CPB Netherlands Bureau for
Economic Policy Analysis

Entries and Regional Growth: The Role of Relatedness

If new businesses and establishments are more closely related to existing economic activities in a region, is this associated with higher regional growth a few years later? We investigate how the relatedness of entrants relates to the growth of Dutch regions using detailed microdata on all Dutch firms, establishments and workers. We find a positive association between this related entry and employment growth, but no effect on productivity growth.

Although this positive association is statistically significant, its economic magnitude is small. Hence, it is probably difficult to substantially influence regional growth through related entry, for example by trying to attract related newcomers to a region.

CPB Discussion Paper

Tijl Hendrich, Jennifer Olsen, Judith Bayer

January 2022

Doi: https://doi.org/ 10.34932/jpdz-jt86

Entries and Regional Growth: The Role of Relatedness*

Tijl Hendrich[†], Jennifer Olsen[‡], Judith Bayer[§]

January 18, 2022

This study examines how the degree of relatedness between entering and incumbent establishments affects regional growth. To this end, we construct a rich panel dataset on establishment entry using administrative microdata on establishments, workers and NUTS-3 regions in the Netherlands. We use revealed skill-relatedness between sectors to define a novel measure for related entry of establishments on the sector-regional and regional level. On the sector-regional level, we find no evidence that entry of establishments is directly associated with employment growth in related sectors. Related entry is negatively associated with short term sector-regional productivity growth. In contrast, we do find robust evidence that an increase of related entry is tied to employment growth on the regional level. This suggests that related entry generates spillovers between groups of related sectors driving aggregate employment growth.

Keywords: Firm Entry, Relatedness, Related Entry, Agglomeration externalities, Regional growth, Skills
JEL codes: L26, M13, O18, R11

*We thank Ron Boschma, Erik Stam, Michael Fritsch, Rob Euwals, Harro van Heuvelen, Benedikt Vogt, Andrei Dubovik, participants of the EconGeo seminar at Utrecht University and participants of the CPB seminars for their valuable comments. We thank Anet Weterings for providing us with the dataset on skill-relatedness between Dutch sectors. The usual disclaimer applies. Our results are based on our own calculations using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl. The analysis code is available upon request.

[†]CPB Netherlands Bureau of Economic Policy Analysis, corresponding author, email: t.j.m.hendrich@cpb.nl

[‡]CPB Netherlands Bureau of Economic Policy Analysis

[§]Laterite - from data to policy

1. Introduction

Within the European Union, regions have been diverging in terms of GDP per capita and employment since the mid-2000s, both between and within countries (Iammarino et al., 2018). Place-based policies, intended to stimulate growth, often aim at attracting new businesses to an area, be it directly through subsidies and enterprise zoning or by making the region itself more attractive to new entrants (Ehrlich and Overman, 2020). Entrants can stimulate growth through a process of *creative destruction*, driving less productive incumbent firms out of the market or forcing them to produce more efficiently (Schumpeter, 1942). New entrants can also introduce new ideas and activities into a region, generating knowledge spillovers and enriching local agglomeration economies (Fritsch, 2015).

A growing body of literature has investigated effects of new entrants on regional productivity and employment growth (Fritsch, 2013). This literature highlights that the process of starting firms and establishments takes place within a local context (Michelacci and Silva, 2007; Rosenthal and Strange, 2020; Kemeny and Storper, 2015; Stam, 2007), resulting in heterogeneous effects of entries across regions with different qualities such as population density and share of highly educated workers (Fritsch and Schroeter, 2011). Beyond the local context, entries clearly also have a sectoral context, and a large body of literature is also devoted to the role entrants play in sectoral productivity (Fritsch, 2013; Aghion et al., 2009).

However, little research has been devoted to *intersectoral* spillover effects of entries on the regional level, of which Fritsch (2013) provides an overview. For instance, new entrants may induce a greater variety of ideas, since they are more inclined to introduce new, innovative products than incumbent establishments, stimulating economic development. The same line of reasoning holds for innovations in the production process.[1]

Meanwhile, the regional development literature has progressed further in the investigation of intersectoral relations, hypothesizing that local spillovers between sectors (and resulting co-agglomeration forces) play a significant role in regional growth (Boschma and Frenken, 2011; Content and Frenken, 2016). Recent advances have focused on defining *revealed relatedness* measures which quantify these intersectoral spillovers. There is evidence that in general, the degree of relatedness between sectors in the regional portfolio may generate growth by an increase in the quality and quantity of spillovers between related sectors (Frenken et al., 2007; Content and Frenken, 2016; Boschma, 2017). Spillovers between related firms have also been put forward as a driver of regional employment growth (Content and Frenken, 2016), plant level productivity growth (Timmermans and Boschma, 2014), plant survival (Neffke et al., 2012; Cappelli et al., 2019) and regional diversification (Boschma, 2017).

In this paper we aim to expand on the regional entry literature by exploring whether the relatedness between entering and incumbent establishments impacts the effect that

[1] Andersson and Noseleit (2011) find significant cross-sectoral effects of entrants in one sector on regional employment growth in other sectors (both negative and positive), indicating that interactions between sectors are a relevant area for further study. This study expands on this research by explicitly modeling the relationship between sectors

entries have on regional growth. Is the effect of entry on regional growth moderated by how related the entry is to the region? We hypothesize that the degree of relatedness between entrants and the incumbent activities in a region positively affects regional growth. In order to test our hypothesis, we aggregate Dutch administrative microdata on firms, establishments and workers to perform a panel analysis of 181 3-digit (SBI) sectors in 39 NUTS-3 (COROP) regions over a period of 11 years (2007 - 2018).

Our contributions are threefold. First, we take intersectoral relatedness into account while studying the entry of new establishments as a mechanism for growth of regions and sectors within these regions. The effects of entry on regional growth have been researched in the past, but mainly using aggregate, regional variables as inputs. We aim to capture spillover effects between sectors-in-regions by modeling intersectoral relations explicitly through a novel measure of *related entry*. We use a measure of *revealed skill-relatedness*, based on observed labour flows, to determine intersectoral relatedness (Neffke and Henning, 2013; Neffke et al., 2017). We isolate the intersectoral effects of related entries, by controlling for entry within the own sector and region. We also control for regional relatedness between *incumbent* firms. This simultaneously allows us to explore these variables as alternative mechanisms for regional growth.

Secondly, we are able to evaluate entry effects at a detailed level by using administrative microdata from Statistics Netherlands. In previous efforts, it has proven difficult to distinguish actual firm entries from administrative events, such as mergers and restructurings (Bosma et al., 2011), or to separate starting firms from incumbent firms opening a new branch (Andersson and Noseleit, 2011). These data restrictions increase the likelihood that firm entry figures are overestimated (See for instance Dejardin (2011); Neffke et al. (2011); Asturias et al. (2017)). The administrative microdata used in this study contains information about the reason why a newly appearing firm is started or terminated, which enables us to separate real entry events from administrative events. It also allows us to distinguish true start-ups (new firms) from other entrants (e.g. new branches of existing firms).

Finally, we capture variation in related entry and growth both at the sector-regional and regional level. This allows us to evaluate whether sector-regional spillovers lead to aggregate effects on the regional level. Usually only one of these two dimensions are analyzed (Fritsch, 2013; Content and Frenken, 2016), while the choice of sectoral and geographical aggregation is known to impact results on the meso level (Beaudry and Schiffauerova, 2009). For example, contrasting effects on the sector-regional level may cancel each other out on the regional level, leading to insignificant aggregate effects. Conversely, interactions between sector-groups may lead to significant regional effects where no effects are observed on the individual sector-regional level. This is known as the ecological fallacy: what is true for an individual is not necessarily true for a group. By analyzing both levels of aggregation we are able to see whether these mechanisms are at play. Using data from Statistics Netherlands we are able to measure sector-regional entry and growth rates, which allows detailed analysis of intrasectoral, intersectoral and aggregate regional effects.

We arrive at two main findings. First, we find that related entries have a small negative impact on short-term productivity growth, with no significant effect on employment at the sector-regional level. This suggests that relatedness *does not* moderate the relationship between entries and growth of individual sectors within regions. If we look at relatedness of existing firms, a higher local presence of related *incumbent* firms is positively associated to productivity growth of a given sector-region, but dampens its employment growth. Our results on intra-sectoral and regional entry largely align with literature.

Second, in contrast with our sector-regional findings, we do observe a significant and robust increase in regional employment growth following an increase in related entry on the aggregate regional level. This suggests that the effects of related entry may be driven by spillovers between multiple sectors within a region. For instance, spillovers could occur between one (entering) sector and a specific combination of other (incumbent) sectors, a dynamic that our individual sector-region analysis cannot capture. This result also provides a nice illustration of the pitfalls of the ecological fallacy: The reaction of the entire group of sectors active in a region does not merely appear to represent the sum of individual reactions from individual sectors. Instead, the interactions between sectors appear to create a whole that is greater than the sum of its parts. Hence, exploring the group dynamics of related entry on different sector-clusters may be a fruitful direction for follow-up research.

The paper is structured as follows. In section 2, we discuss the literature on firm entry, relatedness and regional growth. Subsequently section 3 outlines our empirical framework, while section 4 describes the data. The results are presented in section 5 and discussed in section 6.

2. Theoretical Background

Our paper combines two separate strands of literature relating to regional and sectoral growth. The first strand of literature deals with the effects of firm entry on growth, whereas the second studies the influence of agglomeration economies on growth. In this paper we bridge these two strands of literature by accounting for effects of firm entries on the growth of *related* sectors in a region.

2.1. Entry and Regional growth

New market entries, whether it be start-ups or new branches of existing firms, can have direct as well as indirect effects on regional employment and productivity growth. Direct effects accrue to the entering firm itself: A new entry increases employment through providing new jobs. However, effects on regional employment and productivity also depend on the indirect effects entries have on incumbent firms in the region. In this paper we focus on these indirect effects. Specifically, we consider the effects entries may have on incumbent firms in related industries. The use of microdata allows us to look at the effects of firm entry on a deeper level, explicitly modeling relations between firms.

Both the theoretical and the empirical evidence point to a positive relationship between higher rates of entry and *productivity* growth on the regional- and sectoral- level, through a process of creative destruction (Asturias et al., 2017; Aghion et al., 2009;

Fritsch, 2013). According to the creative destruction hypothesis, the competitive pressure from new entrants forces incumbent firms to become more productive in order to stay afloat, pushing less productive firms out of the market (Schumpeter, 1942). However, gains may take time to materialize, with some studies pointing to a negative short-term effect (as firms transition to new innovations) followed by a larger positive effect over time (Andersson et al., 2011; Carree and Thurik, 2010). Effects are also heterogeneous across countries and sectors: Bartelsman et al. (2004) finds large between-country differences in the rate of entries and exits and in their impact on productivity. More localized, within country analyses show that entrants tend to have a more pronounced impact on services than on manufacturing (Andersson et al., 2011; Bosma et al., 2011; Dejardin, 2011).

The effect of the creative destruction process on *employment* growth is less clear cut. In theory, productivity improvements may reduce the amount of labour needed for constant output, negatively impacting employment. On the other hand, employment growth can be realized if surviving firms can capture latent demand or create new demand through the introduction of new innovations. Regional employment may also benefit from local sectors becoming more competitive in extra-regional markets, increasing their export potential (Jacobs, 1970; Kirchhoff and Phillips, 1988). Using German data, Fritsch and Schroeter (2011) find evidence that entries indeed tend to boost regional employment, although Fritsch and Mueller (2004) demonstrate that effects show a wave-shaped pattern over time. Newcomers generate intra-industry employment growth in the first two years after entering the market, followed by a period of contraction (crowding-out), before generating more growth after about 6 years. Positive effects on growth appear to peak after 7-8 years (Fritsch and Mueller, 2004). van Stel and Suddle (2008) find similar results for the Netherlands, although they conclude that positive effects of entry peak after about 5 years, and fade away around 7 years after entry. Furthermore they find that the overall effect of entries are negative in a subset of rural Dutch areas. Thus, effects on employment can be both negative or positive, and may depend on both the scope of time being analyzed and on regional characteristics.

Finally, entries may also contribute to the local agglomeration benefits of a given region. First, they may contribute to an increase in the scale (and associated urbanization externalities) of a region. Secondly they may also shape the industrial composition of the region, making it more diverse or specialized, thereby creating new opportunities for innovation (Fritsch, 2013; Content and Frenken, 2016).

2.2. Agglomeration, Relatedness and Regional growth

Besides firm dynamics, agglomeration externalities are an important determinant of regional growth (Glaeser et al., 1992; Ellison and Glaeser, 1997; Combes, 2000; Braunerhjelm and Borgman, 2004; Ellison et al., 2010). Duranton and Puga (2004) present three broad economic benefits of agglomeration: (1) *sharing* common resources; (2) facilitating a larger pool of workers for firms to choose from and vice versa, ensuring better *matches* between employees and employers or between producers and suppliers; (3) and the occurrence of knowledge spillovers, i.e. *learning*.

In the past decades, two distinct views have emerged on what type of local sectoral composition is most beneficial for regional agglomeration economies, and resulting growth. According to Marshall (1890), specialization or the spatial concentration of firms *in the same industry* provides the best basis for regional growth. Specialization generates industry-specific agglomeration economies, as firms can share a specialized pool of labour and suppliers, and exchange intra-industry knowledge (*localization externalities*). A contrary view, most often associated with the work of Jacobs (1970), emphasizes the importance of a *diverse industrial mix*: Through novel combinations of ideas, diversity is argued to stimulate innovation and growth. Diversity is also argued to reduce risks, by reducing the exposure of an area to industry-specific shocks. Although the relative importance of specialization and diversity has been the subject of many empirical studies, the evidence on these competing theories has not been conclusive (Beaudry and Schiffauerova, 2009; Kemeny and Storper, 2015; Groot et al., 2016). [2]

More recently this dichotomy has been partially reconciled by adopting the concept of *related variety*, which takes into account the *relatedness* between different sectors (Frenken et al., 2007). Related variety captures the idea of a local industrial composition featuring a diverse set of industries that are related closely enough to facilitate knowledge spillovers (i.e. learning). Industries can be related to each other through using similar technologies, intermediate inputs, production processes, logistical networks, or skill sets, all of which can be avenues of fruitful knowledge exchange. Crucially, differences between related sectors may enhance the diversity, and thus the quality of knowledge spillovers that can take place (Nooteboom, 2000; Neffke and Henning, 2013; Boschma et al., 2012; Boschma, 2017). Arguably, firms that are from related sectors may also benefit as much from sharing and matching as firms that are from the same sector, thus generating co-localization benefits.

