
The Oxford Martin Working Paper Series on
Technological and Economic Change



**Demand-pull and technology-push: What drives the direction of technological change?
An empirical network-based approach**

Kerstin Hötte
Working Paper No. 2021-4



Disclaimer: This is a working paper and represents research in progress. This paper represents the opinions of the authors and does not represent the position of the Oxford Martin School or other institutions or individuals.

For more information on the Oxford Martin Programme on Technological and Economic Change, please visit: <https://www.oxfordmartin.ox.ac.uk/technological-economic-change/>

For more information on the Oxford Martin Programme on the Future of Work, please visit: <https://www.oxfordmartin.ox.ac.uk/future-of-work/>

Demand-pull and technology-push: What drives the direction of technological change?

An empirical network-based approach

Kerstin Hötte*

*Oxford Martin School, University of Oxford
Institute for New Economic Thinking, University of Oxford
Faculty of Business Administration and Economics, University of Bielefeld*

December 17, 2021

Demand-pull and technology-push are drivers of technological change and policy-makers need to understand how both interact and differ by impact.

I introduce two concepts of demand-pull and technology-push measured by a two-layer network of input-output (market) and patent citation (innovation) links between 307 NAICS 6-digit US manufacturing industries in 1977-2012: (1) Demand-pull and technology-push are cross-layer spillovers when demand shocks in the market pull innovation and innovation pushes growth in the market. (2) Demand-pull may also arise from downstream links within the same layer when output users trigger upstream growth. Push effects, in contrast, spill over from up- to downstream industries.

The results show that innovation is a driver of market growth which is factor-biased in favor of capital against labor. I also find support for demand-pull within the market: industries with a strong customer network grow faster, invest more, and grow by labor productivity. Upstream centrality exhibits the opposite effect which may indicate input constraints in the market. Innovation evolves differently: knowledge spillovers give rise to increasing returns as driver of concentration and clustering.

Despite limitations related to data and classifications, the results enable a differentiated view on the drivers of technological change and its consequences which is essential for policy to shape the technological evolution.

*Mail: kerstin.hotte@oxfordmartin.ox.ac.uk

Contents

1 Introduction	1
2 Background and literature	5
3 Conceptual framework: Technological and economic change in a two-layered network	7
3.1 The economy as a two-layer network	7
3.2 Technological similarity, spillovers and centrality	9
3.2.1 Technological similarity	9
3.2.2 Spillovers	9
3.2.3 Centrality	10
3.3 Technology-push and demand-pull	12
3.4 Technological change	12
4 Data	13
5 Results	14
5.1 Descriptive analysis	14
5.2 Demand-pull, technology-push, and technological change	20
6 Discussion	26
6.1 Demand-pull or technology-push?	28
6.2 Limitations and research implications	29
7 Conclusions	31
A Methods	40
A.1 Cosine similarity	40
A.2 Data compilation	40
A.2.1 Input-output data	40
A.2.2 Patent data	41
A.2.3 Supplementary data and processing	42
A.3 Regression data issues	45
A.4 Network plots	46
B Additional descriptive information	46
C Additional regression results	49
C.1 Additional results	49
C.1.1 Relationships between growth rates	49
C.2 Robustness checks	51
C.2.1 Degree-based regressions	51
C.2.2 Strength-based regressions	54

SI.1 Detailed information on data processing	57
SI.1.1 Input-output data	57
SI.1.1.1 Overview	57
SI.1.1.2 Processing steps in detail	58
SI.1.1.3 Technical and conceptual issues	59
SI.1.2 Patent data	61
SI.1.2.1 Processing steps in detail	61
SI.2 Supplementary results	62
SI.2.1 Descriptives	62
SI.2.1.1 Network plots	62
SI.2.1.2 Industry rankings	65
SI.2.2 Regressions	67
SI.2.2.1 Relationships between growth rates	67
SI.2.2.2 Robustness check using non-citation weighted patents	69

1. Introduction

Shaping the direction of technological change is high on the political agenda to cope with the challenges of the 21st century such as climate change or digitization (IPCC, 2018; Brynjolfsson and McAfee, 2012). Demand-pull and technology-push are drivers of technological change (Schmookler, 1966; Myers and Marquis, 1969; Mowery and Rosenberg, 1979; Von Hippel, 1976; Di Stefano et al., 2012) that can be stimulated by different policies (Rosenberg, 1982; Nemet, 2009). Effective policy-making requires an understanding of how these mechanisms interact and how they differ by their impact on the economy. Here, I disentangle both mechanisms and study their impact on various indicators of technological change in US manufacturing.

Technology-push arises when technological and scientific advances enable the development and commercialization of novelties that diffuse if they outperform incumbent solutions. Demand-pull emerges from market needs if customers ask for improvement or face technical limitations while using existing technologies. When inventors recognize this, they may adapt R&D efforts in response to the perceived market potential (Di Stefano et al., 2012; Von Hippel, 1976). Technology-push and demand-pull are interdependent because R&D objectives can be demand-selected and market-needs may arise in response to innovation (Kline and Rosenberg, 1986; Cohen, 2010; Nemet, 2009; Mowery and Rosenberg, 1979) which is one mechanism behind the co-evolution of markets, technology, and industrial structures (Nelson, 1994; Saviotti and Pyka, 2013).

Previous studies of demand-pull and technology-push mostly relied on market shares and sizes as proxies for demand-pull and innovation outputs such as patents for technology-push to study their impact on growth, innovation, and productivity (e.g. Jaffe, 1988; Cohen, 2010).

Here, I used similar measures to study demand-pull and technology-push in a two-layer network that captures market and innovation interactions. Industries are connected in the first layer, called “market layer”, by cross-industrial flows of intermediate goods, and in the second layer, called “innovation layer”, by patent citations. I studied two conceptually different types of technology-push and demand-pull:

Type 1: Technology-push effects are encoded in the dynamics of the innovation layer and demand-pull effects arise from shocks in the market. It is tested how these effects spill over across layers and how they interact with a series of indicators of qualitative and quantitative technological change.

Type 2: Demand-pull and technology-push are associated with the supply and demand side of an industry, i.e. up- and downstream links in the network. Demand-

pull effects are present when the connections to downstream customers in the market and users of innovations induce technological change in the upstream industry. Technology-push effects are the opposite when changes in the availability of inputs induce downstream technological change.

These concepts are operationalized using a panel of 307 NAICS 6-digit US manufacturing industries during 1977-2012 available in quinquennial time steps. The industry level input-output network is inferred from accounting data provided by the Bureau of Economic Analysis (BEA). The patent citation network is compiled on the basis of patents granted by the US Patent and Trademark Office (USPTO) where I mapped patents via their technology classes to NAICS codes (Goldschlag et al., 2020). The data are supplemented with the productivity database from the National Bureau of Economic Research (NBER) and US Census Bureau’s Center for Economic Studies (CES) (Bartlesman and Gray, 1996; Becker et al., 2013).¹

Technological change is measured by changing industry sizes in both layers (output of goods and patents), productivity (labor productivity, TFP), and the evolution of factor input requirements (labor demand, capital and investment intensity). These indicators reflect rises and declines of industries by innovation performance and dominance in the market. The productivity indicators reflect the production efficiency (OECD, 2001), and the factor use indicators inform about the bias of technological change and its consequence for employment (Acemoglu, 2002). These measures are studied descriptively and regressions are used to identify the impact of demand-pull and technology-push on innovation, markets, and technology.²

The results show that both push and pull are drivers of technological change, but their impact varies across types. First, technology-push as a spillover from innovation to markets (Type 1) is a driver of growth in the market, but it is factor-biased in favor of capital. Innovation captured by the centrality in the patent network is positively related to market growth, and exhibits reduction in the demand for labor and an increase in the intensity of capital use. The factor bias is moderated when industries benefit knowledge spillovers which are positively related to labor demand, and are also a driver of innovation. Demand-pull as spillover from markets to innovation (Type 1) does not show any significant impact.

¹The data used in this paper are available on request. An outdated version used for an earlier draft of this paper is available at (Hötte, 2021; Hötte, 2021). An updated version will become published soon.

²To clarify the terminology, throughout this paper, “innovation” is proxied by citation-weighted patent counts, “markets” are reflected in the output of an industry, and “technology” is described by a bundle of indicators such as labor productivity, TFP, and the composition of factor input use.

Second, when push and pull are conceptualized as within-layer dynamics (Type 2), the results offer significant support for demand-pull from customer links in the market. A higher downstream centrality in the market shows a positive relation to market growth, labor productivity, and investments, but it does not show any impact on innovation. Downstream centrality measures the importance of an industry as input provider to other industries. Upstream centrality in the market, which can be interpreted as indicator of input availability, shows the opposite relationship, i.e. lower market growth, productivity and investments. But there is a weak indication that it may be an incentive for innovation.³

The dynamics in the patent network are different: up- and downstream connections are highly correlated and both qualitatively show the same effect. While centrality in the innovation layer does not show an impact on innovation, the results offer support for the presence of knowledge spillovers from innovations in technologically related industries. This supports the idea of technology-push as driving force within the innovation layer.

Both layers are subject to path-dependence, but they evolve differently: While the innovation layer becomes increasingly connected and clustered, and the size distribution becomes increasingly skewed, the evolution of the market is rather fluctuating without any clear trend. The regression results offer one candidate explanation: Innovation benefits from spillovers from related industries which may be one source of increasing returns which drive clustering and concentration, but growth in the market is constrained by the availability of production inputs. This reflects the conceptual difference between tangible goods and intangible knowledge: while goods traded in the market are exhaustible, knowledge is non-rival and does not diminish when it is increasingly used.

Three major limitations exist. First, patent data as an indicator of innovation suffer from a series of well-known limitations (e.g. Jaffe and De Rassenfosse, 2019; Fontana et al., 2013; Kogan et al., 2017). Second, classification systems change over time which hampers the study of long-term technology-industry links (Lafond and Kim, 2019; Yuskavage et al., 2007). Finally, drivers of technological change differ across

³Downstream centrality is an indicator of supply-side market power: It measures that an industry supplies goods that are essential inputs for many other industries (who themselves are important suppliers to others). Upstream centrality is an indicator of specialization in upstream industries on few customers. The results are robust across different up- and downstream centrality measures that serve as indicators for the importance of an industry as input provider. In the patent layer, the distinction between up- and downstream centrality is not possible due to the high correlation of up- and downstream links, but here the PageRank shows a high correlation (> 80%) with the innovation output measured by the stock of citation-weighted patents.

firm, industries, time periods, and industrial maturity (Walsh, 1984; Pavitt, 1984). This research is limited to US manufacturing. Some of these limitations are addressed methodologically, others can not be resolved within this paper but offer promising directions of future research.

Nevertheless, this study offers a series of relevant contributions and insights for research and policy. To my knowledge, the two-layer network approach is a methodologically and conceptually novel approach which allows to capture interactions among patented inventions and goods traded in the market simultaneously. The framework and the data that accompany this study offer a rich basis for future research that aims to understand how innovation shapes markets and vice versa (see also Sec. 6). This understanding is essential as the societal challenges of today require an understanding of how innovation can be used to influence the evolution of markets for goods and labor, and how market forces can be mobilized to shape technological change.

The results show that the impact of technological change may be dependent on its driver which is important for policy design. For example, the results suggest that the debate regarding whether or not capital replaces or complements human labor should be made very carefully: technological change pushed by innovation shows a bias in favor of capital, while demand-pull effects show the opposite relationship. Further, knowledge spillovers across industries do not only foster innovation, but also increase the demand for labor. This is informative for innovation policy design when making the choice between demand-pull or technology-push instruments, and measures that facilitate knowledge spillovers. This study does not provide causal evidence, but reveals a series of - so far - uncovered correlations that provide a rich basis for future investigations.

The results show that market and innovation dynamics differ. While markets are subject to resource constraints with a dampening effect on clustering and concentration, increasing returns from non-rival intangible knowledge gives rise to the emergence of technology clusters. Increasing returns to innovation are engines of growth (Romer, 1990) but can be also a reason for technological lock-in effects (Arthur, 1989). The analysis offers weak indication that changing market conditions may lead to a redirection of R&D activities.

The remainder of the paper is structured as follows: The next section provides an overview of the related literature. The theoretical framework is explained in Sec. 3. dSec. 4 introduces the data. Sec. 5 summarizes the results, Sec. 6 offers a discussion, and 7 concludes.

2. Background and literature

Demand-pull and technology-push as drivers of technological and economic change are the subject of a long-lived debate (Schmookler, 1966; Mowery and Rosenberg, 1979; Pakes and Schankerman, 1984; Cohen, 2010; Saviotti and Pyka, 2013). Demand-pull suggests that R&D activities follow the market: the perceived commercial potential of innovations offers an incentive for targeted R&D. Technology-push effects arise from technological opportunities that enable the development and commercialization of new products and processes, but also further innovation. The two theories differ by the assumptions made about the incentives that influence the decision where to allocate R&D and production efforts, and also by their assumption about the sources of ideas for technological improvements. While demand-pull emphasizes the role of users and customers, technology-push builds on external and internal research (Cohen and Levinthal, 1989; Kline and Rosenberg, 1986; Di Stefano et al., 2012). A bibliometric study by Di Stefano et al., 2012 showed that more recent innovation studies abandon the traditional juxtaposition of demand-pull and technology-push, and increasingly focus technological capabilities to respond to push and pull effects.

Previous studies found support for both effects (see Cohen, 2010; Di Stefano et al., 2012, for an overview). For example, using aggregate time series over the business cycle, Geroski and Walters (1995) studied interactions between manufacturing outputs and innovative activity. They observed causal effects from outputs to innovation, but no support for the other way round. Though they also highlighted the critical role of stochastic determinants which they interpret as supply shocks that point to technology-push from radical innovation. This is conceptually in line with Walsh (1984) whose study of the chemical industry showed that technology-push from radical breakthroughs drives growth in the market which creates demand-pull effects that induce incremental innovation.

The analysis in this paper is conceptually close to the seminal work by Jaffe (1988) who operationalized demand-pull through the market shares of a firm in different industries and technology-push as effect that arises from innovation outputs in technologically related fields. Jaffe's (1988) work relies on aggregate industry and technology classifications, and the author found that pull and push effects can not be empirically distinguished when explaining TFP growth.

This paper builds on an empirical two-layer network approach composed of an input-output (IO) layer (market) and a patent citation layer (innovation). The two layers are used to study how innovation and markets co-evolve (cf. Nelson, 1994). A

co-evolution occurs when dynamics in the market is correlated with innovation (see Saviotti and Pyka, 2013).

In input-output (IO) and patent citation networks, technology can be qualitatively described by the network position of an industry, firm, or patent. The position is determined by the bundle of input links pointing to the types of physical production inputs or pre-existing technological knowledge encoded in patent citations that enables subsequent innovation. Two firms, industries, or patents are said to be technologically similar if they have many overlapping in- and output links which indicates the capability to make use of similar inputs and to serve the needs of similar users (e.g. Carvalho and Voigtländer, 2014; Antony and Grebel, 2012; Acemoglu et al., 2016; Cai et al., 2017; Huang, 2018; Atalay et al., 2011). Similarity is also a measure for the absorptive capacity which is the capability to absorb external technology (Jaffe and De Rassenfosse, 2019; Cohen and Levinthal, 2000; Cohen and Levinthal, 1989).

Using patent citation similarity as a measure for absorptive capacity, Antony and Grebel (2012) showed that knowledge spillovers can explain firms' productivity growth. Kay et al. (2014) used a patent-overlay mapping to illustrate evolving similarities as a manifestation of technological change. Interpreting technological progress as expansion of technologies across fields, Acemoglu et al. (2016) used a citation network of US patents to show that upstream innovation creates positive spillovers on downstream inventions. Huang (2018) found that firms with many patents innovate more often and are more likely to innovate in technology fields that are similar to their pre-existing technology. Taalbi (2020) studied the evolution of the network structure span by supply and use of significant innovations in Swedish industries. He found that path-dependence embodied in technological proximity shapes the evolution of the network. At the region level, Buerger et al. (2012) documented a positive relationship between innovation outputs and subsequent employment growth and R&D investments, but they also highlighted heterogeneity across industries.

In IO networks, the characteristics of an industry's production technology are reflected in the bundle of input used and outputs produced. Using an empirical IO network, Carvalho and Voigtländer (2014) observed that industries in the US tend to adopt new inputs if they are similar to their pre-existing portfolio of inputs. Boehm et al. (2019) studied product line development of multi-product firms reflected in the evolution of output links. Firms enter those product markets where they have core capabilities measured by firms' IO relationships. Carvalho (2014) showed that relatedness through IO links can be a moderating factor of output fluctuations.

Here, I combine both types of networks. The mapping from patents to industries

has proven challenging in the past, but a series of available concordances exist. They mostly rely on the industrial classification of the firms that own patents in specific technological fields (e.g. Kortum and Putnam, 1997; Schmoch et al., 2003; Dorner and Harhoff, 2018; Van Looy et al., 2014; Lybbert and Zolas, 2014; Goldschlag et al., 2020). The concordances make it possible to study interactions between the evolution of patented technology and industries. Proving the economic validity of their concordance, Goldschlag et al. (2020) showed that industry-technology relationships are relatively stable and changes in the technological composition of industries correlate with occupational change.

This study is not the first that simultaneously considers the market and innovation network positions. The two studies that are methodologically and conceptually closest are Jaffe (1988) introduced above and Bloom et al. (2013).

Bloom et al. (2013) build on a similar framework but study the role of market rivalry and knowledge spillovers on firms' performance. They measured market rivalry by competitors' cumulative R&D efforts weighted by the closeness of firms in the product space. Knowledge spillovers were calculated as cumulative R&D efforts of firms that are technologically similar. They found that market rivalry exhibits a negative effect, while firms benefit from positive knowledge spillovers.

3. Conceptual framework: Technological and economic change in a two-layered network

Technology is the capability of firms and industries to transform a bundle of inputs into outputs. Technological change occurs when the quality and/or quantity of inputs or outputs change (Saviotti, 1997; Ruttan, 1959). This analysis investigates to what extent this process of change is driven by technology-push and demand-pull dynamics where technological change is observed in changes in the size ranking of industries, changes in the amount of outputs of goods and patents, factor productivity, and factor inputs.

3.1. The economy as a two-layer network

Qualitative information about the technology used in an industry $i \in N$ is revealed by its IO connections in the market and patent citation patterns with N as set of industries. Industries use intermediate goods as input to production and build on

pre-existing knowledge proxied by patents to innovate. Further, industries use capital and labor as factor inputs.

For the patent citation data, it needs to be noted that citations do not represent a direct flow of knowledge when the inventor of a new patent was inspired by a pre-existing invention encoded in the cited patent. Citations in USPTO patents are a legal requirement to describe prior art and to limit the scope of the claims of the new patent. However, a patent citation can be still seen as an indicator of technological relatedness that reveals that the knowledge encoded in the cited patent contributed to the technological basis onto which a patent builds (Jaffe and De Rassenfosse, 2019; OECD, 2009).

The IO and patent citation relationships span a weighted, directed two-layered network. A node in the network represents an industry i which is connected with other industries $j \in N$ through patent citation links in the innovation layer τ and IO links in the market layer μ . The layers $\alpha = \tau, \mu$ are linked as a duplex network where each industry has a representation in each layer.

The links in the layers are given by the flow of goods $flow_{ij,t}^\mu$ and patent citations $flow_{ij,t}^\tau$ from an industry j to i with $i, j \in N$ in time t . These flows reveal two types of information about the technology used by i and j : an input flow from j to i indicates that i has the capability to use the outputs produced by j . It also reveals information about j 's capabilities as this industry is able to produce outputs that are valuable for i . Hence, the bundle of upstream (input) links and downstream (output) links reveals qualitative information about the technology that is used in these industries.

For the analyses in this paper, the flows of goods and citations are transformed into input shares $w_{ij,t}^{\alpha,in}$ dividing the monetary flow (patent citation count) $flow_{ij,t}^\alpha$ by the sum of input flows $\sum_j flow_{ij,t}^\alpha$. Analogously, output shares $w_{ij,t}^{\alpha,out}$ are obtained by dividing $flow_{ij,t}^\alpha$ by the sum of all outputs produced by industry i , i.e. $\sum_k flow_{ki,t}^\alpha$. Note that $w_{ij,t}^{\alpha,in} \neq w_{ji,t}^{\alpha,out}$ due to the different weighting. They also reflect different concepts: the input share $w_{ij,t}^{\alpha,in}$ reflect j 's relevance as input provider for i while $w_{ij,t}^{\alpha,out}$ reflect j 's relevance as a customer or knowledge user of i . The normalization to shares improves the comparability of different data types (monetary flows, patents) and of industries that are very heterogeneous by size.

Each network layer can be represented as a quadratic, asymmetric $|N| \times |N|$ matrix $W_t^{\alpha,d} = \{w_{ij,t}^{\alpha,d}\}_{i,j \in N}$ with positive non-zero entries if a link from i to j exists in time t . The superscript $d = in, out$ indicates the direction of the links, i.e. $w_{ij,t}^{\alpha,in}$ ($w_{ij,t}^{\alpha,out}$) indicates an input (output) link. The two-layer network is given by the set of both matrices: one representing the input-output network, called "market layer", and the

other representing the patent citation network, called “innovation layer”.

3.2. Technological similarity, spillovers and centrality

The network data are used to derive a number of network-based indicators that describe the characteristics of an industry.

3.2.1. Technological similarity

Industries can be technologically similar in different ways: (1) Two industries i and j are similar if they have similar capabilities to use inputs, i.e. if i and j rely on similar physical production inputs and cite similar patents. (2) Industries can be also similar by outputs if they are capable to serve the needs of similar downstream customers or knowledge users.

