of regional diversification in Europe. *Journal of Economic Geography*, 17(6):1179–1208.

Crescenzi, R., Luca, D., and Milio, S. (2016). The geography of the economic crisis in Europe: national macroeconomic conditions, regional structural factors and short-term economic performance. *Cambridge Journal of Regions, Economy and Society*, 9(1):13–32.

Cuadrado-Roura, J. R., Martin, R., and Rodríguez-Pose, A. (2016). The economic crisis in Europe: urban and regional consequences. *Cambridge Journal of Regions, Economy and Society*, 9(1):3–11.

Dawley, S. (2014). Creating New Paths? Offshore Wind, Policy Activism, and Peripheral Region Development. *Economic Geography*, 90(1):91–112.

Diodato, D. and Weterings, A. B. (2015). The resilience of regional labour markets to economic shocks: Exploring the role of interactions among firms and workers. *Journal of Economic Geography*, 15(4):723–742.

Dosi, G. (1984). *Technical change and industrial transformation: the theory and an application to the semiconductor industry*. The Macmillan Press LTD, London and Basingstoke.

Dosi, G., Freeman, C., Nelson, R., Silverberg, G., and Soete, L. (1988). *Technical change and economic theory.* Pinter, London.

Duijn, J. V. (1983). *The long wave in economic life.* George Allen & Unwin, London.

Duranton, G. and Puga, D. (2000). Diversity and Specialisation in Cities: Why, Where and When Does it Matter? *Urban Studies*, 37(3):533–555.

Essletzbichler, J. (2007). Diversity, stability and regional growth in the United States, 19752002. In Frenken, K., editor, *Applied evolutionary economics and economic geography*, chapter 10. Edward Elgar, Cheltenham.

Essletzbichler, J. (2015). Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas. *Regional Studies*, 49(5):752–766.

Filippetti, A. and Archibugi, D. (2011). Innovation in times of crisis: National systems of innovation, structure, and demand. *Research Policy*, 40(2):179–192.

Fingleton, B., Garretsen, H., and Martin, R. (2012). Recessionary shocks and regional employment: Evidence on the resilience of u.k. regions. *Journal of Regional Science*, 52(1):109–133.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):27–38.

Fratesi, U. and Perucca, G. (2018). Territorial capital and the resilience of European regions. *The Annals of Regional Science*, 60(2):241–264.

Freeman, C., Clark, J., and Soete, L. (1982). *Unemployment and technical innovation: a study of long waves and economic development.* Pinter, London.

Freeman, C. and Perez, C. (1988). *Structural crises of adjustment: business cycles.* Pinter, London.

Frenken, K., van Oort, F. G., and Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5):685–697.

Geroski, P. A. and Walters, C. F. (1995). Innovative Activity over the Business Cycle. *The Economic Journal*, 105(431):916.

Glaeser, E. and Ponzetto, G. A. (2007). Did the Death of Distance Hurt Detroit and Help New York? Technical report, National Bureau of Economic Research, Cambridge, MA.

Glaeser, E. L. (2005). Reinventing Boston: 1630-2003. *Journal of Economic Geography*, 5(2):119–153.

Grabher, G. (1993). The weakness of strong ties: the lock-in of regional development in the Ruhr area. In Grabher, G., editor, *The Embedded Firm*, pages 255–277. Routledge, London.

Griliches, Z. (1981). Market value, R&D, and patents. *Economics Letters*, 7(2):183–187.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *NBER working paper series, 8498*.

Hall, P. A. and Soskice, D. (2001). *Varieties of Capitalism. The Institutional Foundations of Comparative Advantage*. Oxford University Press, New York.

Harding, D. and Pagan, A. (2002). Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics*, 49(2):365–381.

Hassink, R. (2005). How to unlock regional economies from path dependency? From learning region to learning cluster. *European Planning Studies*, 13(4):521–535.

Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43(1):59–74.

Hidalgo, C. A., Kilinger, B., Barabási, A.-L., and Hausmann, R. (2007). The Product Space Conditons the Develpment of Nations. *Science*, 317(July):482–487.

King, G. and Zeng, L. (2001). Logistic Regression in Rare Events Data. *Political Analysis*, 9(02):137–163.