Although several studies point to a positive relationship between related variety and regional growth, it is not clear whether relatedness has a greater impact on employment or productivity. Theoretically, Frenken et al. (2007) argue that spillovers between diverse industries are more likely to lead to radical innovation creating entirely new economic activity and thereby also generating new jobs. However, the empirical insights are mixed, with some studies indicating that regional inter-industry relatedness is associated with a higher productivity growth, but not employment growth - and other studies finding the opposite result (Content and Frenken, 2016). Besides growing, regions with a higher related variety are also more resilient to negative employment shocks (Diodato and Weterings, 2015; Cainelli et al., 2019). Regions also tend to diversify into industries that are related to the regional portfolio (Neffke et al., 2011; Content and Frenken, 2016; Boschma, 2017), dubbed "the principle of relatedness" for its strong robustness across multiple dimensions (Hidalgo et al., 2018).

On a more granular level, Timmermans and Boschma (2014) and Cainelli et al. (2019) find that a higher inflow of workers from related sectors leads to higher survival rates

[2] This is partly because of strong context specificity (there does not appear to be an industrial profile that is equally suited to all regions), and partly due to substantive conceptual and methodological challenges.

and productivity growth on the plant level, whereas Boschma et al. (2014) find that a larger transfer of workers between related firms generates productivity growth on the regional level. Neffke et al. (2012) also find that plants are more likely to survive in the proximity of related activities.

2.3. Related entry

To date, the regional entry literature has mainly been preoccupied with *intra-sectoral* effects of firm entries (Fritsch, 2013). However, in light of the recent insights summarized above, it seems intuitive that entries should also stimulate *inter-sectoral* spillover effects. For instance, new entrants may induce a greater variety of products or production processes into a region, since they are more inclined to innovate than incumbent establishments (Fritsch, 2015), stimulating knowledge spillovers to related firms.[3]

Little research has been devoted to *intersectoral* spillover effects of entries on the regional level, and to our knowledge, no previous study has looked at the regional impact of entries on related sectors. Andersson and Noseleit (2011) study inter-industry entry dynamics of startups, and find significant positive and negative employment effects of entry in one sector on other sectors. Nonetheless, they do so only for a classification of three broad sectors, and do not account for the relatedness between the sectors. Fritsch and Schroeter (2011) consider different regional characteristics affecting the impact of entries, but do not consider sectoral dynamics. The study finds a positive regional employment effect of firm entries which is larger in regions that are densely populated, have a higher share of medium-skilled workers and a higher share of R&D employees.

An analysis by Bosma et al. (2011) on regional Dutch entry data comes closest to our study. They explore effects of relatedness and firm entry on regional growth, although they do not measure how related each entry is to the region. Instead they interact the regional entry rate with *related variety*, which is a measure capturing how related incumbent firms in the area are to one an other. Bosma et al. (2011) find a positive relationship between entries interacted with regional relatedness, and regional growth.

Based on the previous literature, our initial hypothesis is that a higher degree of entry from related firms may positively benefit regional productivity- and employment-growth in related sectors, ultimately leading to a higher aggregate level of growth for the region. However, in the short run it is also possible that related entries negatively impact employment in related sectors by crowding out the local labour market. The size of the regional labour market is inelastic in the short run, so that entry of more related firms may increase local competition for labour between firms. Consistent with the findings of Andersson et al. (2011) and Carree and Thurik (2010), it is also possible that we find temporary negative effects on productivity over the short term as incumbent firms adjust their production processes to absorb new innovations.

[3]Indirectly, this notion of intersectoral spillovers from entering firms also links to the theory of industry clustering by spin-off formation (Klepper and Simons, 2000; Klepper, 2007, 2010). The idea is that highly productive incumbents spawn more and better performing spin-off firms in the same industry, which would increase the propensity of firms in related industries to diversify into the former. See also Boschma (2015).

3. Estimation Strategy

The hypothesis we wish to test is whether the effect of establishment entries on local growth is moderated by the level of skill-relatedness between entering firms and the incumbent economic activities in a region. Is entry in related sectors, j, associated with employment and productivity growth of sector i in region r, given the degree of relatedness of j to the sector i?[4] Furthermore, on the aggregate regional level, we wish to see whether regional growth is influenced by the relatedness of entrants with the regional portfolio of incumbent sectors.

Our aggregate regional analysis explores whether (related) entries are associated with aggregate growth on the regional level, and captures group-level dynamics which may not be observable in an analysis of individual sector-regions. This is ultimately the result which may be of most interest to policy-makers, as regional growth in one sector is not very helpful if it simply displaces growth in others. The benefit of additionally analyzing individual sector-regions is that it provides us with insights into the sectoral dynamics which may be driving our regional results (Andersson and Noseleit, 2011). An analysis at the sector-region level also provides us with a larger dataset, and allows us to explore the impact of sector-level covariates such as the sectoral skill composition.

3.1. Dependent variables

In our analysis we define growth in terms of employment and labour productivity.[5] We obtain labour productivity by measuring the value added per hour worked and employment as the total number of hours worked in a region in a given year. Because employment and productivity distributions both have a long right tail, we define growth rates as the difference in log-levels over time:

$$growth_{irt} = \frac{1}{2}\left(\ln level_{irt} - \ln level_{ir,t-2}\right) \tag{1}$$

For small values, equation 1 can be interpreted as the percentage growth of the variables. To reduce the influence of year-on-year volatility, we use average growth over two years.

3.2. Explanatory variables

The explanatory variable of primary interest is our novel measure of related entry, $relentry_{irt}$. This measure is a combination of relatedness and entry, capturing how

[4] In the remainder of the paper, we will use subscripts i and j for sectors (sector pairs), r for COROP regions and t for time in years. To indicate (degree of) relatedness, we use the convention origin-destination, so that subscript ij defines a relation from i to j. For expressions that are equivalent on the regional and sector-regional level, we will only report the sector-regional variant for brevity.

[5] The downside of using labour productivity as opposed to more broad productivity measures such as Total Factor Productivity (TFP), is that capital is not taken into account. However, we are unable to capture the dispersion of capital on the sector-regional level in order to construct a TFP variable at this level of aggregation.

related sector i is to regional entries from all other sectors. However, we are also interested in observing the effects of intra-industry entry rates within a region, $entry_{irt}$. In other words, we distinguish between regional entry in the own sector ($entry_{irt}$) and regional entry in all other sectors, weighted by relatedness ($relentry_{irt}$).

By choosing entry rates as explanatory variables, we follow related literature on regional employment and productivity developments in relation to entrepreneurship.[6] Fritsch (2013) argues that new business formation is more a cause than a symptom of growth, given that the effect of business formation on economic development in subsequent periods is more salient than the other way around.

We measure entries as a percentage of existing firms within that sector-region (equation 2).[7] This way, we account for the fact that new entries may have a larger impact on relatively small industries.

$$entry_{irt} = \frac{N_{entries,irt}}{N_{firms_{irt}}} \tag{2}$$

Subsequently, we use skill-relatedness as our measure of relatedness between any two three-digit sectors within the Netherlands. The concept of skill-relatedness was coined by Neffke and Henning (2013) as a way to measure how related two industries are in terms of the required set of skills, revealed through inter-industry labor flows. If a larger labor flow is present from industry i to industry j, this is an indication that there is a stronger overlap between the skill requirements in i and j. This overlap may make it easier for switching workers to transfer knowledge and technologies across sectors, thereby promoting knowledge spillovers. For further details on skill-relatedness, see appendix B.

We use the *closeness* of each sector to other sectors in a region to define the degree of relatedness rel_{ir} between sector i and region r (equation 3). Neffke et al. (2011) equate the concept of closeness with the concept of relatedness, arguing that skill-relatedness captures the potential of forming inter-industry networks.[8] We define a sector to be strongly related to another if the skill-relatedness exceeds a cut-off value. In this study we apply a cutoff-value of 0.5 (this roughly corresponds to the 20% most related sectors).[9]

[6] See e.g. Fritsch (1997); van Stel and Suddle (2008); Aghion et al. (2009); Andersson and Noseleit (2011); Bosma et al. (2011); Fritsch and Schroeter (2011); Delfmann and Koster (2016)

[7] Earlier studies vary in how they measure entry rates. In Bosma et al. (2011); Dejardin (2011) entry rates are defined using the ecological approach (the number of entries divided by the number of existing establishments, Audretsch and Fritsch (1994, 1999)), while Andersson and Noseleit (2011) uses the labour market approach (change in the number of establishments with zero or one employee divided by existing employment). We opt to use the ecological approach, for two reasons. First, since employment is one of our dependent variables, the labour market approach could generate confounding effects between our dependent and explanatory variables. Second, our analysis includes entry of new branches, which will not necessarily start off with one or zero employees.

[8] The closeness definition in 3 is equivalent to the concept of network centrality in graph theory. Formally, the sector structure of each region is interpreted a separate weighted graph, with the sector-regions as nodes and their pairwise skill-relatedness values as edges.

[9] As a robustness check, we repeat our analyses using the 10th percentile of skill-relatedness as cutoff value (many sectors weakly related) as well as the 90th percentile (few sectors strongly related), see appendix D. Neffke et al. (2011) use c = 0.25, but show that in fact this specification of closeness is

Then we define rel_{ir} as the number of sectors in a region that are related to sector i, divided by the total number of sectors present in a region (N_r). [10]

$$rel_{ir} = \frac{1}{N_r - 1} \sum_{j \in PF_r, j \neq i} I(SR_{ji} \geq c \equiv 0.5) \tag{3}$$

In equation 3, I(...) is the indicator function, N_r is the number of sectors active in r and c is the cut-off value of skill-relatedness above which sectors are considered to be related. Since skill-relatedness does not depend on time, rel_{ir} is also time-independent in our specification. We write SR_{ji}, because the skill-relatedness is defined bidirectionally (origin and destination) and we are interested in the potential of a sector j to induce co-localization externalities such as knowledge spillovers towards i by supplying workers with related skills.

Knowing the portfolios of related sectors for each sector-region, we define related entry $relentry_{irt}$ as the regional entry rate of sectors that are strongly related to i (equation 4).[11] If the results follow our hypothesis, then a higher $relentry_{irt}$ will be associated with a higher regional growth.

$$relentry_{irt} = \sum_{j \in PF_r, j \neq i} \frac{N_{entries,jrt}}{N_{est,jrt}} \tag{4}$$

Besides relatedness of each sector to other sectors, we include the regional employment share of each 3-digit sector in our starting year as a measure of specialization on the sector region-level, $spec_{ir}$. Similarly on the regional level, we use the highest 2-digit regional employment share as a measure for regional specialization $spec_r$.

In constructing the regional counterpart of $relentry_{irt}$, $entry_{irt}$ and rel_{ir}, we account for regional sector structure by weighting each sectoral contribution by its regional employment share at the start of our observation period (equations 5-7). Thus the regional related entry measure ($relentry_{rt}$) increases if the largest sectors in the region experience more related entries, the regional specialized entry variable ($specentry_{rt}$) increases if its largest sectors experience more entries from their own sector, and the regional relatedness measure (rel_r) increases if larger sectors (in terms of employment) are more related to other sectors that are present in the region. We also look at the effects of (unweigted) entry at the regional level ($entry_{rt}$). This is simply measured as the number of entering establishments divided by the number of existing establishments in the region.

robust to a wide range of values.

[10]The regional sector portfolio size varies substantially between COROP regions, from 41 3-digit sectors in "Delfzijl and surroundings" to 150 in "West-North-Brabant"

[11]Note that entrants are only included in our analysis if they already belong to the portfolio of existing activities in a region, PF_r, otherwise there is a zero-denominator in equation 4

$$relentry_{rt} = \sum_{i=1}^{N_r} \left(\frac{emp_{ir,2007}}{emp_{r,2007}} \right) relentry_{irt} \tag{5}$$

$$specentry_{rt} = \sum_{i=1}^{N_r} \left(\frac{emp_{ir,2007}}{emp_{r,2007}} \right) entry_{irt} \tag{6}$$

$$rel_r = \sum_{i=1}^{N_r} \left(\frac{emp_{ir,2007}}{emp_{r,2007}} \right) rel_{ir,2007} \tag{7}$$

Note that in our baseline analyses we define all new establishments in a region as entrants, even if they represent a new branch of an existing firm. After all, a new branch may bring new activities to a region despite being part of an existing firm. However we also conduct a separate analysis for start-ups to see whether this group of entrants has a different impact on regional growth. An entry is defined as a start-up if it truly represents a new firm (note that we do not consider self-employed as entrants in our analysis).

3.3. Estimation models

As a first stage to our main analysis, we estimate the direct relation between more related establishment entries and subsequent sector-regional growth using a panel fixed-effects regression. Since we measure growth over two years (from year $t-2$ to year t), we measure our time-varying explanatory variables in year $t-3$. Thus we avoid any overlapping periods of observed entry and observed growth in our analyses.

Equation 8 compares the contributions of related entry ($relentry_{ir,t-3}$) to those of entry in the own sector-region ($entry_{ir,t-3}$), regional entry ($entry_{r,t-3}$) and regional entry given relatedness of the sector to the incumbent activities of a region ($entry_{r,t-3} \times rel_{ir}$). We include the latter interaction term to ensure that our $relentry_{irt}$ variable is not biased by any general effects that regional entries may have on sectors that are more related to their regions.

$$\begin{aligned} growth_{irt} = \beta_0 &+ \beta_1 relentry_{ir,t-3} + \beta_2 entry_{ir,t-3} \\ &+ \beta_3 entry_{r,t-3} + \beta_4 \left(entry_{r,t-3} \times rel_{ir} \right) \\ &+ \beta_5 rel_{ir} + \beta_6 spec_{ir} \\ &+ \gamma \mathbf{X}_{irt} + \theta_t + \theta_{ir} + \epsilon_{irt} \end{aligned} \tag{8}$$

In equation 8, $\mathbf{X}_r$ is a vector of controls, θ indicates a fixed effect and ϵ_{irt} is the error term. All other parameters are defined as in sections 3.1 and 3.2. Coefficient β_1 is the main parameter of interest. Given our theoretical framework, we expect $\beta_1 > 0$, at least in the medium- to long- term.