The technological similarity is measured by the cosine similarity $\sigma_{ij,t}^{\alpha,d}$ which measures the angle between the input or output vectors of a pair of industries. It is calculated separately for each layer α and separately for the up- and downstream network $d = in, out$ (see for more detail A.1). The cosine similarity is a commonly used measure in network analysis and classification methods (Kay et al., 2014; McCune et al., 2002; Leydesdorff, 2005; Mikolov et al., 2013).

3.2.2. Spillovers

Industries are interconnected in the networks and shocks in one industry may spill over to industries that are sufficiently close by their up- and downstream linkages (see e.g. Carvalho and Voigtländer, 2014; Acemoglu et al., 2016; Bloom et al., 2013).⁴

Here, closeness in the network is measured $\sigma_{ij,t}^{\alpha,d}$ which moderates the extent to which changes in $A_{j,t}^{\alpha}$ in up- and downstream industries j spill over to i . These spillovers are calculated as

$$Spill(A)_{i,t}^{\alpha,d} = \sum_{j \neq i}^N \sigma_{ij,t}^{\alpha,d} \cdot A_{j,t}^{\alpha} \quad (1)$$

with $\alpha = \mu, \tau$, $d = in, out$. $A_{j,t}^{\tau}$ is the amount of patents and $A_{j,t}^{\mu}$ is the amount of goods produced by j in t .

⁴Bloom et al. (2013) also offer a discussion of the theoretical foundations that underpin the spillover metrics.

Spillovers reveal different types of information dependent on the layer and whether they emerge from up- or downstream linkages. The level of $Spill(A)_{i,t}^{\alpha,d}$ changes either by an output shock in related industries $A_{j,t}^{\alpha}$ or by a change in the similarity $\sigma_{i,j,t}^{\alpha,d}$.

Upstream similarity in the market $\sigma_{i,j,t}^{\mu,in}$ indicates that a pair of industries relies on a similar bundle of intermediate goods as production inputs. An increase in $Spill(A)_{i,t}^{\mu,in}$ indicates a rise in the competition for these inputs as either, competitors that use the same inputs grew ($A_{j,t}^{\mu} \uparrow$) or the extent to which input requirements overlap increased ($\sigma_{i,j,t}^{\mu,in} \uparrow$).

Downstream similarity in the market $\sigma_{i,j,t}^{\mu,out}$ measures the overlap of i 's and j 's customer links which indicates that the outputs of i and j serve similar customer needs. This can be also an indicator for competition if the outputs produced by i and j are substitutes, but it may also indicate demand synergies if the outputs are complements.

Spillovers in the innovation layer are conceptually different. Knowledge created in R&D processes has public good characteristics and it is difficult to exclude others from using it once a new discovery is made. Hence, knowledge produced by j may spill over to i if i is technologically sufficiently similar to make use of j 's knowledge (Cohen and Levinthal, 2000). Again, up- and downstream similarity have slightly different interpretations. Upstream similarity $\sigma_{i,j,t}^{\tau,in}$ indicates that both, i and j rely on similar knowledge inputs to innovation. An increase in upstream spillovers may indicate an increase in the availability of knowledge resources for both. Downstream similarity in contrast, indicates that the patents of both are useful for downstream innovation.

3.2.3. Centrality

An industry is central in the network if it is well connected with other (important) industries. Network centrality is an indicator for the relevance of an industry as input provider or output user (cf. Jackson, 2008; Carvalho, 2014).

An industry can be central in two ways: (1) Downstream centrality indicates that an industry is the supplier of goods or patents that are used by a great number of other industries and/ or that account for a large share in the input bundle of other industries. (2) Upstream centrality indicates that an industry is a critical customer of other industries. Hence, it is a customer of many other industries and/or accounts for a major share in other industries' output vector. If the upstream centrality mainly comes from the weights of the output links of upstream industries, it can be also an indicator of vertical specialization if industries exclusively produce to serve the needs of i .

The former is associated with supply-side, the latter with demand-side market power.

Different approaches exist to measure centrality. The simplest measure is given by the *degree* which simply counts the number of different customers (downstream) or input sources (upstream). An industry is more central by the degree if it has more customer (downstream) or suppliers (upstream). The next complex measure is the *strength* which is a weighted count of in- and output links, i.e. a downstream link receives a higher weight if the input flowing from i to j accounts for a high share in j 's input mix. The analogue holds for the upstream strength. An industry i is more central by the strength if it has many customers that receive a high share of their inputs from i .

I used the PageRank $PR_{i,t}^{\alpha,d}$ as centrality measure which additionally accounts for the centrality of the up- and downstream connected industries. It is an algorithm that assigns ranks to industries by the number and quality of links, where the quality of a link is higher when it comes from an industry that is itself ranked high by the PageRank. The PageRank is calculated through a recursive algorithm. It also takes account of the weighted and directed nature of the links (Csardi, Nepusz, et al., 2006). Originally, the PageRank was developed by Brin and Page (1998) and used in the first versions of the Google search engine to rank websites by their relevance to the users based in-going links that are weighted by the relevance of the websites from which the links are coming.

I used the PageRank as centrality measure because it can be computed on the basis of up- and downstream links and shows a sufficiently high variation. Additional results using the degree and strength as indicators are provided in C.2. The results prove fairly consistent and correlation statistics show a high correlation between different measures of network centrality (see A.3).

It holds for all centrality measures that an increase in the downstream centrality in the market (innovation layer) indicates that an industry i became more important as provider of goods (source of knowledge) in the network. This is associated with demand-pull of Type 2 in industry i because it reflects an increase in the reliance of downstream industries on i 's outputs. In the market, this reflects an improvement of i 's competitiveness as many other industries depend on i 's goods produced by i . In the innovation layer, this reflects the importance of the patents of i .

In contrast, an increase in the upstream centrality indicates that an industry is a critical customer. Many industries sell a high share of their outputs to i . If the centrality increases, the upstream market power of i increases which does not necessarily mean that the diversity of input sources increases. For the strength and the PageRank, an increase in the centrality may also reflect a pattern of vertical specialization when

upstream industries become more specialized to supply goods to i .

3.3. Technology-push and demand-pull

Ex-ante, there is no canonical way to operationalize push and pull in empirical data. Here, I study two different types as already outlined in the introduction (Sec. 1):

Type 1: Technology-push is a spillover from patented innovations to market dynamics, and demand-pull is the opposite: positive shocks in the market that trigger patented innovation. In the analyses below (Sec. 5), I interpret all cross-layer effects of innovation on the market as technology-push, and demand-pull are all effects in the opposite direction.

Type 2: Technology-push and demand-pull are interactions between up- and downstream industries within the same network layer: Technology-push is associated with upstream dynamics when changes in upstream industries induce technological change in downstream industries. Demand-pull is the opposite: Positive shocks in downstream industries indicate an increasing use of the outputs produced by i (goods and knowledge). This may induce technological change in the upstream layer. In the analyses below (Sec. 5), technology-push is reflected in the contributions of upstream industries to downstream technological change, and demand-pull is reflected in the contribution of shocks that occur downstream.

3.4. Technological change

Technological change is a change in the production function that maps a set of inputs to outputs (Ruttan, 1959). At the industry level, this is reflected in a changing composition of in- and outputs in the production of goods and in R&D processes which leads to changes in industry sizes: Some industries grow, others shrink in relative and absolute terms. Technological change is also reflected in changing use of production factors which can be measured by productivity indicators, labor requirements, capital and investment intensity. Here, technological change is analyzed in the following ways:

1. I describe the evolution of the industry size ranking by output of goods $A_{i,t}^\mu$ and patents $A_{i,t}^\tau$ illustrating the rise and decline of industries. Additional descriptive statistics are used to study how the characteristics of the networks evolved over time.
2. Next, regression analyses are used to disentangle the drivers of technological change as reflected in industry growth $gr(A_{i,t}^\mu)$ and $gr(A_{i,t}^\tau)$, a series of pro-

ductivity indicators (value added per employee $(VA/L)_{i,t}$ as measure of labor productivity, total factor productivity $(TFP_{i,t})$), and a series of indicators about the use of different production factors (employment $L_{i,t}$, capital intensity $(K/L)_{i,t}$, and investment intensity $(I/L)_{i,t}$).

4. Data

The two-layer network is inferred from two different data sets on the US economy covering the period from 1977 to 2012. The data are available in five year snapshots. The market layer is compiled on national account data provided by the Bureau of Economic Analysis (BEA). The data are combined with data on patents granted by the US Patent and Trademark Office (USPTO) which are classified by the Cooperative Patent Classification (CPC) system. Using the concordance tables by Goldschlag et al. (2020), 4-digit CPC codes are mapped to industries and aggregated into five year windows in accordance with the IO data. This enables the compilation of a cross-industry patent citation network for different periods.

The networks are given by symmetric matrices where the entries represent the flow of goods $flow_{ij,t}^\mu$ and patent citations $flow_{ij,t}^\tau$. The cross industrial flow data are harmonized to input and output shares. The network data are used to compile the similarity matrices $\Sigma_t^{\alpha,d}$ and the industry variables: citation-weighted patent stocks $A_{i,t}^\tau$, industry output $A_{i,t}^\mu$, the centrality measures PageRank $PR_{i,t}^{\alpha,d}$, degree $D_{i,t}^{\alpha,d}$, strength $S_{i,t}^{\alpha,d}$, and spillovers $Spill(A)_{i,t}^{\alpha,d}$.

The network data are complemented with data from the NBER Manufacturing Productivity Database (Becker et al., 2013; Bartlesman and Gray, 1996). From these data, I extracted employment $L_{i,t}$, labor costs per employee $Wage_{i,t}$, the share of production workers $(L_{prod}/L)_{i,t}$, value added per employee $(VA/L)_{i,t}$ which is used as a proxy of labor productivity, investment per employee $(I/L)_{i,t}$, capital intensity $(K/L)_{i,t}$, and five and four factor productivity $TFP4_{i,t}$, $TFP5_{i,t}$ as additional variables.⁵

The final data consist of a balanced panel of 307 6-digit manufacturing industries. More aggregate data are used for additional robustness checks. The most important steps of the data compilation are summarized in A.2. Additional detail is provided in SI.1.

⁵ $TFP5_{i,t}$ differentiates between energy and non-energy material inputs. This distinction is not made in $TFP4_{i,t}$.

5. Results

This section begins with a descriptive analysis of the two network layers and their overlap. It follows a series of regressions to identify push and pull effects.

5.1. Descriptive analysis

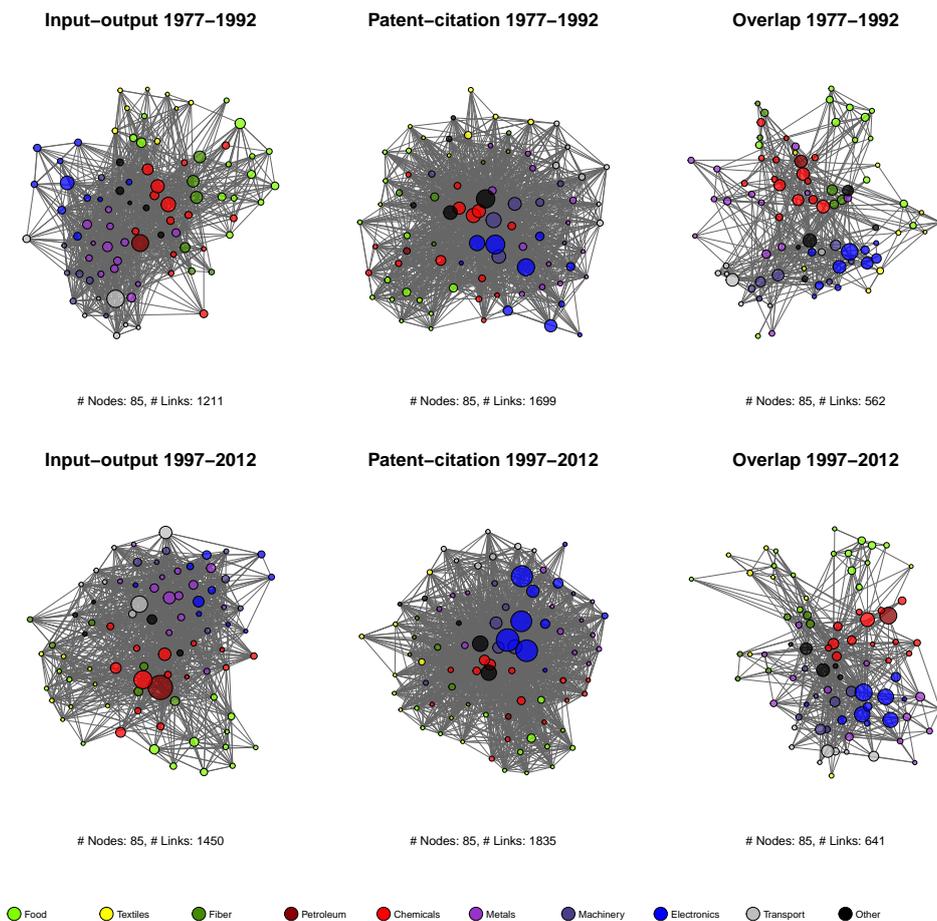
Fig. 1 shows a series of upstream network plots at the 4-digit level for the first and second half of the time period covered by this study. A link between two industries i and j is shown if the connecting weight $w_{ij,t}^{\alpha,in}$ is sufficiently strong and the overlap network shows connections between two industries if the supply relationship is strong in both layers. The node sizes are proportional to the industry size $A_{i,t}^{\alpha}$.⁶ The algorithm that generates the plots tends to group strongly connected industries together. The node colors indicate the broad industry class of an industry.

The overlap network is least dense which must be true by definition. The size distribution across broad industry groups is more balanced in the IO layer compared to the innovation layer where electronics (blue), machinery (dark blue) and chemical manufacturing (red) visually dominate by size.

In the IO network, groups of industries with similar color tend to cluster together. The position of the clusters interacts with the position in the supply chain (or "trophic level" as McNerney et al. (2018) call it). Industries that are close to more primary resource inputs (food processing (greenish), textiles (yellow)) and to final consumers (food, textiles, electronics (blue), transportation (gray)) are located at the margins, while chemicals (red), metals (violet) and petroleum products (brown) with more intermediate positions between primary input providers and end users take central positions. This pattern is different in the innovation layer where also electronics (blue) takes very central positions, i.e. being an important provider and user of innovations. It is also clearly visible that the electronics sector (including computer industries) grew strongly over time as indicated by the increasing size of blue nodes. The large size of the residual category "Other" in black color is partly an artifact from the patent concordance.

These observations are confirmed by the statistics in Tables 1, 2 and Table SI.1. Table 1 shows the properties of the two network layers and of the cosine similarity matrices $\Sigma_{ij,t}^{\alpha,d}$ for each decade which can be interpreted as a symmetric network where the similarity $\sigma_{ij,t}^{\alpha,d}$ is the weight of a link. The first lines of each section in the table

⁶Technical detail is provided in A.4.



Notes: These figures show the network of upstream links (suppliers) at the 4-digit level for two different time periods. A link between a pair of industries i and j is shown if j is a sufficiently important supplier to i , i.e. if the average of the weight $w_{ij,t}^{in,\alpha}$ during time periods 1977-1992 and 1997-2012 exceeds a threshold level given by the average weight over all industry pairs and all periods plus one standard deviation ($\text{mean}_{i,j,t}(w_{ij,t}^{in,\alpha}) + \text{sd}_{i,j,t}(w_{ij,t}^{in,\alpha})$). The overlap network shows nodes as being connected if they are connected on both layers, i.e. links are compiled on the basis of weights averaged across both layers. The size of the nodes is proportional to the size of an industry $A_{i,t}^\alpha$ in the respective layer, and in the overlap network to the weighted mean of both layers ($0.5 \cdot (A_{i,t}^H + A_{i,t}^T)$). Plots of the downstream network are available in [SI.2.1](#). Self-citations and within-sector IO flows are not shown. The colors indicate broad industrial categories given by groups of 3-digit level industries, i.e. Food (311-312), Textiles (314-316), Fiber (321-323), Petroleum (324), Chemicals (325-327), Metals (331,332), Machinery (333), Electronics (334-335), Transport (336), Other (337-339).

Figure 1: Upstream networks of 4-digit level industries for different periods

show the network density which measures the connectivity. It shows the ratio of actual over potential connections in the network. Both layers are rather sparsely connected. The innovation layer is denser and shows an increase in the up- and downstream density over time, while connectivity in the market does not exhibit any clear trend. The average degree (second line of each block in the table) indicates the average number of industries to which an industry is connected. On average, an industry is connected to 20-30 customers and suppliers in the market and cites patents from 44-51 other industries. Both network layers show a negative valued assortativity: larger and more connected industries tend connect more often to smaller and less connected industries.

The density in the cosine similarity networks is an indicator of technological convergence measuring whether or not industries became more similar on average. In line with the connectivity trends, similarity patterns in the market fluctuate but do not show any clear trend, while there is a trend towards an increasing similarity of industries by patent citations.

The overlap of both network shows only a very low density of 2% which is also roughly stable. However, the cosine similarity computed on the basis of concatenated vectors of market inputs and patent citation (outputs and passive citations) is almost as high as the similarity in the innovation layer. This indicates that an industry pair that is similar by IO relations tends to be also similar by patent citations. The similarity shows an increasing trend.

Table 2 shows the Top-10 ranking by output and patents. Petroleum Refineries (dark red color in the plots) rank persistently on the top position in the market. We also observe high ranks for Iron & Steel, Plastics Material & Resin, and Semiconductors. Over time, industries associated with natural product processing (Paper Mills, Newsprint, Paperboard, Slaughtering) gradually disappeared from the top ranks. This was accompanied with the rise of certain machinery and electronics industries (Machine Shops, Aircraft, Vehicle Air-Conditioning, Automobile).

The top ranks in the innovation layer are dominated by metal, machinery and electronics manufacturing as indicated by the leading 2-digit code 33. Only two chemical industries rank high but declined throughout the time period covered (2-digit code 32). The patent ranks show an increasing dominance of ICT-related industries and specialist instruments (Semiconductors, Electronic Computers, Medical & Optical Instruments, Watch & Clock, Wireless Communication). The time trends in the patent ranking are more monotonous over time.

The bottom lines in the tables show the quartile distribution of the industry sizes. As the data is normalized such that the average size in each time period equals one, a

median value that deviates from one indicates a skewed distribution. Both layers show a skewed distribution with concentration at the top ranks. The table reveals a higher and increasing concentration in the innovation layer while market concentration does not show any clear trend. These observations are consistent across different levels of aggregation (see [B](#) and [SI.2.1](#)).

	Input-output				Patent				Overlap			
	1977-1982	1987-1992	1997-2002	2007-2012	1977-1982	1987-1992	1997-2002	2007-2012	1977-1982	1987-1992	1997-2002	2007-2012
<i>Flow matrix - upstream network</i>												
Density	0.07	0.07	0.07	0.10	0.14	0.15	0.16	0.17	0.02	0.02	0.02	0.02
Avg. degree	22.02	21.86	20.08	29.02	44.43	46.39	49.84	51.33	5.92	5.85	5.66	7.43
Avg. weight	0.87	0.83	0.88	0.75	0.36	0.35	0.34	0.34	2.27	2.04	1.97	1.72
Reciprocity	0.15	0.14	0.12	0.33	0.36	0.34	0.32	0.29	0.14	0.14	0.09	0.17
Transitivity	0.34	0.34	0.30	0.34	0.42	0.43	0.45	0.46	0.17	0.18	0.14	0.19
Diameter	3.00	3.00	4.00	2.00	2.00	2.00	2.00	2.00	4.00	3.00	4.00	2.00
Mean dist.	1.43	1.41	1.43	1.49	1.07	1.04	1.02	1.02	1.47	1.44	1.44	1.49
Assort. by degree	-0.25	-0.24	-0.19	-0.33	-0.12	-0.10	-0.07	-0.06	-0.24	-0.23	-0.19	-0.33
Assort. by size	-0.00	-0.02	-0.03	-0.02	-0.02	-0.01	-0.00	-0.00	-0.01	-0.01	-0.01	0.00
<i>Flow matrix - Downstream network</i>												
Density	0.08	0.07	0.07	0.09	0.14	0.15	0.16	0.16	0.02	0.02	0.02	0.02
Avg. degree	24.15	22.50	20.91	28.39	43.72	45.42	48.95	50.33	5.76	5.81	5.12	7.19
Avg. weight	0.87	0.83	0.88	0.75	0.36	0.35	0.34	0.34	4.08	3.83	3.94	1.15
Reciprocity	0.17	0.17	0.13	0.30	0.36	0.35	0.33	0.30	0.15	0.15	0.12	0.16
Transitivity	0.39	0.39	0.34	0.34	0.42	0.42	0.45	0.46	0.24	0.24	0.24	0.20
Diameter	3.00	3.00	4.00	2.00	2.00	2.00	2.00	2.00	4.00	3.00	4.00	2.00
Mean dist.	1.43	1.41	1.43	1.49	1.07	1.04	1.02	1.02	1.47	1.44	1.44	1.49
Assort. by degree	-0.25	-0.24	-0.19	-0.33	-0.12	-0.10	-0.07	-0.06	-0.24	-0.23	-0.19	-0.33
Assort. by size	-0.00	-0.02	-0.03	-0.02	-0.02	-0.01	-0.00	-0.00	-0.01	-0.01	-0.01	0.00
<i>Cosine similarity - upstream network</i>												
Density	0.28	0.27	0.28	0.31	0.38	0.40	0.40	0.40	0.35	0.36	0.37	0.39
Avg. degree	86.33	83.22	85.20	94.89	116.73	121.32	122.52	123.42	107.76	111.41	113.01	117.97
Avg. weight	12.84	11.60	12.20	9.37	23.69	26.35	28.40	31.50	1683.99	1843.93	1885.33	1877.13
Transitivity	0.67	0.68	0.75	0.63	0.72	0.74	0.75	0.74	0.66	0.67	0.67	0.68
<i>Cosine similarity - downstream network</i>												
Density	0.25	0.24	0.23	0.29	0.39	0.40	0.41	0.41	0.37	0.37	0.37	0.38
Avg. degree	76.85	73.10	71.95	88.92	118.84	121.95	124.19	125.52	112.27	113.84	114.18	116.61
Avg. weight	9.59	8.82	7.72	9.47	25.35	26.52	29.41	30.93	1725.32	1792.20	1732.52	2025.42
Transitivity	0.61	0.62	0.58	0.67	0.72	0.74	0.74	0.73	0.70	0.72	0.72	0.68

The upper part of the table shows a series of network statistics compiled at the basis of the up- and downstream links in the market- and innovation layer for different time windows. The links in these time windows are averaged, i.e. $flow_{ij,T}^{\alpha,d} = |T|^{-1} \sum_{t \in T} flow_{ij,t}^{\alpha,d}$ with $T = \{1977, 1982\}, \{1987, 1992\}, \{1997, 2002\}, \{2007, 2012\}$. The lower parts of the table summarize the network characteristics of the cosine similarity network. The network is given by the symmetric $N \times N$ cosine similarity matrix $\Sigma_t^{\alpha,d}$ where the pairwise similarities $\sigma_{ij,t}^{\alpha,d}$ are the weights of a link connecting i and j . For the purpose of calculating aggregate network statistics (and plotting), two industries i, j are shown as being connected in t if their pairwise similarity is higher than the average of similarity of all industries N and all periods T , i.e. $\sigma_{ij,t}^{\alpha,out} > (|N|(|N| - 1) \cdot |T|)^{-1} \sum_{i \in N} \sum_{j \in N, j \neq i} \sum_{t \in T} \sigma_{ij,t}^{\alpha,out}$. The metrics are compiled using the R-package *igraph* (Csardi, Nepusz, et al., 2006). For an introduction to the use of these metrics see also Jackson (2008).

