Spillovers from High Growth Firms: Evidence from Hungary

Francesca de Nicola

The World Bank

Balázs Muraközy[1]

University of Liverpool
Management School

CERS Hungarian Academy of
Sciences

Shawn W. Tan

The World Bank

Abstract

How do High-Growth Firms (HGFs) affect the rest of the economy? We explore this question using Hungarian administrative microdata. Relying on the Birch definition of HGFs, we find evidence for stronger productivity growth for firms operating in industries with more HGFs and for firms supplying industries with more HGFs. Knowledge spillovers or the surge of HGFs' demand for intermediate inputs could explain these positive associations. Firms with intermediate productivity levels seem most likely to benefit from this effect, while we find no differences by age or export status. The results hold irrespective of the level of spatial aggregation and are robust to alternative definitions of HGFs as well as different measures of productivity or spillover.

1. Introduction

Promoting job creation is a policy priority for many governments. Owing to their potential contribution to this goal, High-Growth Firms (HGF) – firms with a significant ability to grow rapidly – have increasingly attracted the attention of policymakers. Programs catering to these firms tend to rely on broad policy instruments, similar to those used to support SMEs. Autio et al. (2007) and the World Bank (2018) offer a review of the evidence on these measures from advanced and developing countries, respectively. However, framing such policies only in terms of job creation is simplistic, because the welfare effects of these policies also depend fundamentally on the overall effects of HGFs on the economy.

Much attention in the literature and among policymakers has focused on the performance of HGFs and the jobs created by them. But HGFs may also generate externalities for other firms or agents in the economy, which should also be taken into account when designing policies. Following the literature on

---

[1] Corresponding author. Address: Chatham St, Liverpool L69 7ZH, Murakozy@liverpool.ac.uk

foreign direct investment (FDI) (Keller, 2010), we use the term *spillovers*[2] for these effects, which includes knowledge spillovers, linkages and other effects.[3]

The sign and magnitude of these spillovers is ambiguous theoretically. On the one hand, HGFs may exert competitive pressures on input and output prices, crowding out other firms. They may poach workers from competitors, put upward pressure on labor and other input costs, or drive output prices down by exploiting economies of scale. On the other hand, HGFs may generate positive spillovers. HGFs may develop or possess knowledge of technology, business organization or markets that can be transferred to other firms through the movement of employees or demonstration. Knowledge transfer is likely to occur over relatively short distances, through face-to-face contact with clients or suppliers, or within the local labor market. In addition, HGFs' high demand for inputs may generate opportunities for suppliers to exploit economies of scale (Javorcik, 2004), and/or facilitate technology upgrading. Because of their advanced production, management, marketing or other skills, HGFs may sell high-quality or innovative inputs, benefitting downstream firms using those inputs.

Assessing whether it is the positive or negative spillovers that dominate is the empirical question we explore in this paper. We rely on detailed firm-level administrative data from Hungary and exploit regional and time variation in the share of HGFs to infer their connection with the performance of other related firms. We distinguish between horizontal (within the same 2-digit industry) and vertical (across industries) spillovers. We identify HGFs as firms in the top 5 percentile according to the Birch Index (Birch, 1981), which is preferable to the alternative OECD definition as the Birch definition provides a less skewed identification of HGFs and is not subject to the macro cycle.

In our empirical strategy, we carefully adapt the methods in the FDI spillover literature to the HGF environment. From a technical point of view, a crucial difference is that the HGF phase is more transient than the FDI one, which, in turn, limits the application of fixed effects methodologies employed in the FDI literature. Instead, we follow Javorcik (2004), who uses first differences regression to difference out the firm fixed effects. To be specific, we regress log differences of firm outcomes on the level of spillover measures and firm controls. We also run separate regressions on the sample restricted to non-HGFs and interact the HGF status with the spillover measures to investigate the effects on non-HGFs. Relatedly, in contrast with the FDI case, the share of HGFs is mechanically linked to the growth distribution, generating challenges discussed in the peer effect literature (see, for example, Angrist, 2014). We handle this problem by excluding firms in the same narrowly defined (4-digit) industry from the calculation of spillover measures. Lastly, unobserved demand shocks can confound the relationship between high growth and productivity increase in related industries. While we employ a large set of fixed effects to control for these shocks, our results can be interpreted as a correlation between the high share of HGFs and productivity growth rather than a causal relationship.

---

[2] Our paper – similarly to many others in the spillover literature – does not provide unambigous evidence for causal effects, even if the term 'spillover' may itself suggest directionality and causality. Still, we use this term with the caveat that our results can mostly be interpreted as correlations.
[3] Spillovers are externalities in the sense that they are unlikely to be priced in when the HGF makes the decision whether to grow or not, but they may have been priced in for firms at the receiving end (e.g. if wages increase because of the higher demand for workers by HGFs in a given sector/locality) or not (e.g. if a technology is adopted by other firms as workers move from HGFs to other firms).

We find evidence of positive spillovers. Rather than attracting workers from or reducing profits of non-HGFs, HGFs' presence is linked with improvements in productivity, income and employment of other local firms either in the same industry or in supplier industries. By estimating markups following De Loecker and Warzynski (2012), we show that the gains in revenue TFP seem to result from a combination of higher markups and higher physical productivity.

Spillovers vary by sector, the absorptive capacity of firms, and the HGF status itself. Firm age and export status do not matter for the effects. Manufacturing firms compared to services firms gain more from HGFs presence in the same industry and downstream industries. Firms with medium productivity levels appear to gain most from HGFs' presence in downstream sectors. This may be explained by the low absorptive capacity of low productivity firms, and the looser relations between local markets and the most productive firms, which tend to have more outside options. HGF presence affects both HGFs and other firms, but slightly differently: only non-HGFs appear to benefit from vertical linkages.

Our results hold, albeit to a lesser extent, when we use the OECD definition of HGF: within-industry spillovers are robust to different definitions of HGFs, while across-industry spillovers are not found. The results are robust to alternative definitions of HGF spillover measures, productivity measures, sample and different sets of fixed effects and industry controls.

We contribute to several strands of the literature. First, we build on and add to the literature on the role of entrepreneurship in development and growth. This literature emphasizes that while entrepreneurship in general is a key determinant of economic growth and employment (Carree and Thurik, 2010), a subset of entrepreneurs, often called high-impact entrepreneurs, contribute disproportionately to employment and economic growth (Autio et al., 2007). The literature also emphasizes some key functions of high-impact entrepreneurs: they apply existing knowledge in innovative ways (Audretsch, 2005) and operate where demand or supply characteristics were unknown (Ács, 2010). Both activities yield new knowledge, at least locally, thus having the potential to create externalities. Our paper contributes to a better understanding of the effects generated by such entrepreneurs.

The entrepreneurship literature is partly motivated by the literature on the role of industrial policy in development. Hausmann and Rodrik (2003) interpret development as a process of discovering a country's cost structure when producing different goods. In general, this type of self-discovery is under-provided because of public good problems (Rodrik, 2005). High impact entrepreneurs can play a key part in providing such self-discovery, as they are likely to investigate production functions and generate externalities via learning by doing for firms in similar activities. Our analysis of inter-industry spillover effects resulting from HGFs may also inform industrial policy.

Second, we expand the growing literature on HGFs by examining the relationship between the presence of HGFs and the performance of other firms. Thus far, the literature has focused mostly on the identification of these fast-growing firms and their characteristics. HGFs tend to be more productive, younger, engaged in scientific R&D and computer programming, but are not concentrated in any given sector (Coad et al., 2014a; OECD, 2016). Du and Temouri (2015) study the relationship between high growth and productivity and find that: (i) HGFs are more likely to exhibit stronger TFP growth; and (ii) firms are more likely to increase their TFP after experiencing a high-growth episode. Nonetheless, a small but growing literature is exploring the spillover effects generated by HGFs. Coad et al. (2014b)

investigate the type of workers HGFs tend to hire. They find that HGFs only hire a significant number of workers from other firms after realizing some rapid growth; in the earlier phases they hire younger, less educated workers who have often experienced a fairly long unemployment period. The industry-level analysis by Bos and Stam (2011) finds that the presence of HGFs is associated with stronger subsequent industry growth in the Netherlands. We contribute to the literature by providing firm-level evidence of the relationship between the presence of HGFs and the performance of other firms.

Third, we extend the productivity spillover literature, which has focused on the effects of foreign firms. The literature has often found positive spillovers for firms supplying sectors with more foreign firms (or backward spillovers).[4] Other studies show that positive spillovers may result from intentional or non-intentional technology transfer from foreign firms to suppliers arising from higher-quality or productivity requirements (Blalock and Gertler, 2008), or owing to access to a higher quality, larger variety or lower priced pool of inputs (Javorcik, 2004). Reyes (2017) focuses specifically on FDI spillover effects on local HGFs in developing countries and finds that multinational enterprises integrated into global value chains are likely to generate stronger spillovers than FDI aimed at serving the local market. We contribute to this literature by investigating a different source of spillovers. We also discuss and propose solutions for new methodological challenges which emanate from the more transient nature of the HGF phase relative to FDI.[5] This contributes to the generalizability of the FDI spillover methodology to other problems.

The paper is organized as follows. Section 2 provides a simple conceptual framework to situate the discussion. Section 3 describes the data used, the construction of the main variables of interest and the empirical strategy. Section 4 presents our results. Finally, Section 5 concludes.

2. Conceptual framework

Our study investigates the association between HGFs and the productivity, growth and income of firms in the economy in several ways. We borrow from the FDI literature to comprehend the spillover effects of HGFs. The two main possible channels are knowledge spillovers and pecuniary externalities working via supplier-buyer linkages. While the FDI spillover literature has focused mainly on the former, recent work has emphasized that most studies, often unintentionally, measure both effects (Havranek and Irsova, 2011; Keller, 2010).