As opposed to direct effects, indirect effects of entrants on incumbent firms may take longer to materialize. Fritsch (2013) therefore recommends analyzing effects of entries over a period of up to ten years, noting that external effects of firm entries on other firms in a region tend to peak after a period of seven-eight years (Fritsch and Mueller, 2004). However, this would leave us with only one year of observed results, seeing as our data-set only contains 11 years of observations (2007-2018). Furthermore, a previous study on intra-industry effects of entry in the Netherlands has found that effects peak sooner, at around 5 years (van Stel and Suddle, 2008). We therefore opt for the use time-lags of up to 6 years (equation 9). This allows us to capture some of the lagged indirect effects of entries on growth, whilst retaining three years of longer-term outcomes to analyze (2015-2018)[12]

$$
\begin{aligned}
growth_{irt} = \beta_0 &+ \beta_1 \sum_{k=3}^{8} \frac{relentry_{ir,t-k}}{6} \\
&+ \beta_2 \sum_{k=3}^{8} \frac{entry_{ir,t-k}}{6} + \beta_3 \sum_{k=3}^{8} \frac{entry_{r,t-k}}{6} \\
&+ \beta_4 rel_{ir} \times \left(\sum_{k=3}^{8} \frac{entry_{r,t-k}}{6} \right) \\
&+ + \beta_5 rel_{ir} + \beta_6 spec_{ir} \\
&+ \gamma \mathbf{X}_{irt} + \theta_t + \theta_{ir} + \epsilon_{irt}
\end{aligned} \tag{9}
$$

Note that the main explanatory variable in this analysis is the average related entry rate between years t-8 and t-3 (see figure 1). In other words, this specification looks at the accumulative effect of a higher average entry rate over time, allowing us to analyze the effect of a sustained period of entry on subsequent growth. This should also capture any lagged effects that occur within the observed time frame, but does not distinguish between the year-specific effects of entries in year t-8 and entries in year t-3.[13] Capturing such year-specific effects is beyond the scope of this study, as it would require a longer period of observation in order to produce robust results.[14]

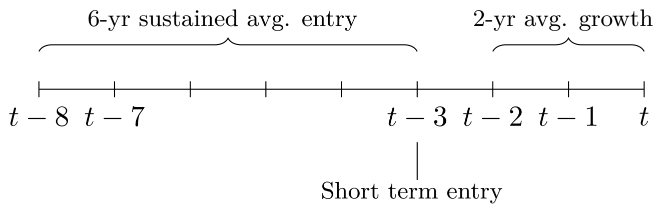
Although these specifications utilize entry rates that *precede* growth, it can not be taken for granted that any observed relationship is causal. It is easy to imagine how entries and subsequent growth may be co-determined by omitted, unobserved factors. For example, investments into regional- or sector-specific infrastructure might stimulate both entries and growth. In order to mitigate the risk of such omitted variables bias, we include a number of controls and fixed effects, as outlined in section 3.4. However, there may still be unobserved factors co-determining entries and growth on the sector-region

[12]Since we observe growth over a period of two years, applying a six-year time lag on the entry variables requires 8 years of data. For example entries from 2007-2012 are used to analyze growth between 2013-2015

[13]Note that year-specific effects in opposite directions may cancel each other out in this cumulative analysis

[14]Capturing such year-specific effects of entry in has been the subject of previous studies, such as van Stel and Suddle (2008) and Fritsch and Mueller (2004). Analyzing year-specific lagged effects of related entry may be an interesting avenue for future research.

Figure 1: Positioning of lagged entry variables and growth in time (in years)



level. This is particularly important to keep in mind when observing the relationship between growth and entries from the same sector, as these variables are likely to be co-determined by unobserved sector dynamics. However, with related entries, this concern is less pressing. Once controlling for entries from the same sector, there is no obvious reason why *related* entries and sector-regional growth should be co-determined by unobserved (sector-specific) dynamics. Still, some risk of omitted variable bias persists, and as such our results should be interpreted as suggestive evidence.

In subgroup-analyses we analyse the effects of related entry in a single cluster of economic activities on all other sectors in our sample. To do this, we divide our sector-regions into manufacturing, high-skilled and low-skilled services following Andersson and Noseleit (2011) (for the categorization of the sectors see appendix E). To avoid introducing omitted variable bias by construction, we control for entry within the own sector group in these analyses. We also conduct these heterogeneous analyses for start-ups to see whether this group of entrants has a different impact on growth.

3.4. Controls

In our study we control for a number of factors known to affect both sector-regional growth and entry. First, we want to control for general urbanization externalities that accrue to larger and denser regions, independent of their industrial composition. To this end we include both the address density and the absolute population size as regional controls. We also control for the skill composition of the region, as this may influence how the region responds to entries (Fritsch and Schroeter, 2011). Furthermore, controlling for skill composition corrects for sorting of highly skilled individuals towards productive cities, which can bias productivity benefits upwards (Combes et al., 2008). We include the share of high-skilled workers, and the share of technical workers, as well as the interaction between these two variables as our measures of the regional skill composition.

Next, growth of a sector-region can be affected by sector dynamics on the country level. To control for this, we include the national turbulence rate of the sector (entries + exits), which in turn may affect how sectors respond to new entries in an area (Fritsch, 1996). Additionally, we account for the sectors' national maturity level using the sector net entry rate (entries - exits). A larger net entry rate signals that a sector is less mature and may thus be more likely to develop and grow (Klepper and Graddy, 1990).

In related literature (such as van Stel and Suddle (2008) and Delfmann and Koster (2016)) controls are often added to account for spatial autocorrelation between regions.

In order to investigate whether our dependent growth variables are spatially correlated we perform Moran's I tests. The outcome of these tests can be seen in appendix D. Based on these tests, it appears that our measures of employment- and productivity-growth are not spatially autocorrelated. Consequently, we do not control for growth in adjacent areas in our regressions.

To explore whether employment- and productivity- growth may be serially autocorrelated we perform the Portmanteau test (Inoue and Solon, 2006) (see appendix D). The Portmanteau test does find evidence of serial correlation on the sector-region level, but not at the aggregate regional level. In order to account for this serial correlation we cluster all standard errors at the sector-region (ir) level[15]. In our aggregate, regional, analyses, errors are clustered on the regional level. [16]

Finally, we include time-dummies to control for any time trends that may be driving results. In alternative specifications of the model we also include sector-fixed effects, region-fixed effects or sector-region-fixed effects. The latter specification (with time- and sector-region-fixed effects) contains the most complete set of controls, but it leaves little variation in the data, particularly in the analyses with longer time-lags. Therefore our preferred specification includes time- and sector- fixed effects along with our regional controls.[17]

4. Data

We construct a data set of productivity, employment and firm-entry at the location- and sector level using Dutch administrative records maintained by Statistics Netherlands (CBS) covering the years 2007 - 2018. As our spatial units, we use the Dutch COROP regions (NUTS-3 level), with each COROP centered around an economic cluster.

4.1. A Unified Data Set on Firm Establishments, Regions and Workers

We use the General Business Register (GBR) and its local counterpart (LGBR) to collect microdata on the sector classification and location (municipality) of all local branches

[15]We also check whether our results are robust to an alternative specification with twoway region- and sector- clustered errors. Sector-regional clustering rules out all error correlations except those between year observations in a given sector and region, which may be overly restrictive. We have opted not to cluster on regions only, given the large heterogeneity between sectors, both nationally and regionally. The alternative specification does not affect any of our main results (results available upon request).

[16]Taking a closer look at the autoregressive processes of the residuals within sector-regions, we see that the vast majority of the observed residuals do not have a significant statistical relationship to their first- and second-order lags. Less than 10% of the observations appear significantly correlated with their lags. Hence we do not see the need to expand our estimation to account for autoregressive dynamics of our dependent variable over time (such as applying GMM estimation, Arellano and Bond (1991)).

[17]An alternative would be doing a multi-level analysis (mixed-effects) to account for variation at different levels as in van Oort et al. (2012), but results become more complex to interpret and become incomparable with the current literature, which does not use this type of analysis.

of all Dutch firms structured as a (public) limited liability (n = 536,209 in 2007).[18] Firms founded with the intention to grow and bring a product or service to the market are generally started with said legal structure. We exclude self-employed firms since generally these do not have the intention of hiring. Consequently, self-employed firms are unlikely to benefit from spillover effects from incoming workers.

Statistics Netherlands assigns each firm and local branch a sector code according to the Dutch coding system (SBI 2008) of which the first two digits correspond to the international NACE rev.2 classification. We use 3-digit sectors as our unit of analysis (see appendix E for a detailed list). Local branches may receive a different sector code than their parent firm if their main economic activities differ. For subgroup analyses we split our sector classification into three broad sector groups *Manufacturing*, *Low-skilled services* and *High-skilled services*, following Bosma et al. (2011), Dejardin (2011) and Andersson and Noseleit (2011). The low-skilled services group contains sectors such as ”463 Wholesale of food, beverages and tobacco”, while the high-skilled services revolves around knowledge intensive sectors like ”620 Computer programming, consultancy and related activities” and ”702 Management consultancy activities”.[19]

Next, we define both entering *startups* and *new local branches* of existing firms as regional entries. Matching detailed firm events data from Statistics Netherlands with the GBR at the firm level, we distinguish ”true entries” (a new firm or branch) from ”false” ones (mergers, restructurings, acquisitions etc.).[20] To pinpoint establishment entries of existing firms, we dismiss new entries in the LGBR that correspond to administrative events on the parent-firm level. We also dismiss entries if there is no corresponding increase in the number of local branches of the parent firm in the region.

Subsequently, we obtain data on the number of hours worked by merging our firm sample with microdata on work municipality (GEMSTPLTAB) and administrative wage records for all Dutch workers (SPOLIS). To calculate labor productivity, we use the microdata on financial statements including balance sheets and income statements of non-financial firms (NFO). Since there are no separate financial statements for a firm’s different local branches, we assign the total value added of each firm to its local branches using in a procedure similar to Brakman et al. (2021): We calculate the share of each branch in the wages summed on firm level, and allocate value added according to that

[18] We correct for municipality redivisions between 2007 and 2018, and take 2018 as our reference year. Since COROP ”South West Friesland” consists of only two municipalities in 2018 due to mergers, and some of the underlying towns were distributed across multiple neighbouring municipalities in different COROP areas, we exclude this area from our analysis, arriving at 39 out of 40 COROP areas. Since ”South West Friesland” houses about 0.5% of Dutch population, this should not affect our results.

[19] We differentiate between high-skilled and low-skilled services by looking at the concentration of highly skilled employees in a sector. We label workers as highly skilled if they have completed a bachelors (or other university level-) degree (Dutch abbreviation: WO), or if they have completed a higher vocational degree (HBO). Our selection of high-skilled services sectors represents the top 10% of services sectors with the highes share of highly educated workers.

[20] Making this distinction can be problematic, for instance Dejardin (2011) and Bosma et al. (2011) both had to infer entries from VAT registrations and chamber of commerce data, and cannot distinguish between new entry and administrative events such as mergers and acquisitions. We on the other hand can observe entry directly through administrative datasets created by Statistics Netherlands.

share.

We exclude sectors-regions in our final dataset if they have less than 10 workers on average throughout the sample period. These sector-regions are highly volatile and may not represent actual economic activity. Also, we drop sector-regions that are not active in all years, as we wish to focus on firm entry on the intensive margin, not entry of entirely new industries. For productivity figures, we require that each sector-region has at least 10 establishments for which financial statement data is available to avoid sampling bias, since the financial statement data is partly based on surveys. Finally, we winsorize sector-region productivity levels if they are above the 99th percentile (or below the 1st percentile) with respect to the national sectoral mean.

Finally, we use the skill-relatedness matrix constructed by Diodato and Weterings (2015) in order to measure the relatedness between industries for each 3-digit industry pair in the Netherlands. They compute skill-relatedness on the national level for the years 2009 to 2011, which fits reasonably well with our time period of 2007 to 2018.[21] Diodato and Weterings (2015) scale the skill-relatedness variable to take on values between 0 and 1 (where 0 is completely unrelated, and 1 is perfectly related) for all 3-digit industry pairs in the Netherlands (see appendix B for the calculation method). As discussed in the estimation section, we define sectors with a relatedness above 0.5 as highly related.

4.2. Descriptive Statistics

First, we explore the characteristics of our skill-relatedness measure. With examples, this rather abstract measure becomes intuitive: sector 206 (Production of synthetic fibers) is strongly related to 132 (Weaving of textiles) and 551 (Hotels) is virtually unrelated to 431 (Demolishment of buildings). Of all sector pairs, 44% have a zero skill-relatedness meaning less than 10 job switches occur between them nationally. As a sanity check for our relatedness measure (equation 3), we check if sectors have a higher relative probability of being active in regions if they are more strongly related to the regional economic portfolio (Neffke et al., 2011). Indeed, figure 2 shows that this is the case for the Netherlands. Once again, the principle of relatedness as formulated in Hidalgo et al. (2018) is confirmed: Sectors are more likely to be active near related economic activities.

Table 1 shows the descriptive statistics of all variables defined in section 3 on the sector region level, totaling N = 84,708 (or 39 regions × 12 years × 181 sectors). Note that the number of growth observations is much lower than many of our explanatory variables, since we use two year average growth and since in about 43% sector-region-years there is no economic activity (meaning zero employment or productivity). We also observe less productivity growth relative to employment growth, because we only include regions with at least 10 NFO records. Related entries, on the other hand, are observed for the full sample of possible sector-regions, as all sectors have at least one related entry in every region anually. Finally, figure 2 shows descriptive statistics on the regional level ($N = 39 \times 11 = 468$).

[21] Additionally, Diodato and Weterings (2015) remove job flows involving payroll workers, workers who are not between 18 and 65 years and workers in jobs smaller than 20 hours or 400 euros per week. Also, intersectoral job flows were removed if they contained less than 10 job switchers.