Kleinknecht, A. (1981). Observations on the Schumpeterian swarming of innovations. *Futures*, 13(4):293–307.

Kleinknecht, A. (1987). *Innovation patterns in crisis and prosperity: Schumpeter's long cycle reconsidered*. Macmillan, Basingstoke, Hampshire.

Kogler, D. F., Rigby, D. L., and Tucker, I. (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. *European Planning Studies*, 21(9):1374–1391.

Kosmidis, I. and Firth, D. (2009). Bias reduction in exponential family nonlinear models. *Biometrika*, 96(4):793–804.

Krugman, P. (1993). *Lessons of Massachusetts for EMU,Adjustment and growth in the European Monetary Union, CEPR.* Cambridge University Press, Cambridge.

Martin, R. (2012). Regional economic resilience, hysteresis and recessionary shocks. *Journal of Economic Geography*, 12(1):1–32.

Martin, R. and Sunley, P. (2015). On the notion of regional economic resilience: conceptualization and explanation. *Journal of Economic Geography*, 15(1):1–42.

Mensch, G. O. (1975). *Das technologische Patt.* Umschau, Frankfurt.

Neffke, F., Hartog, M., Boschma, R., and Henning, M. (2018). Agents of Structural Change: The Role of Firms and Entrepreneurs in Regional Diversification. *Economic Geography*, 94(1):23–48.

Neffke, F., Henning, M., Boschma, R., Lundquist, K. J., and Olander, L. O. (2011). The Dynamics of Agglomeration Externalities along the Life Cycle of Industries. *Regional Studies*, 45(1):49–65.

Perez, C. (1983). Structural change and assimilation of new technologies in the economic and social systems. *Futures*, 15(5):357–375.

Petralia, S. G., Balland, P.-A., and Rigby, D. L. (2016). Unveiling the geography of historical patents in the United States from 1836 to 1975. *Scientific data*, 3:1–14.

Pike, A., Dawley, S., and Tomaney, J. (2010). Resilience, adaptation and adaptability. *Cambridge Journal of Regions, Economy and Society*, 3(1):59–70.

Pinheiro, F. L., Alshamsi, A., Hartmann, D., Boschma, R., and Hidalgo, C. A. (2018). Shooting Low or High: Do Countries Benefit from Entering Unrelated Activities? *Papers in Evolutionary Economic Geography*, 1807.

Rigby, D. L. (2015). Technological Relatedness and Knowledge Space: Entry and Exit of US Cities from Patent Classes. *Regional Studies*, 49(11):1922–1937.

Rocchetta, S. and Mina, A. (2017). Technological Coherence and the Adaptive Resilience of Regional Economies. *Papers in Evolutionary Economic Geography*, 17(13):1–39.

Rosenberg, N. (1982). *Inside the black box: technology and economics.* Cambridge University Press, Cambridge.

Rosenberg, N. and Frischtak, C. R. (1983). Long Waves and Economic Growth: A Critical Appraisal.

Rothwell, J., Lobo, J., Strumsky, D., and Muro, M. (2013). Patenting prosperity: invention and economic performance in the United States and its metropolitan areas. Technical report, Brookings Institution, Washington.

Scherer, F. M. (1982). Demand-Pull and Technological Invention: Schmookler Revisted. *The Journal of Industrial Economics*, 30(3):225.

Schmookler, J. (1966). *Invention and Economic Growth.* Harvard University Press, Cambridge.

Schumpeter, J. A. (1939). *Business cycles.* McGraw-Hill, New York.

Simmie, J. and Martin, R. (2010). The economic resilience of regions: towards an evolutionary approach. *Cambridge Journal of Regions, Economy and Society*, 3(1):27–43.

Sotarauta, M., Beer, A., and Gibney, J. (2017). Making sense of leadership in urban and regional development. *Regional Studies*, 51(2):187–193.

Steijn, M. P. A. (2018). Improvement on the association strength: implementing a probabilistic measure based on combinations without repetition. [Manuscript in Preparation].

Uhlbach, W., Balland, P.-A., and Scherngell, T. (2017). R&D Policy and Technological Trajectories of Regions: Evidence from the EU Framework Programmes. *Papers in Evolutionary Economic Geography*, 17(22):1–21.