Both channels may be relevant in the HGF case. There is certainly a potential for knowledge spillovers if HGFs possess more advanced knowledge relative to other firms; this takes place when workers in HGFs leave and bring along their specialized knowledge (labor mobility) or non-HGFs imitate HGFs (demonstration effect). While the case for a potential technological advantage is clearer in the FDI spillover case, HGFs may possess very relevant knowledge about technological, organizational or marketing solutions which are crucial for operating more effectively in the specific context. While HGFs are no more common in innovative industries, HGFs are more innovative in their industry, which allows them to grow faster than their industry peers (Moreno and Coad, 2015).

---

[4] See, for example, Görg and Greenaway (2004) and Harrison and Rodríguez-Clare (2010) for a summary of the literature on spillover effects from foreign firms.
[5] Though there is evidence that FDI has a transitory effect on suppliers' productivity; see Merlevede et al. (2014).

Indeed, HGFs in Hungary are more productive than non-HGFs. To illustrate technological differences, **Error! Reference source not found.** shows the distribution of TFP (relative to industry-year mean) for HGFs and non-HGFs separately. We find that the productivity distribution of HGFs stochastically dominates that of non-HGFs. The means of the two distributions differ by 36 log points and this difference is highly significant (t=36). Similar differences are present for other performance measures, too. These patterns suggest that HGFs use more productive technologies, in line with a potential to knowledge spillovers.

Figure 1 The distribution of TFP (relative to industry-year mean) for HGFs and non-HGFs
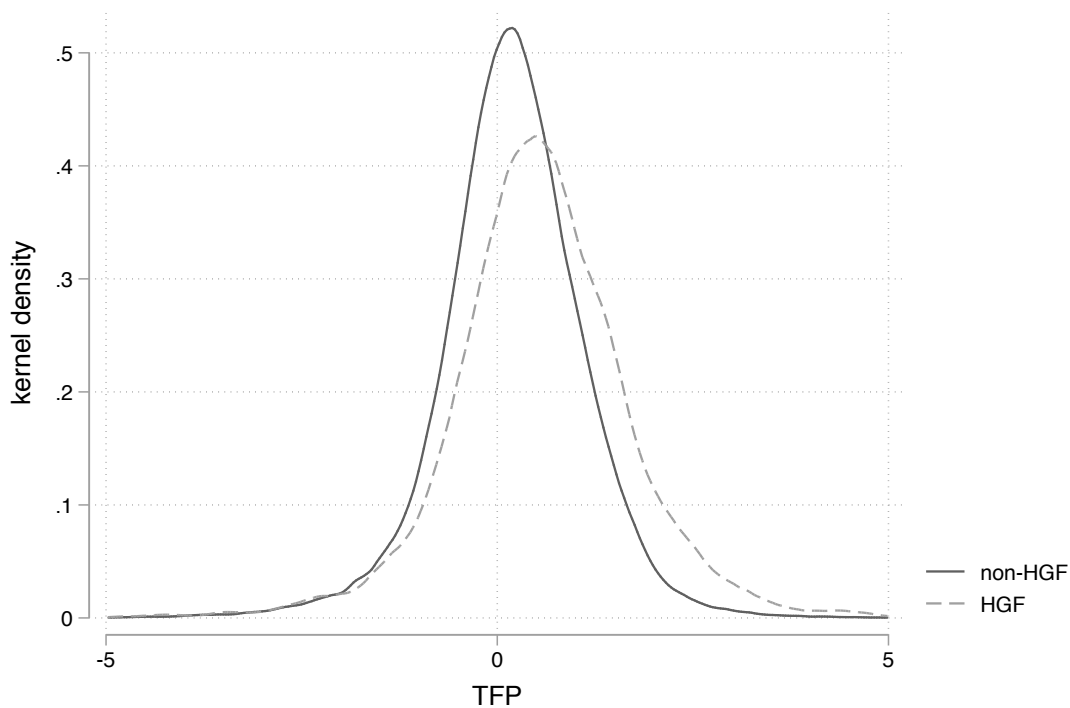
Note: This figure shows kernel density (log) TFP for HGFs and non-HGFs separately. The TFP is estimated with the ACF procedure, and demeaned at the industry-year level; therefore, it shows how productive firms are relative to the industry mean in the given year.

HGFs may also generate pecuniary externalities via their supplier-buyer linkages. HGFs can generate spillovers by simply generating more demand. Larger demand may help firms in further exploiting economies of scale or investing in technological upgrading. The larger demand of HGFs can also result in increased competition in the output or input markets, exerting negative spillovers. Stronger competition in the output markets may result in lower prices and quantities, reducing the opportunities of non-HGFs firms to utilize economies of scale or grow. In the input markets, HGFs can potentially poach the best workers and drive up the wages for all firms in the same region. Coad et al. (2014b) show, however, that there is little evidence that HGFs would poach the best workers during their high growth phase.

While our empirical framework does not allow us to perfectly identify whether non-HGFs are benefiting due to changes in economies of scale or technology investment, we control for changes in the economies of scale to capture the latter effect to some degree[6] -- we follow Javorcik (2004) by controlling for the demand of upstream industries. Next, we recognize that our main productivity measure (TFP) is revenue-based (De Loecker and Warzynski, 2012; De Loecker et al., 2016), which conflates changes in physical productivity with potential changes in markups. This issue is highly relevant not only because HGFs themselves are likely to gain market power during their growth phase, but also because the increased demand or supply generated by these firms can drive up markups of suppliers and buyers. We follow the methodology of De Loecker and Warzynski (2012) to estimate the relationship between HGF presence and markups, and include the markups as control in the productivity regressions.

Lastly, the FDI literature also shows that the spillover effects are heterogeneous across firms. A key concept in the FDI literature is the absorptive capacity of receiving firms (Girma, 2005; Zhang et al., 2010). Firms with a stronger knowledge base or absorptive capacity may be more able to learn from HGFs, more likely to become suppliers of those firms and are more likely to react with technology upgrading to increased competition rather than, say, cutting costs. As common in the spillover literature, we will proxy absorptive capacity by the TFP level of the receiving firm. We also explore if firm age and exporter status can affect the relationships between the presence of HGFs and the firm performance of non-HGFs.

3. Data and Empirical Strategy

3.1. Data

Our main source of information is the firm-level corporate income tax statements during the period 1998-2014 from the Hungarian National Tax Authority (NAV). The dataset has almost universal coverage as it includes all firms that require double-entry bookkeeping.[7] The sample covers more than 95 percent of employment and value added of the business sector and about 55 percent of the full economy in terms of GDP. The dataset includes information on a wide-range of matters such as ownership, financial information, employment, and industry at the NACE 4-digit code level and the location of their headquarters. The Centre for Economic and Regional Studies, Hungarian Academy of Science (CERS-HAS) has extensively cleaned and harmonized the data.

Given the scope of our analysis, we restrict the data in several ways. First, we exclude non-profit organizations. Second, we drop firms that operate either in agriculture or in the non-market service sectors of the economy.[8] Third, we drop all firm-year observations with less than 5 employees.[9] Finally,

---

[6] Note that this inability to distinguish between these two channels is present in most FDI spillover studies. Recently, Fons-Rosen et al. (2017) isolated the channels of knowledge spillovers and demand effects that foreign firms have on domestic firms. They use firm-level patent data to construct a measure of knowledge spillovers from FDI.

[7] Only sole traders and very small firms are not required to use double-entry bookkeeping in Hungary.

[8] Non-market services sectors are sectors 53, 58, 75, 84 to 94, and 96 to 99, based on NACE rev 2.

[9] We do this to avoid overestimating the incidence of HGFs. Besides, these micro-sized firms are likely to include individual entrepreneurs who operate a formal firm to reduce their tax burden (Semjén et al., 2009), and many small firms are likely to underreport their sales, earning and wages paid (Tonin, 2011).

we drop all firm-year observations with missing, negative or zero employment, turnover, material cost, capital or value added. As we will discuss in detail in section 3.3, our regressions will be run for 3-year intervals, which further reduces the number of observations. We also restrict the sample to surviving firms, given our primary interest in TFP change. Finally, we exclude firms which operate in the capital, Budapest.[10] Our final regression sample consists of more than 140,000 firm-year observations.[11]

The data include the most important balance sheet items. Nominal variables were deflated by the appropriate 2-digit industry level deflators from OECD STAN.[12] We rely on two proxies of productivity: labor productivity measured as value added over employment and total factor productivity (TFP). We estimate TFP using the semi-parametric methods in Ackerberg et al. (2015) but will also present results based on other methods (Wooldridge, 2009, and with fixed effects) to check for robustness. We also estimate markups following De Loecker and Warzynski (2012) and use them as dependent variables and as control in the productivity regressions. The dependent variables in our regressions are 3-year changes in these productivity measures, markups, employment, sales, average wage and return on assets (ROA). We winsorize these long differences at the 5th and 95th percentiles of their distribution, calculated separately for each 2-digit industry-year pair.

As for the construction of our main explanatory variable, there are multiple definitions that exist for HGFs in the literature. A key difference between these definitions is whether they rely exclusively on (i) relative or (ii) absolute growth rates. The former tends to identify HGFs as mostly small firms, while the latter provides a less skewed identification. As all theoretical spillover channels – either via vertical linkages or the labor market – are likely to be stronger when they originate from larger firms, the latter measure is more likely to capture these externalities. Hence, that is our preferred approach, but we also show the results under an alternative definition. Specifically, we choose to identify HGFs based on the Birch index (Birch, 1981), computed for all firms with at least 5 employees in year t:

$$Birch_{it} = (emp_{t+3} - emp_t)\frac{emp_{t+3}}{emp_t}$$

We rank firms according to their $Birch_{it}$ values for each year and classify a firm as HGF if their Birch index is above the 95th percentile in that year. The constant and absolute threshold generates variation across industries, regions and years.

---

[10] We do this for two reasons. First, many larger firms register their headquarters in Budapest even though most of their activities are conducted elsewhere. Second, Budapest, with 20% of firms and being a very large geographic entity compared to other counties or microregions, represents an outlier. Our results are robust to including Budapest in the sample.

[11] The full sample includes a total of 1.7 million observations for our regression years (1998, 2001, 2004, 2007, 2009, 2011); 320,000 firms have at least 5 employees and 260,000 firms are observable for at least 3 years. From these, we have all the variables to estimate TFP for 210,000 firms, from which we exclude 70,000 because they operate in Budapest.