Figure 2: Relative probability of each sector of being active in a region conditional on degree of relatedness

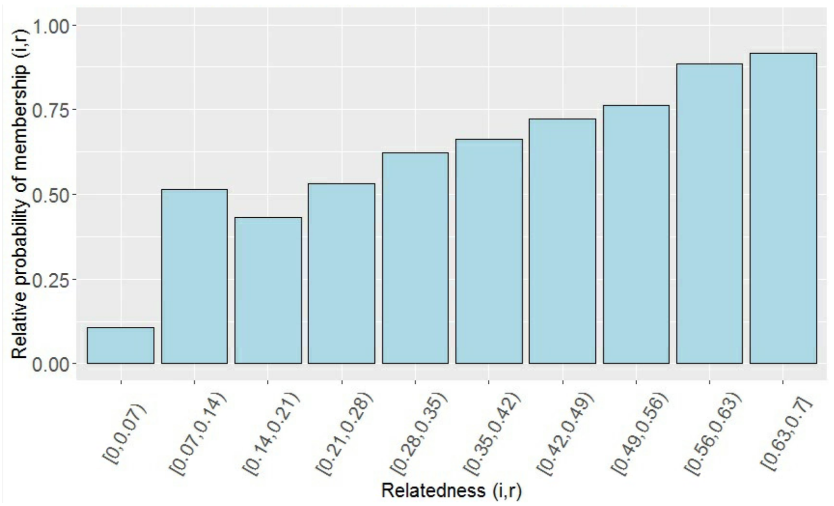


Table 1: Descriptive statistics on sector-region level, 2007 - 2018

Variable	Description	N	mean	sd	p1	p5	p95	p99
$g_{emp,irt}(log)$	Average 2-year employment growth in sector i, region r, year t (logged) with employment measured as the total number of hours worked (eq. 1)	42,050	-0.00996	0.271	-0.946	-0.369	0.310	0.789
$g_{prod,irt}(log)$	Average 2-year labour productivity growth in sector i, region r, year t (logged) with labour productivity measured as the value added per hour worked	19,668	0.00571	0.106	-0.281	-0.154	0.163	0.299
$relentry_{irt}$	Related entry, all establishments (eq. 4)	84,708	2.151	2.132	0	0	5.837	7.456
$relentry_{irt}(startups)$	Related entry, startups (eq. 4)	84,708	1.378	1.388	0	0	3.826	4.910
$entry_{irt}(log)$	Log sector-regional entry rate, all establishments	36,845	-2.200	0.609	-3.792	-3.258	-1.253	-0.875
$entry_{irt}(startups, log)$	Log sector-regional entry rate, startups	32,017	-2.565	0.707	-4.344	-3.777	-1.427	-1.030
$entry_{rt}(log)$	Log regional entry rate, all establishments	84,708	-2.255	0.168	-2.642	-2.500	-1.955	-1.853
$entry_{rt}(startups, log)$	Log regional entry rate, startups	84,708	-2.717	0.184	-3.102	-2.993	-2.380	-2.294
rel_{ir}	Relatedness of sector i to region r (eq. 3)	84,708	0.314	0.110	0.0444	0.125	0.488	0.570
$spec_{ir}$	Regional employment share of sector i in 2007	50,460	-5.543	1.417	-8.856	-7.961	-3.325	-2.844
$netentry_{it}(startups)$	National entries minus exits for sector i, startups	84,708	0.0134	0.0308	-0.0565	-0.0255	0.0692	0.113
$turbulence_{it}(startups, log)$	Log national entries plus exits for sector i, startups	71,136	-2.192	0.575	-4.190	-3.135	-1.360	-1.093
$high\_skill_{rt}$	Regional share of working population between 18 and 67 years of age with vocational education (bachelor) or higher	84,708	0.290	0.0620	0.163	0.181	0.402	0.437
$tech\_skill_{rt}$	Regional share of working population with technical education ("Natural Sciences, Mathematics & Statistics", "Information & Communication Technologies", "Engineering, Manufacturing & Construction")	84,708	0.172	0.0251	0.115	0.130	0.215	0.229
$adress\_density_{rt}$	Log mean number of addresses within a 1 km radius of each given address in the region	84,708	7.194	0.487	6.358	6.439	8.249	8.421
$total\_population_{rt}$	Size of total working population in region, logged	84,708	12.71	0.713	10.78	11.59	14.05	14.16

Table 2: Descriptive statistics on the region level, 2007 - 2018

Variable	Description	N	mean	sd	p1	p5	p95	p99
$g_{emp,rt}(log)$	Average employment growth over 2 years in region r, at time t (logged), with employment measured as the total number of hours worked	390	-0.00418	0.0256	-0.0712	-0.0446	0.0380	0.0576
$g_{prod,rt}(log)$	Average 2-year labour productivity growth in region r, at time t (logged) with labour productivity measured as the value added per hour worked	390	0.00889	0.0338	-0.0689	-0.0433	0.0585	0.0978
$relentry_{rt}$	Related entry, all establishments	468	3.577	1.187	1.226	1.782	5.652	6.578
$relentry_{rt}(startups)$	Related entry, startups	468	2.295	0.796	0.798	1.101	3.707	4.300
$specentry_{rt}$	Entries weighted by regional employment share of entering sector, all establishments	468	0.0920	0.0205	0.0505	0.0620	0.127	0.153
$specentry_{rt}(startups)$	Entries weighted by regional employment share of entering sector, startups	468	0.0572	0.0140	0.0300	0.0368	0.0829	0.0974
$entry_{rt}(log)$	Logged regional entry rate (unweighted), all establishments	468	-2.255	0.169	-2.642	-2.500	-1.955	-1.853
$entry_{rt}(startups, log)$	Logged regional entry rate (unweighted), startups	468	-2.717	0.184	-3.102	-2.993	-2.380	-2.294
rel_r	Relatedness of incumbent sectors in the region in 2007	468	0.356	0.0224	0.300	0.318	0.392	0.417
$spec_r$	Regional specialization: The employment share of the largest 2-digit sector in the region in 2007	468	0.158	0.0410	0.112	0.114	0.229	0.332
$high\_skill_{rt}$	The share of high-skilled workers in region r, at time t	468	0.290	0.0621	0.163	0.181	0.402	0.437
$tech\_skill_{rt}$	The share of technically schooled workers in region r, at time t	468	0.172	0.0251	0.115	0.130	0.215	0.229
$address\_density_{rt}$	Logged address density of region r, at time t	468	7.194	0.488	6.358	6.439	8.249	8.421
$total\_population_{rt}$	Logged total population (in absolute terms) of region r, at time t	468	12.71	0.714	10.78	11.59	14.05	14.16

5. Results

First we present regression results on short-term employment and productivity growth on the sector-region level (section 5.4). Next, we perform a regression variant that considers the effects of an average sustained higher (related) entry across 6 years (section 5.2). Third, we perform a heterogeneous analysis, exploring whether start-ups, or entrants from different sub-groups (high-skilled services, low-skilled services and manufacturing) have different impacts on growth in related sectors (section 5.3). Finally, section 5.4 addresses results on the regional level. Results of various robustness checks are included in appendix D.

5.1. Sector-Region Employment and Productivity Growth

5.1.1. Employment growth

Table 3 explores the short-term relationship between related entries in year t-3 and average employment growth from year t-2 to t. Columns 1, 2, 4 and 5 include sector fixed effects to account for unobserved heterogeneity between sectors. In all columns, time fixed effects absorb common time trends, among which is the Great Recession of 2008. Columns 3 and 6 include sector-region fixed effects so that only year-on-year variation within sector-regions remains. As a result, the amount of variance explained becomes very low, but coefficient estimates remain stable. In all specifications regional controls, as specified in section 3.4, have been included[22] Coefficient estimates remain relatively stable across specifications.

We observe no significant relationship between related entry and subsequent employment growth. Entry in the own sector ($entry_{irt}$) is positively related to subsequent employment growth, in line with regional development literature (Andersson and Noseleit, 2011; Dejardin, 2011).[23]

It appears that a standard deviation increase in the regional entry variable (change in $log(entry_{rt})$ of 0.168) has a *negative* association with employment growth (for a sector-region with an average $rel_{ir} = 0.314$), provided we account for the negative interaction term between regional entry and the relatedness of a sector with a region. I.e. although the coefficient of $entry_{rt}$ itself is positive (and weakly significant), the combined effect is negative once we account for the interaction term [24].

Additionally, we find that the relatedness between a sector and the incumbent sectors of a region at the beginning of the observed period has a strongly significant *negative* association with subsequent employment growth across the entire sample period. As does the specialization (employment share) of the own sector in a region at the beginning of

[22]We include the following controls: Address density, absolute population size, the share of high-skilled workers, the share of technical workers, and the interaction between these two skill-variables.

[23]This finding may be mainly driven by direct growth of the entering firms in the years after entry: it conforms to the well-known S-shape curve for firm entry employment dynamics across time (Fritsch, 2015)

[24]This is calculated as $\beta_4 \times \sigma_{entry_{rt}} + \beta_5 \times \sigma_{entry_{rt}} \times mean_r el_{ir} = 0.0882 \times 0.168 + (-0.346 \times 0.168 \times 0.314) = -0.003$

Table 3: Panel Fixed-Effects estimates of related entry, entry and relatedness on employment and productivity growth, 2007 - 2018

	Employment growth			Productivity growth		
	(1)	(2)	(3)	(4)	(5)	(6)
$relentry_{irt}$		0.00191	-0.00106		-0.00344***	-0.00338*
		(0.00262)	(0.00329)		(0.00130)	(0.00188)
$entry_{ir,t-3}$ (log)	0.0136***	0.0135***	0.0176***	-0.00108	-0.00101	-0.00197
	(0.00403)	(0.00403)	(0.00444)	(0.00177)	(0.00177)	(0.00200)
$entry_{r,t-3}$ (log)	-0.235	0.0882**	0.158***	0.00521	-0.0524***	-0.0859***
	(0.0204)	(0.0374)	(0.0491)	(0.00947)	(0.0160)	(0.0275)
$entry_{r,t-3} \times rel_{ir}$		-0.346***	-0.452***		0.0203***	0.314***
		(0.0982)	(0.118)		(0.0217)	(0.103)
rel_{ir}	0.262***	-0.566**		-0.0414*	0.464***	
	(0.0599)	(0.246)		(0.0217)	(0.103)	
$spec_{ir}$ (log)	-0.0175***	-0.0178***		0.00207**	0.00230**	
	(0.00279)	(0.00279)		(0.00106)	(0.00106)	
$netentry_{it}$ (startups)	0.187*	0.189*	0.153	-0.0569	-0.0574	-0.0563
	(0.107)	(0.107)	(0.108)	(0.0681)	(0.0681)	(0.0683)
$turbulence_{it}$ (startups, log)	-0.0747***	-0.00746***	-0.00754***	0.0263***	0.0267***	0.0272***
	(0.0131)	(0.0131)	(0.0135)	(0.00541)	(0.00542)	(0.00543)
Constant	-0.703**	-0.390**	0.0174	0.125**	-0.0782	-0.879
	(0.0131)	(0.162)	(2.240)	(0.0558)	(0.0692)	(1.202)
Observations	27,574	27,574	27,574	17,001	17,001	17,001
Number of sector-regions	4,052	4,052	4,052	1,967	1,967	1,967
Region controls	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES
Sector FE	YES	YES	NO	YES	YES	NO
Sector-Region FE	NO	NO	YES	NO	NO	YES
R_2 (overall)	0.0387	0.0391	0.000847	0.0720	0.0732	0.00137

Employment growth is obtained from administrative wage microdata (SPOLIS) from Statistics Netherlands, combined with work location microdata (GEMSTPLTAB). Labour productivity growth is taken from firm balance microdata (NFO). Added value totals are distributed across establishments using shares of the total wage earned on firm level. We include 39 NUTS-3 regions (COROP) and 181 sectors at 3-digit SBI level. Logged variables are indicated with '(log)'. Standard errors in parentheses, all errors are clustered on sector-regional level (ir). All models include year fixed-effects. Significance reported on 1% (***), 5% (**) and 10% (*) levels.

the observed period. In other words sectors seem to grow if there is a larger flow of entries from their own sector, but do not appear to benefit from own-sector incumbents. This negative association with specialization could signal that employment growth of mature sectors within a region slows down towards a certain peak level. This finding is supported by the positive and significant coefficient of net startup entry in column 3, which indicates that on a national level, more mature sectors (i.e. with a lower net startup entry) experience lower employment growth.

It is interesting to note that the negative relationship between relatedness and growth only becomes apparent once controlling for related entry. Without considering effects of related entry, the relatedness of a sector with the incumbent activities of a region appears to impact employment positively. This indicates that the omission of entry dynamics present a source of omitted variable bias when analyzing inter-sectoral dynamics. I.e. sectors that are more related to a region grow more, however if you consider sector-regions with the same frequency of related entrants, sectors that are more related to a region grow less.

The negative effect of related entry observed in column 3 could potentially arise from three different channels.[25] First of all, if a sector-region is strongly related to other sectors in the region, this increases the the potential number of suitable jobs outside the sector from a worker perspective. In the short run this can lead to a decrease in sector-regional employment growth since labour supply is inelastic. Second, a higher degree of relatedness on the sectoral level can potentially lead to a less diverse regional portfolio, up to the extreme (although hypothetical) case that all sectors in the portfolio are perfectly related. Diodato and Weterings (2015) argue that a more diverse portfolio makes a region less prone to negative employment shocks.[26] Third, once controlling for related entry, we find that a higher skill-relatedness between sector-regions is associated with a higher productivity growth (effects on productivity growth can be seen in columns 4-6). In turn, this productivity growth can make some labour redundant (Boschma et al., 2014).