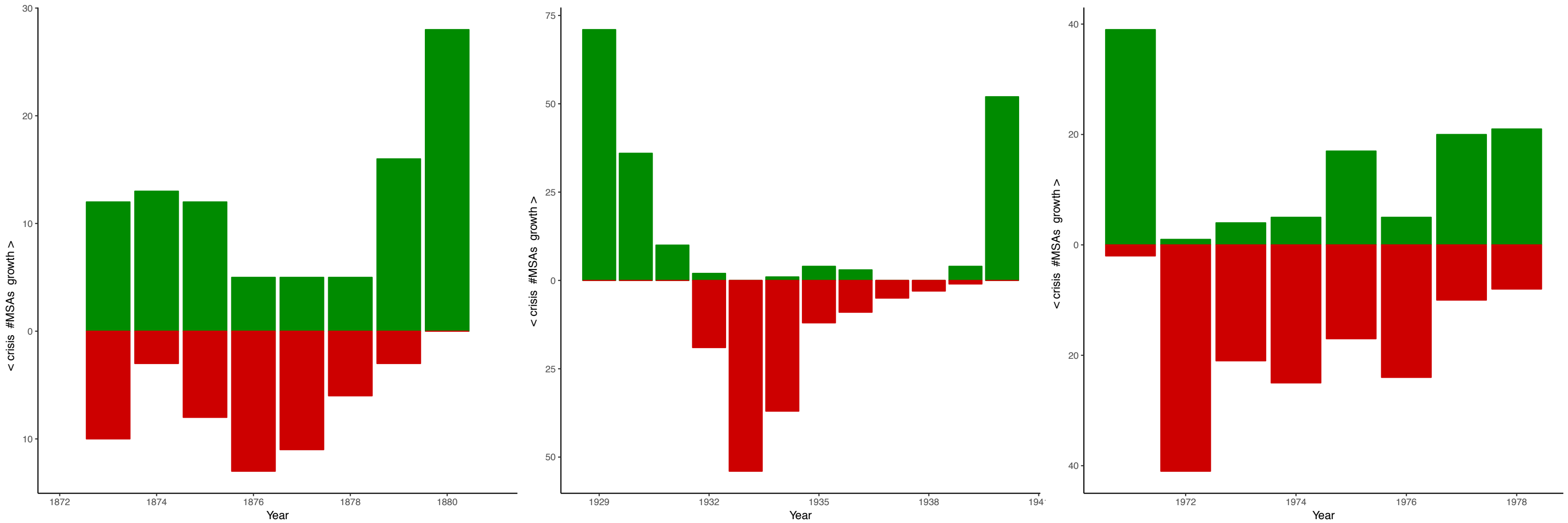
van Eck, N. J. and Waltman, L. (2009). How to Normalize Cooccurrence Data? An Analysis of SomeWell-Known Similarity Measures. *Journal of the American Society for Information Science*, 60(8):1635–1651.

Von Hippel, P. (2015). Linear vs. Logistic Probability Models: Which is Better, and When? [Blog post].

Webber, D. J., Healy, A., and Bristow, G. (2018). Regional Growth Paths and Resilience: A European Analysis. *Economic Geography*, pages 1–21.

Xiao, J., Boschma, R., and Andersson, M. (2018). Resilience in the European Union: the effect of the 2008 crisis on the ability of regions in Europe to develop new industrial specializations. *Industrial and Corporate Change*, 27(1):15–47.

Appendix A



(a) Long Depression (b) Great Depression (c) Oil Crisis

Figure A1 – Number of MSAs starting a period of growth (green) versus a period of crisis (red)

TABLE A1 – REGRESSION RESULTS (HYPOTHESIS 1 AND 2) - ROBUSTNESS CHECK

	Dependent variable:		
	entry		
	0-33.3%	33.3-66.7%	66.7-100%
Crisis	-0.563^{***}	-0.364^{***}	-0.216^{***}
	(0.146)	(0.063)	(0.031)
Population	-0.141^{*}	-0.290^{***}	-0.063^{***}
	(0.074)	(0.058)	(0.010)
Present*$\mathbf{W}$	0.556^{***}	0.421^{***}	0.314^{***}
	(0.036)	(0.019)	(0.011)
Constant	-5.564^{***}	-2.491^{***}	-2.069^{***}
	(0.630)	(0.443)	(0.687)
Time Fixed Effects	Yes	Yes	Yes
Technology Fixed Effects	Yes	Yes	Yes
MSA Fixed Effects	Yes	Yes	Yes
Observations	144,966	203,202	236,718
Log Likelihood	−7,312.256	−19,140.410	−52,671.230
Akaike Inf. Crit.	15,684.510	39,502.820	106,662.500
Note:	*p<0.1; **p<0.05; ***p<0.01		