Table 1: Aggregate network statistics over time at the 6-digit level.

<i>Top 10 industries by Aggr. output ($A_{i,t}^{\mu}$)</i>									
1977-1982			1987-1992		1997-2002		2007-2012		
1	Petroleum Refineries	324110 37.43	Petroleum Refineries	324110 19.09	Petroleum Refineries	324110 18.94	Petroleum Refineries	324110 40.72	
2	Animal Slaughter.	311611 10.01	Plastics Mat. & Resin	325211 8.14	Semiconductor & Device	334413 12.12	Iron & Steel Mills	331111 9.74	
3	All Petrol. & Coal Prod.	324199 8.69	Animal Slaughter.	311611 7.83	Iron & Steel Mills	331111 10.97	Utility Vhcl.	336112 7.61	
4	Iron & Steel Mills	331111 8.59	Semiconductor & Device	334413 6.71	Plastics Mat. & Resin	325211 7.90	Plastics Mat. & Resin	325211 7.38	
5	Plastics Mat. & Resin	325211 6.52	All Petrol. & Coal Prod.	324199 5.92	Plastics Plumb. Fixture	326191 4.86	Aircraft Mnft.	336411 5.29	
6	Paper Mills	322121 4.73	Iron & Steel Mills	331111 5.81	Machine Shops	332710 4.78	Motor Vhcl. Air-Cond.	336391 4.71	
7	Newsprint Mills	322122 4.73	Paper Mills	322121 5.10	Sawmills	321113 4.56	Semiconductor & Device	334413 4.45	
8	Metal Can Mnft.	332431 4.32	Newsprint Mills	322122 5.10	Elect. Circuit Assembly	334418 4.06	Automobile Mnft.	336111 3.84	
9	Paperboard Mills	322130 3.94	Paperboard Mills	322130 5.10	Paperboard Mills	322130 3.85	Gum & Wood Chemicals	325191 3.37	
10	Paint & Coat. Mnft.	325510 3.76	Paint & Coat. Mnft.	325510 4.12	Paper Mills	322121 3.70	Basic Organic Chem.	325199 3.37	
Quartiles:									
0.19, 0.55, 1.15			0.2, 0.615, 1.345		0.22, 0.555, 1.2175		0.325, 0.57, 0.93		
<i>Top 10 industries by Patent stock ($A_{i,t}^{\sigma}$)</i>									
1977-1982			1987-1992		1997-2002		2007-2012		
1	Adhesive Mnft.	325520 16.23	Semiconductor & Device	334413 19.33	Semiconductor & Device	334413 25.97	Electr. Computer Mnft.	334111 30.12	
2	Misc. Chem. Prepar.	325998 15.31	Adhesive Mnft.	325520 15.52	Electr. Computer Mnft.	334111 23.77	Semiconductor & Device	334413 29.65	
3	Semiconductor & Device	334413 13.73	Electr. Computer Mnft.	334111 14.87	Medical Instrum.	339112 14.09	Medical Instrum.	339112 15.08	
4	Fastener, Button & Pin	339993 11.62	Misc. Chem. Prepar.	325998 14.60	Fastener, Button & Pin	339993 13.60	Optical Instrum. & Lens	333314 13.34	
5	Electr. Computer Mnft.	334111 9.47	Fastener, Button & Pin	339993 14.00	Optical Instrum. & Lens	333314 12.48	Elctrmed. Apparatus	334510 11.79	
6	Optical Instrum. & Lens	333314 9.05	Optical Instrum. & Lens	333314 11.59	Adhesive Mnft.	325520 12.33	Fastener, Button & Pin	339993 11.27	
7	Power Transm. Equ.	333613 9.01	Medical Instrum.	339112 9.07	Misc. Chem. Prepar.	325998 11.48	Watch & Clock Mnft.	334518 11.14	
8	Pump & Equ. Mnft.	333911 6.08	Power Transm. Equ.	333613 8.04	Elctrmed. Apparatus	334510 11.02	Broadcast. & Wireless Communic.	334220 11.04	
9	Speed Changer & Gear	333612 5.99	Elctrmed. Apparatus	334510 7.61	Watch & Clock Mnft.	334518 9.47	Adhesive Mnft.	325520 9.98	
10	Watch & Clock Mnft.	334518 5.67	Watch & Clock Mnft.	334518 6.57	Telephone Mnft.	334210 8.02	Misc. Chem. Prepar.	325998 9.23	
Quartiles:									
0.0575, 0.33, 0.99			0.06, 0.28, 0.8625		0.06, 0.22, 0.7375		0.05, 0.19, 0.6925		

Notes: Industries are ranked by output (patent stock) $A_{i,t}^{\alpha}$ averaged across the time window indicated in the column header in decreasing order, i.e. showing the largest industries on top. The values $A_{i,t}^{\alpha}$ were normalized before through division by the economy-wide average output (patent stock) in t , i.e. the mean value for each period equals one. The bottom lines of each sub-table show the quartile values as indicators for the skewness of the distribution. Deviations of the median from the average indicate skewness.

Table 2: Top-10 ranking of industries by output and patent stock at the 6-digit level.

5.2. Demand-pull, technology-push, and technological change

Here, I present the results of regressions investigating the drivers of market growth, innovation, and changes in productivity, employment, and factor use.

The first set of results stem from a regression of industry sizes $A_{i,t}^\alpha$ and growth rates on lagged size $A_{i,t-1}^\alpha$, spillovers $Spill(A)_{i,t-1}^{\alpha,d}$, network centrality indicators $PR_{i,t-1}^{\alpha,d}$ from both layers $\alpha = \mu, \tau$ and industry level controls. Spillovers and the PageRank are computed on the basis on up- and downstream linkages. Both are included but in separate regression runs because they suffer from high multicollinearity in the innovation layer (see also B).

To control for the characteristics of an industry, lagged labor force $L_{i,t-1}$, wages $Wage_{i,t-1}$, capital intensity $(K/L)_{i,t-1}$, per capita investment $(I/K)_{i,t-1}$, the share of production labor $(Lprod/L)_{i,t-1}$, labor productivity $(VA/L)_{i,t-1}$, and lagged 5-factor productivity $TFP5_{i,t-1}$ are included. These variables capture industrial production inputs that are not reflected in the intermediate goods input-output relations and the characteristics of intangible assets as captured by patent citations.

The regression equations are given by

$$Y_{i,t} = \sum_{\alpha=\mu,\tau} \left[\beta_A^\alpha A_{i,t-1}^\alpha + \beta_{PR}^\alpha PR_{i,t-1}^{\alpha,d} + \beta_S^\alpha Spill(A)_{i,t-1}^{\alpha,d} \right] + \beta' \mathbf{X}_{i,t-1} \quad (2)$$

where $Y_{i,t} \in \{A_{i,t}^\alpha, gr(A_{i,t}^\alpha)\}_{\alpha=\mu,\tau}$ and $d = in, out$ and $\mathbf{X}_{i,t-1}$ is a vector of industry level controls listed above. The regressions with the industry sizes as explanatory variables are indicative for the direction of change and show whether an industry grows or not. The growth rate regressions inform about the pace of change and show whether an industry grows or shrinks at an increasing or decreasing rate. For example, a positive coefficient in the size regressions and a negative coefficient in the growth rate regressions indicate a pattern of growth but at a decreasing rate.

The regressions include two-ways industry and time fixed effects (FE) and standard errors are clustered. To cope with skewness, all variables (except for $(Lprod/L)_{i,t-1}$) were first log-linearized and subsequently outliers were removed. A detailed description of the data transformations are available in A.3.

Table 3 shows the first set of results. Within both layers, lagged industry size $A_{i,t-1}^\alpha$ is negatively associated with growth in the same layer $gr(A_{i,t}^\alpha)$ and positively with industry size $A_{i,t}^\alpha$. This indicates a positive auto-correlation which is an indicator of path-dependence in the evolution of industry sizes: growing industries grow, but at a decreasing rate. Note that the regressions include two-ways FE, i.e. the regressors reflect

	$gr(A_{i,t}^\mu)$		$A_{i,t}^\mu$		$gr(A_{i,t}^\tau)$		$A_{i,t}^\tau$	
	up	down	up	down	up	down	up	down
$A_{i,t-1}^\mu$	-0.4963*** (0.0293)	-0.4952*** (0.0254)	0.4366*** (0.0331)	0.4169*** (0.0296)	0.0013 (0.0073)	0.0044 (0.0068)	0.0012 (0.0096)	0.0089 (0.009)
$A_{i,t-1}^\tau$	-0.1178 (0.0618)	-0.071 (0.0585)	-0.1304 (0.071)	-0.0715 (0.0657)	-0.2441*** (0.0255)	-0.259*** (0.029)	0.7449*** (0.0267)	0.7289*** (0.0306)
$PR_{i,t-1}^{\mu,d}$	-0.0425 (0.035)	0.15*** (0.0346)	-0.0971* (0.0427)	0.2079*** (0.0417)	0.0154 (0.0098)	-0.0114 (0.0094)	0.0279 (0.0159)	-0.0215 (0.0154)
$PR_{i,t-1}^{\tau,d}$	0.5277** (0.1679)	0.3841* (0.1811)	0.5602** (0.1917)	0.4142* (0.2084)	-0.1082 (0.0566)	-0.0535 (0.0575)	-0.0941 (0.075)	-0.0196 (0.0734)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.002 (0.0337)	-0.0175 (0.0317)	0.012 (0.0395)	0.0144 (0.0366)	0.0112 (0.0088)	-0.0013 (0.0106)	0.0165 (0.0138)	-0.0121 (0.015)
$Spill(A)_{i,t-1}^{\tau,d}$	-0.0148 (0.1313)	-0.075 (0.1309)	-0.0047 (0.1508)	-0.1303 (0.1424)	0.1569*** (0.0427)	0.0966 (0.0522)	0.2119*** (0.0562)	0.074 (0.0607)
$L_{i,t-1}$	-0.1895* (0.0963)	-0.2128* (0.1001)	-0.2129 (0.1187)	-0.2472* (0.1237)	-0.0065 (0.0236)	-0.0059 (0.0243)	-0.0325 (0.0312)	-0.0313 (0.0326)
$Wage_{i,t-1}$	0.4116 (0.4129)	0.4756 (0.4072)	0.7706 (0.4618)	0.8305 (0.4546)	0.0881 (0.093)	0.0805 (0.0953)	0.1929 (0.1182)	0.1693 (0.1202)
$(K/L)_{i,t-1}$	-0.6859** (0.2185)	-0.6998** (0.2189)	-0.7432** (0.2462)	-0.7605** (0.2471)	-0.0658 (0.0517)	-0.0561 (0.0498)	-0.0986 (0.0713)	-0.074 (0.0686)
$(Lprod/L)_{i,t-1}$	-0.5913 (0.5831)	-0.5524 (0.5797)	-1.204 (0.6665)	-1.164 (0.6662)	-0.088 (0.1822)	-0.1535 (0.1759)	-0.0946 (0.2349)	-0.2067 (0.2326)
$(I/L)_{i,t-1}$	-0.0788 (0.0882)	-0.0691 (0.0896)	-0.0359 (0.1009)	-0.0158 (0.1013)	0.0436* (0.0201)	0.0468* (0.0201)	0.0385 (0.0295)	0.0419 (0.0298)
$(VA/L)_{i,t-1}$	0.3025 (0.1566)	0.247 (0.1512)	0.1628 (0.1852)	0.0918 (0.1792)	0.0544 (0.0353)	0.0542 (0.0369)	0.0318 (0.0474)	0.0337 (0.0507)
$TFP5_{i,t-1}$	-0.2847 (0.1833)	-0.2551 (0.178)	-0.4068 (0.2211)	-0.3865 (0.2093)	0.0886 (0.06)	0.0845 (0.0595)	0.1479* (0.0707)	0.1407* (0.068)
R^2	0.2822	0.289	0.1779	0.1878	0.2157	0.205	0.5954	0.5882
N	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.3638	0.3638	7.683	7.683	0.0524	0.0524	8.45	8.45

Notes: The regressions aim to explain the factors that influence the evolution of industry sizes in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^\alpha$, $gr(A_{i,t}^\tau)$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformations and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table 3: Regression results: The evolution of industry sizes

deviations from the industry-internal average. A similar pattern of auto-correlation can be observed in an additional regression analysis using growth rates as regressors (see Table C.4). These results also show a significant negative auto-correlation among growth rates in the market, which is less significant in the innovation layer. This suggests a pattern of growth moderation in the market which is weaker in innovation.

Another interesting observation from the results in Table 3 is the ambiguous role of centrality in the market layer dependent on the direction. Downstream $PR_{i,t-1}^{\mu,out}$ shows a positive association with the growth rate and industry size in the market. Its counterpart compiled on upstream links shows the opposite relationship, even though less significant. Further, upstream centrality $PR_{i,t-1}^{\mu,in}$ shows a weakly significant positive association with innovation.

Additional results in Tables C.5-C.8 show that this pattern also holds for alternative network centrality measures (degree and strength). Hence, a central position in the network of customer links is positively related to industrial growth, while centrality in the input network shows a weakly negative effect. Note that $PR_{i,t-1}^{\mu,in}$ can be also high if an industry has many diverse suppliers which would be reflected in the in-degree $D_{i,t-1}^{\mu,in}$. The robustness checks (see C.2) confirm that $D_{i,t-1}^{\mu,in}$ is not the driver as it does not show the same negative relationship as the strength and the PageRank.

Centrality in the innovation layer $PR_{i,t-1}^{\tau,d}$ in both directions exhibits a positive association with market size and growth, but it is not significant to explain growth in the innovation layer. Additional results about of a correlation analysis that explores the interactions between growth rates indicate that an increasing patent centrality is positively correlated with innovation outputs $A_{i,t}^{\tau}$ (see C.4).

The next noteworthy observation is the role of spillovers. Spillovers are not significant in the market, but in the innovation layer. In line with previous research (e.g. Jaffe, 1988; Bloom et al., 2013; Acemoglu et al., 2016), the results confirm a positive role of up- and downstream knowledge spillovers as drivers of innovation. This observation is robust across a series of robustness checks (see C.1 and SI.2.2).

The results show that industries with both, a high use of labor and a high capital-labor ratio tend to shrink, while a higher investment intensity and TFP are positively related to innovation.

	$gr((VA/L)_{i,t})$		$(VA/L)_{i,t}$		$gr(TFP4_{i,t})$		$TFP4_{i,t}$		$gr(TFP5_{i,t})$		$TFP5_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$A_{i,t-1}^{\mu}$	0.0167** (0.0063)	0.0055 (0.0057)	0.0177** (0.0065)	0.0061 (0.0058)	4e-04 (5e-04)	3e-04 (5e-04)	0.004 (0.0026)	0.0039 (0.0021)	3e-04 (5e-04)	2e-04 (5e-04)	0.004 (0.0026)	0.0039 (0.0021)
$A_{i,t-1}^{\tau}$	-0.0186 (0.0187)	-0.0099 (0.0192)	-0.0188 (0.0188)	-0.0094 (0.0193)	-5e-04 (0.0014)	-3e-04 (0.0014)	4e-04 (0.0056)	0.0035 (0.0062)	-5e-04 (0.0014)	-3e-04 (0.0014)	3e-04 (0.0056)	0.0033 (0.0062)
$PR_{i,t-1}^{\mu,d}$	-0.0364** (0.0125)	0.029** (0.0104)	-0.0379** (0.0129)	0.0299** (0.0107)	-0.0011 (0.001)	-4e-04 (7e-04)	-0.0016 (0.006)	0.0064 (0.0037)	-0.001 (0.001)	-4e-04 (7e-04)	-0.0017 (0.006)	0.0065 (0.0037)
$PR_{i,t-1}^{\tau,d}$	0.0271 (0.0451)	-0.0283 (0.0575)	0.0203 (0.0478)	-0.038 (0.0616)	0.0041 (0.0041)	0.0041 (0.0047)	0.0129 (0.0152)	0.0071 (0.0177)	0.0041 (0.0041)	0.0043 (0.0047)	0.013 (0.0152)	0.0074 (0.0177)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0037 (0.0089)	0.0084 (0.0095)	-0.0036 (0.0091)	0.0079 (0.0097)	-0.0011 (9e-04)	-0.0015 (8e-04)	-2e-04 (0.0041)	6e-04 (0.0035)	-9e-04 (9e-04)	-0.0014 (8e-04)	-3e-04 (0.0041)	7e-04 (0.0035)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0344 (0.034)	0.0277 (0.0373)	0.0299 (0.0385)	0.022 (0.0354)	-2e-04 (0.0027)	-0.0039 (0.003)	-0.0119 (0.0154)	-0.0205 (0.0171)	-4e-04 (0.0027)	-0.0042 (0.003)	-0.0119 (0.0154)	-0.0203 (0.0171)
$L_{i,t-1}$	-0.0183 (0.0245)	-0.0191 (0.0253)	-0.0182 (0.0247)	-0.0193 (0.0255)	2e-04 (0.0021)	2e-04 (0.0021)	-0.0039 (0.0105)	-0.0051 (0.0102)	1e-04 (0.0021)	1e-04 (0.0021)	-0.004 (0.0105)	-0.0052 (0.0102)
$Wage_{i,t-1}$	0.0971 (0.1382)	0.1068 (0.1393)	0.1121 (0.1404)	0.1221 (0.1417)	-0.0094 (0.0099)	-0.0103 (0.0098)	-0.0157 (0.0373)	-0.0154 (0.0367)	-0.0102 (0.0099)	-0.0112 (0.0098)	-0.0159 (0.0373)	-0.0156 (0.0367)
$(K/L)_{i,t-1}$	0.1087 (0.085)	0.1127 (0.086)	0.1017 (0.0856)	0.1056 (0.0867)	0.0205*** (0.0053)	0.0215*** (0.0053)	0.1496*** (0.0272)	0.1499*** (0.0261)	0.0203*** (0.0053)	0.0213*** (0.0052)	0.1499*** (0.0271)	0.1502*** (0.026)
$(Lprod/L)_{i,t-1}$	0.1237 (0.1664)	0.105 (0.1677)	0.1528 (0.1774)	0.1348 (0.1784)	-0.0051 (0.017)	-0.0067 (0.0169)	0.0453 (0.0808)	0.0475 (0.0802)	-0.0068 (0.0167)	-0.0084 (0.0166)	0.0428 (0.0813)	0.0451 (0.0807)
$(I/L)_{i,t-1}$	-0.016 (0.0251)	-0.0135 (0.0245)	-0.0185 (0.0254)	-0.0161 (0.0248)	-6e-04 (0.0024)	-9e-04 (0.0024)	-0.029** (0.0098)	-0.0285** (0.0098)	-4e-04 (0.0024)	-7e-04 (0.0023)	-0.0289** (0.0098)	-0.0284** (0.0097)
$(VA/L)_{i,t-1}$	-0.4396*** (0.0515)	-0.4436*** (0.0517)	0.5576*** (0.053)	0.5534*** (0.0533)	-0.0147*** (0.004)	-0.0142*** (0.0039)	-0.0546* (0.0212)	-0.0567** (0.0203)	-0.0146*** (0.004)	-0.0141*** (0.0039)	-0.0549** (0.0212)	-0.0571** (0.0203)
$TFP4_{i,t-1}$					-0.0574*** (0.0091)	-0.0579*** (0.0087)	0.9737*** (0.0825)	0.9737*** (0.0826)				
$TFP5_{i,t-1}$	0.1003 (0.0722)	0.1007 (0.0738)	0.1006 (0.0717)	0.1009 (0.0735)					-0.0573*** (0.009)	-0.0578*** (0.0086)	0.9736*** (0.0823)	0.9736*** (0.0824)
R^2	0.178	0.1763	0.3309	0.3295	0.1811	0.183	0.7084	0.7095	0.1807	0.183	0.7083	0.7095
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.2664	0.2664	4.48	4.48	0.6937	0.6937	0.6754	0.6754	0.6938	0.6938	0.6756	0.6756

Notes: The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP4_{i,t}$, $TFP5_{i,t}$, $gr((VA/L)_{i,t})$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3. The difference between four and five factor TFP is that $TFP5_{i,t}$ is a disaggregation of material inputs into energy and non-energy materials which is not done for $TFP4_{i,t}$ (see Bartlesman and Gray (1996)).