[12] The OECD STAN includes value added, capital, intermediate input and output price deflators. When a variable has a specific deflator, we deflate it with that, while we deflate all other variables with the output price deflator. Naturally, this process does not control for within-industry price differences.

As a robustness check, we estimate the regressions with the OECD definition of HGFs, whereby a firm is high-growth if it employs at least 5 employees at the beginning of the period and has an annualized employment growth rate of at least 20 percent for a period of 3 years (OECD, 2010).

By definition, the share of Birch-HGFs is 5 percent for the full sample in each year. There are some differences across industries, with the majority of HGFs in high-tech manufacturing industries.[13] The OECD definition shows similar patterns across industries, but by construction it also varies across years, following the macro cycle (Table A1 in the Appendix presents the share of HGFs by different types of sectors according to the Birch index and OECD definitions).[14]

Importantly, the dataset allows us to exploit the detailed geographical information of each firm. The data on the municipality of the firms' headquarters are used to construct our spillover measures. Municipalities are the smallest geographical unit of measurement in Hungary: they can be aggregated up to 174 microregions, where each microregion tends to include one city and the neighboring smaller settlements. The microregions can be grouped into 20 counties, including Budapest, which can be further aggregated into 7 regions, representing the NUTS-2 geographic aggregation.[15]

The choice of the geographical unit to create the spillover variables is essential to reduce spurious results. A highly aggregated geographical unit may result in spurious correlations between the spillover and firm performance variables.[16] We consider the spillovers from HGFs at the level of counties and microregions as they can suggest different spillover channels. Counties are likely to involve many supplier-buyer relationships as firms within a value-chain tend to co-locate near each other.[17] Microregions are the relevant unit of observation of local labor markets, and thus may be the appropriate unit for the analysis of knowledge spillovers via labor mobility.

3.2. Constructing the spillover measures

HGFs can affect non-HGFs in their industry or in sectors that buy from or sell to them. Without detailed firm-to-firm transaction data, the buyer and seller linkages are established through national input-output tables, following the vertical spillover literature (Javorcik, 2004). To construct the spillover

---

[13] We rely on Eurostat's indicators on high-tech industry and knowledge-intensive services. Manufacturing industries are grouped in four categories: high-tech, medium-high-tech, medium-low-tech and low tech. Services are organized in knowledge-intensive services and other services.

[14] The relatively large HGF share in high-tech manufacturing over many years is somewhat in contrast with findings in the literature (Daunfeldt et al., 2016;Henrekson and Johansson, 2010). The World Bank (2018) shows that while there may be more HGFs in high-tech sectors in Hungary, there is no consistent pattern of HGFs in high-tech sectors across many developing countries. The table shows that this is partly driven by the peculiarities of the Birch definition, but the share of HGFs is still large in these industries according to the OECD definition, suggesting that high-tech manufacturing was indeed quite dynamic in this period in Hungary.

[15] The NUTS (Nomenclature of Territorial Units for Statistics) system of classifying geographical units is the standard statistical system used for all European Union member countries.

[16] This issue has been identified as the Modifiable Areal Unit Problem (MAUP). See Briant et al. (2010) for a discussion.

[17] Hillberry and Hummels (2008) show that spatial frictions reduce trade drastically, and attribute this fall to the co-location of firms within a value-chain.

measures, we first calculate the share of HGFs in the industry. Let HS_{jrt} be the employment-weighted share of HGFs in (2-digit) industry j, region (county or microregion) r, and year t:[18]

$$HS_{jrt} = \frac{\sum_{i \in jr} emp_{it} HGF_{it}}{\sum_{i \in jr} emp_{it}}$$

where $\sum_{i \in jr} emp_{it}$ is the total number of employees in industry j in region (county or microregion) r and year t, and HGF_{it} is an indicator variable that equals 1 if firm i is an HGF. HS_{jrt} accounts for the firms with at least 5 employees and an HGF phase starting in year t, $t-1$ or $t-2$. The results are robust to using an alternative definition of HGFs where sales are used as weights.

The upstream and downstream measures are based on the average share of HGFs in supplier and buyer industries in the region, weighted by the volume of intermediate-good flows across other industries. Importantly, when calculating these measures for a given firm, we omit said firm's 2-digit industry from the computation to capture only genuine inter-industry spillovers. Buyer-seller connections are identified using information from the 2-digit (symmetric harmonized) input-output matrices for Hungary provided by the OECD:[19]

$$Upstream\ HGF_{jrt} = \sum_{m \neq j} \alpha_{mj} HS_{mrt}$$

and

$$Downstream\ HGF_{jrt} = \sum_{m \neq j} \alpha_{jm} HS_{mrt}$$

where α_{mj} are the normalized coefficients from the symmetric input-output matrix representing (domestic) intermediate good flows from industry m to j, and α_{jm} are coefficients representing intermediate goods flows from industry j to m.

We also control for within-industry spillovers by calculating the share of HGFs in the *firms' 2-digit industry*; using data on all entities with the same 2-digit code, save for the ones that belong to the firm's own 4-digit industry[20] (denoted by g). Thus, this measure is calculated as:

$$Within\text{-}industry\ HGF_{grt} = \frac{\sum_{i \in jr} HGF_{it} - \sum_{i \in gr} HGF_{it}}{\sum_{i \in jr} emp_{it} - \sum_{i \in gr} emp_{it}}$$

Note that $Within\text{-}industry\ HGF_{gjrt}$ is very close to the `traditional' horizontal spillover measure, as defined in, for example, Javorcik (2004). The difference is that we omit the firm's 4-digit industry, to

---

[18] For consistency with the regression sample, we calculate the spillover measures based on the subsamples of firms with at least 5 employees.

[19] Source: http://stats.oecd.org/Index.aspx?DataSetCode=IOTS. These tables are available only between 2000-2011, so we have used the 2011 weights for the years after 2011 and the 2000 weights for the years before 2000.

[20] The sub-industry is narrowly defined as it describes one specific economic activity. An example is subindustry 10.11, namely, the processing and preserving of meat. The second measure is the share of HGFs in the firm's *industry* defined as the 2-digit NACE code. The industry is more broadly defined since it captures many economic activities. Using the same example, the processing and preserving of meat will be grouped together with the production of meat products under industry 10, and the manufacture of food products.

avoid generating spurious correlation. If all firms – including HGFs – are in the sample, then the high growth of the HGF will show up both in the HGF measure and on the left-hand side. Omitting HGFs from the regressions would introduce endogenous selection based on the dependent variable.[21] We also report results restricted to non-HGFs and interact the spillover variables with the HGF status of the firm, and these results are similar to the results on the full sample, which are in line with HGFs generating externalities for other firms.[22]

Table 1 shows summary statistics of our key spillover measures and firm outcomes. The employment share of Birch HGFs was 14.5% on average in a 2-digit industry-country combination, with substantial variation. The average value of the vertical measures is higher, reflecting the fact that sectors with more HGFs have stronger linkages with other sectors.[23]

Table 1 Summary statistics for the regression sample

Variable	Obs	Mean	Std. Dev.	Min	Max
Downstream HGF	140,271	0.234	0.050	0.000	0.507
Upstream HGF	140,271	0.215	0.049	0.000	0.373
Within-industry HGF	140,271	0.145	0.173	0.000	0.987
Labor prod (3-year change)	140,271	0.042	0.536	-1.064	1.183
TFP ACF (3-year change)	140,271	0.068	0.670	-1.257	1.571
Log employment (3-year change)	140,271	-0.077	0.423	-1.099	0.647
log sales (3-year change)	140,271	-0.036	0.519	-1.174	0.940
log average wage (3-year change)	140,271	0.169	0.274	-0.337	0.746
ROA (3-year change)	140,271	-0.026	0.169	-0.408	0.322

Notes: This table shows summary statistics of the main variables for the regression sample. HGF spillover measures are based on the 3-year Birch employment-based definition and weighted by the number of employees. TFP estimation is based on the Ackerberg et al. (2015) approach, at the 2-digit industry level, using the sample of firms with at least 5 employees.

3.3. Empirical strategy

The specification for our baseline regression model is:

$$\Delta y_{it} = \sum_m \beta^m * spillover_{grt}^m + \gamma X_{it} + \mu_{st} + \rho_{Rt} + \varepsilon_{it}$$

where i indexes firms, g (4-digit) industries, r counties, s 1-digit sectors, R NUTS2 regions and t years. Δy_{it} is the three-year growth rate (between t and $t+3$) of the main firm performance variables: labor

---

[21] Note that the situation differs from the FDI spillover case in two respects. First, running the regression on always domestic-owned firms makes sense in that literature; and second, the spillover variable is not derived directly from the dependent variable in the FDI case. These differences motivate our deviation from that literature.

[22] Note that there are alternatives to this solution, for example, subtracting the given firm's status from the HGF measure. This, however, would generate a number of other problems, explored by the peer effects literature (Angrist, 2014). We think that our approach is a clear and transparent one.

[23] This broadly reflects that manufacturing has a larger weight in input-output tables and the employment share of Birch HGFs is larger in manufacturing.

productivity, TFP, employment, sales, average wages (all of these in logs), and ROA.[24] $spillover_{grt}^m$ are the three different measures of spillovers (m={ downstream, upstream and within-industry}) and β^m are the corresponding spillover relationships. X_{it} captures three firm level characteristics, all related to productivity growth: (log) employment, labor productivity and capital intensity (or log(fixed assets/employee)). Note that these variables are included as levels in year t, an initial value before the growth measured by the dependent variable. Following Javorcik (2004) we also control for the change in demand in buyer industries in all specifications. Together with differencing the dependent variable, they are likely to capture a rich set of heterogeneous firm-level characteristics.