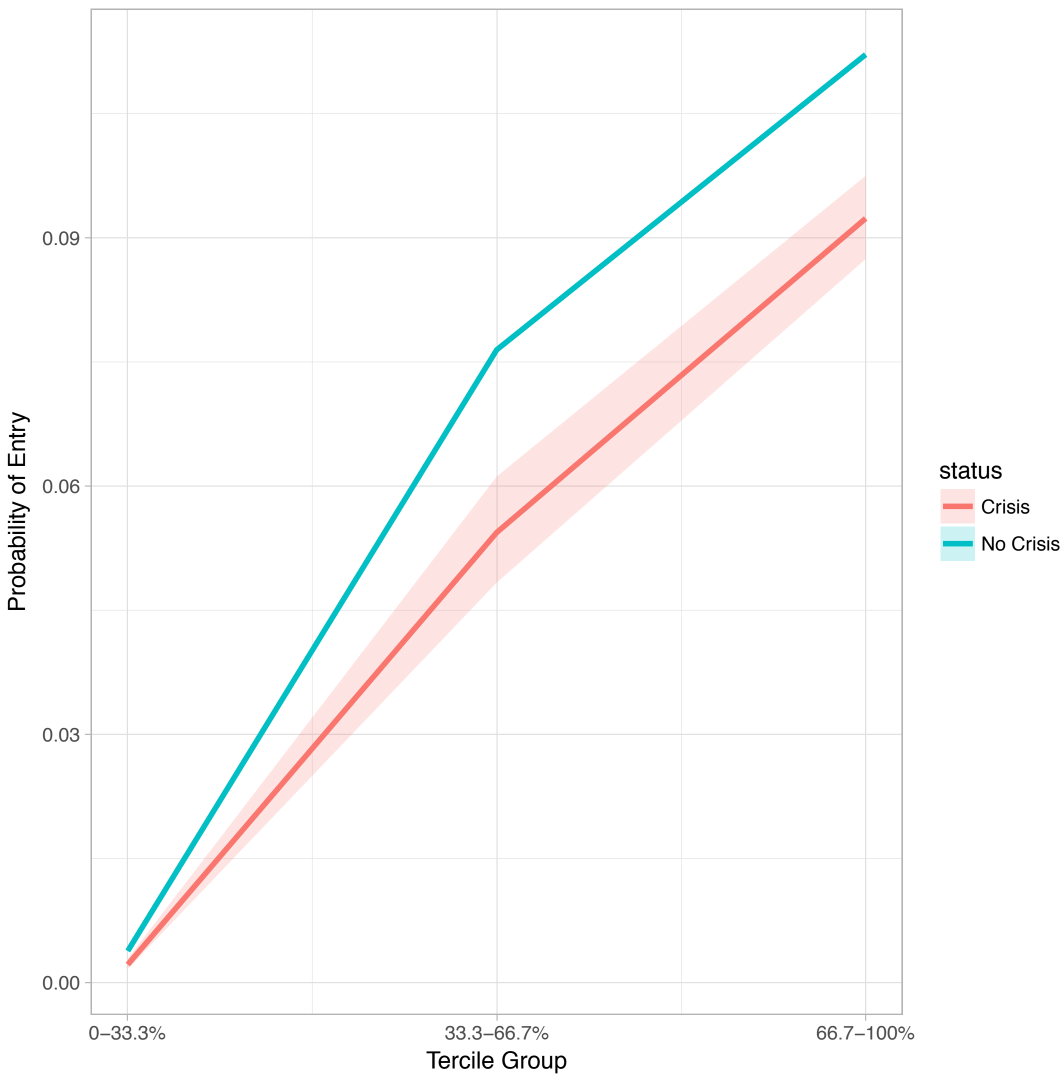


FIGURE A2 – PROBABILITY OF ENTRY ACCORDING TO CRISIS STATUS

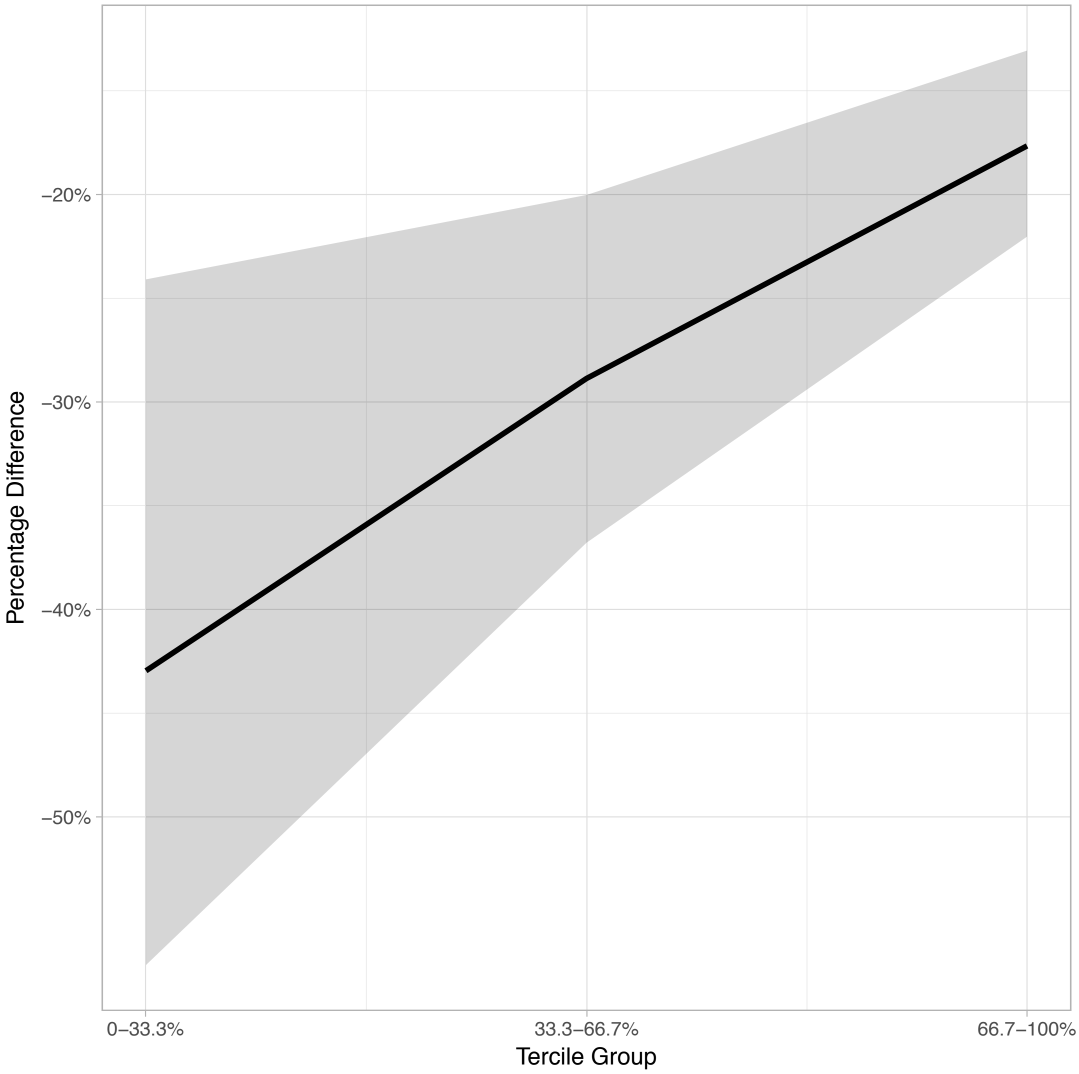


Figure A3 – Percentage difference in probability of entry between crisis and no crisis across tercile groups

Table A2 – Marginal effects (Hypothesis 3)