5.1.2. Productivity growth

In Table 3, columns 4 to 6 specify three models analogous to columns 1 to 3, but with the two-year average growth in (labour) productivity as the dependent variable. We find that in the short term, entry in related sectors is associated with a *decrease* in productivity growth (columns 5 and 6). This may seem counter-intuitive, although this effect is short lived and does not occur in our *sustained entry* analyses (see section 5.2). A speculative explanation is that following a shock in related entry, highly skilled, productive workers in a given sector have a higher propensity to transfer to the related entrants compared to low-skilled workers (for instance if on average these related entrants are more attractive to the former in terms of innovativeness or career opportunities) temporarily lowering productivity in the sector they leave behind. Another theory could be that related entrants trigger a process of creative destruction, which leads to losses on the short term while the sector transitions to new methods. We note that these results are consistent

[25]Note that this result is not necessarily tied to a decrease in regional employment growth, as Table 3 considers each sector-region as a single unit of observation irrespective of employment size.

[26]There is a limit to this: Frenken et al. (2007) proposes that job creating externalities decrease if regional portfolio variety is too large.

with the findings of Andersson and Noseleit (2011), who also observe short term negative effects of entry on productivity.

In contrast to related entry, table 3 shows that own-sector entry is not significantly related to productivity growth. Just as with employment, the change in the relatedness coefficient between column 4 and 5 indicates that the omission of entry dynamics present a source of omitted variable bias. Once controlling for the effects of related entry, the relatedness of a sector with a region appears to boost productivity. The specialization (employment-share) of a sector in a region also has a positive coefficient. This result implies that productivity benefits from both localization externalities (from specialization) and co-localization externalities (from relatedness) on the sector-region level.

Regional entry coefficients in Table 3 are consistently negative, but those of rel_{ir} and the interaction between the two variables are positive, yielding a positive value overall (this is analogous to the negative effect of regional entry on employment). I.e. when viewed together with the interaction term, it appears that a standard deviation increase in the regional entry variable (change in ln_entry of 0.168) has a *positive* effect on productivity growth in sectors with an average relatedness to their region (average $rel_{ir} = 0.314$). Viewed in combination with the negative effects of entry on employment, this provides suggestive evidence that productivity gains come at the cost of employment gains.

5.2. Growth after a sustained period of startup entry

Following Fritsch and Mueller (2004), we wish to account for the possibility that the effects of related entry rates take longer to materialize by regressing growth on a 6 year period of sustained entry. Note that this analysis can only be performed for the last three years of analysis, 2015-2018, in order to be able to observe the 6-year average entry rates and subsequent 2-year average growth rates.

The effects from entry observed in our short-term analyses become generally insignificant in this sustained analysis, or weakly significant but not robust across changes in specification (Table 4). This is true of both employment and productivity growth analyses. This points to a lack of significant (cumulative) effects over the medium term [27]. However it should also be noted that the decrease in the number of observations also decreases the statistical power of this regression relative to our short-term analysis.

5.3. Heterogeneous analyses on three sector groups

To investigate whether specific sectors are driving our baseline results, we consider three sector groups separately: manufacturing low-skilled services and high-skilled servicesẆe also consider the flow of start-ups, as opposed to all entrants (including new branches) to see whether this group generates any particular effects. Table 5 displays the short-term heterogeneous analysis for start-ups for a equivalent to our analysis in section 5.1,

[27] We found similar results when we repeated the analysis using average entry across 5 instead of 6 year window. Results of this analysis are available upon request.

Table 4: Panel Fixed-Effects estimates of 6-year sustained related entry, 2007 - 2018

	Employment growth		Productivity growth	
	(1)	(2)	(3)	(4)
$relentry_{ir,t-8,t-3}$	0.00775	-0.00107	-0.00400	-0.00258
	(0.00671)	(0.0320)	(0.00317)	(0.0201)
$entry_{ir,t-8,t-3}$	-0.00243	-0.0841**	0.00308	0.0183
	(0.00874)	(0.0409)	(0.00608)	(0.0215)
$entry_{r,t-8,t-3}$ (log)	0.0237	0.135	-0.0617	0.267
	(0.0872)	(0.443)	(0.0406)	(0.255)
$entry_{r,t-8,t-3}$ (log) * $\mathrm{rel_r}$	-0.314	-0.161	0.210*	-0.234
	(0.251)	(1.060)	(0.109)	(0.572)
rel_{ir}	0.00775		0.442	
	(0.00671)		(0.285)	
$spec_{ir}$ (log)	-0.00243		0.000636	
	(0.00874)		(0.00254)	
Observations	11,978	11,978	5,901	5,901
Number of Sector-Regions	4,055	4,055	1,967	1,967
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Sector FE	YES	NO	YES	NO
Sector-Region FE	NO	YES	NO	YES
R_2 (overall)	0.148	0.000725	0.232	0.00360

Logged variables are indicated with '(log)'. Standard errors in parentheses, all errors are clustered on sector-regional level (ir). All models include year fixed-effects. Significance reported on 1% (***), 5% (**) and 10% (*) levels.

whereas table 7 (appendix C) shows the same analysis over the short term (equivalent to our analysis in section 5.2).

Table 5: Panel fixed-effects estimates of related start-ups from three sector groups manufacturing, low-skill services and high-skill services, 2007 - 2018

	Employment growth				Productivity growth			
	Baseline	MF	LS	HS	Baseline	MF	LS	HS
$\text{relentry}_{ir,t\text{-}3}$	-0.0016	-0.0019	0.0088	-0.0223	-0.0027	-0.0035	-0.0015	-0.0064
(startups)	(0.0042)	(0.0048)	(0.0083)	(0.015)	(0.0023)	(0.0029)	(0.0046)	(0.0074)
$\text{entry}_{ir,t\text{-}3}$	0.0132***	0.0132***	0.0135***	0.0133***	-0.0002	-0.0001	-0.0004	-0.0003
(startups)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
$\text{entry}_{r,t\text{-}3}$	0.129***	0.0619*	0.0741*	0.0233	-0.0729***	-0.0232	-0.0412*	0.0092
(startups, log)	(0.0434)	(0.0324)	(0.0416)	(0.0302)	(0.0237)	(0.0176)	(0.0236)	(0.0158)
$\text{entry}_{r,t\text{-}3}$ * rel_r	-0.313***	-0.325**	-0.354*	0.0644	0.262***	0.281***	0.298***	0.110
(startups, log)	(0.102)	(0.133)	(0.182)	(0.464)	(0.0520)	(0.0714)	(0.0977)	(0.239)
Observations	24,060	24,060	24,060	24,060	15,909	15,909	15,909	15,909
# Sector-Regions	3,930	3,930	3,930	3,930	1,966	1,966	1,966	1,966
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Sector-Region FE	YES	YES	YES	YES	YES	YES	YES	YES
R^2 (overall)	0.0004	0.0005	0.0004	0.0003	0.0011	0.0008	0.0030	0.0046

Related entry, entry and relatedness are measured at t-3, growth is measured from t-2 to t. Employment growth is obtained from administrative wage microdata (SPOLIS) from Statistics Netherlands, combined with work location microdata (GEMSTPLTAB). Labour productivity growth is taken from firm balance microdata (NFO). Added value totals are distributed across establishments using shares of the total wage earned on firm level. Logged variables are indicated with '(log)'. Standard errors in parentheses, all errors are clustered on sector-regional level (*ir*). Significance reported on 1% (***), 5% (**) and 10% (*) levels.

Over the short term, we find no significant effects of related entry from any of the three sub-groups. We do find positive and strongly significant coefficients of entering firms on employment growth in their own sector in all three groups (as in the baseline analysis).

In the sustained analysis (see appendix C) we find weak evidence that related entry of high-skilled services startups positively affects employment growth of incumbent firms in related sectors. The coefficient of related entry is strongly significant and positive in column 3 of Table 7. However there is no significant relationship if we look at the full set of high-skill entrants, instead of only considering start-ups (i.e. it is not robust to the inclusion of new regional branches of existing firms)[28]. Still, it provides a weak indication that high-skill start-ups may induce particular benefits for related sectors in the region. This finding aligns with Andersson and Noseleit (2011), who argue that high-skilled services entrants positively affect other sectors in the long run due to innovations and supply side influences (although they also find a short term negative influence due to initial displacement effects). Arguably innovative spillovers take time to realize and would be more likely to originate from new startups, as opposed from new branches of incumbent firms, as the latter may be aimed at at capitalizing on scale instead of innovation. This could explain why the relationship is only found for this particular subset of entrants. However we also note that this result is also not robust to a stricter definition of the relatedness cutoff (see D). Although this can be a matter of reduced precision (as the standard errors grow markedly with a less strict cut-off value for relatedness) it leaves us with the overall impression that the observed relationship is not very robust to alternative specifications.

5.4. Aggregate regional results

At the sector-regional level, we have considered the average response of individual sectors within a region to entries in the group of related sectors. In our region-level analyses, we observe the cumulative response of all sectors within the region. We note that interactions between sectors in a region could lead to different results at these two levels of aggregation. On the one hand, growth in one sector may come at the cost of growth in another, causing a displacement of growth within the region. On the other hand, growth in one sector may spill over to other sectors (for example by feeding in to local sector-specific agglomeration economies), causing the regional response to be larger than responses on the sector-regional level.

We find some evidence that related entry is positively associated with regional employment growth after 3 years at the 1%-level (Table 6). This is interesting because no such relation was found on the sector-regional level. This implies that there is some positive interaction between sectors, which is not captured in our analysis of individual sector-regions. In our analyses of sustained entries, the same holds but the association is not robust to the addition of region-fixed-effects.[29] However, it should be noted that the

[28]results of the heterogeneous analysis for the full set of entrants (not only start-ups) are available on request

[29]If we increase the cut-off value for rel_{ir} from 0.5 to 0.66 (10% most related sector pairs are 'strongly

statistical power is limited in this last estimation due to a lower number of observations.

Surprisingly, the significance of related entry dominates over specialized entry (more entries from the largest sectors in a region) and unweighted regional entry.[30] The impact of the sector-structure of the incumbent sectors of a region (rel_r) also appears to be insignificant for growth at the regional level. This implies that observed effects on the sector-regional level causes a displacement of regional growth from one sector to another, without leading to gains at the regional level. For productivity, we either find insignificant or non-robust regional growth associations with (related) entry.

While related entry is significantly associated to employment growth, the magnitude of the association is small.[31]. A quick evaluation yields that for an average region, an increase of 1 standard deviation in related entry is associated with a 1.2 percent increase in employment growth.[32] From a practical viewpoint this 1 standard deviation shock could imply for example a 33% increase in the entry rate across all sectors in a region, as can be seen by combining equations 4 and 5.[33] As a more intuitive example, we consider the impact of a shock in the entry rate of the single 3-digit sector 620 ("Computer programming, consultancy and related activities"). This sector is relatively large nationally, it has a highly skilled worker population and it is strongly related to many other sectors (46 out of 181). If we increase the entry rate of sector 620 with 33%, the associated increase in related entry for the average region is only 0.33%. In addition, entry rate increases of this magnitude do not occur often: they fall beyond the 95th percentile in Table 1, $entry_{irt}$. Furthermore, high entry rates are commonly found in sector-regions with smaller employment shares, which carry less weight in our definition of related entry.

Since our regional level findings contrast with those on sector-regional level, we test both specifications to ensure that we are not dealing with a statistical artifact.[34] First, we check for heteroskedasticity by repeating the sector-region analysis while assigning each observation a weight according to its employment share using Weighted Least Squares. Second, we repeat the sector-regional regression for a subsample of large sectors[35] to evaluate if they drive said result. In both cases, our findings are similar to tables 3 and 4. Another source of bias could be composition effects, as sectors have strongly varying employment shares in each region. However, we deem it unlikely that these variations underlie the significance of $relentry_{rt}$ in 6, since they constitute additional

related'), the related entry coefficient in column 3 remains significant. See appendix D.3.

[30] In itself, the absence of a (strong) regional association between entry and employment growth fits with existing literature on the Netherlands (van Stel and Suddle, 2008; Koster, 2011; Delfmann and Koster, 2016) and Belgium (Dejardin, 2011)

[31] Recent advances like Oster (2019) allow us to gain some insight into sign and magnitude of possible omitted variable bias, particularly in 6 We test coefficient bias for $relentry_{rt}$ and find that according to the Oster (2019) procedure this bias is relatively small. Also, it occurs in the same direction as the coefficient change when including (observed) controls (see D.4)

[32] This is calculated as $\Delta g_{emp,rt} = \beta_1 * SD_{relentry} = 0.00980 \times 1.187 = 1.2\%$.

[33] A change of one standard deviation in $relentry$ is equivalent to $SD_{relentry_{rt}} / \mu_{relentry_{rt}} = 1.187/3.58 =$ 33% of its mean.

[34] Significance levels and R^2 of regressions on aggregated and disaggregated levels are known to differ, see for instance Garret (2003)

[35] We include sectors with more than 1000 workers on national level

source of noise in the accompanying regressions. Also, we do not find similar differences between significance of sectoral entry ($entry_{irt}$, Table 3) and aggregated sectoral entry ($specentry_{rt}$, Table 6). In conclusion, the differences on the sector-regional and regional level do not seem to be driven by heteroskedasticity across sectors of different sizes.[36]

As a last check, we also explore whether the regional result can be driven by entrants themselves growing more in related regions. This would not be captured by the related entry variable in our sector-region analyses, because it always considers entrants from *related* sectors. We explore this by performing a version of our sector-region analysis which includes an interaction term between entries from the own sector and the relatedness of a sector with its region. This specification does not affect our results (i.e. it is not a source of omitted variable bias). We also see no significant effect from the added interaction term [37].

[36]Results of this analysis are available upon request.

[37]Results of this analysis are available upon request.