We also include two types of fixed effects to capture industry-level and regional shocks. First, we include industry-year fixed effects, μ_{st}, to control for sectoral macro-shocks. Second, we include region-year fixed effects, ρ_{Rt}, to control for regional macro-shocks. Both types of fixed effects are measured at a somewhat more aggregated level than the spillover variables themselves, to allow for some variation across industries within sectors, and across counties within regions.[25]

Given that both the dependent variable and the spillover variables are calculated for 3-year periods, running the regression at the annual level would inflate the number of observations artificially. Therefore, we run the regressions for 3-year periods (with a small modification at the end of the sample).[26] The periods included are 1998-2001, 2001-2004, 2004-2007, 2007-2010, 2009-2012, and 2011-2014.

Our identification strategy relies on variation from three main sources: geographic variation within regions, industry variation within sectors and time variation (and the interactions of these variables). We show that the results are robust to alternative specifications. Across-industry cross-sectional variation is important for our results. This is because within industry variation is quite limited for several reasons. First, there is limited variation of links in input-output tables and HGF shares in 2-digit industries. Second, many periods are involved when measuring the dependent and the spillover variables. The dependent variable captures firm growth between periods t and $t + 3$, which depends on changes during this period, including firms starting their high growth period in any of these years.[27] The explanatory variables consider firms which start their high growth period between $t - 2$ and t. Consequently, in the empirical specification one 'year' observation consists of changes across 5-6 years, even not considering delayed effects of high growth phases. Useful time variation is quite limited.

To clarify further our empirical strategy, we contrast it with the usual method employed in the FDI spillovers literature. From a technical point of view, an important difference between the two empirical

---

[24] We winsorize these 3-year changes at the 5th and 95th percentiles. While this 'winsorizes' HGFs from the left-hand side, we find this sensible given that the aim is to estimate the relationship between HGF activity and the growth of the average firm, and winsorizing outliers helps with estimating a more precise effect on the mean. Winsorizing does not affect our results in a significant manner.

[25] For the fixed effects: industries are defined as sections, which are 2-digit divisions aggregated to the level identified by an alphabetical code in NACE revision 2; and regions include typically 3 counties, at the level most of our variables are measured.

[26] As the full sample covers 17 years, not divisible by 3, the three last periods intersect.

[27] Consider firms which start their high growth phase in T+2, which may clearly affect firm growth between t and t+3. These firms will even be considered HGFs in t+5.

exercises is that HGF status is much more transient than foreign ownership: most firms that are foreign or become foreign owned remain in that state. As a result, in typical FDI-spillover equations (e.g. Harrison and Rodríguez-Clare, 2010), the dependent variable is the *level* of productivity and the equation is estimated with firm fixed effects. This method is less appropriate here as HGF status is less persistent. It is more intuitive in our setting to check whether productivity *growth* is higher when more firms are experiencing a high-growth phase in vertically related industries. The empirical specification is analogous to the differenced regressions in Javorcik (2004), which differences out time-invariant firm characteristics.

A second difference with respect to the FDI literature emerges concerning the horizontal measures of spillovers. In the FDI spillover literature, the horizontal measure includes the foreign firm but the regression is conducted on the sample of domestic firms. This method is hard to apply to our setting. Restricting the sample to non-HGFs would lead to the endogenous exclusion of firms with the strongest growth performance. Therefore, we run our main regression on the full sample, but report in Table 5 that we find similar results when the sample is restricted to non-HGFs and when we interact the spillover variables with the HGF status of the firm.

In summary, our identification strategy attempts to use sensible temporal assumptions together with a set of fixed effects. Still, our results are conservatively interpreted as correlations. It is possible that there are unobserved industry characteristics which may partly drive this correlation. Also, reverse causality is possible: stronger productivity growth of suppliers may help their buyers to grow rapidly. Given that regressions aimed at predicting HGF status usually have a low explanatory power, it is unlikely that one can find credible instruments for regional HGFs presence. Consequently, we deem our strategy a relatively transparent and credible way to document the correlation between HGFs and productivity growth in related industries.

4. Results

We first report the main results, and then discuss the heterogeneous effects and robustness checks. Importantly, we report beta coefficients in all tables because the spillover measures do not have a very intuitive unit of measurement. The original coefficients can be calculated with the help of the standard deviations in Table 1, on which we sometimes rely when interpreting the results.

4.1. Main firm-level regression results

Our main results are reported in Table 2. First, productivity growth is positively correlated with HGF share in the same and vertically related industries (columns 1 and 2). The beta coefficients show that one-standard-deviation increase in HGF share is associated with a 1.5-2 percent of a standard deviation larger increase in productivity in the same industry and a 2-3 percent of a standard deviation larger increase in supplier industries. These relationships are significant also from an economic perspective, representing about a 1-1.5 percentage points of additional productivity growth in the same industry and 1.5-2 percent higher growth in supplier industries in three years. These results are consistent with positive productivity spillovers both within subsectors and from buyers to sellers.

Second, the presence of HGFs, especially within the same 2-digit industry and buyer industries, is associated with stronger growth of other firms (columns 3 and 4). The presence of HGFs in downstream

industries is associated with higher sales growth, but no extra growth in employment. This is consistent with HGFs generating additional demand, which may drive the premium in productivity growth.[28]

Table 2 Main regression results

Dependent:	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity		Size		Income	
					av. wage	
	Labor prod.	TFP	Emp. (ln)	Sales (ln)	(ln)	ROA
Downstream HGF share	0.032***	0.023***	-0.000	0.017***	0.022***	0.012***
	(7.397)	(4.370)	(-0.039)	(3.824)	(5.246)	(2.656)
Upstream HGF share	-0.002	-0.002	0.001	0.001	0.003	0.001
	(-0.330)	(-0.302)	(0.110)	(0.112)	(0.435)	(0.106)
Within-industry HGF share	0.015***	0.020***	0.008**	0.012***	-0.004	0.012***
	(3.866)	(3.523)	(2.411)	(2.751)	(-0.773)	(2.875)
Demand	0.003	0.011*	0.008*	0.013***	0.011**	-0.003
	(0.808)	(1.918)	(1.898)	(2.604)	(2.298)	(-0.586)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	140,271	140,271	140,271	140,271	140,271	140,271
R-squared	0.215	0.152	0.085	0.070	0.162	0.095

Note: This table investigates the relationship between the presence of HGFs in upstream and downstream industries as well as in the same sector, and 3-year change in the different dependent variables. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. Firms are classified as HGFs based on the 3-year Birch employment definition.

Third, potential productivity gains will be distributed between the two key stakeholders of firms: employees (through higher wages, column 5) and owners (through higher return on assets, ROA, column 6). In fact, both workers and owners benefit from HGFs in supplier industries, while the higher productivity associated with higher HGF share in the same industry seems to benefit only owners.

Table 3 attempts to shed some light on what drives the productivity increase. The first two columns examine whether the TFP estimation method matters. In column 1 we estimate productivity with the simple fixed effects estimator rather than the ACF procedure. In column 2 we address the concern that HGF status may be associated with an increase in market share and, therefore, an increase in market power. We address this issue by including the firm's market share in the 4-digit industry and county into the first step of the productivity estimation procedure. Results remain mostly unchanged in the different specifications.

An even more consequential question is whether the increase in the revenue-based TFP is driven by an increase in prices and markups or by an increase in physical productivity. In column 3 the dependent

---

[28] As the estimation uses a revenue-based productivity measure, the higher productivity levels can also be a result of higher product prices or profits that are driven up by the additional demand from HGFs.

variable is the firm-year level markup. We indeed find that HGF presence is associated with higher markups. HGFs seem to drive up the prices both of firms in the same subsector and in supplier industries, in line with the increased demand hypothesis. In columns 4 and 5 we re-estimate the main productivity equations controlling for the change in markup. As expected, revenue-based productivity is strongly positively associated with markups. However, we find evidence that HGF presence is associated with higher productivity even when controlling for markups, though its effect is halved. These results suggest that a substantial portion of the positive externality of HGFs is explained by higher prices, but physical productivity may also increase.[29]

Table 3 Other productivity measures and markups

Dependent:	(1) TFP (FE)	(2) TFP (ACF, MS)	(3) Markup	(4) Lab. Prod	(5) TFP (ACF)
Downstream HGF share	0.023***	0.023***	0.022***	0.017***	0.011**
	(5.007)	(4.294)	-5.387	(6.120)	(2.462)
Upstream HGF share	-0.001	-0.003	-0.005	0.001	0.000
	(-0.181)	(-0.378)	(-0.861)	(0.384)	(0.072)
Within-industry HGF share	0.024***	0.020***	0.019***	0.002	0.009**
	(4.451)	(3.726)	-5.037	(0.886)	(1.982)
Demand	0.008	0.016**	-0.003	0.005*	0.013**
	(1.608)	(2.475)	(-0.609)	(1.769)	(2.488)
Markup (3-year change)				0.679***	0.569***
				(182.291)	(150.074)
Controls	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes
Observations	140,271	140,271	140,271	140,271	140,271
R-squared	0.171	0.150	0.229	0.571	0.401

Note: The table replicates Table 2, column 2 but relying on different productivity measures. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. Firms are classified as HGFs based on the 3-year Birch employment definition. FE stands for Fixed Effects, and MS for Market Share (included in the first estimation step). Markups are estimated based on the De Loecker-Warzynski (2012) method.

Taken together, the results indicate that HGF presence in the same industry is associated with higher growth in productivity and size. Downstream productivity spillovers are also positive, resulting from increased sales with a similarly sized workforce. We find that the increase in measured TFP results from a combination of increased markups and physical productivity.

---

[29] Even though the estimated markup is significant and meaningful, we decided against including it as a control in the remaining tables because we think that including would lead to important econometric problems, such as simultaneity.

4.2. Which firms benefit?

The spillover effects can vary along several observable dimensions, which may shed some light on the underlying mechanisms, especially on the role of knowledge spillovers in these associations. First, we examine whether the relationships differ by industry, age and export status. Second, we investigate whether HGFs are more likely to benefit from the presence of other HGFs than other firms. Third, we study whether absorptive capacity, proxied by the TFP level of the receiving firm, matters. Finally, we check the importance of geographical proximity in these relationships.