	Dependent variable:			
	entry			
	0-25%	25-50%	50-75%	75-100%
RDI	-0.003^{**}	-0.007^{***}	-0.010^{***}	-0.084^{***}
	(0.002)	(0.002)	(0.003)	(0.007)
Relatedness dens.	0.038	0.075	0.185^{***}	0.133^{***}
	(0.158)	(0.059)	(0.035)	(0.009)
Population	-0.001	0.001^{**}	-0.00002	-0.001^{**}
	(0.002)	(0.0004)	(0.001)	(0.001)
Present***W**	0.002^{***}	0.004^{***}	0.008^{***}	0.023^{***}
	(0.0004)	(0.0004)	(0.001)	(0.001)
Time Fixed Effects	No	No	No	No
Technology Fixed Effects	No	No	No	No
MSA Fixed Effects	No	No	No	No
Observations	19,951	25,164	28,182	28,966
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A3 – Regression results (Hypothesis 3) - Robustness check

	Dependent variable:	
	entry	
	0-50%	50-100%
RDI	-2.544^{***}	-1.600^{***}
	(0.940)	(0.267)
Relatedness dens.	-0.982	-0.075
	(0.861)	(0.057)
Population	0.948^{***}	0.331^{***}
	(0.108)	(0.031)
Present***W**	0.948^{***}	0.331^{***}
	(0.108)	(0.031)
Time Fixed Effects	Yes	Yes
Technology Fixed Effects	Yes	Yes
MSA Fixed Effects	Yes	Yes
Observations	9,310	52,587
Log Likelihood	-907.028	$-7,998.234$
Akaike Inf. Crit.	2,300.056	17,148.470
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A4 – Regression results (Hypothesis 3) - Grouped RDI

	Dependent variable:			
	entry			
	0-25%	25-50%	50-75%	75-100%
RDI 0.8-1	-3.840^{***}	-0.533	-0.324^{**}	-0.333^{***}
	(1.184)	(0.326)	(0.159)	(0.067)
RDI 1-1.2	-2.011^{***}	-1.092^{***}	-0.643^{***}	-0.682^{***}
	(0.728)	(0.338)	(0.160)	(0.080)
RDI 1.2-1.4	-2.167^{***}	-0.928^{***}	-0.452^{***}	-0.882^{***}
	(0.711)	(0.316)	(0.163)	(0.117)
RDI 1.4-1.8	-2.941^{***}	-1.155^{***}	-0.748^{***}	-1.329^{***}
	(0.714)	(0.315)	(0.177)	(0.170)
Relatedness dens.	12.819	11.471	10.650^{***}	2.513^{***}
	(64.774)	(8.278)	(1.988)	(0.164)
Population	-1.129	0.108^{*}	-0.019	0.018^{*}
	(0.924)	(0.056)	(0.044)	(0.010)
Present$*\mathbf{W}$	0.739^{***}	0.613^{***}	0.475^{***}	0.421^{***}
	(0.123)	(0.046)	(0.031)	(0.018)
Constant	-3.725^{***}	-4.486^{***}	-4.614^{***}	-3.375^{***}
	(0.562)	(0.337)	(0.210)	(0.070)
Time Fixed Effects	No	No	No	No
Technology Fixed Effects	No	No	No	No
MSA Fixed Effects	No	No	No	No
Observations	19,951	25,164	28,182	28,966
Log Likelihood	-321.898	$-1{,}008.195$	$-2{,}394.856$	$-6{,}194.191$
Akaike Inf. Crit.	659.796	2,032.391	4,805.712	12,404.380

Note 1: The reference category is RDI 0.3-0.8.

Note 2: *p<0.1; **p<0.05; ***p<0.01

Table A5 – Regression results (Hypothesis 4)