Table 6: Regressions of regional growth on related, sectorally aggregated and regional entry of starting firms, 2007 - 2018

	Employment, 3yr		Employment, 6yr avg		Productivity, 3yr		Productivity, 6yr avg	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{relentry}_{\text{r,t-k}}$	0.00980***	0.00887**	0.0184***	-0.0163	-0.00173	-0.00586	-0.000740	0.0350
	(0.00305)	(0.00438)	(0.00571)	(0.0432)	(0.00288)	(0.00432)	(0.00596)	(0.0447)
$\text{specentry}_{\text{r,t-k}}$	-0.0593	-0.225*	-0.467*	-1.663	0.131	0.122	0.0396	-2.008
	(0.0848)	(0.113)	(0.266)	(2.193)	(0.133)	(0.226)	(0.252)	(1.441)
$\text{entry}_{\text{r,t-k}}$ (log)	0.0984	0.171	0.195	-2.473	-0.128	-0.143	0.0165	0.227
	(0.0799)	(0.155)	(0.221)	(3.248)	(0.101)	(0.185)	(0.239)	(3.599)
$\text{entry}_{\text{r,t-k}}$ * rel_{r} (log)	-0.366	-0.502	-0.562	6.038	0.362	0.386	-0.167	-0.404
	(0.223)	(0.432)	(0.568)	(8.601)	(0.265)	(0.483)	(0.642)	(9.505)
rel_{r}	-0.983*		-1.224		0.843		-0.730	
	(0.530)		(1.400)		(0.624)		(1.556)	
spec_{r} (log)	0.00101		-0.0654*		-0.0341*		0.00918	
	(0.0259)		(0.0328)		(0.0187)		(0.0289)	
Observations	351	351	117	117	351	351	117	117
Number of Regions	39	39	39	39	39	39	39	39
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Region FE	NO	YES	NO	YES	NO	YES	NO	YES
R^2 (overall)	0.463	0.0518	0.415	0.000475	0.610	0.0269	0.512	1.05e-05

Related entry, entry, relatedness and specialization variables are constructed using start-of-period sector-regional employment shares. The entry variables are taken at $t-3$ for the short-term models (1, 2, 5 and 6) and as the average from $t-8$ to $t-3$ for their sustained counterparts (3, 4, 7 and 8). Logged variables are indicated with '(log)'. Standard errors in parentheses, all errors are clustered on regional level (r), fixed effects as indicated. Significance reported on 1% (***), 5% (**) and 10% (*) levels.

6. Discussion & Conclusions

In this paper, we have investigated the association between firm entry in related sectors and growth of employment and labour productivity. We do this in both a disaggregated (sector-regional) and aggregated (regional) setting to investigate the possibility of a direct influence of related sectors on one another as well as the aggregate effects. We single out these associations, by also accounting for general effects of entry and of the sector structure of a region.

Contrary to our expectations, we find little general evidence for a direct association between related entry and either productivity or employment growth of local sectors. Although spillovers from new entrants to related sectors could take longer to materialize than influences on the own sector, we find a similar lack of association after a sustained 6-year period of higher related entry. We do, however, find that the sector-structure of a region at the beginning of the observed period is significantly related to subsequent growth: The degree of relatedness or specialization of a sector with a region at the start of the period appears to positively impact productivity growth, at the cost of employment growth.

From these results, it seems that the regional presence of incumbent firms from related sectors play a larger role than the flow of related entrants, at least in our 11 year timeframe. This could imply that beneficial knowledge-spillovers between sectors mainly take place within incumbent networks of firms from related sectors, and not by new ideas from entrants "trickling through" to related firms. Seen in this light, both our relatedness- and specialization- measures may be thought of as indicators of regional comparative advantages that have been shaped over the long term, which are largely path dependent and may not be easily changed.

Our results provide weak additional support for the special role of high skilled service startups in regional development that is frequently referred to in the literature. A higher sustained 6-year period of high-skilled services startups is associated with employment growth in related sectors, however this result is not robust to alternative specifications. Speculatively, two channels could be involved. First, high-skill entrants may be more likely to induce employment-generating knowledge spillovers with related sectors. Another channel could be the supplier channel, where new innovations in the sectors with more entrants also spur increased activity downstream or upstream sectors (e.g. producing and distributing new and complex product).

In contrast to our sector-regional findings, we find evidence that related entries positively impact employment growth on the regional level. The result is robust in our short-term analysis, but does not hold in our sustained 6-year analysis once we control for region fixed effects. This regional result is different from what one would expect by simply aggregating sector-regional results to the regional level. Through a number of robustness tests, we exclude our method of aggregation as a source of differences between regional and sector-regional results. Heterogeneous effects across sectors of different sizes is similarly excluded as a driver of the differing results.

Based on these findings, we believe that this discrepancy could at least in part be due to the way in which spillover effects materialize. In the sector-regional analysis we consider how entrants impact each related sector individually, whereas the regional analysis considers how entrants impact the entire group of related sectors that are present in a region. Speculatively, the reaction of a given sector to a related entry could hinge on regional presence of a specific group of related sectors. Since these group dynamics will probably differ across different sectoral combinations, they could also be hard to capture empirically.[38]

Thus, group interactions between different sector clusters merit further research. On a more abstract level, our analysis reaffirms the importance of considering the disaggregated (sector-regional) level together with the aggregated (in our case regional) level. As our results aptly demonstrate a relationship observed at the individual level does not necessarily hold on the group level, and vice versa. Whereas related literature is almost exclusively focused on the regional level, more granular analyses help to investigate (or rule out) candidate mechanisms that help explain aggregate developments.

From a policy perspective, our results emphasize the complexity of intersectoral spillovers and their influence on regional growth. Policymakers should be aware that even if related entry emerges as a robust determinant for regional growth, any employment growth effects of policy measures aimed at the concept are likely to be small. Furthermore, by nature related entry is not a concept that is suited for direct manipulation. The measure is a whole composed of many parts, and from our results it becomes apparent that the presence of a larger group of related sector plays a role in explaining the association with employment growth found in this paper.

References

Aghion, P., Blundell, R., Griffith, R., Howitt, P., and Prantl, S. (2009). The Effects of Entry on Incumbent Innovation and Productivity. *The Review of Economics and Statistics*, 91(1):20–32.

Andersson, M., Braunerhjelm, P., and Thulin, P. (2011). Creative Destruction and Productivity – entrepreneurship by type, sector and sequence. Technical Report 256, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.

Andersson, M. and Noseleit, F. (2011). Start-ups and employment dynamics within and across sectors. *Small Business Economics*, 36(4):461–483.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277–297.

[38]Tentatively, group dynamics may depend for example on the presence of certain specific skills or occupations in the workforce of several related sectors. Classifications of these are available, like the ESCO (European Skills, Competences, Qualifications and Occupations) system, although microdata with this degree of granularity are hard to find and are not available for the Netherlands.

Asturias, J., Hur, S., Kehoe, T. J., and Ruhl, K. J. (2017). Firm Entry and Exit and Aggregate Growth. Working Paper 23202, National Bureau of Economic Research. DOI: 10.3386/w23202.

Audretsch, D. B. and Fritsch, M. (1994). On the measurement of entry rates. *Empirica*, 21(1):105–113.

Audretsch, D. B. and Fritsch, M. (1999). The Industry Component of Regional New Firm Formation Processes. *Review of Industrial Organization*, 15(3):239–252.

Bartelsman, E. J., Haltiwanger, J., and Scarpetta, S. (2004). Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries. SSRN Scholarly Paper ID 612230, Social Science Research Network, Rochester, NY.

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2):318–337.

Boschma, R. (2015). Do spinoff dynamics or agglomeration externalities drive industry clustering? A reappraisal of Steven Klepper's work. *Industrial and Corporate Change*, 24(4):859–873.

Boschma, R. (2017). Relatedness as driver of regional diversification: a research agenda. *Regional Studies*, 51(3):351–364.

Boschma, R., Eriksson, R. H., and Lindgren, U. (2014). Labour Market Externalities and Regional Growth in Sweden: The Importance of Labour Mobility between Skill-Related Industries. *Regional Studies*, 48(10):1669–1690.

Boschma, R. and Frenken, K. (2011). The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, 11(2):295–307.

Boschma, R., Minondo, A., and Navarro, M. (2012). Related variety and regional growth in Spain*. *Papers in Regional Science*, 91(2):241–256.

Bosma, N., Stam, E., and Schutjens, V. (2011). Creative destruction and regional productivity growth: evidence from the Dutch manufacturing and services industries. *Small Business Economics*, 36(4):401–418.

Brakman, S., Hendrich, T., Olsen, J., and Marrewijk, C. v. (2021). The Revealed Comparative Advantages of Dutch Cities. *CPB Discussion Paper series.*

Braunerhjelm, P. and Borgman, B. (2004). Geographical Concentration, Entrepreneurship and Regional Growth: Evidence from Regional Data in Sweden, 1975-99. *Regional Studies*, 38(8):929–947.

Cainelli, G., Ganau, R., and Modica, M. (2019). Industrial relatedness and regional resilience in the European Union. *Papers in Regional Science*, 98(2):755–778.

Cappelli, R., Boschma, R., and Weterings, A. (2019). Labour mobility, skill-relatedness and new plant survival across different development stages of an industry. *Environment and Planning A: Economy and Space*, 51(4):869–890.

Carree, M. A. and Thurik, A. R. (2010). The Impact of Entrepreneurship on Economic Growth. In Acs, Z. J. and Audretsch, D. B., editors, *Handbook of Entrepreneurship Research: An Interdisciplinary Survey and Introduction*, International Handbook Series on Entrepreneurship, pages 557–594. Springer New York, New York, NY. DOI: 10.1007/978-1-4419-1191-9_20.

Combes, P.-P. (2000). Economic Structure and Local Growth: France, 1984–1993. *Journal of Urban Economics*, 47(3):329–355.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742.

Content, J. and Frenken, K. (2016). Related variety and economic development: a literature review. *European Planning Studies*, 24(12):2097–2112.

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

Delfmann, H. and Koster, S. (2016). The effect of new business creation on employment growth in regions facing population decline. *The Annals of Regional Science*, 56(1):33–54.

Diodato, D. and Weterings, A. B. R. (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.

Duranton, G. and Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In Henderson, J. V. and Thisse, J.-F., editors, *Handbook of Regional and Urban Economics*, volume 4 of *Cities and Geography*, pages 2063–2117. Elsevier. DOI: 10.1016/S1574-0080(04)80005-1.

Ehrlich, M. v. and Overman, H. G. (2020). Place-Based Policies and Spatial Disparities across European Cities. *Journal of Economic Perspectives*, 34(3):128–149.

Ellison, G. and Glaeser, E. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, 100(3):1195–1213.

Frenken, K., Oort, F. V., and Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, 41(5):685–697.

Fritsch, M. (1996). Turbulence and growth in West Germany: a comparison of evidence by regions and industries. *Review of Industrial Organization*, 11(2):231–251.

Fritsch, M. (1997). New Firms and Regional Employment Change. *Small Business Economics*, 9(5):437–448.

Fritsch, M. (2013). New Business Formation and Regional Development: A Survey and Assessment of the Evidence. *Foundations and Trends® in Entrepreneurship*, 9(3):249–364.

Fritsch, M. (2015). Chapter 15: Methods of analyzing the relationship between new business formation and regional development. In *Handbook of Research Methods and Applications in Economic Geography*. Edward Elgar Publishing.

Fritsch, M. and Mueller, P. (2004). Effects of New Business Formation on Regional Development over Time. *Regional Studies*, 38(8):961–975.

Fritsch, M. and Schroeter, A. (2011). Why does the effect of new business formation differ across regions? *Small Business Economics*, 36(4):383–400.

Garret, T. (2003). Aggregated versus disaggregated data in regression analysis: implications for inference. *Economics Letters*, 81(1):61–65.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, 100(6):1126–1152.

Groot, H. L. F. d., Poot, J., and Smit, M. J. (2016). Which Agglomeration Externalities Matter Most and Why? *Journal of Economic Surveys*, 30(4):756–782.

Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., and Zhu, S. (2018). The Principle of Relatedness. In Morales, A. J., Gershenson, C., Braha, D., Minai, A. A., and Bar-Yam, Y., editors, *Unifying Themes in Complex Systems IX*, Springer Proceedings in Complexity, pages 451–457. Springer International Publishing.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, 317(5837):482–487.

Iammarino, S., Rodriguez-Pose, A., and Storper, M. (2018). Regional inequality in Europe: evidence, theory and policy implications. *Journal of Economic Geography*, 19(2):273–298.

Inoue, A. and Solon, G. (2006). A Portmanteau Test for Serially Correlated Errors in Fixed Effects Models. *Econometric Theory*, 22(5):835–851.

Jacobs, J. (1970). *The Economy of Cities*. Vintage Books. Google-Books-ID: I6XjDAAAQBAJ.

Kemeny, T. and Storper, M. (2015). Is specialization good for regional economic development? *Regional Studies*, 49(6):1003–1018.

Kirchhoff, B. A. and Phillips, B. D. (1988). The effect of firm formation and growth on job creation in the United States. *Journal of Business Venturing*, 3(4):261–272.

Klepper, S. (2007). Disagreements, Spinoffs, and the Evolution of Detroit as the Capital of the U.S. Automobile Industry. *Management Science*, 53(4):616–631.

Klepper, S. (2010). The origin and growth of industry clusters: The making of Silicon Valley and Detroit. *Journal of Urban Economics*, 67(1):15–32.

Klepper, S. and Graddy, E. (1990). The Evolution of New Industries and the Determinants of Market Structure. *The RAND Journal of Economics*, 21(1):27–44.

Klepper, S. and Simons, K. L. (2000). Dominance by birthright: entry of prior radio producers and competitive ramifications in the U.S. television receiver industry. *Strategic Management Journal*, 21(10-11):997–1016.

Koster, S. (2011). Individual foundings and organizational foundings: their effect on employment growth in The Netherlands. *Small Business Economics*, 36(4):485–501.

Marshall, A. (1890). *Principles of Economics*. Macmillan and Co., London, 8 edition.

Michelacci, C. and Silva, O. (2007). Why So Many Local Entrepreneurs? *The Review of Economics and Statistics*, 89(4):615–633.

Neffke, F. and Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3):297–316.

Neffke, F., Henning, M., and Boschma, R. (2011). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography*, 87(3):237–265.

Neffke, F. M., Otto, A., and Weyh, A. (2017). Inter-industry labor flows. *Journal of Economic Behavior & Organization*, 142:275–292.

Neffke, F. M. H., Henning, M., and Boschma, R. (2012). The impact of aging and technological relatedness on agglomeration externalities: a survival analysis. *Journal of Economic Geography*, 12(2):485–517.

Nooteboom, B. (2000). Institutions and Forms of Co-ordination in Innovation Systems. *Organization Studies*, 21(5):915–939.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Rosenthal, S. S. and Strange, W. C. (2020). How Close Is Close? The Spatial Reach of Agglomeration Economies. *Journal of Economic Perspectives*, 34(3):27–49.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy.* Harper & Brothers. Google-Books-ID: MRg5crpAOBIC.