Table 4 splits the sample by sector, age and export status. We find evidence of a positive relationship between productivity growth and within-industry HGF presence for both manufacturing and services firms. Vertical linkages are more important in manufacturing, where input-output linkages may be more pronounced.[30] We do not find heterogeneity in terms of age, indicating that not only young, but also established firms can learn from HGFs. We also find that exporting and non-exporting firms can benefit from HGF presence to a similar extent. Even though exporters tend to be more productive than non-exporters, being exposed to more HGFs downstream and within the industry is associated with higher productivity growth.

Table 4 Results with different subsamples

Dependent:	Labor productivity growth between t and t+3					
Sample:	Manufacturing	Services	Young	Old	Non-exporter	Exporter
Downstream HGF share	0.049***	-0.011	0.020***	0.037***	0.031***	0.029***
	(6.823)	(-1.500)	(3.315)	(7.669)	(6.987)	(4.335)
Upstream HGF share	0.034***	0.001	-0.005	-0.000	-0.004	0.008
	(3.061)	(0.299)	(-0.656)	(-0.075)	(-0.759)	(0.851)
Within-industry HGF share	0.012*	0.013**	0.014***	0.016***	0.017***	0.015***
	(1.950)	(2.300)	(2.710)	(3.447)	(4.443)	(2.636)
Demand	0.015**	-0.002	0.011*	0.001	0.002	0.003
	(2.031)	(-0.409)	(1.655)	(0.192)	(0.497)	(0.467)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	34,419	105,852	38,935	101,336	105,533	34,733
R-squared	0.185	0.224	0.262	0.190	0.229	0.193

Note: The table splits the sample to Manufacturing/Services, Young (less than 5-year-old)/Old firms and exporters/non-exporters and re-runs the main regressions on these subsamples. The explanatory variables show the share of HGFs in different industries in the same county, by using the 2-digit input-out matrix from OECD STAN. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. The firm controls are log employment, labor productivity and log capital

---

[30] These distinctions may capture the different breadth of industries in manufacturing and services, rather than substantive differences between the two groups.

per labor ratio. The definition of HGF is the 3-year Birch employment definition.

One may also ask whether HGFs will benefit more or differently from the presence of other HGFs than other firms. While HGFs may have a stronger interest in learning about growth opportunities or new technologies, non-HGFs, which are probably more strongly anchored to the local market may benefit more from the expanding demand of HGFs.

Table 5 investigates this question in two ways. Panel A restricts the sample to non-HGFs, and the results are very similar to those in Table 2 showing that we indeed identify the effect of HGFs on other firms in our main specification. Methodologically, it also shows that the main results are not driven by the mechanical relationship of HGFs having higher growth rates. In Panel B we allow for the spillover effects to differ between HGFs and non-HGFs by introducing interactions between the spillover variables and the receiving firms' HGF status. As in Panel A, the results for non-HGFs are very similar to the main specification, while we also find that HGFs only benefit from within industry spillovers. Therefore, horizontal spillovers seem to affect both HGFs and non-HGFS, while vertical linkages are more likely to benefit non-HGF firms.

Table 5 Main results for non-HGF firms

PANEL A: Restricted to non-HGF

	Productivity			Size	Income	
Dependent:	Labor prod.	TFP	Emp. (ln)	Sales (ln)	av. wage (ln)	ROA
Downstream HGF share	0.033***	0.024***	-0.006	0.012***	0.022***	0.012**
	(7.629)	(4.446)	(-1.448)	(2.734)	(5.160)	(2.557)
Upstream HGF share	-0.004	-0.003	0.003	0.002	0.002	-0.000
	(-0.675)	(-0.424)	(0.599)	(0.280)	(0.312)	(-0.078)
Within-industry HGF share	0.014***	0.019***	0.007**	0.011**	-0.005	0.011***
	(3.588)	(3.244)	(1.970)	(2.409)	(-0.903)	(2.661)
Demand	0.003	0.012*	0.009**	0.013**	0.010**	-0.004
	(0.682)	(1.887)	(1.968)	(2.538)	(2.054)	(-0.794)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	133,601	133,601	133,601	133,601	133,601	133,601
R-squared	0.211	0.148	0.083	0.064	0.161	0.096

PANEL B: Interact with HGF status

	Productivity			Size	Income	
Dependent:	Labor prod.	TFP	Emp. (ln)	Sales (ln)	av. wage (ln)	ROA
Non HGF X Downstream HGF share	0.048***	0.034***	-0.008	0.018***	0.030***	0.017***
	(7.755)	(4.502)	(-1.383)	(2.911)	(5.076)	(2.713)

Non HGF X Upstream HGF share	-0.005	-0.005	0.005	0.003	0.003	-0.000
	(-0.672)	(-0.506)	(0.706)	(0.378)	(0.350)	(-0.054)
Non HGF X Within-industry HGF share	0.013***	0.017***	0.009***	0.012***	-0.005	0.011***
	(3.397)	(3.109)	(2.626)	(2.689)	(-0.937)	(2.708)
HGF X Downstream HGF share	0.015	0.021	-0.018**	0.006	0.035***	-0.005
	(1.197)	(1.560)	(-2.343)	(0.603)	(2.922)	(-0.426)
HGF X Upstream HGF share	0.026*	0.016	-0.013*	-0.001	0.008	0.021
	(1.944)	(1.061)	(-1.651)	(-0.040)	(0.617)	(1.520)
HGF X Within-industry HGF share	0.014***	0.020***	-0.012***	-0.004	0.004	0.005
	(4.540)	(5.188)	(-5.502)	(-1.435)	(1.309)	(1.497)
Demand	0.003	0.011*	0.009**	0.014***	0.011**	-0.003
	(0.807)	(1.916)	(2.140)	(2.741)	(2.296)	(-0.573)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	140,271	140,271	140,271	140,271	140,271	140,271
R-squared	0.217	0.154	0.179	0.116	0.162	0.095

Note: This table investigates whether the effect of HGF presence differs between HGFs and other firms. Panel A restricts the sample to non-HGFs while Panel B interacts the spillover measures with the HGF and non-HGF dummies. The explanatory variables show the share of HGFs in different industries in the same microregion and county, by using the 2-digit input-out matrix from OECD STAN. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. Firm controls are log employment, labor productivity and log capital per labor ratio. The definition of HGF is the 3-year Birch employment definition.

Next, we investigate whether the strength of the relationship depends on the absorptive capacity of the non-HGFs (Table 6). We proxy the absorptive capacity with the productivity level of the firm in the initial year (t). More productive firms are more likely to attract the business of rapidly growing buyers or may benefit more from knowledge spillovers because of their better absorptive capacity. Specifically, to test these hypotheses, we split the sample into four productivity quartiles within each industry-year combination in the initial year and run the main specifications in these subsamples.[31] Therefore, for example, column 1 in Table 6 shows the main regression run on the sample of firms which were in the lowest quartile of the productivity distribution. We find some evidence of heterogeneity, as productivity effects are stronger for firms with a medium level productivity, with significant differences between the middle quartiles on the one hand, and the lower and upper quartile on the other. Firms with very low productivity levels may lack the capabilities to learn from HGFs or supply them effectively, while highly

---

[31] The quartiles are created based on labor productivity. For each firm we calculate the quartile for each period, and take the modal quartile to ensure that firms do not move across subsamples.

productive firms may already have easy access to domestic and international markets and, therefore, their performance is less dependent on local HGFs. This difference is robust to taking deciles instead of quartiles, to using the other two TFP measurements, and to employing the OECD definition of HGF instead of the one developed in Section 3.

Table 6 HGF presence for firms in different initial productivity level quartiles

Panel A:	Labor productivity growth between t and t+3			
Sample: productivity quartile in year t	1st	2nd	3rd	4th
Downstream HGF share	0.012*	0.040***	0.046***	0.018***
	(1.809)	(6.228)	(7.540)	(2.746)
Upstream HGF share	-0.006	0.019***	-0.013*	-0.013*
	(-0.748)	(2.596)	(-1.845)	(-1.696)
Within-industry HGF share	0.001	0.011**	0.010*	0.005
	(0.131)	(1.969)	(1.712)	(0.858)
Demand	-0.014*	0.010	-0.006	0.007
	(-1.822)	(1.493)	(-1.002)	(1.025)
Controls	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
Observations	28,509	33,030	36,884	41,848
R-squared	0.271	0.258	0.268	0.228

Panel B:	TFP growth between t and t+3			
Sample: productivity quartile in year t	1st	2nd	3rd	4th
Downstream HGF share	0.002	0.029***	0.033***	0.008
	(0.338)	(3.884)	(4.701)	(1.117)
Upstream HGF share	-0.003	0.018**	-0.006	-0.021**
	(-0.324)	(2.013)	(-0.688)	(-2.453)
Within-industry HGF share	0.005	0.015**	0.016***	0.008
	(0.722)	(2.292)	(2.662)	(1.150)
Demand	-0.014*	0.017**	0.010	0.016*
	(-1.756)	(2.044)	(1.373)	(1.937)
Controls	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
Observations	28,509	33,030	36,884	41,848
R-squared	0.229	0.173	0.180	0.169

Note: The table splits the sample into 4 quartiles based on the initial productivity level of firms and re-runs the main regressions in Table 2 on these subsamples. The dependent variables are changes between year t and

t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. The explanatory variables show the share of HGFs (based on 3-year Birch employment definition) in different industries at the same county, by using the 2-digit input-out matrix from OECD STAN. The firm controls are log employment, labor productivity and log capital per labor ratio. We split the sample to productivity quartile within industry-year and take the model quartile for each firm.

Finally, we attempt to measure the geographic span of spillovers by constructing the variables at different levels of geographic aggregation (Table 7). Transferring tacit knowledge, for example, may require face-to-face interactions through personal relationships or labor mobility. Supplier-buyer relationships may operate over somewhat longer distances. The previous regressions include county-level measures, which are relatively broad geographic aggregates. In contrast, microregions are very narrowly defined, often interpreted as local labor markets. Recall that there are 20 counties and 174 microregions in Hungary. Consequently, knowledge spillovers via labor mobility are likely to show up at the microregion level, while broader supplier-buyer relationships may be present at the county level.[32]

First, we examine the spillovers at a more detailed geographic level and include only the microregion-level spillover measures (columns 1 and 3). Second, we include the microregion-level measures together with the county-level measures (columns 2 and 4).