Table 4: Regression results: Determinants of productivity growth

The next set of results is a regression of productivity indicators ($(VA/L)_{i,t}$, $TFP4_{i,t}$, $TFP5_{i,t}$) on industry characteristics. The results presented in Table 4 show that interactions between the evolution of productivity and the patent network are absent. For the market, it is again found that the downstream centrality $PR_{i,t-1}^{\mu,out}$ exhibits a positive relationship with the technological performance of an industry: it is significantly positively associated with level and growth rate of value added. Further, a weakly significant positive association with $TFP4_{i,t}$ and $TFP5_{i,t}$ can be observed. In contrast, market centrality by input links exhibits the opposite relationship showing a negative signed coefficient for value added. Again, these results are consistent across a series of robustness checks using alternative centrality measures (see C.1 and SI.2.2).

All productivity measures are auto-correlated by value and negatively auto-correlated in their growth rate (again, controlling for industry and time FE) which is an indicator of path-dependence. A higher capital intensity is positively associated with TFP and TFP growth. Labor productivity measured as value added per employee shows a negative correlation with TFP and TFP growth.

	$gr(L_{i,t})$		$L_{i,t}$		$gr((K/L)_{i,t})$		$(K/L)_{i,t}$		$gr((I/L)_{i,t})$		$(I/L)_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$A_{i,t-1}^{\mu}$	0.0025 (0.0066)	-0.0014 (0.0057)	0.0052 (0.0085)	5e-04 (0.0076)	-1e-04 (1e-04)	-1e-04 (1e-04)	-0.0026 (0.0044)	-0.0019 (0.0038)	0.0097 (0.0117)	-0.0055 (0.0106)	0.0093 (0.01)	-0.004 (0.0093)
$A_{i,t-1}^{\tau}$	-0.0304** (0.0108)	-0.0368*** (0.0106)	-0.0347* (0.0136)	-0.0417** (0.013)	2e-04 (2e-04)	3e-04 (2e-04)	0.0146 (0.0089)	0.0153 (0.0082)	-0.0298 (0.0242)	-0.0336 (0.0256)	-0.0268 (0.0222)	-0.0278 (0.0231)
$PR_{i,t-1}^{\mu,d}$	-0.0096 (0.0086)	0.0146 (0.008)	-0.0129 (0.0106)	0.0211 (0.0108)	1e-04 (1e-04)	-1e-04 (2e-04)	0.0051 (0.0055)	-0.0055 (0.0052)	-0.0335* (0.0163)	0.0348** (0.0131)	-0.0286 (0.0147)	0.0332** (0.0118)
$PR_{i,t-1}^{\tau,d}$	-0.0039 (0.0362)	0.0269 (0.0443)	-0.0176 (0.0445)	0.0123 (0.0553)	6e-04 (6e-04)	5e-04 (7e-04)	0.0147 (0.0235)	0.0271 (0.0299)	0.0889 (0.0707)	0.1586 (0.0842)	0.0875 (0.0671)	0.1345 (0.0787)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0203* (0.0092)	0.0104 (0.008)	-0.0203 (0.0128)	0.0111 (0.0105)	0 (2e-04)	1e-04 (2e-04)	0.0107 (0.0063)	1e-04 (0.0053)	-0.02 (0.0157)	0.0325* (0.0139)	-0.0135 (0.0144)	0.0301* (0.0128)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0761** (0.0278)	0.0849** (0.0273)	0.0853* (0.035)	0.1003** (0.0348)	-7e-04 (5e-04)	-9e-04 (5e-04)	-0.0121 (0.0193)	-0.0385 (0.0207)	0.0811 (0.0562)	0.0535 (0.0562)	0.078 (0.0511)	0.0489 (0.0533)
$L_{i,t-1}$	-0.1912*** (0.0233)	-0.1925*** (0.0239)	0.7963*** (0.0306)	0.7936*** (0.0313)	0.001 (5e-04)	0.001* (5e-04)	0.0202 (0.0172)	0.0197 (0.0173)	-0.0846* (0.0389)	-0.0903* (0.0389)	-0.0955** (0.0348)	-0.0999** (0.0347)
$Wage_{i,t-1}$	-0.0442 (0.0769)	-0.0393 (0.0768)	-0.0187 (0.0991)	-0.0083 (0.0979)	2e-04 (0.0015)	1e-04 (0.0015)	0.0504 (0.0597)	0.0407 (0.058)	0.2743 (0.1433)	0.2629 (0.1459)	0.3491* (0.137)	0.3402* (0.1399)
$(K/L)_{i,t-1}$	-0.195*** (0.0496)	-0.1937*** (0.0471)	-0.1693** (0.0644)	-0.1708** (0.0612)	-0.0064*** (0.0011)	-0.0063*** (0.0011)	0.7889*** (0.0391)	0.7903*** (0.0376)	-0.1742 (0.0976)	-0.1723 (0.0961)	-0.2352** (0.0903)	-0.2322** (0.089)
$(Lprod/L)_{i,t-1}$	0.8074*** (0.1648)	0.7936*** (0.1619)	0.947*** (0.2156)	0.936*** (0.2115)	-0.0105** (0.0036)	-0.0104** (0.0036)	-0.5101*** (0.1239)	-0.5119*** (0.1227)	0.5702* (0.2598)	0.5613* (0.2557)	0.6169* (0.2531)	0.6024* (0.2495)
$(I/L)_{i,t-1}$	0.0356 (0.02)	0.0361 (0.0198)	0.0485* (0.0239)	0.0497* (0.024)	0.0022*** (4e-04)	0.0022*** (4e-04)	0.1206*** (0.017)	0.1218*** (0.017)	-0.764*** (0.0418)	-0.7591*** (0.0409)	0.3043*** (0.0431)	0.3094*** (0.0424)
$(VA/L)_{i,t-1}$	0.0833* (0.0376)	0.082* (0.036)	0.0601 (0.0487)	0.0561 (0.0472)	6e-04 (9e-04)	6e-04 (9e-04)	-0.0081 (0.0283)	-0.008 (0.0275)	0.1056 (0.0659)	0.1017 (0.0657)	0.113 (0.0606)	0.1082 (0.06)
$TFP5_{i,t-1}$	-0.131* (0.0619)	-0.141* (0.06)	-0.1901* (0.0755)	-0.1992** (0.072)	0.0031** (0.0011)	0.0031** (0.001)	0.1326** (0.0436)	0.1326** (0.0417)	0.1535 (0.1152)	0.131 (0.1123)	0.1157 (0.1093)	0.0984 (0.1063)
R^2	0.1555	0.1539	0.6834	0.6837	0.2154	0.2159	0.6544	0.655	0.3216	0.3232	0.1477	0.1505
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	-0.0354	-0.0354	3.166	3.166	0.0018	0.0018	0.6928	0.6928	0.2482	0.2482	1.865	1.865

Notes: The regressions aim to identify the drivers of technological change reflected in changing input factor use in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $gr(L_{i,t})$, $Wage_{i,t}$, $(K/L)_{i,t}$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table 5: Regression results: Patterns of changing input factor use

The last set of results explores the interaction between the two-layered network and the structural characteristics of production, i.e. labor inputs, capital intensity, and investments. These characteristics are indicative about a potential factor bias of technological change.

The results show a significant negative relationship between innovation encoded in $A_{i,t-1}^{\tau}$ and labor, both in levels and growth rates. This effect is moderated by knowledge spillovers in the innovation layer which show a positive correlation with labor. Further, the capital intensity is negatively associated with labor and the share of production workers $(L_{prod}/L)_{i,t-1}$ positively. The results also show that labor productivity $(VA/L)_{i,t-1}$ and investments per employee $(I/L)_{i,t-1}$ have a positive association with labor.

Again, an interesting difference between up- and downstream centrality in the market is observed: a higher centrality in the customer network $PR_{i,t-1}^{\mu,out}$ shows a positive association with industrial labor demand and per capita investments, while the opposite holds for upstream centrality. The effects on labor are statistically less significant, but qualitatively robust. Again, the findings are consistent across a series of robustness checks (see [C.1](#) and [SI.2.2](#)).

6. Discussion

The key results from this study can be wrapped up as follows:

1. During period of study, both network layers exhibit different dynamics: The innovation layer became increasingly connected and the size distribution of industries increasingly skewed. In the market, connectivity is higher and more stable. Industries became more similar and increasingly connected by patent citation patterns, but there is no clear trend in the market. The increasing patent citation similarity may indicate a convergence of industrial R&D. Generally, the size ranking in the innovation layer is clearly dominated by ICT-related and other electronics industries, especially during the last decades. These sectors are also most central by their citation patterns. In contrast, the picture of the top ranks in the market is more divers. The top rank is clearly dominated by the petroleum industry, but the ranks below frequently change, and also many chemical and natural product manufacturing industries score high by market size.
2. The results reveal patterns of path-dependence in the evolution of market size and innovation outputs: growing industries continue to grow even though the

growth rate is diminishing. Path dependence is also reflected in the evolution of productivity indicators and input use. However, growth in the innovation layer are more strongly auto-correlated compared to growth in the market. This is in line with the changing concentration pattern.

3. The regressions show a differential role of up- versus downstream centrality in the market. The results show that upstream centrality is negatively associated with growth in the market, labor productivity, and investments. However, there is a weak indication that it may be positively associated with innovation. In contrast, downstream centrality in the market $PR_{i,t}^{\mu,out}$ is positively associated with growth in the market, labor productivity, and investments per capita. A positive but less significant relationship is also found for TFP and labor demand. The positive effects of $PR_{i,t-1}^{\mu,out}$ on industry performance can be interpreted as support for demand-pull (Type 2) from the market.
4. Centrality in the innovation layer $PR_{i,t}^{\tau,d}$ exhibits different effects. It does not show any qualitative difference between up- and downstream centrality. The patent centrality shows a positive association with growth in the market and an only weakly significant positive correlation with investments. This can be interpreted as empirical support for technology-push (Type 1) when patented innovation triggers growth in the market. However, the results do not indicate any effects of patent centrality on productivity or other variables of interest. Generally, the results indicate that more innovation $A_{i,t-1}^{\tau}$ is associated with factor biased technological change reflected in a lower demand for labor and higher capital intensity.
5. The next noteworthy observation is the role of knowledge spillovers from innovation in technologically similar industries. The results provide empirical support for the existence of positive knowledge spillover from upstream industries as driver of innovation. Also knowledge spillovers from downstream industries are supportive for innovation, though less significant. This is in line with previous research (e.g. Acemoglu et al., 2016; Antony and Grebel, 2012; Jaffe, 1986). Knowledge spillovers are also positively associated with labor demand, but do not show any significant interaction with productivity.
6. The results also reveal further noteworthy interactions between factor inputs and patterns of technological change. For example, investments and TFP exhibit a positive correlation with innovation.

6.1. Demand-pull or technology-push?

The answer to this question is complex, and as many other authors suggested before (e.g. Kline and Rosenberg, 1986; Saviotti and Pyka, 2013; Mowery and Rosenberg, 1979; Cohen, 2010): the evolution of markets and technology is interdependent and both, demand-pull and technology-push can be drivers of change.

The analysis offers support for technology-push of Type 1: industries that become more central in the innovation layer grow in the market. This type of technology-push is factor biased as higher innovation outputs are associated with a lower demand for labor and a higher capital intensity. It should be noted that a higher patent centrality does not necessarily imply more patents but both are strongly correlated. The factor bias of technology-push from the patent network may be moderated or even offset if industries benefit from knowledge spillovers which stimulate further innovation.

Across layers, there is no empirical support that demand-push from the market has an impact on innovation outputs. There is only a weak indication that industries investing in physical production capital grow by their output of patents.

The second type of pull and push builds on the distinction between up- and downstream links in the network layers and measures pull and push within the same layer. The results offer support for the existence of demand-pull (Type 2) that arises from customer links in the market. A central position in the downstream network is positively associated with growth in the market, labor productivity, labor demand, and investments, but it shows no effect on innovation. The downstream centrality in the market indicates that an industry has many important customers who themselves have many important customers. An increase in an industry's downstream centrality indicates that it became more relevant as input supplier.⁷

Upstream centrality in the market shows a negative effects on market growth, productivity, and investment. These upstream dynamics can not be directly interpreted as technology-push effect but may be informative about the availability of inputs and diversification of upstream industries. The negative effect of the upstream centrality in the market indicates that the specialization of upstream suppliers has a weakly negative association with downstream growth.

Dynamics in the patent network are different: both up- and downstream connections show qualitatively similar effects. This is not surprising as up- and downstream

⁷Note that this paper does not make an assessment of market rivalry as Bloom et al. (2013) did. Their study relies on a similar measure of proximity in the product space (here $\sigma_{i,t}^{\mu, out}$) and the authors used this to evaluate the impact of competitors' R&D efforts on firm performance. Here, downstream spillovers $Spill(A)_{i,t}^{\mu, out}$ in the market measure what happens if the industries that serve the needs of similar customers grow but this does not show any significant effect.

centrality measures are highly correlated. The results show that innovation outputs are positively associated with up- and downstream spillovers, though the downstream spillovers are only poorly significant. Spillovers are also positively associated with labor demand and weakly negatively with the capital intensity. Changes in the level of upstream spillovers may be interpreted as a technology-push effect (Type 2) within the patent network as they indicate the availability of knowledge and technological opportunities that are available from upstream industries.

Downstream spillovers, in contrast, show that industries producing knowledge that is useful for similar purposes grow. This can be broadly understood as a demand-pull effect as similarities by user pattern indicates an increasing use of the knowledge produced an industry. However, it should be noted that citations in a patent do not necessarily indicate the actual use of a patent, but also reflect the need to declare the state of prior art as a legal requirement of a patent (Jaffe and De Rassenfosse, 2019).

One important caveat related to the use of up- and downstream centrality measures needs to be mentioned: the network data is limited to manufacturing industries and does not incorporate the network structure of the rest of the economy (e.g. input links to primary resource providers and output links to end users and customers). Controlling for the supply chain position of an industry is beyond the scope of this analysis. It may be captured by the FE approach, but it would be an important test in future analyses using similar data.

6.2. Limitations and research implications

This study is subject to three major blocks of limitations.

First, patents as a measure of innovation are imperfect: the use of patents to protect IP varies across industries (Fontana et al., 2013; Arundel and Kabla, 1998; Cohen et al., 2000), and patents vary greatly by value and not every patent indicates a technological breakthrough (Trajtenberg, 1990; Kogan et al., 2017). Sometimes patents are only filed for defensive purposes to protect a pre-existing, but not a new invention (Granstrand, 1999). Over time citation practices may have changed, not least because of the improved computer-assisted search techniques (Hall et al., 2005; Marmor, 1980). These limitation are partly addressed by restricting the sample to manufacturing where patents are a common means of IP protection (Blank and Kappos, 2012), by controlling for industry and time FE, and by using citation-weighted patents.

Second, studying innovation and industrial evolution over time is challenging because of non-static classification systems (Marmor, 1980; Yuskavage et al., 2007; Lafond

and Kim, 2019). This analysis relies on industry codes that are purposely designed to describe industries by their production processes. NAICS is designed as a means for the description of industries and regularly (quasi-endogenously) updated to meet this purpose. This can be one explanation for the less skewed sector-size distribution, and possibly also for the higher stability of the IO network. In the regressions, I controlled for industry and time FE hoping to capture potential distortions.

Aim of this analysis is the economic study of technological change. This justifies the choice of NAICS codes rather than patent classes. Patents needed to be mapped to industries based on their technological classification. Inferring from patents to industrial dynamics is a challenging endeavor (Antonelli, 2014; Dosi and Nelson, 2010), not least because the industry where a patent is filed is not necessarily the same industry where the patented invention is used. A variety of concordances that allow the mapping from patents to industries have been proposed (e.g., Lybbert and Zolas, 2014; Van Looy et al., 2014; Goldschlag et al., 2020; Dorner and Harhoff, 2018). A systematic, dynamic comparison between these concordances and their implications for economic research is an interesting and methodologically valuable avenue for future research. It would be also interesting to compare the results of this study with an approach using patent-classes as means of description: classifying IO flows by their correspondence in patent-classes can be insightful to understand the impact of demand-pull on the dynamics of patented innovations. But this is beyond the scope of this paper.

Finally, here, I studied demand-pull and technology-push at the aggregate level. But patterns of innovation, knowledge sources, and IP practices differ across firms, industries, and technology fields (Pavitt, 1984; Carlsson and Stankiewicz, 1991; Blank and Kappos, 2012). Walsh (1984) documented that whether demand-pull or technology-push dominate may be a matter of industry maturity. The static dimension of sector heterogeneity is captured by the FE approach in the regressions and the analysis is limited to industries that have non-zero patent counts in all periods. Walsh (1984) highlighted in a study of the chemical industry that technology-push from radical breakthroughs may drive growth in the market which in turn creates demand-pull effects that induce incremental innovation. Future research within a similar methodological framework may take account of the distinction between radical and incremental innovation and the chronology of the technology cycle.

It should be also noted that this study is limited to manufacturing. Various studies have documented the decline of US manufacturing since the 1980s (e.g. Elsby et al., 2013; Fort et al., 2018). It would be important to verify whether the observed patterns are general or unique for the US manufacturing sector during the four decades of study.

Another avenue for future research may address the role of input scarcity as driver of innovation. The analysis in this paper provides only weak support. A preceding version of this paper that included a section on link formation processes suggested that input scarcity may be a determinant of cross-industrial link formation Hötte (2021). This is important because it may improve our understanding of the innovation and market impact of policies like a carbon tax that operate through the channel of input costs.

7. Conclusions

In this paper, I introduced two conceptual types of technology-push and demand-pull as drivers of technological change. The first type builds on interactions from patented innovations to the evolution of markets as technology-push, and vice versa as demand-pull. The second type builds on the distinction of up- and downstream connections in the coupled market and innovation network.

The results indicate that both, demand-pull and technology-push are drivers of technological change, but they are differently important for the evolution of markets, factor input use, and innovation. Technology-push from patented innovation can be a driver of market growth, but this shows a factor bias in favor of capital. Further, the results indicate that within the patent network, spillovers across industries are supportive for innovation which may explain an increasing concentration and clustering in the patent network. Patent spillovers are associated with higher labor demand which may moderate the labor-saving effect of technology-push.

Further, I find support for demand-pull within the market associated with industrial growth, labor productivity, and investment. Nevertheless, market growth may be constrained by the availability of production inputs. Qualitatively, this is in line with rather fluctuating evolution of relative industry sizes over time, which contrasts with the clear trend towards concentration in the innovation layer.

The results bear important insights for the debate whether or not technological change is labor-saving. The results indicate that the sources of technological change, i.e. whether it is pushed by new technological opportunities or pulled from customers in the market. I find that innovation driven by patented innovation tends to be labor-saving but this effect may be moderated or even offset in the presence of knowledge spillovers. In contrast, industry growth that is driven by demand-push in the market is associated with a higher demand for labor, more investments, and higher labor productivity.

The results in this study should not be overstated as causal evidence in face of conceptual and methodological limitations whose solution is beyond the scope of this

study. Rather, this study reports a series of - until now - unrevealed patterns of correlation which provide a rich basis for future research.

Acknowledgements

I want to thank Angelo Secchi and Antoine Mandel who facilitated this work at an early stage. Further, many thanks to Herbert Dawid, the participants of the Annual GENED meeting 2020, the INET Complexity group meeting, and the OMPTEC-FoW seminar for helpful comments. Further gratitude is owed to Cord Wiljes and the Competence Center for Research Data at Bielefeld University for data management support, and to Angelos Theodorakopoulos for scientific advice and encouraging words. I also want to acknowledge very valuable feedback by three anonymous reviewers. Moreover, I gratefully acknowledge the financial support by the German Academic Foundation, Deutsch-Französische Hochschule, Bielefeld Graduate School of Economics and Management, and from the EU Horizon 2020 Research and Innovation programme under grant agreement No 822330 TECHNEQUALITY.