	Dependent variable:				
	entry				
	0-20%	20-40%	40-60%	60-80%	80-100%
RDI 0.8-1	−1.469***	−1.292***	−0.770***	−0.471***	−0.408***
	(0.095)	(0.080)	(0.058)	(0.039)	(0.027)
RDI 1-1.2	−2.391***	−1.847***	−1.169***	−0.710***	−0.663***
	(0.117)	(0.085)	(0.061)	(0.043)	(0.038)
RDI 1.2-1.4	−2.607***	−2.147***	−1.561***	−1.163***	−1.135***
	(0.114)	(0.087)	(0.067)	(0.054)	(0.059)
RDI 1.4-1.8	−2.903***	−2.426***	−2.032***	−1.549***	−1.408***
	(0.119)	(0.092)	(0.079)	(0.070)	(0.090)
Crisis	−1.617***	−1.032***	−1.556***	−0.554***	−0.233***
	(0.610)	(0.310)	(0.264)	(0.091)	(0.036)
Relatedness density	−26.501	6.969	9.038***	4.201***	1.706***
	(32.971)	(4.840)	(1.735)	(0.564)	(0.060)
Population	0.185***	0.049	−0.015	−0.031**	0.010**
	(0.034)	(0.038)	(0.030)	(0.015)	(0.004)
Present***W**	0.839***	0.523***	0.539***	0.446***	0.338***
	(0.034)	(0.020)	(0.014)	(0.010)	(0.007)
RDI 0.8-1 * Crisis	−0.900	0.013	0.950***	−0.072	0.047
	(1.176)	(0.415)	(0.302)	(0.133)	(0.072)
RDI 1-1.2 * Crisis	1.492**	0.301	0.763**	−0.041	−0.074
	(0.741)	(0.415)	(0.308)	(0.130)	(0.091)
RDI 1.2-1.4 * Crisis	1.675**	0.616	1.366***	0.504***	0.157
	(0.682)	(0.378)	(0.296)	(0.143)	(0.143)
RDI 1.4-1.8 * Crisis	0.925	0.936***	1.276***	0.711***	−0.134
	(0.688)	(0.354)	(0.313)	(0.163)	(0.228)
Constant	−2.825***	−2.988***	−3.374***	−3.251***	−2.764***
	(0.062)	(0.085)	(0.084)	(0.065)	(0.026)
Time Fixed Effects	No	No	No	No	No
Technology Fixed Effects	No	No	No	No	No
MSA Fixed Effects	No	No	No	No	No
Observations	144,950	144,950	144,951	144,950	144,951
Log Likelihood	−4,854.524	−7,728.097	−12,705.440	−21,631.490	−39,694.050
Akaike Inf. Crit.	9,735.049	15,482.190	25,436.890	43,288.970	79,414.100
Note 1:					The reference category is RDI 0.3-0.8.
Note 2:					*p<0.1; **p<0.05; ***p<0.01

Table A6 – Regression results (Hypothesis 4) - Robustness check

	Dependent variable:		
	entry		
	0-33.3%	33.3-66.7%	66.7-100%
RDI 0.8-1	-0.534^{***}	-0.622^{***}	-0.257^{***}
	(0.114)	(0.058)	(0.031)
RDI 1-1.2	-0.954^{***}	-0.749^{***}	-0.282^{***}
	(0.137)	(0.068)	(0.043)
RDI 1.2-1.4	-1.192^{***}	-1.121^{***}	-0.548^{***}
	(0.148)	(0.077)	(0.059)
RDI 1.4-1.8	-1.284^{***}	-1.311^{***}	-0.785^{***}
	(0.150)	(0.087)	(0.080)
Crisis	0.127	-0.659^{***}	-0.125^{***}
	(0.351)	(0.175)	(0.040)
Relatedness density	25.099^{***}	11.225^{***}	2.678^{***}
	(6.081)	(0.816)	(0.059)
Population	-0.157^{**}	-0.347^{***}	-0.056^{***}
	(0.065)	(0.061)	(0.011)
Present***W**	0.551^{***}	0.410^{***}	0.293^{***}
	(0.036)	(0.019)	(0.011)
RDI 0.8-1 * Crisis	-1.813^{***}	0.326	-0.047
	(0.685)	(0.211)	(0.068)
RDI 1-1.2 * Crisis	-0.228	0.187	-0.245^{***}
	(0.467)	(0.209)	(0.079)
RDI 1.2-1.4 * Crisis	-0.293	0.721^{***}	0.015
	(0.431)	(0.208)	(0.110)
RDI 1.4-1.8 * Crisis	-0.525	0.480^{**}	-0.038
	(0.401)	(0.217)	(0.149)
Constant	-6.580^{***}	-2.488^{**}	-3.374^{***}
	(1.175)	(1.119)	(0.715)
Time Fixed Effects	Yes	Yes	Yes
Technology Fixed Effects	Yes	Yes	Yes
MSA Fixed Effects	Yes	Yes	Yes
Observations	193,453	227,784	239,775
Log Likelihood	$-7,803.078$	$-19,271.320$	$-51,718.460$
Akaike Inf. Crit.	16,858.160	39,910.630	104,830.900

Note 1: The reference category is RDI 0.3-0.8.