Microregion-level variables have very similar coefficients to county-level ones, suggesting that the channels at small distances are not systematically different. Importantly, both sets of variables remain significant when both the microregion- and county-level variables are included, suggesting that HGF externalities are stronger at short distances, which is in line with the possibly tacit nature of knowledge spillovers.

This exercise also serves as a robustness check. Our results are robust to different geographic aggregation levels, and the forces we quantify act both at short and longer distances.

Table 7 Microregion-level measures

Variable:	Aggregation	Labor prod.	Labor prod	TFP	TFP
			Dependent variable:		
			Growth between t and t+3		
Downstream HGF share	microregion	0.028***	0.017***	0.022***	0.015***
		(7.415)	(4.279)	(5.108)	(3.471)
Upstream HGF share	microregion	0.005	0.011***	-0.000	0.004
		(1.337)	(2.876)	(-0.102)	(0.851)
Within-industry HGF share	microregion	0.011***	0.007**	0.019***	0.014***
		(3.684)	(2.091)	(5.453)	(3.772)
Demand	microregion	0.001	0.000	0.003	0.002
		(0.291)	(0.139)	(0.733)	(0.649)
Downstream HGF share	county		0.025***		0.017***
			(5.165)		(2.992)

---

[32] Note that HGFs in the microregions are already included in the county-level measure; hence, their effect is captured to some extent by those variables.

Upstream HGF share	county		-0.007		-0.004
			(-1.181)		(-0.533)
Within-industry HGF share	county		0.011***		0.013**
			(2.665)		(2.125)
Demand	county		0.003		0.011*
			(0.704)		(1.781)
Controls		yes	yes	yes	yes
region-year FE		yes	yes	yes	yes
sector-year FE		yes	yes	yes	yes
Firm controls		yes	yes	yes	yes
Observations		140,271	140,271	140,271	140,271
R-squared		0.215	0.216	0.152	0.152

Note: The table shows the main regressions in Table 2 with spillovers measures defined at a finer level of geographic aggregation. The explanatory variables show the share of HGFs in different industries at the same microregion and county, by using the 2-digit input-out matrix from OECD STAN. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. Firm controls are log employment, labor productivity and log capital per labor ratio. The definition of HGF is the 3-year Birch employment definition.

Let us summarize these findings under two themes. First, they support the hypothesis that knowledge transfer plays a role in the relationships we find. In particular, only firms above a certain degree of absorptive capacity seem to be able to benefit from HGF spillovers, while very productive firms may not have much to learn from HGFs. The finding that spillovers are stronger within very short distances is also in line with the hypothesis that, in line with the tacit nature of knowledge, personal interactions may be important in the associations we observe. Second, we find that even established, exporting or HGF firms can benefit from HGF spillovers, and we only find a negative relationship for firms with very low productivity levels. The presence of HGFs seems to be associated with positive productivity growth for a large set of firms.

4.3. Robustness checks

A key question is the extent to which the results are specific for the HGF definition we have chosen. Table A2 re-runs the regressions with the OECD HGF definition instead of the Birch definition. We find that the within-industry spillovers are robust to this change in definition, while the vertical spillover effects disappear. A key difference between the two definitions is that the OECD (relative growth) definition favors smaller firms, while the Birch definition classifies many larger firms as HGFs. Therefore, larger HGFs, i.e. quickly expanding buyers, seem to be needed for vertical spillovers, while the size of HGFs appears to be less important for horizontal effects. This may be in line with the greater importance of linkage effects in vertical spillovers and that of less size-dependent knowledge effects in horizontal relationships. Relatedly, Lenaerts and Merlevede (2015) show that small foreign firms do not generate FDI spillover effects.

The next robustness check relates to concerns about the weights used in calculating the spillover measures. Recall that in our main specification we weight each HGF with its employment share. These weights may create an upward bias in the spillover measures and establish a spurious relationship

between the spillover measures and firm outcomes. Table A3 presents the results when we do not weight the HGFs when calculating the spillover measures. As expected, we find a similar but somewhat weaker relationship compared to the main results.

Given the temporal assumptions in our empirical specification, a concern may arise that these effects differ over time. In Table A4, we investigate whether the measured relationships change over time by running cross-sectional regressions for the different periods. We find relatively similar results and no obvious trends or structural breaks.

We also conduct further robustness checks to investigate whether the results are robust to including different sets of fixed effects in the regressions (Table A5). Recall that our baseline regression includes region-year and sector-year fixed effects and identify from variation across counties in regions and across industries within sectors. We control for finer geographic heterogeneity with microregion-year fixed effects (Table A5, columns 1 and 4). This set of fixed effects can capture very fine local demand or supply shocks. This change does not affect the results substantially. Still, it is possible that the results are identified from industry-specific demand shocks. To capture this, we include county-sector-year level fixed effects to find that our results are robust to different ways of controlling for local demand shocks (Table A5, columns 2 and 5).

An additional concern is that unobserved firm-level heterogeneity may bias the results. Recall, however, that firm-level heterogeneity is already handled in the original specification by differencing the dependent variable and controlling for key firm characteristics. Also, the explanatory variables only vary at a more aggregated level. Still, we report specifications with firm fixed effects (Table A5, columns 3 and 6) to control fully for unobserved firm-level heterogeneity (representing "fixed trends"). The results are somewhat sensitive to this, but we still find a significant positive relationship between the share of HGFs and TFP growth.

Lastly, a relevant concern about our empirical specification is that the share of HGFs can also be correlated with other industry characteristics. For example, the share of HGFs in an industry may be strongly correlated with the share of foreign firms. If foreign firms generate productivity spillovers, it is possible that our regressions can capture that. We thus construct downstream/upstream/within-industry measures from several relevant industry characteristics and control for them (Table A6). These measures are constructed analogously to the HGF spillover measures described in 3.2, but, rather than aggregating the HGF dummies, we aggregate up the foreign-owned dummies or sales growth rates.

HGFs may simply be more prevalent in industries with higher sales growth, and this sales growth, rather than the high-growth firms themselves, drives the increased productivity. To investigate this, we calculate aggregate sales for each industry-county combination between t and $t + 3$, and use them to create the downstream, upstream, and within-industry measures.[33] Controlling for changes in industry-level sales does not change our results significantly: HGF activity itself, rather than sales growth seems to be associated with productivity change (Table A6, column 1).

---

[33] Note that the "demand" variable already present in the main specification captures demand growth in downstream industries.

We also check whether our results are driven by a correlation between the share of globally integrated firms and HGFs. We control for the share of foreign firms (Table A6, column 2). Note that this is indeed the FDI spillover equation in Javorcik (2004) with the additional HGF spillover measures. As with the FDI spillover literature, we find effects consistent with positive downstream and horizontal spillovers. These variables are correlated with HGF share and reduce the point estimates for the HGF measures, yet the effects remain significant. As an additional proxy for internationalization, we control for the share of exporters (Table A6, column 3). This is, naturally, strongly correlated with the FDI share at the industry level, and its inclusion yields very similar results to column 2.

HGFs may simply be more prevalent in more innovative or more high-tech sectors. We thus control for the share of innovative firms at the industry-level (Table A6, column 4). We rely on data from the Community Innovation Survey. We find stronger productivity growth in industries with more innovative firms and in industries supplying more innovative industries. The regressions yield positive and mostly significant coefficients for the HGF measures, but they are substantially smaller than in the baseline model. Finally, we control for the industry's technology by including the average productivity of the vertically related sectors (Table A6, column 5). We find that firms that are in more productive industries or are supplying industries with more productive firms experience higher productivity growth. The coefficient for the within-industry HGF measure again becomes smaller but remains significant.

These results show that the HGF share in industries is indeed correlated with technology level and international integration. These industry-level variables explain a significant part of the association between HGF share in vertically related industries and productivity growth, but even conditional on these factors, a higher HGF share seems to be associated with a premium in productivity growth.

5. Conclusions

HGFs have a disproportionate ability to create jobs and generate output. But there is less evidence and understanding about whether their contribution to economic growth outweighs the costs they impose on other firms. In order to examine the impact of HGFs on the economy, we exploit detailed administrative microdata from Hungary. Productivity growth is stronger when firms operate in industries with more HGFs or supply such industries. These results are in line with firms benefiting from knowledge spillovers or increased demand from HGFs. The increase in productivity results from a combination of increased markups and physical productivity. Not only do HGFs generate positive spillovers, but they also benefit more firms in the middle of the productivity distribution. Firms in the bottom quartile of the distribution are not productive enough to take more of and advantage of HGFs' demand. Firms in the top quartile of the distribution have access to wider markets and do not depend on HGFs as much. Moreover, manufacturing firms are better able to benefit from the presence of HGFs than services firms. The relationship between HGF presence and firm performance is not heterogeneous by firm age or export status. Lastly, the spillover effects are still present at short distances, and the effects are stronger, suggesting that personal interactions may be important for our results.

The results are robust to alternative definitions of HGF spillovers measures, productivity measures, samples, and different sets of fixed effects and industry controls. As discussed in the paper, establishing causality presents significant econometric challenges. The fact that the relationship between HGF shares

and productivity gains continues to hold when controlling for a plethora of factors and is not sensitive to several robustness checks gives us confidence in the results.