References

- Acemoglu, Daron (2002). “Directed technical change”. In: *The Review of Economic Studies* 69.4, pp. 781–809. DOI: [10.1111/1467-937X.00226](https://doi.org/10.1111/1467-937X.00226).
- Acemoglu, Daron, Ufuk Akcigit, and William R Kerr (2016). “Innovation network”. In: *Proceedings of the National Academy of Sciences*, p. 201613559. DOI: [10.1073/pnas.1613559113](https://doi.org/10.1073/pnas.1613559113).
- Antonelli, Cristiano (2014). *The economics of innovation, new technologies and structural change*. Routledge.
- Antony, Jürgen and Thomas Grebel (2012). “Technology flows between sectors and their impact on large-scale firms”. In: *Applied Economics* 44.20, pp. 2637–2651. DOI: [10.1080/00036846.2011.566191](https://doi.org/10.1080/00036846.2011.566191).
- Arthur, W Brian (1989). “Competing technologies, increasing returns, and lock-in by historical events”. In: *The Economic Journal* 99.394, pp. 116–131. DOI: [10.2307/2234208](https://doi.org/10.2307/2234208).
- Arundel, Anthony and Isabelle Kabla (1998). “What percentage of innovations are patented? Empirical estimates for European firms”. In: *Research Policy* 27.2, pp. 127–141. DOI: [10.1016/S0048-7333\(98\)00033-X](https://doi.org/10.1016/S0048-7333(98)00033-X).
- Atalay, Enghin, Ali Hortacsu, James Roberts, and Chad Syverson (2011). “Network structure of production”. In: *Proceedings of the National Academy of Sciences* 108.13, pp. 5199–5202. DOI: [10.1073/pnas.1015564108](https://doi.org/10.1073/pnas.1015564108).
- Bartlesman, Eric and Wayne B Gray (1996). *The NBER manufacturing productivity database*. National Bureau of Economic Research Technical Working Paper 0205. DOI: [10.3386/t0205](https://doi.org/10.3386/t0205).
- Becker, Randy, Wayne Gray, and Jordan Marvakov (2013). *NBER-CES manufacturing industry database: Technical notes*. National Bureau of Economic Research Working Paper No. 5809. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.368.671&rep=rep1&type=pdf>.
- Blank, Rebecca M and David J Kappos (2012). *Intellectual property and the US economy: Industries in focus*. Economics and Statistics Administration & US Patent and Trademark Office.
- Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013). “Identifying technology spillovers and product market rivalry”. In: *Econometrica* 81.4, pp. 1347–1393. DOI: [10.3982/ECTA9466](https://doi.org/10.3982/ECTA9466).

- Boehm, Johannes, Swati Dhingra, and John Morrow (2019). *The comparative advantage of firms*. CEPR Discussion Paper No. DP13699. Accessed on 17 Dec. 2020. Centre for Economic Policy Research. URL: <https://ssrn.com/abstract=3379615>.
- Brin, Sergey and Lawrence Page (1998). “The anatomy of a large-scale hypertextual web search engine”. In: *Computer networks and ISDN systems* 30.1-7, pp. 107–117. DOI: [10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X).
- Brynjolfsson, Erik and Andrew McAfee (2012). *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Digital Frontier Press.
- Buerger, Matthias, Tom Broekel, and Alex Coad (2012). “Regional dynamics of innovation: Investigating the co-evolution of patents, research and development (R&D), and employment”. In: *Regional Studies* 46.5, pp. 565–582. DOI: [10.1080/00343404.2010.520693](https://doi.org/10.1080/00343404.2010.520693).
- Cai, Jie, Nan Li, and Ana Maria Santacreu (2017). *Knowledge Diffusion, Trade and Innovation across Countries and Sectors*. Working Paper 2017-029A. Federal Reserve Bank of St. Louis Research Division. DOI: [10.20955/wp.2017.029](https://doi.org/10.20955/wp.2017.029).
- Carlsson, Benny and Rikard Stankiewicz (1991). “On the nature, function and composition of technological systems”. In: *Journal of Evolutionary Economics* 1.2, pp. 93–118. DOI: [10.1007/BF01224915](https://doi.org/10.1007/BF01224915).
- Carvalho, Vasco M (2014). “From micro to macro via production networks”. In: *Journal of Economic Perspectives* 28.4, pp. 23–48. DOI: [10.1257/jep.28.4.23](https://doi.org/10.1257/jep.28.4.23).
- Carvalho, Vasco M and Nico Voigtländer (2014). *Input diffusion and the evolution of production networks*. Tech. rep. National Bureau of Economic Research. DOI: [10.3386/w20025](https://doi.org/10.3386/w20025).
- Cohen, Wesley M (2010). “Fifty years of empirical studies of innovative activity and performance”. In: *Handbook of the Economics of Innovation*. Vol. 1. Elsevier, pp. 129–213.
- Cohen, Wesley M and Daniel A Levinthal (1989). “Innovation and learning: the two faces of R & D”. In: *The economic journal* 99.397, pp. 569–596. DOI: [10.2307/2233763](https://doi.org/10.2307/2233763).
- (2000). *Absorptive capacity: A new perspective on learning and innovation*. DOI: [10.2307/2393553](https://doi.org/10.2307/2393553).
- Cohen, Wesley M, Richard Nelson, and John P Walsh (2000). *Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)*. National Bureau of Economic Research Working Paper No. 7552. DOI: [10.3386/w7552](https://doi.org/10.3386/w7552).

- Csardi, Gabor, Tamas Nepusz, et al. (2006). “The igraph software package for complex network research”. In: *InterJournal, complex systems* 1695.5, pp. 1–9. URL: <https://igraph.org>.
- Cysouw, Michael (2018). *qlcMatrix: Utility Sparse Matrix Functions for Quantitative Language Comparison*. R package version 0.9.7. URL: <https://CRAN.R-project.org/package=qlcMatrix>.
- Di Stefano, Giada, Alfonso Gambardella, and Gianmario Verona (2012). “Technology push and demand pull perspectives in innovation studies: Current findings and future research directions”. In: *Research Policy* 41.8, pp. 1283–1295. DOI: [10.1016/j.respol.2012.03.021](https://doi.org/10.1016/j.respol.2012.03.021).
- Dorner, Matthias and Dietmar Harhoff (2018). “A novel technology-industry concordance table based on linked inventor-establishment data”. In: *Research Policy* 47.4, pp. 768–781. DOI: [10.1016/j.respol.2018.02.005](https://doi.org/10.1016/j.respol.2018.02.005).
- Dosi, Giovanni and Richard R Nelson (2010). “Technical change and industrial dynamics as evolutionary processes”. In: *Handbook of the Economics of Innovation*. Vol. 1. Elsevier, pp. 51–127. DOI: [10.1016/S0169-7218\(10\)01003-8](https://doi.org/10.1016/S0169-7218(10)01003-8).
- Elsby, Michael WL, Bart Hobijn, and Ayşegül Şahin (2013). “The decline of the US labor share”. In: *Brookings Papers on Economic Activity* 2013.2, pp. 1–63. DOI: [10.1353/eca.2013.0016](https://doi.org/10.1353/eca.2013.0016).
- Fontana, Roberto, Alessandro Nuvolari, Hiroshi Shimizu, and Andrea Vezzulli (2013). “Reassessing patent propensity: Evidence from a dataset of R&D awards, 1977–2004”. In: *Research Policy* 42.10, pp. 1780–1792. DOI: [10.1016/j.respol.2012.05.014](https://doi.org/10.1016/j.respol.2012.05.014).
- Fort, Teresa C, Justin R Pierce, and Peter K Schott (2018). “New perspectives on the decline of US manufacturing employment”. In: *Journal of Economic Perspectives* 32.2, pp. 47–72. DOI: [10.1257/jep.32.2.47](https://doi.org/10.1257/jep.32.2.47).
- Geroski, Paul A and Chris F Walters (1995). “Innovative activity over the business cycle”. In: *The Economic Journal* 105.431, pp. 916–928. DOI: [10.2307/2235158](https://doi.org/10.2307/2235158).
- Goldschlag, Nathan, Travis J Lybbert, and Nikolas J Zolas (2020). “Tracking the technological composition of industries with algorithmic patent concordances”. In: *Economics of Innovation and New Technology* 29.6, pp. 582–602. DOI: [10.1080/10438599.2019.1648014](https://doi.org/10.1080/10438599.2019.1648014).
- Granstrand, Ove (1999). *The economics and management of intellectual property*. Edward Elgar Publishing.
- Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg (2005). “Market value and patent citations”. In: *RAND Journal of economics*, pp. 16–38. URL: <https://www.jstor.org/stable/1593752>.

- Horowitz, Karen and Mark Planting (2006). *Concepts and Methods of the U.S. Input-Output Accounts. Measuring the Nation's Economy*. Bureau of Economic Analysis (BEA), U.S. Department of Commerce: Input-Output Manual. URL: https://apps.bea.gov/papers/pdf/I0manual_092906.pdf.
- Hötte, Kerstin (2021). *Data publication: Demand-pull and technology-push: What drives the direction of technological change? Empirical data on a coupled two-layer input-output and patent-citation network*. Data Publication. Bielefeld University. DOI: [10.4119/unibi/2952814](https://doi.org/10.4119/unibi/2952814).
- Hötte, Kerstin (2021). *Demand-pull and technology-push: What drives the direction of technological change?—An empirical network-based approach*. Tech. rep. arXiv preprint arXiv:2104.04813.
- Hötte, Kerstin, Anton Pichler, and François Lafond (2021). *Data Publication: The scientific knowledge base of low carbon energy technologies*. Data Publication. Bielefeld University. DOI: [10.4119/unibi/2950291](https://doi.org/10.4119/unibi/2950291).
- Hötte, Kerstin, Anton Pichler, and François Lafond (2021). “The rise of science in low-carbon energy technologies”. In: *Renewable and Sustainable Energy Reviews* 139, p. 110654. DOI: [10.1016/j.rser.2020.110654](https://doi.org/10.1016/j.rser.2020.110654).
- Huang, Jingong Jim (2018). “Technology Network, Innovation and Growth”. In: *2018 Meeting Papers*. 178. Society for Economic Dynamics.
- IPCC (2018). “Summary for Policymakers”. In: *An IPCC Special Report on the impacts of global warming of 1.5 degree C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, Summary for Policymakers*. Ed. by V Masson-Delmotte, P Zhai, HO Pörtner, D Roberts, J Skea, P R Shukla, A Pirani, W Moufouma-Okia, C Péan, R Pidcock, S Connors, J B R. Matthews, Y Chen, X Zhou, M I Gomis, E Lonnoy, T Maycock, M Tignor, and T Waterfield. World Meteorological Organization, Geneva, Switzerland.
- Jackson, Matthew O (2008). *Social and economic networks*. Princeton University Press.
- Jaffe, Adam B (1986). “Technological Opportunity and Spillovers of R&D: Evidence from Firms’ Patents, Profits, and Market Value”. In: *The American Economic Review* 76.5, pp. 984–1001. URL: <http://www.jstor.org/stable/1816464>.
- (1988). “Demand and supply influences in R&D intensity and productivity growth”. In: *The Review of Economics and Statistics*, pp. 431–437. DOI: [10.2307/1926781](https://doi.org/10.2307/1926781).
- Jaffe, Adam B and Gaétan De Rassenfosse (2019). “Patent citation data in social science research: Overview and best practices”. In: *Research handbook on the economics of*

- intellectual property law*. Edward Elgar Publishing. DOI: [10.4337/9781789903997.00043](https://doi.org/10.4337/9781789903997.00043).
- Kay, Luciano, Nils Newman, Jan Youtie, Alan L Porter, and Ismael Rafols (2014). “Patent overlay mapping: Visualizing technological distance”. In: *Journal of the Association for Information Science and Technology* 65.12, pp. 2432–2443. DOI: [10.1002/asi.23146](https://doi.org/10.1002/asi.23146).
- Kline, Stephen J and Nathan Rosenberg (1986). “An overview of innovation”. In: *The positive sum strategy: Harnessing technology for economic growth*. The National Academy of Science, USA.
- Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017). “Technological innovation, resource allocation, and growth”. In: *The Quarterly Journal of Economics* 132.2, pp. 665–712. DOI: [10.1093/qje/qjw040](https://doi.org/10.1093/qje/qjw040).
- Kortum, Samuel and Jonathan Putnam (1997). “Assigning patents to industries: tests of the Yale technology concordance”. In: *Economic Systems Research* 9.2, pp. 161–176. DOI: [10.1080/09535319700000011](https://doi.org/10.1080/09535319700000011).
- Kymn, Kern O (1990). “Aggregation in input–output models: a comprehensive review, 1946–71”. In: *Economic Systems Research* 2.1, pp. 65–93. DOI: [10.1080/09535319000000008](https://doi.org/10.1080/09535319000000008).
- Lafond, François and Daniel Kim (2019). “Long-run dynamics of the US patent classification system”. In: *Journal of Evolutionary Economics* 29.2, pp. 631–664. DOI: [10.1007/s00191-018-0603-3](https://doi.org/10.1007/s00191-018-0603-3).
- Leydesdorff, Loet (2005). “Similarity measures, author cocitation analysis, and information theory”. In: *Journal of the American society for Information Science and Technology* 56.7, pp. 769–772. DOI: [10.1002/asi.20130](https://doi.org/10.1002/asi.20130).
- Lybbert, Travis J and Nikolas J Zolas (2014). “Getting patents and economic data to speak to each other: An ‘algorithmic links with probabilities’ approach for joint analyses of patenting and economic activity”. In: *Research Policy* 43.3, pp. 530–542. DOI: [10.1016/j.respol.2013.09.001](https://doi.org/10.1016/j.respol.2013.09.001).
- Marmor, Alfred C (1980). “The Approach of the United States Patent and Trademark Office to Finding Prior Art”. In: *Journal of Chemical Information and Computer Sciences* 20.1, pp. 6–9. DOI: [10.1021/ci60021a003](https://doi.org/10.1021/ci60021a003).
- McCune, Bruce, James B Grace, and Dean L Urban (2002). *Analysis of ecological communities*. Vol. 28. MjM software design Gleneden Beach, OR.
- McNerney, James, Charles Savoie, Francesco Caravelli, and J Doyne Farmer (2018). *How production networks amplify economic growth*. Tech. rep. arXiv preprint arXiv:1810.07774.

- Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean (2013). “Efficient estimation of word representations in vector space”. In: *arXiv preprint arXiv:1301.3781*.
- Mowery, David and Nathan Rosenberg (1979). “The influence of market demand upon innovation: a critical review of some recent empirical studies”. In: *Research Policy* 8.2, pp. 102–153. DOI: [10.1016/0048-7333\(79\)90019-2](https://doi.org/10.1016/0048-7333(79)90019-2).
- Myers, Sumner and Donald G Marquis (1969). *Successful industrial innovation*, pp. 17–69.
- Nelson, Richard R (1994). “The co-evolution of technology, industrial structure, and supporting institutions”. In: *Industrial and corporate change* 3.1, pp. 47–63. DOI: [10.1093/icc/3.1.47](https://doi.org/10.1093/icc/3.1.47).
- Nemet, Gregory F (2009). “Demand-pull, technology-push, and government-led incentives for non-incremental technical change”. In: *Research Policy* 38.5, pp. 700–709. DOI: [10.1016/j.respol.2009.01.004](https://doi.org/10.1016/j.respol.2009.01.004).
- OECD (2001). *Measuring productivity-OECD Manual: Measurement of Aggregate and Industry-Level Productivity Growth*. Tech. rep. Organisation for Economic Cooperation and Development (OECD). DOI: [10.1787/9789264194519-en](https://doi.org/10.1787/9789264194519-en). URL: <http://dx.doi.org/10.1787/9789264194519-en>.
- (2009). “Chapter 6: The Use and Analysis of Citations in Patents”. In: *OECD Patent Statistics Manual*. OECD, pp. 105–123. DOI: [10.1787/9789264056442-en](https://doi.org/10.1787/9789264056442-en).
- Pakes, Ariel and Mark Schankerman (1984). “9. An Exploration into the Determinants of Research Intensity”. In: *R&D, Patents and Productivity*. Ed. by Zvi Griliches. University of Chicago Press. DOI: [10.7208/9780226308920-012](https://doi.org/10.7208/9780226308920-012).
- Pavitt, Keith (1984). “Sectoral patterns of technical change: Towards a taxonomy and a theory”. In: *Research Policy* 13, pp. 343–373. DOI: [10.1016/0048-7333\(84\)90018-0](https://doi.org/10.1016/0048-7333(84)90018-0).
- Romer, Paul M (1990). “Endogenous technological change”. In: *Journal of political Economy* 98.5, Part 2, pp. 71–102. DOI: [10.1086/261725](https://doi.org/10.1086/261725).
- Rosenberg, Nathan (1982). *Inside the black box: technology and economics*. Cambridge University Press.
- Ruttan, Vernon W (1959). “Usher and Schumpeter on invention, innovation, and technological change”. In: *The Quarterly Journal of Economics*, pp. 596–606. DOI: [10.2307/1884305](https://doi.org/10.2307/1884305).
- Saviotti, Pier Paolo (1997). “Black boxes and variety in the evolution of technologies”. In: *Economics of structural and technological change*. Routledge, pp. 195–223.

- Saviotti, Pier Paolo and Andreas Pyka (2013). “The co-evolution of innovation, demand and growth”. In: *Economics of Innovation and New technology* 22.5, pp. 461–482. DOI: [10.1080/10438599.2013.768492](https://doi.org/10.1080/10438599.2013.768492).
- Schmoch, Ulrich, Francoise Laville, Pari Patel, and Rainer Frietsch (2003). *Linking technology areas to industrial sectors*. Final Report to the European Commission, DG Research. European Commission.
- Schmookler, Jacob (1966). *Invention and economic growth*. Harvard University Press. DOI: [10.4159/harvard.9780674432833](https://doi.org/10.4159/harvard.9780674432833).
- Taalbi, Josef (2020). “Evolution and structure of technological systems-An innovation output network”. In: *Research Policy* 49.8, p. 104010. DOI: [10.1016/j.respol.2020.104010](https://doi.org/10.1016/j.respol.2020.104010).
- Trajtenberg, Manuel (1990). “A penny for your quotes: patent citations and the value of innovations”. In: *The Rand journal of economics*, pp. 172–187. DOI: [10.2307/2555502](https://doi.org/10.2307/2555502).
- Van Looy, Bart, Caro Vereyen, and Ulrich Schmoch (2014). *Patent Statistics: Concor- dance IPC V8–NACE Rev. 2*. Report. Eurostat, European Commission.
- Von Hippel, Eric (1976). “The dominant role of users in the scientific instrument innovation process”. In: *Research Policy* 5.3, pp. 212–239. DOI: [10.1016/0048-7333\(76\)90028-7](https://doi.org/10.1016/0048-7333(76)90028-7).
- Walsh, Vivien (1984). “Invention and Innovation in the Chemical Industry: Demand-pull or Discovery-push?” In: *Research Policy* 13.4, pp. 211–234. DOI: [10.1016/0048-7333\(84\)90015-5](https://doi.org/10.1016/0048-7333(84)90015-5).
- Yuskavage, Robert E et al. (2007). “Converting historical industry time series data from SIC to NAICS”. In: *The Federal Committee on Statistical Methodology 2007 Research Conference. 5-7 November 2007*.

APPENDIX

A. Methods

A.1. Cosine similarity

The cosine similarity used in the main text is given by the normalized dot product of the vectors $\mathbf{w}_{k,t}^{\alpha,d} = (w_{k1,t}^{\alpha,d}, w_{k2,t}^{\alpha,d}, \dots, w_{kN,t}^{\alpha,d})$ of $k = i, j \in N$. It measures the angle between the two in- or output share vectors $\mathbf{w}_{i,t}^{\alpha,d}$ and $\mathbf{w}_{j,t}^{\alpha,d}$ on layer α in time t normalized to the length one. It is given by the formula:

$$\sigma_{ij,t}^{\alpha,d} = \frac{(\mathbf{w}_{i,t}^{\alpha,d})^T \mathbf{w}_{j,t}^{\alpha,d}}{\sqrt{((\mathbf{w}_{i,t}^{\alpha,d})^T \mathbf{w}_{i,t}^{\alpha,d})((\mathbf{w}_{j,t}^{\alpha,d})^T \mathbf{w}_{j,t}^{\alpha,d})}}.$$

where the T in the superscript indicates that the vector is transposed. For the analysis, additional test with alternative similarity metrics such as the inverse of the Euclidean and the Canberra distance. The cosine was finally chosen because tests indicated a high explanatory power and because it is very commonly used in network-based technology studies (e.g. Hötte et al., 2021; Kay et al., 2014; McCune et al., 2002; Leydesdorff, 2005; Mikolov et al., 2013).

A.2. Data compilation

A.2.1. Input-output data

BEA provides detailed current and historical benchmark IO tables in a quinquennial frequency dating back to 1947.⁸ I used the most disaggregate data at the 6-digit level. The data are accounting data which show monetary flows between industries including final demand, and dummy positions that ensure the financial closure. Accounting positions are largely but not perfectly compatible with NAICS or Standard Industrial Classification (SIC) codes. I converted the data step-wise into a time-consistent and convenient format. First, the data are transformed from accounting positions into industry codes, i.e. SIC codes for the 1977-1987 data and into NAICS for later periods.

⁸The data were downloaded from <https://www.bea.gov/industry/benchmark-input-output-data> and <https://www.bea.gov/industry/historical-benchmark-input-output-tables> [Both accessed in Oct 2021].

The industry codes are harmonized to the NAICS 2002 version using concordance tables provided by BEA.⁹

After harmonizing the data, I obtained for each period a matrix of monetary flows between 1179 distinct 6-digit NAICS industries. The entries of the matrix are input flows $flow_{ij,t}^{\mu,in}$ indicating the monetary value of the inputs that i buys from j in time t . Division of the flows by the row sums $\sum_j flow_{ij,t}^{\mu,d}$ gives the input shares $w_{i,t}^{\mu,in}$. The output shares $w_{i,t}^{\mu,out}$ are obtained by division by column sums. Note that some rows and columns are empty for some t . This results from the harmonization procedure to uniform NAICS codes and can happen when the classification changes. Industries can emerge or disappear over time. For example, industries associated with computer technologies were less granular in the 70s compared to the 90s. This is often associated with a split (merge) of pre-existing industries.