Stam, E. (2007). Why Butterflies Don‘t Leave: Locational Behavior of Entrepreneurial Firms. *Economic Geography*, 83(1):27–50.

Timmermans, B. and Boschma, R. (2014). The effect of intra- and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows. *Journal of Economic Geography*, 14(2):289–311.

van Oort, F. G., Burger, M. J., Knoben, J., and Raspe, O. (2012). Multilevel approaches and the firm-agglomeration ambiguity in economic growth studies. *Journal of Economic Surveys*, 26(3):468–491.

van Stel, A. and Suddle, K. (2008). The impact of new firm formation on regional development in the Netherlands. *Small Business Economics*, 30(1):31–47.

Appendix A Correlation Matrices

A.1 Sector-regional level

Figure 3: Correlation matrix on the sector-regional level

	$productivity_{irt}$ (log)	$employment_{irt}$ (log)	rel_{ir}	$high\_skill_{rt}$	$tech\_skill_{rt}$	$netentry_{it}$ (startups)	$spec_{ir}$ (log)	$adres\_density_{rt}$ (log)	$total\_pop_{rt}$ (log)	$turbulence_{it}$ (startups, log)	$relentry_{irt-3}$	$entry_{irt-3}$ (log)	$entry_{rt-3}$ (log)
$productivity_{irt}$ (log)	1												
$employment_{irt}$ (log)	-0.172***	1											
rel_{ir}	0.0139	0.00617	1										
$high\_skill_{rt}$	-0.0131	0.0191*	-0.0682***	1									
$tech\_skill_{rt}$	0.00316	0.00581	0.0665***	-0.572***	1								
$netentry_{it}$ (startups)	-0.0343***	0.137***	-0.120***	0.0764***	-0.0586***	1							
$spec_{ir}$ (log)	0.0274***	-0.0955***	0.188***	-0.129***	0.0723***	-0.234***	1						
$adres\_density_{rt}$ (log)	-0.00655	0.0110	-0.0806***	0.693***	-0.686***	0.0687***	-0.166***	1					
$total\_pop_{rt}$ (log)	-0.0123	0.0165*	-0.0742***	0.465***	-0.398***	0.0651***	-0.315***	0.584***	1				
$turbulence_{it}$ (startups, log)	0.0371***	0.0247**	-0.144***	0.0687***	-0.0702***	0.516***	-0.222***	0.0722***	0.0383***	1			
$relentry_{irt-3}$	-0.00825	0.0312***	0.611***	0.251***	-0.174***	0.00453	-0.0800***	0.315***	0.544***	-0.0270***	1		
$entry_{irt-3}$ (log)	-0.00363	0.0708***	-0.0740***	0.0872***	-0.107***	0.384***	-0.202***	0.142***	0.0992***	0.473***	0.130***	1	
$entry_{rt-3}$ (log)	-0.0195*	0.0404***	-0.0516***	0.412***	-0.426***	0.0453***	-0.138***	0.567***	0.453***	-0.0177*	0.483***	0.277***	1

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Variable suffixes indicate variation in sectors (i) regions (r) and years (t). Growth ('g_') is calculated from t-2 to t. Entry variables are taken at t-3 ('l3_')

A.2 Regional level

Figure 4: Correlation matrix on the regional level

	productivity$_{rt}$ (log)	employment$_{rt}$ (log)	rel$_r$	high_skill$_{rt}$	tech_skill$_{rt}$	spec$_r$	adres_density (log)	total_pop (log)	relentry$_{rt-3}$	specentry$_{rt-3}$	entry$_{rt-3}$ (log)
productivity$_{rt}$ (log)	1										
employment$_{rt}$ (log)	-0.288***	1									
rel$_r$	0.0273	-0.0400	1								
high_skill$_{rt}$	-0.0642	0.0972	-0.710***	1							
tech_skill$_{rt}$	0.0565	0.0202	0.636***	-0.596***	1						
spec_r	-0.0418	-0.00672	-0.0462	0.0512	-0.0580	1					
adres_density (log)	-0.0372	0.0514	-0.699***	0.758***	-0.638***	0.138**	1				
total_pop (log)	-0.0402	0.0805	-0.594***	0.515***	-0.391***	-0.179***	0.599***	1			
relentry$_{rt-}$	-0.0615	0.168**	-0.466***	0.414***	-0.254***	-0.144**	0.495***	0.847***	1		
specentry$_{rt-3}$	-0.0503	0.0972	-0.408***	0.400***	-0.377***	-0.119*	0.480***	0.429***	0.652***	1	
entry$_{rt-3}$ (log)	-0.0665	0.0921	-0.486***	0.422***	-0.385***	-0.0110	0.528***	0.416***	0.680***	0.861***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Variable suffixes indicate variation in regions (r) and years (t). Growth ('g_') is calculated from t-2 to t. Entry variables are taken at t-3 ('l3_')

Appendix B Skill-relatedness calculation

The concept of skill-relatedness was coined by Neffke and Henning (2013) as a way to measure how related two industries are in terms of the required set of skills, revealed through inter-industry labor flows. If a larger labor flow is present from industry i to industry j, this is an indication that there is a stronger overlap between the skill requirements in i and j. This overlap may make it easier for switching workers to transfer knowledge and technologies across sectors, thereby promoting knowledge spillovers. The benefit of relying on such "revealed relatedness" measures (instead of administrative relatedness) is that we base our measure on true spillovers between firms. Using German data Neffke et al. (2017) find that skill-relatedness is more predictive for regional diversification and local industry growth than alternative revealed relatedness measures such as co-location (Hidalgo et al., 2007) and input-output linkages. Other beneficial properties of skill-relatedness are stability over time and a relative independence of the distance across which labour flows occur. (Neffke et al., 2017).[39]

To compute skill-relatedness between two sectors, Neffke et al. (2017) compares the observed numbers of workers between two years $F_{ijt,t+1}$ with an estimate of these flows $\hat{F}_{ijt,t+1}$ in which workers switch industries at random using observed flows as probabilities (eq. 10).[40] Equation assumes that the total outflow from sector i flows proportionally to j, where the proportion is the ratio between labor inflow into j (the numerator) and the grand total of flows in the entire economy (the denominator).

$$\hat{F}_{ijt,t+1} = \sum_j F_{ijt,t+1} \left[\frac{\sum_i F_{ijt,t+1}}{\sum_i \sum_j F_{ijt,t+1}} \right] \tag{10}$$

The skill-relatedness $SR_{ijt,t+1}$ can now be computed on national level by dividing observed labor flows by the ones estimated under the 'random flows' assumption (eq. 11). In the remainder, we assume that skill-relatedness is constant over our time period, similar to Neffke et al. (2017) who establish that skill-relatedness is practically time invariant.

$$SR^*_{ijt,t+1} = \frac{F_{ijt,t+1}}{\hat{F}_{ijt,t+1}} \approx SR_{ij} \tag{11}$$

Since the SR^*_{ij} obtained this way ranges from 0 to ∞ and is generally highly skewed, is centered between -1 and 1 using (eq. 12).[41] Two industries are considered skill-related if $SR_{ij} \geq 0$, where SR_{ij} is measured using job flows on the national level. By definition, $SR_{ii} \equiv 1$.

$$SR_{ij} = \frac{SR^*_{ij} + 1}{SR^*_{ij} - 1} \tag{12}$$

[39] See Diodato and Weterings (2015) for more details on how skill-relatedness is calculated

[40] PM in earlier work, they used a regression, but the current method gives virtually the same results.

[41] So eq. 12 maps $[0, 1]$ onto $[-1, 0]$ and $[1, \infty]$ onto $[0, 1]$.

Appendix C Sustained heterogeneous analysis

Table 7: Panel fixed-effects estimates of sustained related start-ups (over 6 year period)) from three sector groups manufacturing , low-skilled services and high-skilled services , 2007 - 2018

	Employment growth				Productivity growth			
	Baseline	MF	LS	HS	Baseline	MF	LS	HS
$relentry_{ir,t-8,t-3}$	-0.0041	-0.0052	-0.160*	0.379***	0.0182	0.00689	0.0669	-0.0078
(startups)	(0.0430)	(0.0526)	(0.0843)	(0.139)	(0.0254)	(0.0306)	(0.0478)	(0.0704)
$entry_{ir,t-8,t-3}$	-0.0710**	-0.0705**	-0.0689**	-0.0693**	0.0326**	0.0323**	0.0335**	0.0181
(startups)	(0.0339)	(0.0339)	(0.0339)	(0.0337)	(0.0161)	(0.0162)	(0.0162)	(0.0215)
$entry_{r,t-8,t-3}$	0.233	0.0498	0.682**	0.178	0.309	0.224	0.332	0.282
(startups, log)	(0.361)	(0.291)	(0.342)	(0.266)	(0.200)	(0.162)	(0.210)	(0.190)
$entry_{r,t-8,t-3} * rel_r$	-0.0723	1.131	-1.792	-0.816	-0.293	0.0348	-0.808	-3.820
(startups, log)	(0.795)	(0.955)	(1.396)	(3.892)	(0.415)	(0.565)	(0.786)	(2.733)
Observations	11,511	11,511	11,511	11,511	5,896	5,896	5,896	5,901
# Sector-Regions	3,933	3,933	3,933	3,933	1,966	1,966	1,966	1,967
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Sector-Region FE	YES	YES	YES	YES	YES	YES	YES	YES
R^2 (overall)	0.0002	3.39e-05	4.97e-05	0.0001	0.0026	0.0018	0.0019	0.0004
R^2 (within)	0.0078	0.008	0.0092	0.0092	0.0593	0.0591	0.0599	0.0574
R^2 (between)	0.0004	5.45e-05	0.0001	0.0002	0.0036	0.0021	0.0026	0.0002

Entry variables use 6-year averages from $t-8$ up to and including $t-3$. Logged variables are indicated with '(log)'. Standard errors in parentheses, all errors are clustered on sector-regional level (ir). All models include year fixed-effects. Significance reported on 1% (***), 5% (**) and 10% (*) levels.

Appendix D Robustness Checks

D.1 Temporal and spatial autocorrelation

As with any panel study with a geographical component, our regressions may suffer from both serial correlation and spatial autocorrelation, which can both lead to issues with the estimated p-values.

We use the Portmaneau test to check for serial correlation of the residuals for our preferred specifications (following the method of Inoue and Solon (2006)). On the sector-regional level the null-hypothesis of zero serial correlation is rejected. I.e. serial correlation is present in our data. Therefore we always cluster our standard errors at the sector-regional level. On the regional level, serial correlation is weakly rejected. Because it is only weakly rejected we still opt for clustered standard errors on the regional level in our regional analyses.

Residual	N	IS	P-value
$Employment_{ir}$	4,052	309.72	0.000
$Productivity_{ir}$	1,967	425.47	0.000
$Employment_{r}$	39	23.17	0.080
$Productivity_{r}$	39	25.41	0.044

In order to further investigate the observed autocorrelation on the sector-regional level, we simply regress sector-region residuals on their first lag. This rough analysis reveals that residuals are significantly related to their lags in 10% of the observed sector-regions or less. Based on this analysis we see little added value in exploring a model with lagged dependent variables (e.g. Arellano-Bond GMM estimation) as an alternative econometric specification.

Lagged residual (ir)	N	p-value (mean)	p-value (p10)
Employment	23,100	0.46	0.05
Productivity	23,100	0.53	0.11

The Moran's I test can be applied to measure the spatial correlation of data-points at one point in time. On the regional level we therefore investigate spatial correlation by applying Moran's I test to the *average* 2-year employment- and productivity- growth over the observed period, using a distance based spatial weights matrix. The null-hypothesis of no spatial correlation is not rejected. I.e. there is no statistically significant spatial correlation present in the regional data. On the sector-regional level, we perform the Moran's I test separately on the average regional growth rate of each sector. We find that the null hypothesis of no spatial auto-correlation is not rejected for a majority of the sectors. Thus we conclude that spatial correlation does not present a problem in our sector-regional analysis.

Average growth	N	Moran's I	P-value
Employment	39	-0.0321	0.419
Productivity	39	-0.00700	0.230

D.2 Collinearity in Explanatory Variables

On the regional level, another cause for concern is collinearity of our explanatory variables. Essentially every regional entry will show up in all our entry measures, provided that it is related to at least one sector present in the region. However it will be weighted differently in each of the three separate aggregate regional measures of entry: With *related* entry entries are weighted more if they are related to larger incumbent sectors, with *specialized* entry entries are weighted more if they are entering into larger incumbent sectors, whereas regional entry is unweighted. Needless to say there is a fairly high degree of correlation between the three variables (all around, or above 70%, see Table in appendix A.2).

Therefore we always check our regional results for each variable separately. From this analysis we see that the estimated coefficients remain relatively stable across specifications, leading us to be less concerned about collinearity issues.

D.3 Alternative cut-off Skill-Relatedness

We evaluate the influence of our choice for the cut-off value that defines sector pairs as strongly related (see equation 3). Changing the cut-off will mechanically affect variables that incorporate skill-relatedness, namely related entry, relatedness and the interaction between regional entry and relatedness. When we repeat our analyses using a cut-off value of 0.66 (equivalent to the 10% most related sector pairs) instead of 0.50, we find that indeed the coefficients of aforementioned variables increase in absolute value, but remain equal in terms of sign and significance, while other coefficients remain roughly equal (results are available upon request). We conclude that our results are robust to the choice of cut-off.