The paper provides evidence for policymakers who may be concerned about directing scarce resources towards supporting HGFs, instead of using these resources for firms in other parts of the economy. Nevertheless, it also highlights the fact that the attention and resources given to HGFs can benefit other firms in the economy.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, *83*(6), 2411-2451.
 https://doi.org/10.3982/ECTA13408

Ács, Z. (2010). High-Impact Entrepreneurship. in Ács, Z., and Audretsch, D.B. (ed.) *Handbook of Entrepreneurship Research. 165. International Handbook Series on Entrepreneurship 5*, Chapter 7. https://doi.org/10.1007/978-1-4419-1191-9_7

Angrist, J. D. (2014). The perils of peer effects. *Labor Economics*, 30, 98-108. https://doi.org/10.3386/w19774

Audretsch, D. B. (2005). The Knowledge Spillover Theory of Entrepreneurship and Economic Growth. in Vinig, G.T., and Van Der Voort, R.C.W. (ed.) *The Emergence of Entrepreneurial Economics Vol. 9: Research on Technological Innovation, Management and Policy*, 37-54. https://doi.org/10.1016/S0737-1071(05)09003-7

Autio, E., Kronlund, M., and Kovalainen, A. (2007). High-Growth SME Support Initiatives in Nine Countries: Analysis, Categorization, and Recommendations. *Report Prepared for the Finnish Ministry of Trade and Industry. Publications, Industries Department*.

Birch, D. L. (1981). Who creates jobs? *The Public Interest*, (65), 3-14.

Blalock, G., and Gertler, P. J. (2008). Welfare gains from foreign direct investment through technology transfer to local suppliers. *Journal of International Economics*, *74*(2), 402-421. https://doi.org/10.1016/j.jinteco.2007.05.011

Bos, J. W., and Stam, E. (2011). Gazelles, industry growth and structural change. *Discussion Paper Series/Tjalling C. Koopmans Research Institute*, *11*(02).

Briant, A., Combes, P.P., and Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, *67*(3), 287-302. https://doi.org/10.1016/j.jue.2009.09.014

Carree, M. A., and A. Roy Thurik A. R. (2010). The impact of entrepreneurship on economic growth, in Ács, Z., and Audretsch, D. (ed.) *Handbook of Entrepreneurship Research*, 557-594. https://doi.org/10.1007/978-1-4419-1191-9_20

Coad, A., Daunfeldt, S. O., Hölzl, W., Johansson, D., and Nightingale, P. (2014a). High-growth firms: introduction to the special section. *Industrial and Corporate Change*, *23*(1), 91-112. https://doi.org/10.1093/icc/dtt052

Coad, A., Daunfeldt, S. O., Johansson, D., & Wennberg, K. (2014b). Whom do high-growth firms hire?. *Industrial and Corporate Change*, *23*(1), 293-327.

https://doi.org/10.1093/icc/dtt051

Daunfeldt, S. O., Elert, N., and Johansson, D. (2016). Are high-growth firms overrepresented in high-tech industries? *Industrial and Corporate Change, 25*(1), 1-21.
https://doi.org/10.1093/icc/dtv035

De Loecker, J., and Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, *102*(6), 2437-71.
https://doi.org/10.1257/aer.102A2437

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, *84*(2), 445-510.
http://dx.doi.org/10.3982/ECTA11042

Du, J., and Temouri, Y. (2015). High-growth firms and productivity: evidence from the United Kingdom. *Small Business Economics*, *44*(1), 123-143.
https://doi.org/10.1007/s11187-014-9584-2

Fons-Rosen, C., Kalemli-Ozcan, S., Sorensen, B. E., Villegas-Sanchez, C., and Volosovych, V. (2017). Foreign investment and domestic productivity: identifying knowledge spillovers and competition effects. *National Bureau of Economic Research Working Paper,* No. 23643.

Girma, S. (2005). Absorptive capacity and productivity spillovers from FDI: a threshold regression analysis. *Oxford Bulletin of Economics and Statistics*, *67*(3), 281-306.
https://doi.org/10.1111/j.1468-0084.2005.00120.x

Görg, H., and Greenaway, D. (2004). Much ado about nothing? Do domestic firms really benefit from foreign direct investment? *The World Bank Research Observer*, *19*(2), 171-197.
https://doi.org/10.1093/wbro/lkh019

Harrison, A., and Rodríguez-Clare, A. (2010). Trade, Foreign Investment, and Industrial Policy for Developing Countries. in Rodrik, D. and Rosenzweig, M. R. (ed.) *Handbook of Development Economics Vol. 5,* 4039-4214.
http://dx.doi.org/10.1016/B978-0-444-52944-2.00001-X

Hausmann, R., and Rodrik, D. (2003). Economic development as self-discovery. *Journal of Development Economics*, *72*(2), 603-633.
https://doi.org/10.1016/S0304-3878(03)00124-X

Havranek, T., and Irsova, Z. (2011). Estimating vertical spillovers from FDI: Why results vary and what the true effect is. *Journal of International Economics*, *85*(2), 234-244.
https://doi.org/10.1016/j.jinteco.2011.07.004

Hillberry, R. and Hummels, D. (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review*, 52(3), 527-550.
https://doi.org/10.1016/j.euroecorev.2007.03.003

Henrekson M., and Johansson, D. (2010). Gazelles as job creators: a survey and interpretation of the evidence. *Small Business Economics, 35*, 227-244.
https://doi.org/10.1007/s11187-009-9172-z

Javorcik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review*, *94*(3), 605-627. https://doi.org/10.1257/0002828041464605

Keller, W. (2010). International trade, foreign direct investment, and technology spillovers. in Hall, B.H., and Rosenberg, N. (ed.) *Handbook of the Economics of Innovation Vol. 2*, 793-829. https://doi.org/10.1016/S0169-7218(10)02003-4

Merlevede, B., & Lenaerts, K. (2015). Firm size and spillover effects from foreign direct investment: the case of Romania. *Small Business Economics*, 45(3), 595-611.

Merlevede, B. Schoors, K., and Spatareanu, M. (2014). FDI Spillovers and Time since Foreign Entry. *World Development, 56*(C), 108-126. https://doi.org/10.1016/j.worlddev.2013.10.022

Moreno, F. and Coad, A. (2015). High-growth Firms: Stylized Facts and Conflicting Results. *Entrepreneurial growth: individual, firm and region (Advances in Entrepreneurship, Firm Emergence and Growth, Volume 17*), 187-230.

https://doi:10.1108/S1074-754020150000017016

OECD (2010). High-Growth Enterprises: What Governments Can Do to Make a Difference. *OECD Publishing, Paris*. http://dx.doi.org/10.1787/9789264048782-en

OECD (2016). High-growth enterprises rate. in *Entrepreneurship at a Glance 2016, OECD Publishing, Paris.* http://dx.doi.org/10.1787/entrepreneur_aag-2016-en

Reyes, J. D. (2017) FDI Spillovers and High-Growth Firms in Developing Countries. *World Bank Policy Research Working Paper* No. 8243. https://doi.org/10.1596/1813-9450-8243

Rodrik, D. (2005). Growth strategies. in Aghion, P., and Durlauf, S. N. (ed.) *Handbook of Economic Growth*, *1*, 967-1014. https://doi.org/10.3386/w10050

Semjén, A. Tóth, I. J., and Fazekas, M. (2009). Az egyszerűsített vállalkozói adó (eva) tapasztalatai vállalkozói interjúk alapján*.* in Semjén, A., and Tóth, I.J. (ed.) *Rejtett gazdaság. KTI Könyvek (11),*131-149. ISBN 978-963-9796-58-4

Tonin, M. (2011). Minimum wage and tax evasion: Theory and evidence. *Journal of Public Economics*, *95*(11-12), 1635-1651. https://doi.org/10.1016/j.jpubeco.2011.04.005

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economic Letters, 104*(3), 112-114. https://doi.org/10.1016/j.econlet.2009.04.026

World Bank (2018), *High-Growth Firms: Facts and Fiction of High Growth Entrepreneurship in Developing Countries*.

https://doi.org/10.1596/978-1-4648-1368-9

Zhang, Y., Li, H., Li, Y., and Zhou, L. A. (2010). FDI spillovers in an emerging market: the role of foreign firms' country origin diversity and domestic firms' absorptive capacity. *Strategic Management Journal*, *31*(9), 969-989.

https://doi.org/10.1002/smj.856

Appendix

Table A1 Share of firms with HGF status (3-year Birch employment growth definition)

Panel A: Birch definition

	1998	2001	2004	2007	2009	2011	Total
Low-tech mfg	6.2%	5.8%	4.4%	6.4%	6.8%	5.9%	5.9%
Medium-low tech mfg	7.2%	6.8%	6.9%	4.7%	9.5%	8.3%	7.1%
Medium-high tech mfg	8.0%	7.7%	9.4%	7.2%	12.6%	9.6%	9.1%
High-tech mfg	13.6%	12.2%	19.7%	19.4%	22.7%	12.6%	17.0%
KIS	3.2%	3.7%	4.2%	4.8%	4.7%	4.7%	4.4%
LKIS	4.7%	4.8%	3.5%	4.3%	3.2%	3.4%	3.9%
Construction	3.8%	3.7%	3.3%	3.6%	2.4%	3.4%	3.3%
Electricity, water etc	2.4%	5.1%	4.0%	9.6%	8.0%	8.4%	6.6%
Total	5.2%	5.1%	4.4%	4.8%	4.9%	4.6%	4.8%

Panel B: OECD definition

	1998	2001	2004	2007	2009	2011	Total
Low-tech mfg	8.5%	5.2%	5.1%	3.9%	5.5%	5.1%	5.5%
Medium-low tech mfg	8.8%	6.5%	6.6%	2.9%	7.0%	5.9%	6.1%
Medium-high tech mfg	8.6%	7.1%	6.5%	3.6%	4.9%	3.4%	5.7%
High-tech mfg	8.7%	8.1%	13.4%	10.1%	10.5%	4.2%	9.0%
KIS	5.5%	6.1%	7.4%	5.3%	6.2%	5.6%	6.0%
LKIS	9.7%	7.1%	5.9%	4.6%	4.5%	4.8%	5.8%
Construction	9.8%	6.7%	7.3%	4.1%	4.4%	6.5%	6.2%
Electricity, water etc	1.2%	5.4%	4.7%	4.7%	6.4%	7.0%	5.1%
Total	9.0%	6.7%	6.3%	4.3%	5.1%	5.2%	5.9%

Note: The table shows the share of HGFs in the regression sample. Panel A and B show the share of HGFs according to the 3-year definition, based on employment. KIS stands for Knowledge Intensive Services, and LKIS for Less Knowledge Intensive Services. Sector types are defined based on the Eurostat definition.