A.2.2. Patent data

The patent citation layer is taken from a data set compiled for an earlier project (Hötte et al., 2021; Hötte et al., 2021). The data contains a list of USPTO patents including grant year and CPC technology classes. For this project, the Master Classification File from the USPTO from January 2020 is used.¹⁰ This file offers a list that maps individual patents to one or more CPC classes at the most disaggregate level.

To obtain NAICS level patent data, the mapping from patents to CPC classes and the mapping from CPC to NAICS codes are merged. The CPC to NAICS mapping by Goldschlag et al. (2020) is a probabilistic mapping and that comes along with probability weights whenever the one CPC class maps to multiple NAICS. These weights are used when compiling industry level patent stocks and industry-to-industry citation counts. To compile the patent stocks, weighted patents per NAICS class are aggregated for a given time window.

I used the time windows prior to the benchmark year. For example, for the patent stock in 1977, all patents granted in 1973-1977 are summed up. However, one could also argue to use the subsequent time window 1977-1981 to compile patents for 1977. I used granted patents and the time lag between patent application and grant often accounts for a few years. Further, the IO data is a time snapshot of the last year in the time window. Here, I used the earlier time window for three major reasons: (1) Innovation is a dynamic concept comprising the process of invention, innovation and

⁹Detailed explanations of conceptual and technical issues (e.g. changing classification systems, ambiguous mappings) that arose during the compilation are available in SI.1.1.

¹⁰<https://bulkdata.uspto.gov/data/patent/classification/>

commercialization, and diffusion. Using the earlier time window takes account of the diffusion lag. (2) Patents are seen as a proxy for the stock of available technological knowledge and patents that will be granted in future are not yet available as knowledge for current use. (3) This approach is consistent with other research where discounted patent stocks were used as proxies of innovation and technological knowledge (e.g. Antony and Grebel, 2012; Huang, 2018).

The same procedure is applied to the citation data, where both the citing and the cited patent both are mapped to NAICS codes. In numbers, more than 37.66 M citation links between 3.75 M individual patents are first expanded to the CPC 4-digit level and then aggregated into citation counts for each NAICS-to-NAICS pair in the relevant time period.

These NAICS-to-NAICS citation counts are transformed into a symmetric matrix where the entries $flow_{ij,t}^\tau$ correspond to the flow of citations from i to j , i.e. the number of times that i cites patents from industry j . As above, the entries of $flow_t^\tau$ are transformed into input shares $w_{ij,t}^{\tau,d}$ through division by the row sum $\sum_j flow_{ij,t}^{\tau,d}$. Output shares $w_{i,t}^{\tau,out}$ are obtained by division by column sums.

A.2.3. Supplementary data and processing

The data are supplemented by data from the NBER Manufacturing Productivity Database (Becker et al., 2013; Bartlesman and Gray, 1996).¹¹ More details are provided in A.2.3.

For the main analyses, I used 6-digit level data and the subset of manufacturing industries. More aggregate level and additional data on non-manufacturing sectors are used for robustness checks.¹² Robustness checks with more aggregate data aim to cope with concerns about the reliability of the classification approach, as classification systems change over time and many sequential transformations were necessary.

The data are unbalanced panel data, i.e. some industries have no data entry for output flows or patent counts in some periods. For the main analysis, industries with incomplete coverage were removed. The final data is characterized by $A_{i,t}^\alpha > 0 \forall t, \alpha$. This reduces the sample size from 473 to 307 6-digit manufacturing industries.

The networks (both, cross-industrial flow and share matrices) and the raw patent data are used to construct industry level variables. Using the raw patent data, I compiled aggregate citation-weighted patent stocks $A_{i,t}^\tau$ at the industry level. The

¹¹<https://www.nber.org/research/data/nber-ces-manufacturing-industry-database>

¹²The data are available in the accompanying research data publication.

citation weights are used to control for the heterogeneity of patents by value (e.g. Jaffe and De Rassenfosse, 2019). Non-weighted patent stocks are used in robustness checks.

Using the IO flow network, I extracted the sum of output given by the column sum $A_{i,t}^\mu = \sum_k flow_{ki,t}^{\mu,in}$ as measure for the market size. To cope with potential inconsistency across time, additional robustness checks are made with normalized data dividing all entries by the cross-industry average $\frac{1}{|N|} \sum_j A_{j,t}^\alpha$ for each t . The normalized size measures the size relative to other industries in t and is used to illustrate the evolution of size ranking of industries over time (see Sec. 5.1). In the normalized data, the cross-industry average equals one.

The network data $W_t^{\alpha,d}$ are used to compile a series of network centrality measures, such as the PageRank $PR_{i,t}^{\alpha,d}$, degree $D_{i,t}^{\alpha,d}$, and strength $S_{i,t}^{\alpha,d}$ (Jackson, 2008). The PageRank is used in the main analyses and the other measures are used for robustness checks. However, it should be noted that the properties of the network (e.g. centrality, density, clustering) may change in an unsystematic way when the aggregation level changes (Kymn, 1990).

The weight matrices $W_t^{\alpha,d}$ are further used to compute the cosine similarity matrices $\Sigma_t^{\alpha,d} = \{\sigma_{ij,t}^{\alpha,d}\}_{i,j \in N}$ using a sparsity-robust version by Cysouw (2018). The matrices $\Sigma_{i,t}^{\alpha,d}$ and industry sizes $A_{i,t}^\alpha$ are used to compute cross-industry spillovers $Spill(A)_{i,t}^{\alpha,d}$. Growth rates are calculated as $gr(A_{i,t}^\alpha) = \frac{A_{i,t}^\alpha - A_{i,t-1}^\alpha}{A_{i,t-1}^\alpha}$.

In the robustness checks at other aggregation levels, all measures are re-compiled from the network data at the respective aggregation level as network properties (e.g. centrality, density, clustering) may change in an unsystematic way when the aggregation level changes (Kymn, 1990).

A.3. Regression data issues

	Before transformation				After transformation			
	Mean	Min	Max	Median	Mean	Min	Max	Median
A_i^μ	5447	0.30	480728	2498	7.564	0.26	13.08	7.80
A_i^τ	47145	0.02	2741069	6722	8.422	0.02	14.82	8.77
$gr(A_i^\mu)$	4.429	-0.99	3587	0.27	0.3554	-0.69	6.187	0.24
$gr(A_i^\tau)$	0.0659	-0.98	23.71	0.06	0.0453	-0.68	3.207	0.07
$PR_i^{\mu,in}$	0.0033	0.00	0.2095	0.00	1.015	0.40	5.349	0.74
$PR_i^{\mu,out}$	0.0033	0.00	0.1344	0.00	1.091	0.40	4.908	0.84
$PR_i^{\tau,in}$	0.0033	0.00	0.0699	0.00	1.051	0.40	4.261	0.79
$PR_i^{\tau,out}$	0.0033	0.00	0.0697	0.00	1.052	0.40	4.258	0.81
$gr(PR_i^{\mu,in})$	0.86	-0.99	176.7	0.00	0.2733	-6.90	12.08	0.70
$gr(PR_i^{\mu,out})$	0.6922	-1.00	154.8	0.01	0.5625	-6.90	11.95	2.61
$gr(PR_i^{\tau,in})$	-0.0128	-0.41	0.99	-0.01	-0.9099	-6.01	6.899	-2.46
$gr(PR_i^{\tau,out})$	-0.0119	-0.33	0.7176	-0.01	-0.9067	-5.81	6.577	-2.31
$Spill(A)_i^{\mu,in}$	167850	-129887	938204	125987	2.565	0.12	4.552	2.61
$Spill(A)_i^{\mu,out}$	132206	-30240	805997	77047	2.15	-1.39	4.402	2.13
$Spill(A)_i^{\tau,in}$	3983778	274970	16269824	3258570	5.753	3.35	7.395	5.79
$Spill(A)_i^{\tau,out}$	4085870	340967	16429533	3313036	5.788	3.56	7.405	5.80
$Spill(gr(A))_i^{\mu,in}$	49.86	-3587	1040	21.04	0.0046	-0.00	0.0804	0.00
$Spill(gr(A))_i^{\mu,out}$	93.94	-1.54	3620	10.95	0.0062	-0.00	0.0827	0.00
$Spill(gr(A))_i^{\tau,in}$	2.328	-80.71	64.82	12.44	3e-04	-0.01	0.0065	0.00
$Spill(gr(A))_i^{\tau,out}$	3.554	-79.28	66.56	13.80	4e-04	-0.01	0.0066	0.00
$(VA/L)_i$	105.1	10.46	2404	75.76	4.344	2.44	7.785	4.33
$TFP4_i$	3.932	0.03	275.9	0.96	0.6716	0.03	2.428	0.67
$TFP5_i$	4.136	0.04	326.2	0.96	0.6718	0.04	2.428	0.67
$gr((VA/L)_i)$	0.3218	-0.48	4.309	0.30	0.2649	-0.39	1.669	0.26
$gr(TFP4_i)$	1.006	0.80	3.122	1.00	0.6937	0.59	0.8483	0.69
$gr(TFP5_i)$	1.007	0.80	2.973	1.00	0.6937	0.59	0.848	0.69
L_i	36.43	0.74	469.5	22.67	3.188	0.55	6.154	3.15
$(I/L)_i$	7.125	0.20	221.6	4.42	1.771	0.18	5.405	1.68
$(K/L)_i$	117.2	5.15	1958	75.10	0.6647	0.05	3.025	0.56
$(Lprod/L)_i$	0.716	0.29	0.931	0.74	0.7163	0.29	0.9248	0.74
IP_i	0.118	0.00	0.9687	0.07	0.1039	0.00	0.6774	0.07
$Wage_i$	29.85	5.42	101.4	27.85	3.3	1.86	4.629	3.35
$gr(L_i)$	-0.0442	-0.81	1.078	-0.04	-0.0375	-0.59	0.7314	-0.04
$gr((I/L)_i)$	0.3365	-0.84	6.366	0.26	0.2511	-0.61	1.997	0.23
$gr((K/L)_i)$	0.1877	-0.43	3.701	0.14	0.0019	-0.00	0.0363	0.00
$gr((Lprod/L)_i)$	-0.0077	-0.29	0.3589	-0.01	-0.0072	-0.29	0.3589	-0.01
$gr(IP_i)$	4.912	-1.00	3236	0.23	0.2861	-0.69	8.082	0.21
$gr(Wage_i)$	0.2419	-0.06	0.6664	0.20	0.2123	-0.06	0.5107	0.18
Patents and spillovers - if not weighted by citations								
A_i^τ	2557	0.05	141771	504.30	5.849	0.05	11.86	6.20
$gr(A_i^\tau)$	0.1227	-0.73	4.375	0.10	0.0965	-0.55	1.682	0.10
$Spill(A)_i^{\tau,in}$	221787	15452	787660	181621	2.951	0.93	4.379	2.94
$Spill(A)_i^{\tau,out}$	226416	21750	774061	186994	2.982	1.16	4.362	2.97
$Spill(gr(A))_i^{\tau,in}$	9.086	-22.76	56.54	9.39	9e-04	-0.00	0.0056	0.00
$Spill(gr(A))_i^{\tau,out}$	9.827	-22.44	57.84	9.94	0.001	-0.00	0.0058	0.00
N	w/o trade:	2137	w trade:	1301	w/o trade:	2042	w trade:	1264

Notes: This table shows the overview statistics of the variables included in the regression equations before and after data transformation. The last block of rows shows the data entries of the patent counts when the data is not weighted by patent citations.

Table A.1: Overview statistics of regression variables

Table A.1 shows an overview of the variables included in the regression analyses. The columns at the left hand side show the raw data values and the right hand side columns show the values after a series of data transformations that are done to make the data more comparable and to cope with outliers and highly skewed distributions.

The transformation steps include in sequential order:

1. Linear scaling of spillovers (0.0001), PageRank (1000) and capital per labor (0.01) where the value in parentheses shows the scaling factor. This is done to obtain more comparable coefficients.
2. All variables except for the share of production labor $((Lprod/L)_{i,t})$ transformed to log values using the formula $\log(1 + x)$ to cope with < 1 values. The log-linearization is done to cope with highly skewed data.
3. Outliers are removed according a an interquartile range (IQR) based formula. Those values are treated as outliers that are beyond the 25/75% quantile values minus/plus $(a \cdot IQR)$ with $a = 30$ in the baseline models. Robustness checks are made with more restrictive (i.e. $a = 5$ and $a = 10$ IQR) removal rules. The regression results are qualitative consistent with the baseline.

A.4. Network plots

In the network plots shown in Sec. 5.1, links between two industries i and j if j is a sufficiently important input supplier to i and the weight $w_{i,j,t}^{\alpha,in}$ exceeds a threshold level defined by the average of weights across all industry-pairs and time periods plus one standard deviation. In the overlap networks, this rule is applied to the sum of the weights across both layers $w_{i,j,t}^{\mu,in} + w_{i,j,t}^{\tau,in}$. The node size is scaled proportionally to $A_{i,t}^{\alpha}$ which is the average value over the time window . For the overlap, $A_{i,t}^{\alpha}$ is additionally averaged across both layers.

B. Additional descriptive information

Figure B.1 shows the pairwise correlation of different indicators used to describe the network across and within layers. The figure illustrates two observations: (1) the degree is least correlated with other variables, and (2) variable computed on the basis of in- and out-going links are highly correlated. Tables B.2 and B.3 show the Top 10 ranking of industries by up- and downstream centrality as measured by the PageRank.

<i>Top 10 industries by Pagerank ($PR_{i,t}^{\alpha, in}$)</i>									
1977-1982			1987-1992			1997-2002			
1	Petroleum Refineries	324110 0.17	Petroleum Refineries	324110 0.08	Petroleum Refineries	324110 0.10	Copper Refineries	331411 0.05	
2	Copper Refineries	331411 0.05	Plastics Mat. & Resin	325211 0.05	Iron & Steel Mills	331111 0.07	Iron & Steel Mills	331111 0.05	
3	Plastics Mat. & Resin	325211 0.04	Copper Refineries	331411 0.05	Semiconductor & Device	334413 0.05	Automobile Mnft.	336111 0.03	
4	All Petrol. & Coal Prod.	324199 0.04	All Petrol. & Coal Prod.	324199 0.04	Sawmills	321113 0.04	Biological Prod.	325414 0.02	
5	Iron & Steel Mills	331111 0.04	Chem. Preparations	325998 0.03	Plastics Mat. & Resin	325211 0.04	Plastics Mat. & Resin	325211 0.02	
6	Chem. Preparations	325998 0.02	Iron & Steel Mills	331111 0.03	Copper Refineries	331411 0.03	Ship Building & Repair	336611 0.02	
7	Paperboard Mills	322130 0.02	Inorganic Dye & Pigm.	325131 0.03	Gum & Wood Chem.	325191 0.02	Aircraft Mnft.	336411 0.02	
8	Organic Chem.	325199 0.02	Organic Chem.	325199 0.03	Organic Chem.	325199 0.02	Dog & Cat Food Mnft.	311111 0.02	
9	Inorganic Dye & Pigm.	325131 0.01	Fats & Oils Refin.	311225 0.01	Machine Shops	332710 0.02	Petroleum Refineries	324110 0.01	
10	Metal Can Mnft.	332431 0.01	Nitrogen. Fertl. Mnft.	325311 0.01	Print Circuit Assembly	334418 0.02	Semiconductor & Device	334413 0.01	
Quartiles:									
0.01, 0.01, 0.01			0.01, 0.01, 0.01			0.01, 0.01, 0.02			0.01, 0.01, 0.01

<i>Top 10 industries by Pagerank ($PR_{i,t}^{\tau, in}$)</i>									
1977-1982			1987-1992			1997-2002			
1	Adhesive Mnft.	325520 0.05	Adhesive Mnft.	325520 0.05	Semiconductor & Device	334413 0.05	Semiconductor & Device	334413 0.06	
2	Chem. Preparations	325998 0.05	Chem. Preparations	325998 0.05	Adhesive Mnft.	325520 0.05	Adhesive Mnft.	325520 0.04	
3	Semiconductor & Device	334413 0.03	Semiconductor & Device	334413 0.04	Chem. Preparations	325998 0.04	Electr. Computer Mnft.	334111 0.04	
4	Power Transm. Equ.	333613 0.03	Power Transm. Equ.	333613 0.03	Laboratory Apparatus	339111 0.03	Chem. Preparations	325998 0.04	
5	Fastener & Pin	339993 0.02	Fastener & Pin	339993 0.03	Electr. Computer Mnft.	334111 0.03	Optical Instrum. & Lens	333314 0.03	
6	Laboratory Apparatus	339111 0.02	Electr. Computer Mnft.	334111 0.03	Fastener & Pin	339993 0.03	Fastener & Pin	339993 0.03	
7	Speed Changer & Gear	333612 0.02	Laboratory Apparatus	339111 0.02	Optical Instrum. & Lens	333314 0.02	Wireless Communic.	334220 0.02	
8	Electr. Computer Mnft.	334111 0.02	Speed Changer & Gear	333612 0.02	Power Transm. Equ.	333613 0.02	Medical Instrum.	339112 0.02	
9	Boiler & Heat Exch.	332410 0.02	Optical Instrum. & Lens	333314 0.02	Speed Changer & Gear	333612 0.02	Power Transm. Equ.	333613 0.02	
10	Optical Instrum. & Lens	333314 0.02	Boiler & Heat Exch.	332410 0.02	Dental Equ. & Supplies	339114 0.02	Elctrmed. Apparatus	334510 0.02	
Quartiles:									
0.01, 0.01, 0.0125			0.01, 0.01, 0.01			0.01, 0.01, 0.01			0.01, 0.01, 0.02

Notes: Industries are ranked by the PageRank compiled on upstream links $PR_{i,t}^{\alpha, in}$ averaged across the time window indicated in the column header in decreasing order, i.e. showing the largest industries on top. The values $PR_{i,t}^{\alpha, in}$ were normalized before through division by the economy-wide average in t , i.e. the mean value for each period equals one. The bottom lines of each sub-table show the quartile values as indicators for the skewness of the distribution. Deviations of the median from the average indicate skewness.

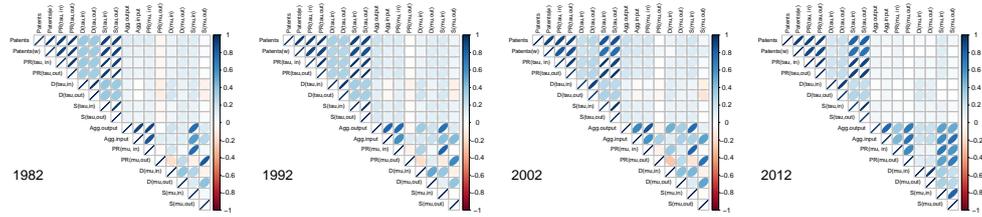
Table B.2: Top-10 ranking of industries by the upstream PageRank at the 6-digit level.

<i>Top 10 industries by Pagerank ($PR_{i,t}^{\alpha, out}$)</i>								
1977-1982			1987-1992			1997-2002		
1	Mobile Home Mnft.	321991 0.07	Motor Home Mnft.	336213 0.10	Aircraft Mnft.	336411 0.04	Petroleum Refineries	324110 0.07
2	Motor Home Mnft.	336213 0.04	Ice & Frozen Dessert	311520 0.04	Mobile Home Mnft.	321991 0.04	Copper Refineries	331411 0.06
3	Frozen Spec. Food Mnft.	311412 0.03	Mobile Home Mnft.	321991 0.04	Automobile Mnft.	336111 0.04	Iron & Steel Mills	331111 0.04
4	Motor Vhcl. Body Mnft.	336211 0.03	Missile & Space Vhcl.	336414 0.03	Motor Home Mnft.	336213 0.03	Die-Cut & Paper Office	322231 0.03
5	Missile & Space Vhcl.	336414 0.03	Aircraft Mnft.	336411 0.03	Ship Building & Repair	336611 0.03	Plastics Mat. & Resin	325211 0.02
6	Ice & Frozen Dessert	311520 0.03	Motor Vhcl. Body Mnft.	336211 0.02	Retail Bakeries	311811 0.03	Soybean Processing	311222 0.02
7	Ship Building & Repair	336611 0.02	Frozen Spec. Food Mnft.	311412 0.02	Commercial Bakeries	311812 0.03	All Misc. Electr. Equ.	335999 0.01
8	Automobile Mnft.	336111 0.02	Ship Building & Repair	336611 0.02	Travel Trailer & Camper	336214 0.02	Graphite Prod.	335991 0.01
9	Light & Utility Truck	336112 0.02	Travel Trailer & Camper	336214 0.02	Oth. Animal Food Mnft.	311119 0.02	Motor Vhcl. Air-Cond.	336391 0.01
10	Heavy Duty Truck Mnft.	336120 0.02	Electr. Computer Mnft.	334111 0.02	Dog & Cat Food Mnft.	311111 0.02	Gum & Wood Chem.	325191 0.01
Quartiles:								
0.01, 0.01, 0.02			0.01, 0.01, 0.02			0.01, 0.01, 0.01		0.01, 0.01, 0.01

<i>Top 10 industries by Pagerank ($PR_{i,t}^{\tau, out}$)</i>								
1977-1982			1987-1992			1997-2002		
1	Adhesive Mnft.	325520 0.05	Adhesive Mnft.	325520 0.05	Semiconductor & Device	334413 0.05	Semiconductor & Device	334413 0.07
2	Chem. Preparations	325998 0.05	Chem. Preparations	325998 0.05	Adhesive Mnft.	325520 0.05	Electr. Computer Mnft.	334111 0.04
3	Semiconductor & Device	334413 0.04	Semiconductor & Device	334413 0.04	Chem. Preparations	325998 0.05	Adhesive Mnft.	325520 0.04
4	Power Transm. Equ.	333613 0.03	Power Transm. Equ.	333613 0.03	Electr. Computer Mnft.	334111 0.03	Chem. Preparations	325998 0.04
5	Fastener & Pin	339993 0.03	Fastener & Pin	339993 0.03	Fastener & Pin	339993 0.03	Optical Instrum. & Lens	333314 0.03
6	Electr. Computer Mnft.	334111 0.02	Electr. Computer Mnft.	334111 0.03	Optical Instrum. & Lens	333314 0.02	Fastener & Pin	339993 0.03
7	Speed Changer & Gear	333612 0.02	Speed Changer & Gear	333612 0.02	Power Transm. Equ.	333613 0.02	Power Transm. Equ.	333613 0.02
8	Urethane & Foam Prod.	326150 0.02	Optical Instrum. & Lens	333314 0.02	Speed Changer & Gear	333612 0.02	Medical Instrum.	339112 0.02
9	Boiler & Heat Exch.	332410 0.02	Boiler & Heat Exch.	332410 0.02	Urethane & Foam Prod.	326150 0.01	Misc. Food Mnft.	311999 0.02
10	Optical Instrum. & Lens	333314 0.02	Urethane & Foam Prod.	326150 0.02	Dental Equ. & Supplies	339114 0.01	Watch & Clock Mnft.	334518 0.02
Quartiles:								
0.01, 0.01, 0.01			0.01, 0.01, 0.01			0.01, 0.01, 0.01		0.01, 0.01, 0.0175

Notes: Industries are ranked by the PageRank compiled on downstream links $PR_{i,t}^{\alpha, out}$ averaged across the time window indicated in the column header in decreasing order, i.e. showing the largest industries on top. The values $PR_{i,t}^{\alpha, out}$ were normalized before through division by the economy-wide average in t , i.e. the mean value for each period equals one. The bottom lines of each sub-table show the quartile values as indicators for the skewness of the distribution. Deviations of the median from the average indicate skewness.