Table 8: Panel Fixed-Effects estimates of related entry, entry and relatedness on employment and productivity growth, 2007 - 2018

	Employment growth				Productivity growth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{relentry}_{\text{r,t-3}}$	0.00583***			0.00980***	3.57e-05			-0.00173
	(0.00206)			(0.00305)	(0.00203)			(0.00288)
$\text{specentry}_{\text{r,t-3}}$		-0.0172		-0.0593		0.0843		0.131
		(0.0679)		(0.0848)		(0.0661)		(0.133)
$\text{entry}_{\text{r,t-3}}$ (log)			-0.00718	0.0984			0.0105	-0.128
			(0.0155)	(0.0799)			(0.0130)	(0.101)
$\text{entry}_{\text{r,t-3}}$ (log) * rel_{r}				-0.366				0.362
				(0.223)				(0.265)
rel_{r}	-0.0956	-0.0755	-0.0816	-0.983*	0.00164	-0.00172	0.00987	0.843
	(0.0977)	(0.106)	(0.106)	(0.530)	(0.0578)	(0.0581)	(0.0630)	(0.624)
spec_{r} (log)	0.00837	0.0102	0.0103	0.00101	-0.0407**	-0.0353**	-0.0392**	-0.0341*
	(0.0218)	(0.0249)	(0.0257)	(0.0259)	(0.0206)	(0.0174)	(0.0191)	(0.0187)
Observations	351	351	351	351	351	351	351	351
Number of Regions	39	39	39	39	39	39	39	39
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Region FE	NO	NO	NO	NO	NO	NO	NO	NO
R^2 (overall)	0.451	0.440	0.440	0.463	0.607	0.608	0.607	0.610
R^2 (within)	0.463	0.458	0.458	0.472	0.622	0.621	0.622	0.626
R^2 (between)	0.308	0.241	0.244	0.357	0.126	0.196	0.156	0.134

Employment growth is obtained from administrative wage microdata (SPOLIS) from Statistics Netherlands, combined with work location microdata (GEMSTPLTAB). Labour productivity growth is taken from firm balance microdata (NFO). Added value totals are distributed across establishments using shares of the total wage earned on firm level. We include 39 NUTS-3 regions (COROP). Logged variables are indicated with '(log)'. Standard errors in parentheses, all errors are clustered on regional level (r). All models include year fixed-effects. Significance reported on 1% (***), 5% (**) and 10% (*) levels.

D.4 Omitted Variable Bias

In our research, we attempt to diminish omitted variable bias (OVB) by choosing appropriate controls and by making use of fixed effects. However, sources of ommitted variable bias are likely to remain, as one cannot expect firm entry to be exogenously determined. Recent advances like Oster (2019) allow us to gain some insight into sign and magnitude of possible OVB, particularly in Table 6.

We test coefficient bias for $relentry_{rt}$ and find that according to the Oster (2019) procedure this bias is relatively small for models 1 and 3 from Table 6, with the Oster procedure suggesting OVB in Table 6 yields an underestimation (see Table 9). Also, the bias occurs in the same direction as the coefficient change when including (observed) controls.

For model 2, the coefficient sign reverses when the R_{max} increment becomes larger than 13% and remains stable before that. This is an intuitive result, given that model 2 includes region and year fixed effects that already absorb all time-invariant factors in these dimensions.

For model 4, where $\beta_{relentry}$ is larger, insignificant and negative, the Oster procedure consistently reports positive and rising β values. This suggests that this model indeed suffers from OVB, where the corrected coefficient is closer to that of models 1,2 and 3.

Table 9: Oster (2019) results for columns 1 to 4 of Table 6 for a range of R_{max} increments with respect to the original R^2 of the controlled regression.

			R_{max} increment			
Table 6 column	R^2 (within)	$\beta_{relentry}$	5%	10%	15%	20%
Column 1	0.472	0.00980***	0.0232	0.0531	0.0881	0.124
Column 2	0.483	0.00887***	0.00874	0.00827	-0.575	-3.02
Column 3	0.421	0.0184***	0.0394	0.0647	0.0912	0.118
Column 4	0.493	-0.0163	0.0481	0.133	0.233	0.341

As parameters, we follow the standard assumption (and upper bound) that selection into treatment on unobserved variables is equal to that of observed variables ($\delta = 1$). For the theoretical maximum R^2 value R_{max} we perform a parameter study where R_{max} is 5, 10, 15 or 20 percentage points higher than the R^2 of the controlled regression to see to what extent the coefficients for $relentry_{rt}$ start to deviate. Note that we do not assume that regional growth can be fully explained even with full knowledge on unobservables ($R_{max} = 1$). Given complexity, explanatory power will at the regional scale always be hampered by measurement errors and all kinds of idiosyncratic variation. In addition, our R^2 values are probably already close to the maximum given that we use both time and region fixed effects.

Appendix E Sector classifications (3-digits)

Table 10 gives and overview of the 3-digit SBI sectors included in our analyses. We exclude activities that depend on geography (fishing, forestry, mining, extraction), utilities (energy, waste), finance activities and holdings, payrolling, temporary work bureaus and job pools and finally public and semi-public sectors (e.g. government, care, education).

Table 10: Three-digit SBI codes included in the analysis (MAN = manufacturing, LS = low-skilled services, HS = high-skilled services)

SBI Code	Description	MAN	LS	HS
011	Growing of non-perennial crops	x		
012	Growing of perennial crops	x		
013	Plant propagation	x		
014	Animal production	x		
015	Mixed farming	x		
016	Support activities to agriculture and post-harvest crop activities	x		
017	Hunting, trapping and related service activities	x		
101	Processing and preserving of meat and production of meat products	x		
102	Processing and preserving of fish, crustaceans and molluscs	x		
103	Processing and preserving of fruit and vegetables	x		
104	Manufacture of vegetable and animal oils and fats	x		
105	Manufacture of dairy products	x		
106	Manufacture of grain mill products, starches and starch products	x		
107	Manufacture of bakery and farinaceous products	x		
108	Manufacture of other food products	x		
109	Manufacture of prepared animal feeds	x		
110	Manufacture of beverages	x		
120	Manufacture of tobacco products	x		
131	Preparation and spinning of textile fibres	x		
132	Weaving of textiles	x		
133	Finishing of textiles	x		
139	Manufacture of other textiles	x		
141	Manufacture of wearing apparel, except fur apparel	x		
142	Manufacture of articles of fur	x		
143	Manufacture of knitted and crocheted apparel	x		
151	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery and harness; dressing and dyeing of fur	x		
152	Manufacture of footwear	x		
161	Sawmilling and planing of wood	x		

Continued on next page

Table 10 – continued from previous page

SBI code	Description	MAN	LS	HS
162	Manufacture of products of wood, cork, straw and plaiting materials	x		
171	Manufacture of pulp, paper and paperboard	x		
172	Manufacture of articles of paper and paperboard	x		
181	Printing and service activities related to printing	x		
182	Reproduction of recorded media	x		
191	Manufacture of coke oven products	x		
192	Manufacture of refined petroleum products	x		
201	Manufacture of basic chemicals, fertilisers and nitrogen compounds, plastics and synthetic rubber in primary forms	x		
202	Manufacture of pesticides and other agrochemical products	x		
203	Manufacture of paints, varnishes and similar coatings, printing ink and mastics	x		
204	Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations	x		
205	Manufacture of other chemical products	x		
206	Manufacture of man-made fibres	x		
211	Manufacture of basic pharmaceutical products	x		
212	Manufacture of pharmaceutical preparations	x		
221	Manufacture of rubber products	x		
222	Manufacture of plastic products	x		
231	Manufacture of glass and glass products	x		
232	Manufacture of refractory products	x		
233	Manufacture of clay building materials	x		
234	Manufacture of other porcelain and ceramic products	x		
235	Manufacture of cement, lime and plaster	x		
236	Manufacture of articles of concrete, cement and plaster	x		
237	Cutting, shaping and finishing of stone	x		
239	Manufacture of abrasive products and non-metallic mineral products n.e.c.	x		
241	Manufacture of basic iron and steel and of ferro-alloys	x		
242	Manufacture of tubes, pipes, hollow profiles and related fittings, of steel	x		
243	Manufacture of other products of first processing of steel	x		
244	Manufacture of basic precious and other non-ferrous metals	x		
245	Casting of metals	x		
251	Manufacture of structural metal products	x		
252	Manufacture of tanks, reservoirs and containers of metal	x		
253	Manufacture of steam generators, except central heating hot water boilers	x		

Continued on next page

Table 10 – continued from previous page

SBI code	Description	MAN	LS	HS
254	Manufacture of weapons and ammunition	x		
255	Forging, pressing, stamping and roll-forming of metal; powder metallurgy	x		
256	Treatment and coating of metals; machining	x		
257	Manufacture of cutlery, tools and general hardware	x		
259	Manufacture of other fabricated metal products	x		
261	Manufacture of electronic components and boards	x		
262	Manufacture of computers and peripheral equipment	x		
263	Manufacture of communication equipment	x		
264	Manufacture of consumer electronics	x		
265	Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks	x		
266	Manufacture of irradiation, electromedical and electrotherapeutic equipment	x		
267	Manufacture of optical instruments and photographic equipment	x		
268	Manufacture of magnetic and optical media	x		
271	Manufacture of electric motors, generators, transformers and electricity distribution and control apparatus	x		
272	Manufacture of batteries and accumulators	x		
273	Manufacture of wiring and wiring devices	x		
274	Manufacture of electric lighting equipment	x		
275	Manufacture of domestic appliances	x		
279	Manufacture of other electrical equipment	x		
281	Manufacture of general-purpose machinery	x		
282	Manufacture of other general-purpose machinery	x		
283	Manufacture of agricultural and forestry machinery	x		
284	Manufacture of metal forming machinery and machine tools	x		
289	Manufacture of other special-purpose machinery	x		
291	Manufacture of motor vehicles	x		
292	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers	x		
293	Manufacture of parts and accessories for motor vehicles	x		
301	Building of ships and boats	x		
302	Manufacture of railway locomotives and rolling stock	x		
303	Manufacture of air and spacecraft and related machinery	x		
304	Manufacture of military fighting vehicles	x		
309	Manufacture of transport equipment n.e.c.	x		
310	Manufacture of furniture	x		
321	Manufacture of jewellery, bijouterie and related articles	x		
322	Manufacture of musical instruments	x		
323	Manufacture of sports goods	x		

Continued on next page

Table 10 – continued from previous page

SBI code	Description	MAN	LS	HS
324	Manufacture of games and toys	x		
325	Manufacture of medical and dental instruments and supplies	x		
329	Manufacturing n.e.c.	x		
331	Repair of fabricated metal products, machinery and equipment		x	
332	Installation of industrial machinery and equipment		x	
411	Development of building projects		x	
412	Construction of residential and non-residential buildings		x	
421	Construction of roads and railways		x	
422	Construction of utility projects		x	
429	Construction of other civil engineering projects		x	
431	Demolition and site preparation		x	
432	Electrical, plumbing and other construction installation activities		x	
433	Building completion and finishing		x	
439	Other specialised construction activities		x	
451	Sale of motor vehicles		x	
452	Maintenance and repair of motor vehicles		x	
453	Sale of motor vehicle parts and accessories		x	
454	Sale, maintenance and repair of motorcycles and related parts and accessories		x	
461	Wholesale on a fee or contract basis		x	
462	Wholesale of agricultural raw materials and live animals		x	
463	Wholesale of food, beverages and tobacco		x	
464	Wholesale of household goods		x	
465	Wholesale of information and communication equipment		x	
466	Wholesale of other machinery, equipment and supplies		x	
467	Other specialised wholesale		x	
469	Non-specialised wholesale trade		x	
471	Retail sale in non-specialised stores		x	
472	Retail sale of food, beverages and tobacco in specialised stores		x	
473	Retail sale of automotive fuel in specialised stores		x	
474	Retail sale of information and communication equipment in specialised stores		x	
475	Retail sale of other household equipment in specialised stores		x	
476	Retail sale of cultural and recreation goods in specialised stores		x	
477	Retail sale of other goods in specialised stores		x	
478	Retail sale via stalls and markets		x	
479	Retail trade not in stores, stalls or markets		x	

Continued on next page

Table 10 – continued from previous page

SBI code	Description	MAN	LS	HS
491	Passenger rail transport, interurban		x	
492	Freight rail transport		x	
493	Other passenger land transport		x	
494	Freight transport by road and removal services		x	
495	Transport via pipeline		x	
501	Sea and coastal passenger water transport		x	
502	Sea and coastal freight water transport		x	
503	Inland passenger water transport		x	
504	Inland freight water transport		x	
511	Passenger air transport		x	
512	Freight air transport and space transport		x	
521	Warehousing and storage		x	
522	Support activities for transportation		x	
531	Postal activities under universal service obligation		x	
532	Other postal and courier activities		x	
551	Hotels and similar accommodation		x	
552	Holiday and other short-stay accommodation		x	
553	Camping grounds, recreational vehicle parks and trailer parks		x	
559	Other accommodation		x	
561	Restaurants and mobile food service activities		x	
562	Event catering and other food service activities		x	
563	Beverage serving activities		x	
581	Publishing of books, periodicals and other publishing activities			x
582	Software publishing			x
591	Motion picture, video and television programme activities			x
592	Sound recording and music publishing activities			x
601	Radio broadcasting			x
602	Television programming and broadcasting activities			x
611	Wired telecommunications activities			x
612	Wireless telecommunications activities			x
613	Satellite telecommunications activities			x
619	Other telecommunications activities			x
620	Computer programming, consultancy and related activities			x
631	Data processing, hosting and related activities; web portals			x
639	Other information service activities			x
691	Legal activities			x
692	Accounting, bookkeeping and auditing activities; tax consultancy			x

Continued on next page

Table 10 – continued from previous page

SBI code	Description	MAN	LS	HS
711	Architectural and engineering activities and related technical consultancy			x
712	Technical testing and analysis			x
721	Research and experimental development on natural sciences and engineering			x
722	Research and experimental development on social sciences and humanities			x
731	Advertising			x
732	Market research and public opinion polling			x
741	Specialised design activities			x
742	Photographic activities			x
743	Translation and interpretation activities			x
749	Other professional, scientific and technical activities n.e.c.			x
750	Veterinary activities			x
771	Rental and leasing of motor vehicles		x	
772	Rental and leasing of personal and household goods		x	
773	Rental and leasing of other machinery, equipment and tangible goods		x	
774	Leasing of intellectual property and similar products, except copyrighted works			x
791	Travel agency and tour operator activities		x	
799	Other reservation service and related activities		x	
801	Private security activities		x	
802	Security systems service activities		x	
803	Investigation activities		x	
811	Combined facilities support activities		x	
812	Cleaning activities		x	
813	Landscape service activities		x	
821	Office administrative and support activities		x	
822	Activities of call centres		x	
823	Organisation of conventions and trade shows			x
829	Business support service activities n.e.c.		x	