Table A2 Main results when using the OECD3 definition

Dependent:	Productivity		Size		Income av. wage	
	Labor prod.	TFP	Emp. (ln)	Sales (ln)	(ln)	ROA
Downstream HGF share	0.004	0.005	-0.000	0.003	0.005	0.003
	(0.987)	(0.903)	(-0.062)	(0.552)	(0.923)	(0.688)
Upstream HGF share	0.007	0.003	0.000	0.003	0.005	0.004
	(1.387)	(0.480)	(0.042)	(0.530)	(0.759)	(0.848)
Within-industry HGF share	0.016***	0.024***	0.002	0.010***	-0.002	0.014***
	(4.878)	(4.539)	(0.624)	(2.844)	(-0.388)	(4.414)
Demand	0.005	0.012**	0.009*	0.014***	0.012**	-0.002
	(1.222)	(2.054)	(1.920)	(2.847)	(2.548)	(-0.483)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	140,271	140,271	140,271	140,271	140,271	140,271
R-squared	0.214	0.151	0.085	0.069	0.161	0.095

Note: The table replicates Table 2 but relying on the 3-year OECD instead of the Birch definition. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. Firms are classified as HGFs based on the 3-year OECD employment definition.

Table A3 Main results when using unweighted spillover measures

Dependent:	Productivity		Size		Income av. wage	
	Labor prod.	TFP	Emp. (ln)	Sales (ln)	(ln)	ROA
Downstream HGF share	0.038***	0.033***	-0.002	0.021***	0.025***	0.018***
	(9.905)	(6.451)	(-0.395)	(4.769)	(5.778)	(4.315)
Upstream HGF share	0.015***	0.010*	0.002	0.012***	0.008	0.004
	(3.969)	(1.764)	(0.629)	(2.669)	(1.568)	(0.951)
Within-industry HGF share	0.002	0.008**	0.013***	0.004	-0.010***	0.008**
	(0.779)	(2.154)	(3.999)	(1.036)	(-3.112)	(2.442)
Demand	0.001	0.010	0.009**	0.012**	0.009*	-0.003
	(0.225)	(1.607)	(1.999)	(2.264)	(1.900)	(-0.753)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	140,271	140,271	140,271	140,271	140,271	140,271
R-squared	0.216	0.152	0.085	0.070	0.162	0.095

Note: This table replicates Table 2, but spillover measures are not weighted by employment but by the number of HGFs out of all firms in a given sector-county pair. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. Firms are classified as HGFs based on the 3-year Birch employment definition.

Table A4 Cross-sectional identification

Dependent:	Labor productivity growth between t and t+3					
Year	1998	2001	2004	2007	2009	2011
Downstream HGF share	0.029***	0.008	0.034***	0.014	0.076***	0.027***
	(3.194)	(0.926)	(3.171)	(1.520)	(5.616)	(3.116)
Upstream HGF share	0.033***	0.006	-0.016	-0.002	-0.008	-0.010
	(2.811)	(0.581)	(-1.329)	(-0.196)	(-0.489)	(-0.819)
Within-industry HGF share	0.021***	0.009	0.017*	0.021***	-0.001	0.022**
	(2.588)	(1.049)	(1.826)	(2.962)	(-0.112)	(2.223)
Demand	0.007	0.035***	0.007	-0.000	-0.012	-0.009
	(0.685)	(3.889)	(0.938)	(-0.028)	(-0.929)	(-0.801)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	16,288	20,376	26,327	26,595	24,910	25,775
R-squared	0.239	0.185	0.250	0.188	0.159	0.160
Dependent:	TFP growth between t and t+3					
Year	1998	2001	2004	2007	2009	2011
Downstream HGF share	0.028*	-0.041***	0.034***	0.048***	0.029**	0.031***
	(1.862)	(-3.219)	(2.751)	(3.558)	(2.567)	(3.011)
Upstream HGF share	0.033	0.004	-0.020	-0.013	0.001	-0.010
	(1.560)	(0.209)	(-1.170)	(-0.748)	(0.043)	(-0.754)
Within-industry HGF share	0.025	0.016	0.017	0.033**	0.005	0.018
	(1.496)	(1.327)	(1.323)	(1.969)	(0.405)	(1.536)
Demand	0.009	0.022*	0.023***	0.010	-0.009	-0.004
	(0.439)	(1.775)	(2.870)	(0.707)	(-0.624)	(-0.325)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	16,288	20,376	26,327	26,595	24,910	25,775
R-squared	0.139	0.134	0.203	0.127	0.129	0.140

Note: This table re-runs the main regressions on the subsamples from different periods. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. The firm controls are log employment, labor productivity and log capital per labor ratio.

Table A5 Different sets of fixed effects

Dependent:	Labor productivity			TFP		
	Growth between t and t+3					
Downstream HGF share	0.034***	0.051***	0.010	0.023***	0.038***	0.107***
	(7.460)	(9.183)	(1.036)	(4.244)	(6.071)	(7.508)
Upstream HGF share	-0.000	0.039**	-0.005	-0.003	0.006	-0.009
	(-0.035)	(2.423)	(-0.353)	(-0.271)	(0.340)	(-0.475)
Within-industry HGF share	0.015***	0.011**	0.006	0.019***	0.018***	0.017***
	(3.934)	(2.415)	(1.281)	(3.425)	(2.579)	(2.675)
Demand	0.006	0.016**	-0.009**	0.020**	0.040***	0.006
	(1.211)	(2.069)	(-2.236)	(2.552)	(4.154)	(1.066)
Controls	yes	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes	yes
microregion-year FE	yes			yes		
County-sector-year FE		yes			yes	
Firm FE			yes			yes
Firm controls	yes	yes	yes	yes	yes	yes
Observations	140,271	140,242	121,292	140,271	140,242	121,292
R-squared	0.224	0.224	0.595	0.160	0.163	0.539

Note: The table extends the main regression in Table 2 with different sets of fixed effects: In columns (1) and (4) it more detailed geographical fixed effects (column 1 and 4); sector-year fixed effects (column 2 and 5); firm fixed effects (column 3 and 6). The explanatory variables show the share of HGFs in different industries at the same microregion and county, by using the 2-digit input-out matrix from OECD STAN. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. The firm controls are log employment, labor productivity and log capital per labor ratio. The definition of HGF is the 3-year Birch employment definition. Small firms are firms with less than 5 employees.

Table A6 Controlling for other industry characteristics

		Labor productivity change between t and t+3			
Control:	Sales growth	Foreign share	Exporter share	Innovative firm share	Average labor prod
Downstream HGF share	0.033***	0.011**	0.006	0.010*	0.007
	(7.408)	(2.084)	(1.049)	(1.825)	(1.612)
Upstream HGF share	-0.001	0.003	0.004	0.000	0.000
	(-0.135)	(0.599)	(0.868)	(0.060)	(0.006)
Within-industry HGF share	0.015***	0.011***	0.012***	0.008**	0.009***
	(3.753)	(2.639)	(3.051)	(2.081)	(2.645)
Demand	0.009*	0.003	0.002	-0.001	-0.002
	(1.852)	(0.615)	(0.503)	(-0.328)	(-0.540)
Control Downstream		0.038***	0.036***	0.056***	0.083***
		(4.954)	(4.848)	(8.430)	(16.512)
Control upstream	-0.011**	-0.005	-0.005	-0.003	0.038***
	(-2.480)	(-0.656)	(-0.653)	(-0.532)	(6.122)
Control same sector	0.006*	0.018***	0.028***	0.033***	-0.006
	(1.671)	(3.391)	(4.399)	(8.996)	(-1.104)
Controls	yes	yes	yes	yes	yes
region-year FE	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes
Observations	140,271	136,771	136,771	140,271	140,271
R-squared	0.215	0.216	0.216	0.216	0.219

Panel B: TFP change between t and t+3

			Exporter	Innovative	Average labor
Added control variable:	Sales growth	Foreign share	share	firm share	prod
Downstream HGF share	0.023***	-0.002	-0.003	0.001	0.003
	(4.376)	(-0.328)	(-0.410)	(0.188)	(0.604)
Upstream HGF share	-0.001	0.005	0.007	-0.001	-0.001
	(-0.186)	(0.647)	(1.076)	(-0.112)	(-0.144)
Within-industry HGF share	0.019***	0.014**	0.014***	0.012**	0.008**
	(3.371)	(2.493)	(2.929)	(2.104)	(2.228)
Demand	0.016**	0.012**	0.012**	0.005	0.002
	(2.123)	(2.025)	(2.040)	(0.963)	(0.488)
Control downstream		0.048***	0.031***	0.057***	0.167***
		(4.765)	(3.568)	(7.108)	(35.277)
Control upstream	-0.009	-0.022**	-0.022**	-0.001	0.019***
	(-1.239)	(-2.352)	(-2.506)	(-0.207)	(2.950)
Control same sector	0.008*	0.028***	0.071***	0.047***	-0.002
	(1.781)	(3.957)	(8.784)	(11.561)	(-0.405)
Controls	yes	yes	yes	yes	yes

region-year FE	yes	yes	yes	yes	yes
sector-year FE	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes
Observations	140,271	136,771	136,771	140,271	140,271
R-squared	0.152	0.153	0.154	0.153	0.164

Note: The table extends the main regressions in Table 2 with controlling for different industry-level characteristics. The explanatory variables show the share of HGFs in different industries at the same microregion and county, by using the 2-digit input-out matrix from OECD STAN. The dependent variables are changes between year t and t+3, the explanatory variables are measured at time t. Beta coefficients are reported and t-values are reported in parentheses. Standard errors are clustered at the county-industry (2-digit NACE)-year level. The firm controls are log employment, labor productivity and log capital per labor ratio. Firms are classified as HGFs based on the 3-year Birch employment definition. In each column a different additional control variable is considered, and its name is reported at the top of the column. Small firms are firms with less than 5 employees