Table B.3: Top-10 ranking of industries by the downstream PageRank at the 6-digit level.



Notes: This figure shows a correlation plot between different pairs of indicators. *PR* is short for PageRank, *D* for degree, *S* for strength. *Patents(w)* is the weighted patent stock. The correlation at the diagonal is by definition equal one. The colors and the shape of the ellipses indicate the strength of correlation. Data: 4-digit, balanced panel.

Figure B.1: Pairwise correlations of network indicators

C. Additional regression results

C.1. Additional results

C.1.1. Relationships between growth rates

Regressions using growth rates of the explanatory variables can be understood as a correlation analysis of industrial growth. This may be indicative for the presence of increasing returns. For example, the observation in Table C.4 that the negative auto-correlation in the growth of patent stocks is less significant than the negative auto-correlation in the market can be a weak indicator for different types of growth patterns. However, also note that the R^2 is very low, i.e. a large part of the variation remains unexplained. For the other technological change indicators, see SI.2.2.1.

	$gr(A_{i,t}^\mu)$		$A_{i,t}^\mu$		$gr(A_{i,t}^\tau)$		$A_{i,t}^\tau$	
	up	down	up	down	up	down	up	down
$gr(A_{i,t-1}^\mu)$	-0.3658*** (0.0253)	-0.3793*** (0.023)	0.2887*** (0.0376)	0.2747*** (0.0356)	0.0037 (0.0098)	0.0027 (0.0086)	-0.0032 (0.0123)	-0.0073 (0.0098)
$gr(A_{i,t-1}^\tau)$	-0.1346 (0.0942)	-0.1346 (0.0972)	-0.0539 (0.1007)	-0.0619 (0.0965)	-0.2013 (0.1197)	-0.2087* (0.1035)	0.1772 (0.1036)	0.1972* (0.0925)
$gr(PR_{i,t-1}^{\mu,d})$	-0.0025 (0.0033)	-3e-04 (0.0036)	-0.0032 (0.0033)	-0.0078 (0.0042)	-2e-04 (0.0012)	0 (0.001)	-8e-04 (0.0017)	-2e-04 (0.0014)
$gr(PR_{i,t-1}^{\tau,d})$	0.0044 (0.0057)	0.0092 (0.0069)	0.0034 (0.0063)	0.0075 (0.007)	-0.0038 (0.003)	-0.0054* (0.0022)	0.0118*** (0.0033)	0.0147*** (0.003)
$Spill(gr(A)_{i,t-1}^{\mu,d})$	3.554 (2.62)	-0.591 (1.377)	4.388 (2.919)	3.143 (1.983)	0.8179 (0.5357)	0.3308 (0.5702)	1.467 (1.013)	0.9381 (1.066)
$Spill(gr(A)_{i,t-1}^{\tau,d})$	-4.532 (18.25)	-7.006 (17.62)	30.66 (21.82)	17.2 (20.3)	3.296 (6.764)	7.848 (7.86)	56.53*** (9.058)	52.02*** (9.049)
$gr(L_{i,t-1})$	-0.0304 (0.2301)	-0.0187 (0.2319)	0.1147 (0.2897)	0.143 (0.2886)	0.2085* (0.0958)	0.2013* (0.0964)	-0.2221 (0.1307)	-0.1969 (0.1297)
$gr(Wage_{i,t-1})$	0.222 (0.5239)	0.1689 (0.5281)	0.9017 (0.624)	0.7713 (0.6231)	-0.0808 (0.1073)	-0.0853 (0.1088)	-0.0655 (0.1901)	-0.1257 (0.1894)
$gr((K/L)_{i,t-1})$	-8.974 (10.97)	-9.59 (11.47)	9.342 (14.06)	6.308 (14.19)	-0.1342 (5.443)	-0.3908 (5.446)	0.2116 (6.959)	0.0456 (7.066)
$gr((Lprod/L)_{i,t-1})$	1.373* (0.6922)	1.349 (0.6947)	0.4584 (0.5678)	0.4821 (0.5666)	-0.0889 (0.1279)	-0.081 (0.1294)	-0.2678 (0.1759)	-0.3057 (0.1826)
$gr((I/L)_{i,t-1})$	0.0985 (0.0871)	0.1022 (0.0873)	0.1199 (0.0913)	0.1319 (0.0912)	0.006 (0.0217)	0.0065 (0.0216)	-0.0353 (0.0251)	-0.0317 (0.0252)
$gr((VA/L)_{i,t-1})$	0.2185 (0.1661)	0.2186 (0.1658)	0.2261 (0.1703)	0.2631 (0.1726)	0.0864 (0.0483)	0.0874 (0.0485)	-0.0088 (0.0756)	-0.0039 (0.0764)
$gr(TFP5_{i,t-1})$	0.0363 (1.667)	-0.0891 (1.635)	2.052 (1.837)	1.898 (1.811)	-0.0244 (0.3673)	-0.0571 (0.3622)	0.806 (0.7075)	0.6859 (0.6818)
R^2	0.1428	0.1423	0.0814	0.0841	0.0805	0.0838	0.0927	0.0975
N	1661	1661	1661	1661	1661	1661	1661	1661
Average	0.3663	0.3663	7.824	7.824	0.0594	0.0594	8.513	8.513

Notes: The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^\alpha$, $gr(A_{i,t}^\alpha)$, $gr(PR_{i,t}^{\alpha,d})$, $Spill(gr(A)_{i,t}^{\alpha,d})$, $gr(L_{i,t})$, $gr(Wage_{i,t})$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $gr((VA/L)_{i,t})$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3. The difference between four and five factor TFP is that $TFP5_{i,t}$ is a disaggregation of material inputs into energy and non-energy materials which is not done for $TFP4_{i,t}$ (see Bartlesman and Gray (1996)).

Table C.4: Regression results: The evolution of industry sizes using growth rates

C.2. Robustness checks

C.2.1. Degree-based regressions

	$gr(A_{i,t}^\mu)$		$A_{i,t}^\mu$		$gr(A_{i,t}^\tau)$		$A_{i,t}^\tau$	
	mnft up	mnft down	mnft up	mnft down	mnft up	mnft down	mnft up	mnft down
$A_{i,t-1}^\mu$	-0.5175*** (0.026)	-0.4939*** (0.0256)	0.3934*** (0.0296)	0.4192*** (0.0295)	0.003 (0.0068)	0.004 (0.0069)	0.0043 (0.0092)	0.0081 (0.0091)
$A_{i,t-1}^\tau$	-0.0425 (0.0569)	-0.0346 (0.0531)	-0.0508 (0.0658)	-0.0368 (0.0603)	-0.2543*** (0.0246)	-0.2655*** (0.0225)	0.7435*** (0.0268)	0.7276*** (0.026)
$D_{i,t-1}^{\mu,d}$	0.017 (0.018)	0.0934*** (0.0203)	0.0251 (0.0221)	0.1442*** (0.0284)	0.0061 (0.0053)	-0.013 (0.0072)	0.0114 (0.0076)	-0.0296* (0.0126)
$D_{i,t-1}^{\tau,d}$	0.0552 (0.3329)	0.8696* (0.3537)	0.0333 (0.3853)	1.039* (0.4181)	-0.0943 (0.1459)	-0.1735 (0.1721)	-0.201 (0.2003)	-0.2021 (0.2069)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0077 (0.0336)	-0.0587 (0.0319)	0.0125 (0.04)	-0.0417 (0.0368)	0.0111 (0.0086)	0.0027 (0.0107)	0.0143 (0.0137)	-0.0058 (0.0154)
$Spill(A)_{i,t-1}^{\tau,d}$	-0.0286 (0.133)	-0.0678 (0.1305)	-0.0117 (0.1526)	-0.1141 (0.1422)	0.1624*** (0.0413)	0.0974 (0.0506)	0.2164*** (0.0544)	0.0712 (0.059)
$L_{i,t-1}$	-0.1388 (0.0938)	-0.1624 (0.0958)	-0.1585 (0.1151)	-0.187 (0.1175)	-0.0174 (0.0261)	-0.0107 (0.0243)	-0.0446 (0.0342)	-0.0342 (0.0325)
$Wage_{i,t-1}$	0.414 (0.4244)	0.4976 (0.418)	0.7624 (0.4773)	0.8536 (0.4663)	0.0811 (0.0952)	0.0709 (0.0995)	0.179 (0.1226)	0.1587 (0.1261)
$(K/L)_{i,t-1}$	-0.6038** (0.22)	-0.6155** (0.2139)	-0.6563** (0.2485)	-0.6661** (0.24)	-0.0866 (0.0534)	-0.0662 (0.0482)	-0.1221 (0.073)	-0.0787 (0.0664)
$(Lprod/L)_{i,t-1}$	-0.8125 (0.5809)	-0.6204 (0.567)	-1.448* (0.6746)	-1.249 (0.6432)	-0.0621 (0.1822)	-0.1588 (0.1791)	-0.0869 (0.2265)	-0.2217 (0.2347)
$(I/L)_{i,t-1}$	-0.0866 (0.0903)	-0.0846 (0.0895)	-0.0458 (0.1038)	-0.0376 (0.1024)	0.0441* (0.0197)	0.0474* (0.0196)	0.038 (0.029)	0.0432 (0.0289)
$(VA/L)_{i,t-1}$	0.3372* (0.1636)	0.3354* (0.1531)	0.2122 (0.1972)	0.2145 (0.183)	0.0495 (0.0356)	0.0465 (0.0357)	0.0265 (0.0467)	0.0198 (0.0482)
$TFP5_{i,t-1}$	-0.2402 (0.1864)	-0.2354 (0.1783)	-0.3709 (0.2265)	-0.3675 (0.2065)	0.0776 (0.06)	0.0817 (0.0582)	0.138* (0.0686)	0.1413* (0.0649)
R^2	0.2778	0.2894	0.1724	0.1889	0.2139	0.2083	0.5958	0.591
N	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.3638	0.3638	7.683	7.683	0.0524	0.0524	8.45	8.45

Notes: The regressions aim to explain the factors that influence the evolution of industry sizes in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^\alpha$, $gr(A_{i,t}^\tau)$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table C.5: Regression results: The evolution of industry sizes

	$gr((VA/L)_{i,t})$		$(VA/L)_{i,t}$		$gr(TFP4_{i,t})$		$TFP4_{i,t}$		$gr(TFP5_{i,t})$		$TFP5_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$A_{i,t-1}^{\mu}$	0.0066 (0.0058)	0.0051 (0.0058)	0.0074 (0.0058)	0.0056 (0.0059)	3e-04 (5e-04)	4e-04 (5e-04)	0.0041* (0.002)	0.0039 (0.0021)	1e-04 (5e-04)	3e-04 (5e-04)	0.0041* (0.002)	0.0039 (0.0021)
$A_{i,t-1}^{\tau}$	-0.0115 (0.0189)	-0.0173 (0.0178)	-0.0128 (0.0193)	-0.0182 (0.018)	7e-04 (0.0013)	2e-04 (0.0013)	0.0065 (0.0063)	0.0034 (0.0058)	8e-04 (0.0013)	2e-04 (0.0013)	0.0064 (0.0064)	0.0032 (0.0058)
$D_{i,t-1}^{\mu,d}$	-0.0032 (0.0048)	0.0136 (0.0073)	-0.0038 (0.0049)	0.0146 (0.0075)	-4e-04 (5e-04)	0 (6e-04)	-0.0011 (0.0019)	8e-04 (0.0028)	-4e-04 (5e-04)	0 (6e-04)	-0.0011 (0.0019)	8e-04 (0.0029)
$D_{i,t-1}^{\tau,d}$	-0.0827 (0.0695)	-0.0246 (0.0724)	-0.0817 (0.0704)	-0.024 (0.0729)	-0.0127* (0.0059)	-0.011 (0.006)	-0.0772** (0.0294)	-0.058 (0.0315)	-0.0133* (0.0058)	-0.0112 (0.006)	-0.0773** (0.0295)	-0.0581 (0.0315)
$Spill(A)_{i,t-1}^{\mu,d}$	6e-04 (0.0093)	0.0026 (0.0089)	0.0011 (0.0095)	0.0021 (0.0091)	-0.001 (8e-04)	-0.0015 (8e-04)	-2e-04 (0.0038)	-8e-04 (0.0033)	-9e-04 (8e-04)	-0.0014 (8e-04)	-3e-04 (0.0038)	-8e-04 (0.0033)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0334 (0.0352)	0.0347 (0.0379)	0.0291 (0.0366)	0.0297 (0.0389)	-7e-04 (0.0027)	-0.0042 (0.003)	-0.0142 (0.0155)	-0.0197 (0.0168)	-9e-04 (0.0027)	-0.0046 (0.003)	-0.0142 (0.0155)	-0.0195 (0.0169)
$L_{i,t-1}$	-0.0196 (0.0255)	-0.016 (0.0251)	-0.0203 (0.0258)	-0.0168 (0.0253)	0 (0.0021)	5e-04 (0.0021)	-0.0057 (0.0102)	-0.0032 (0.01)	0 (0.0021)	4e-04 (0.0021)	-0.0057 (0.0102)	-0.0032 (0.01)
$Wage_{i,t-1}$	0.0861 (0.1457)	0.0961 (0.1457)	0.1009 (0.1484)	0.111 (0.1483)	-0.0104 (0.0098)	-0.0103 (0.0099)	0.0215 (0.0384)	-0.02 (0.0383)	-0.0113 (0.0098)	-0.0112 (0.0098)	-0.0218 (0.0384)	-0.0202 (0.0383)
$(K/L)_{i,t-1}$	0.1101 (0.0911)	0.1125 (0.0898)	0.1021 (0.0921)	0.1041 (0.0908)	0.0206*** (0.0053)	0.0218*** (0.0052)	0.1484*** (0.0262)	0.1515*** (0.026)	0.0203*** (0.0053)	0.0216*** (0.0052)	0.1487*** (0.0261)	0.1518*** (0.0259)
$(Lprod/L)_{i,t-1}$	0.12 (0.1712)	0.0941 (0.1712)	0.1531 (0.1856)	0.1271 (0.186)	-0.006 (0.0173)	-0.0103 (0.017)	0.0408 (0.0817)	0.0302 (0.0809)	-0.0079 (0.0171)	-0.012 (0.0168)	0.0382 (0.0822)	0.0276 (0.0814)
$(I/L)_{i,t-1}$	-0.0171 (0.026)	-0.0168 (0.0252)	-0.0195 (0.0263)	-0.0194 (0.0254)	-7e-04 (0.0024)	-0.0011 (0.0024)	-0.0294** (0.01)	-0.0299** (0.01)	-5e-04 (0.0024)	-8e-04 (0.0024)	-0.0293** (0.01)	-0.0299** (0.0099)
$(VA/L)_{i,t-1}$	-0.4338*** (0.0553)	-0.4289*** (0.0552)	0.5631*** (0.0569)	0.5684*** (0.0568)	-0.0147*** (0.0039)	-0.0142*** (0.0038)	-0.0549** (0.0213)	-0.0531* (0.021)	-0.0145*** (0.0039)	-0.0141*** (0.0038)	-0.0552** (0.0213)	-0.0534* (0.021)
$TFP4_{i,t-1}$					-0.0573*** (0.0087)	-0.0574*** (0.0085)	0.9739*** (0.0804)	0.9746*** (0.0808)				
$TFP5_{i,t-1}$	0.0959 (0.0716)	0.0971 (0.0739)	0.0953 (0.0711)	0.0965 (0.0737)					-0.0572*** (0.0086)	-0.0573*** (0.0084)	0.9737*** (0.0802)	0.9745*** (0.0806)
R^2	0.1697	0.1708	0.3237	0.3248	0.1833	0.1849	0.7109	0.7102	0.1831	0.1848	0.7108	0.7102
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.2664	0.2664	4.48	4.48	0.6937	0.6937	0.6754	0.6754	0.6938	0.6938	0.6756	0.6756

Notes:

The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP4_{i,t}$, $TFP5_{i,t}$, $gr((VA/L)_{i,t})$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3. The difference between four and five factor TFP is that $TFP5_{i,t}$ is a disaggregation of material inputs into energy and non-energy materials which is not done for $TFP4_{i,t}$ (see Bartlesman and Gray (1996)).