Note 2: *p<0.1; **p<0.05; ***p<0.01

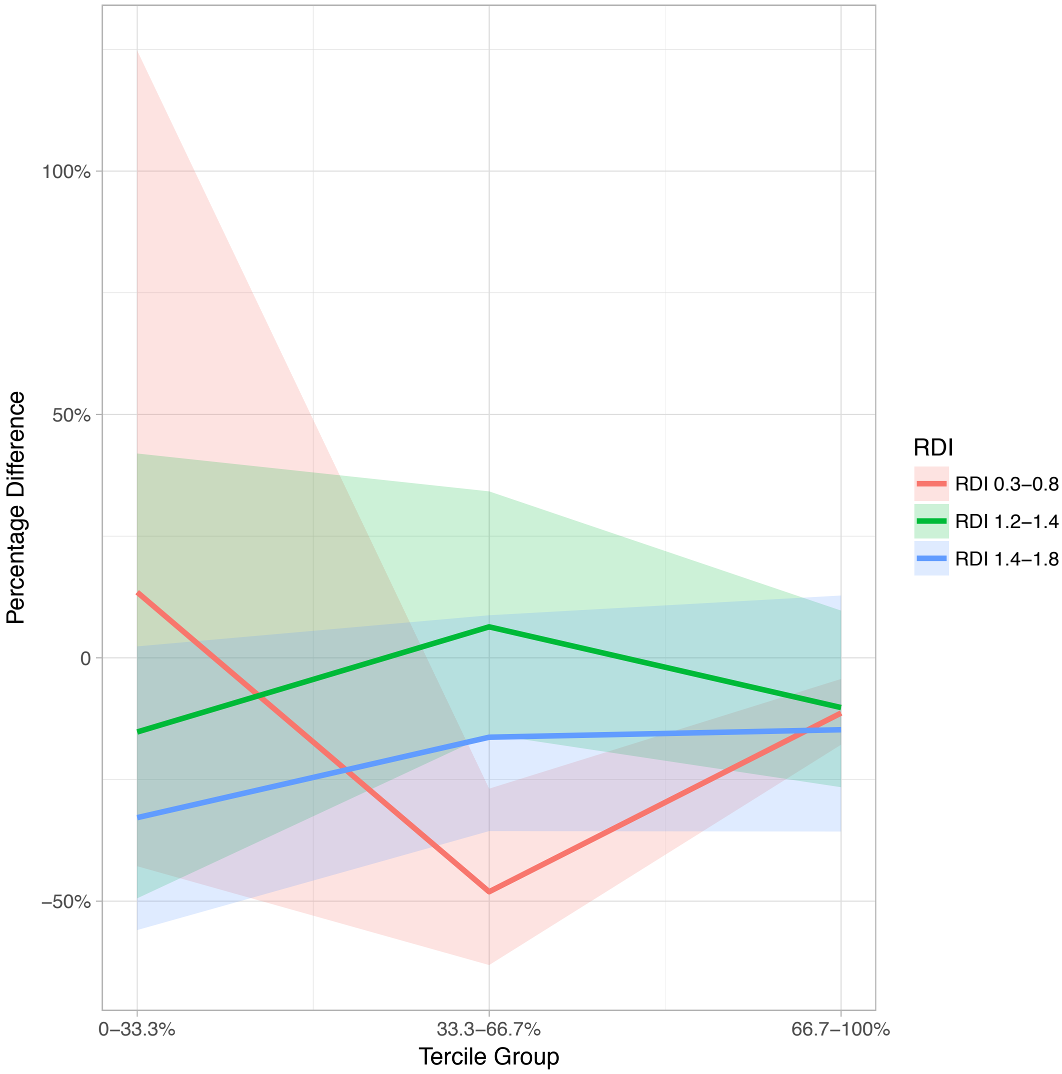


Figure A4 – Percentage difference in probability of entry between crisis and no crisis across tercile groups