Table C.6: Regression results: Determinants of productivity growth

	$gr(L_{i,t})$		$L_{i,t}$		$gr((K/L)_{i,t})$		$(K/L)_{i,t}$		$gr((I/L)_{i,t})$		$(I/L)_{i,t}$	
$A_{i,t-1}^{\mu}$	-0.003 (0.0059)	-0.0023 (0.0057)	-0.0026 (0.0077)	-5e-04 (0.0077)	0 (1e-04)	-1e-04 (1e-04)	0 (0.0037)	-0.0012 (0.0038)	-0.0049 (0.0111)	-0.0046 (0.0106)	-0.0025 (0.0095)	-0.0033 (0.0092)
$A_{i,t-1}^{\tau}$	-0.0185 (0.0148)	-0.0283 (0.0145)	-0.0236 (0.0183)	-0.0372* (0.018)	1e-04 (3e-04)	4e-04 (3e-04)	0.0124 (0.0098)	0.0165 (0.0095)	-0.0148 (0.0295)	-0.0182 (0.0298)	-0.0181 (0.0271)	-0.0205 (0.0272)
$D_{i,t-1}^{\mu,d}$	0.0051 (0.0045)	-0.0065 (0.0073)	0.0074 (0.0058)	-0.0056 (0.0094)	-1e-04 (1e-04)	1e-04 (1e-04)	-0.002 (0.003)	0.0038 (0.0047)	0.0078 (0.0083)	0.0156 (0.011)	0.0052 (0.0073)	0.0146 (0.01)
$D_{i,t-1}^{\tau,d}$	-0.0834 (0.0581)	0.0308 (0.066)	-0.1141 (0.0701)	-0.0131 (0.0742)	0.0025 (0.0016)	0.0025 (0.0014)	0.0264 (0.0408)	0.0325 (0.0397)	-0.0696 (0.1031)	0.0997 (0.1039)	-0.046 (0.095)	0.0764 (0.0961)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0188* (0.009)	0.0079 (0.0075)	-0.0183 (0.0125)	0.0075 (0.0098)	0 (2e-04)	1e-04 (1e-04)	0.0098 (0.0062)	1e-04 (0.0049)	-0.0176 (0.0158)	0.0217 (0.0135)	-0.0119 (0.0145)	0.0199 (0.0123)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0975** (0.0328)	0.0999** (0.0313)	0.1128** (0.0408)	0.1224** (0.0397)	-0.001 (6e-04)	-0.0012* (6e-04)	-0.0209 (0.0217)	-0.0511* (0.0226)	0.0988 (0.0627)	0.0584 (0.0605)	0.0867 (0.0561)	0.0496 (0.0567)
$L_{i,t-1}$	-0.1951*** (0.0224)	-0.1881*** (0.0231)	0.7908*** (0.0295)	0.7987*** (0.0303)	0.0011* (5e-04)	0.001* (5e-04)	0.0224 (0.0165)	0.0206 (0.0168)	-0.0778* (0.0385)	-0.0723 (0.0379)	-0.0874* (0.0344)	-0.0828* (0.0338)
$Wage_{i,t-1}$	-0.0585 (0.0791)	-0.0538 (0.078)	-0.0366 (0.1016)	-0.0284 (0.0995)	5e-04 (0.0015)	3e-04 (0.0014)	0.0564 (0.0601)	0.051 (0.0589)	0.263 (0.1459)	0.263 (0.1431)	0.3414* (0.1387)	0.3397* (0.1363)
$(K/L)_{i,t-1}$	-0.2051*** (0.0496)	-0.189*** (0.0485)	-0.1834** (0.0639)	-0.1677** (0.0633)	-0.0062*** (0.0011)	-0.0063*** (0.0011)	0.7954*** (0.0383)	0.7956*** (0.0387)	-0.1637 (0.0981)	-0.1425 (0.0949)	-0.2202* (0.0911)	-0.203* (0.0881)
$(Lprod/L)_{i,t-1}$	0.8217*** (0.1632)	0.8085*** (0.1604)	0.9654*** (0.2144)	0.9484*** (0.2079)	-0.0106** (0.0035)	-0.0103** (0.0034)	-0.5275*** (0.1233)	-0.5295*** (0.12)	0.5263* (0.2557)	0.5047* (0.2562)	0.5684* (0.2496)	0.5431* (0.2498)
$Invest_{i,t-1}$	0.0361 (0.0203)	0.0367 (0.0203)	0.0489* (0.0241)	0.0496* (0.0243)	0.0022*** (4e-04)	0.0022*** (4e-04)	0.1201*** (0.017)	0.1211*** (0.0171)	-0.7662*** (0.0428)	-0.7641*** (0.0414)	0.3022*** (0.044)	0.3042*** (0.043)
$VA_{i,t-1}$	0.0838* (0.0381)	0.0845* (0.0378)	0.0612 (0.0485)	0.0611 (0.0486)	6e-04 (9e-04)	6e-04 (9e-04)	-0.0076 (0.0278)	-0.0072 (0.0279)	0.1176 (0.0681)	0.1231 (0.0672)	0.1241* (0.0631)	0.1293* (0.0619)
$TFP5_{i,t-1}$	-0.1386* (0.0619)	-0.144* (0.0629)	-0.2007** (0.0752)	-0.2035** (0.0758)	0.0032** (0.001)	0.0031** (0.001)	0.1378** (0.0434)	0.1372** (0.0433)	0.1544 (0.1174)	0.1423 (0.1155)	0.1189 (0.1114)	0.1089 (0.1096)
R^2	0.1555	0.1503	0.6837	0.6821	0.2191	0.2205	0.6536	0.6543	0.3196	0.3198	0.1447	0.1458
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	-0.0354	-0.0354	3.166	3.166	0.0018	0.0018	0.6928	0.6928	0.2482	0.2482	1.865	1.865

Notes: The regressions aim to identify the drivers of technological change reflected in changing input factor use in a balanced panel of 6-digit NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $gr(L_{i,t})$, $Wage_{i,t}$, $(K/L)_{i,t}$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table C.7: Regression results: Patterns of changing input factor use

C.2.2. Strength-based regressions

	$gr(A_{i,t}^\mu)$		$A_{i,t}^\mu$		$gr(A_{i,t}^\tau)$		$A_{i,t}^\tau$	
	mnft up	mnft down	mnft up	mnft down	mnft up	mnft down	mnft up	mnft down
$A_{i,t-1}^\mu$	-0.4998*** (0.0308)	-0.5007*** (0.0252)	0.4328*** (0.0345)	0.4085*** (0.0295)	-4e-04 (0.0078)	0.0052 (0.0065)	-8e-04 (0.0105)	0.0104 (0.0088)
$A_{i,t-1}^\tau$	-0.1056 (0.06)	-0.0516 (0.0568)	-0.12 (0.0684)	-0.0543 (0.0628)	-0.2384*** (0.0251)	-0.2515*** (0.0292)	0.752*** (0.0259)	0.739*** (0.0305)
$S_{i,t-1}^{\mu,d}$	-0.0368 (0.0605)	0.2433*** (0.0443)	-0.1081 (0.0684)	0.3415*** (0.0572)	0.0266 (0.0155)	-0.0147 (0.0152)	0.0443 (0.025)	-0.0329 (0.0248)
$S_{i,t-1}^{\tau,d}$	0.8686** (0.2957)	0.4859 (0.3379)	0.9443** (0.3334)	0.5805 (0.3854)	-0.2865** (0.0996)	-0.2258 (0.1209)	-0.2749* (0.1295)	-0.1971 (0.1477)
$Spill(A)_{i,t-1}^{\mu,d}$	-2e-04 (0.0339)	-0.0237 (0.0315)	0.016 (0.0397)	0.0076 (0.0363)	0.0101 (0.0087)	-0.0019 (0.0106)	0.0149 (0.0138)	-0.013 (0.0149)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0321 (0.1306)	-0.0282 (0.1342)	0.0463 (0.1489)	-0.0668 (0.1442)	0.1391** (0.0448)	0.0784 (0.0557)	0.1946** (0.0593)	0.0546 (0.0628)
$L_{i,t-1}$	-0.1834 (0.0973)	-0.1982* (0.0986)	-0.2051 (0.1192)	-0.2327 (0.1209)	-0.0033 (0.0243)	-0.0024 (0.0245)	-0.0294 (0.0325)	-0.0264 (0.0331)
$Wage_{i,t-1}$	0.3882 (0.4153)	0.4275 (0.4117)	0.739 (0.4668)	0.7615 (0.4574)	0.0966 (0.0939)	0.0902 (0.0951)	0.2026 (0.1193)	0.1829 (0.12)
$(K/L)_{i,t-1}$	-0.6596** (0.2176)	-0.682** (0.2166)	-0.712** (0.2449)	-0.7485** (0.2428)	-0.0654 (0.0523)	-0.0513 (0.0497)	-0.098 (0.0727)	-0.0649 (0.0692)
$(Lprod/L)_{i,t-1}$	-0.6697 (0.5825)	-0.762 (0.5773)	-1.282 (0.6729)	-1.42* (0.6604)	-0.0863 (0.1832)	-0.151 (0.1745)	-0.0971 (0.2336)	-0.2054 (0.2303)
$(I/L)_{i,t-1}$	-0.0769 (0.0888)	-0.0759 (0.0909)	-0.0344 (0.1014)	-0.024 (0.1034)	0.0422* (0.0204)	0.0463* (0.0202)	0.0373 (0.03)	0.0414 (0.03)
$(VA/L)_{i,t-1}$	0.3025 (0.1589)	0.2862 (0.1527)	0.1642 (0.1898)	0.1444 (0.181)	0.0581 (0.0352)	0.0545 (0.036)	0.0354 (0.0471)	0.0319 (0.0494)
$TFP5_{i,t-1}$	-0.2487 (0.1819)	-0.2498 (0.1791)	-0.3772 (0.2188)	-0.3886 (0.2082)	0.0836 (0.0584)	0.0849 (0.0582)	0.1461* (0.0679)	0.1452* (0.0656)
R^2	0.281	0.2893	0.1761	0.1893	0.2184	0.2066	0.5963	0.5888
N	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.3638	0.3638	7.683	7.683	0.0524	0.0524	8.45	8.45

Notes: The regressions aim to explain the factors that influence the evolution of industry sizes in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^\alpha$, $gr(A_{i,t}^\tau)$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table C.8: Regression results: The evolution of industry sizes

	$gr((VA/L)_{i,t})$		$(VA/L)_{i,t}$		$gr(TFP4_{i,t})$		$TFP4_{i,t}$		$gr(TFP5_{i,t})$		$TFP5_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$A_{i,t-1}^{\mu}$	0.021** (0.0066)	0.0038 (0.0058)	0.0218** (0.0067)	0.0043 (0.0059)	3e-04 (5e-04)	3e-04 (5e-04)	0.0045 (0.0027)	0.0036 (0.002)	2e-04 (5e-04)	2e-04 (5e-04)	0.0046 (0.0027)	0.0036 (0.002)
$A_{i,t-1}^{\tau}$	-0.0207 (0.0192)	-0.014 (0.0191)	-0.0214 (0.0194)	-0.0143 (0.0193)	-3e-04 (0.0014)	0 (0.0014)	0.0017 (0.0057)	0.0038 (0.0061)	-3e-04 (0.0014)	0 (0.0014)	0.0016 (0.0057)	0.0037 (0.0062)
$S_{i,t-1}^{\mu,d}$	-0.0654*** (0.0166)	0.0367* (0.0144)	-0.0671*** (0.0172)	0.0378* (0.0148)	-0.0012 (0.0014)	-3e-04 (0.0011)	-0.0042 (0.0081)	0.0089 (0.0053)	-0.0011 (0.0014)	-3e-04 (0.0011)	-0.0043 (0.0082)	0.0089 (0.0054)
$S_{i,t-1}^{\tau,d}$	0.076 (0.0805)	-0.0118 (0.1065)	0.07 (0.0819)	-0.0206 (0.109)	0.0048 (0.0079)	0.0039 (0.0081)	0.0085 (0.0234)	0.0068 (0.0311)	0.0048 (0.0079)	0.0042 (0.0081)	0.0086 (0.0234)	0.0073 (0.0311)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0034 (0.0091)	0.0071 (0.0093)	-0.0032 (0.0093)	0.0067 (0.0096)	-0.001 (9e-04)	-0.0015 (8e-04)	-5e-04 (0.004)	2e-04 (0.0036)	-9e-04 (9e-04)	-0.0014 (8e-04)	-5e-04 (0.004)	3e-04 (0.0036)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0376 (0.0348)	0.0336 (0.0377)	0.033 (0.0363)	0.0279 (0.0391)	0 (0.0028)	-0.0038 (0.0031)	-0.0121 (0.0156)	-0.0191 (0.0174)	-2e-04 (0.0028)	-0.0042 (0.0031)	-0.012 (0.0156)	-0.0189 (0.0174)
$L_{i,t-1}$	-0.0172 (0.0248)	-0.0187 (0.0256)	-0.0174 (0.0251)	-0.0193 (0.0258)	4e-04 (0.002)	3e-04 (0.0021)	-0.0031 (0.0102)	-0.0045 (0.01)	3e-04 (0.002)	2e-04 (0.002)	-0.0031 (0.0102)	-0.0046 (0.01)
$Wage_{i,t-1}$	0.0915 (0.1411)	0.0953 (0.144)	0.1064 (0.1435)	0.1102 (0.1465)	-0.0096 (0.0099)	-0.0101 (0.0098)	-0.0159 (0.0374)	-0.0175 (0.037)	-0.0104 (0.0098)	-0.011 (0.0098)	-0.0161 (0.0374)	-0.0177 (0.0371)
$(K/L)_{i,t-1}$	0.112 (0.0866)	0.1092 (0.0879)	0.1045 (0.0874)	0.1013 (0.0889)	0.0209*** (0.0052)	0.0218*** (0.0052)	0.1513*** (0.0268)	0.1504*** (0.0258)	0.0206*** (0.0051)	0.0216*** (0.0051)	0.1516*** (0.0267)	0.1507*** (0.0257)
$(Lprod/L)_{i,t-1}$	0.1248 (0.1668)	0.0936 (0.172)	0.1557 (0.1793)	0.1254 (0.1852)	-0.0059 (0.0169)	-0.0076 (0.0168)	0.0417 (0.0814)	0.0408 (0.0811)	-0.0076 (0.0167)	-0.0093 (0.0166)	0.0391 (0.0819)	0.0383 (0.0816)
$(I/L)_{i,t-1}$	-0.0161 (0.0253)	-0.0146 (0.025)	-0.0187 (0.0256)	-0.0171 (0.0253)	-6e-04 (0.0024)	-9e-04 (0.0024)	-0.029** (0.0099)	-0.0289** (0.0098)	-5e-04 (0.0024)	-7e-04 (0.0024)	-0.029** (0.0099)	-0.0288** (0.0098)
$(VA/L)_{i,t-1}$	-0.4411*** (0.0525)	-0.4353*** (0.0542)	0.5562*** (0.0541)	0.5619*** (0.0558)	-0.0146*** (0.004)	-0.0143*** (0.0039)	-0.0545* (0.0212)	-0.0548** (0.0209)	-0.0145*** (0.004)	-0.0142*** (0.0039)	-0.0548** (0.0212)	-0.0551** (0.0209)
$TFP4_{i,t-1}$					-0.0572*** (0.0091)	-0.0576*** (0.0087)	0.9748*** (0.083)	0.9734*** (0.0826)				
$TFP5_{i,t-1}$	0.0967 (0.0735)	0.0947 (0.0738)	0.0961 (0.073)	0.0941 (0.0735)					-0.0571*** (0.009)	-0.0575*** (0.0086)	0.9746*** (0.0828)	0.9733*** (0.0824)
R^2	0.1804	0.1734	0.3327	0.3269	0.1802	0.1825	0.7083	0.7093	0.1799	0.1824	0.7083	0.7093
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.2664	0.2664	4.48	4.48	0.6937	0.6937	0.6754	0.6754	0.6938	0.6938	0.6756	0.6756

Notes:

The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP4_{i,t}$, $TFP5_{i,t}$, $gr((VA/L)_{i,t})$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3. The difference between four and five factor TFP is that $TFP5_{i,t}$ is a disaggregation of material inputs into energy and non-energy materials which is not done for $TFP4_{i,t}$ (see Bartlesman and Gray (1996)).

Table C.9: Regression results: Determinants of productivity growth

	$gr(L_{i,t})$		$L_{i,t}$		$gr((K/L)_{i,t})$		$(K/L)_{i,t}$		$gr((I/L)_{i,t})$		$(I/L)_{i,t}$	
$A_{i,t-1}^{\mu}$	0.0054 (0.007)	-0.0019 (0.0057)	0.0091 (0.0092)	-5e-04 (0.0078)	-2e-04 (1e-04)	-1e-04 (1e-04)	-0.0036 (0.0048)	-0.0014 (0.0038)	0.0166 (0.0125)	-0.007 (0.0106)	0.0156 (0.0108)	-0.0055 (0.0092)
$A_{i,t-1}^{\tau}$	-0.0171 (0.0158)	-0.03 (0.0158)	-0.0238 (0.0198)	-0.0397* (0.0195)	2e-04 (3e-04)	4e-04 (3e-04)	0.0111 (0.0115)	0.0131 (0.0103)	-0.0481 (0.0349)	-0.0579 (0.036)	-0.0509 (0.0318)	-0.0557 (0.0322)
$S_{i,t-1}^{\mu,d}$	-0.0236 (0.0131)	0.0158 (0.0122)	-0.0322* (0.0163)	0.0252 (0.0161)	3e-04 (2e-04)	-1e-04 (2e-04)	0.0107 (0.009)	-0.01 (0.0077)	-0.0718** (0.0264)	0.0294 (0.0201)	-0.0635** (0.0242)	0.0331 (0.0184)
$S_{i,t-1}^{\tau,d}$	-0.0443 (0.0734)	0.0371 (0.0928)	-0.0421 (0.089)	0.0644 (0.1129)	4e-04 (0.0012)	-8e-04 (0.0014)	0.0208 (0.0461)	0.0158 (0.058)	0.2696 (0.147)	0.433* (0.1941)	0.2749* (0.138)	0.395* (0.1793)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0208* (0.0091)	0.0093 (0.0078)	-0.0207 (0.0127)	0.0102 (0.0103)	0 (2e-04)	0 (2e-04)	0.0107 (0.0062)	-5e-04 (0.0052)	-0.0202 (0.0156)	0.0292* (0.0138)	-0.0139 (0.0143)	0.0273* (0.0126)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0895** (0.0325)	0.1015** (0.0325)	0.1029* (0.0402)	0.1262** (0.0412)	-9e-04 (6e-04)	-0.0012* (6e-04)	-0.0174 (0.0211)	-0.0494* (0.0226)	0.1044 (0.0622)	0.0928 (0.0599)	0.0949 (0.0551)	0.0809 (0.0551)
$L_{i,t-1}$	-0.191*** (0.0233)	-0.1917*** (0.0238)	0.7957*** (0.0307)	0.793*** (0.0313)	0.0011* (5e-04)	0.0011* (5e-04)	0.0208 (0.0171)	0.022 (0.0173)	-0.0843* (0.0393)	-0.0864* (0.0387)	-0.0946** (0.0351)	-0.0966** (0.0345)
$Wage_{i,t-1}$	-0.0507 (0.0773)	-0.053 (0.0776)	-0.0264 (0.099)	-0.0266 (0.0987)	3e-04 (0.0015)	2e-04 (0.0014)	0.0536 (0.0592)	0.0471 (0.0581)	0.2611 (0.1443)	0.2386 (0.1474)	0.3373* (0.1373)	0.3183* (0.1406)
$(K/L)_{i,t-1}$	-0.1986*** (0.0503)	-0.1952*** (0.048)	-0.1752** (0.0652)	-0.1764** (0.0624)	-0.0063*** (0.0011)	-0.0062*** (0.0011)	0.7928*** (0.0391)	0.7975*** (0.0384)	-0.1685 (0.0963)	-0.1632 (0.0943)	-0.2266* (0.0889)	-0.223* (0.0873)
$(Lprod/L)_{i,t-1}$	0.8303*** (0.162)	0.8008*** (0.1606)	0.9799*** (0.2121)	0.951*** (0.2101)	-0.0109** (0.0035)	-0.0108** (0.0035)	-0.53*** (0.1221)	-0.5312*** (0.1216)	0.5741* (0.2596)	0.5307* (0.2548)	0.611* (0.2521)	0.5685* (0.2481)
$Invest_{i,t-1}$	0.0368 (0.02)	0.0371 (0.0201)	0.0501* (0.0239)	0.0512* (0.0242)	0.0022*** (4e-04)	0.0022*** (4e-04)	0.1199*** (0.017)	0.1204*** (0.0172)	-0.7624*** (0.0422)	-0.7602*** (0.0417)	0.3056*** (0.0436)	0.3082*** (0.0432)
$VA_{i,t-1}$	0.0802* (0.0377)	0.0833* (0.0371)	0.0557 (0.0485)	0.0576 (0.0483)	6e-04 (9e-04)	7e-04 (9e-04)	-0.0061 (0.028)	-0.0068 (0.0279)	0.0998 (0.0664)	0.1082 (0.0667)	0.1082 (0.061)	0.1149 (0.0611)
$TFP5_{i,t-1}$	-0.1362* (0.0618)	-0.1447* (0.0604)	-0.1976** (0.0751)	-0.2058** (0.0722)	0.0032** (0.0011)	0.0032** (0.001)	0.1368** (0.0434)	0.1378** (0.0418)	0.1538 (0.1121)	0.1347 (0.111)	0.1174 (0.1056)	0.1014 (0.1042)
R^2	0.1553	0.1506	0.6835	0.6828	0.2153	0.2159	0.6538	0.6543	0.3246	0.323	0.1523	0.1511
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	-0.0354	-0.0354	3.166	3.166	0.0018	0.0018	0.6928	0.6928	0.2482	0.2482	1.865	1.865

Notes: The regressions aim to identify the drivers of technological change reflected in changing input factor use in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $gr(L_{i,t})$, $Wage_{i,t}$, $(K/L)_{i,t}$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table C.10: Regression results: Patterns of changing input factor use

SUPPLEMENTARY MATERIAL

SI.1. Detailed information on data processing

This analysis builds on two distinct sources of data brought into a consistent form that enables the statistical analyses; this is (1) a series of time snapshots of the network of cross-industrial IO flows and patent citations and (2) a panel data set of aggregate statistical indicators at the industry level. Industries are identified by 6-digit NAICS codes (and later aggregated to the 4-digit level). The time snapshots cover 5-year intervals from 1977 to 2012. Obtaining these data involved a series of steps of re-formatting and harmonization which are explained in detail below.

SI.1.1. Input-output data

The IO data is constructed by the composition and harmonization of the historical benchmark tables provided by [Bureau of Economic Analysis \(BEA\)](#).¹³ Since 1947, BEA publishes IO tables at the detailed industry level every 5 years. The data is collected in BEA's quinquennial Survey of Current Business. A detailed manual on BEA's IO data is provided by Horowitz and Planting (2006)

SI.1.1.1. Overview

The raw data shows monetary transactions between industries. It covers also final demand sectors and public services. For this project, I use tables from 1977-2007. I made a series of conversions and processing steps to harmonize the data. Over time, industrial classification systems and technical methods of data processing, formatting and saving have changed. The earliest tables are only available in text format that was manually edited to make it readable for statistical software. A further challenge arises from changes in the classification system, most pronounced in the conversion from SIC to NAICS.

The final data structure is a series of quadratic matrices for each period that show the monetary transactions between industries in NAICS 2002 codes. The data is also

¹³<https://www.bea.gov/industry/historical-benchmark-input-output-tables> [accessed on Dec 21, 2020]

used to create a panel of industry level indicators, i.e. outputs, inputs, and growth rates.

SI.1.1.2. Processing steps in detail

Here, I explain single steps of data processing. Again, the steps are consistent with the R-code provided in the data publication Hötte (2021).

Step 1: For each period, the IO data are downloaded separately. Some manual harmonization and data conversions were made to obtain machine-readable, harmonized data tables. For example, the very old data is only available in text format which is not ready to be read by statistical software. The more recent tables are Excel files with many macros and text-explanations. All tables were reformatted individually. The scripts are available in the data publication. After this step, all tables have a uniform format which is a long 3-column table with column (1) as producer ID, column (2) as user ID, and column (3) indicating the monetary value of the goods that flow from producer to user.

Step 2: I created large quadratic matrices with rows as producers and columns as users. The entries $flow_{ij,t}^{out,\mu}$ are flows of goods from i to j . Hence, column-wise reading indicates the composition of inputs used by sector j and row-wise reading indicates the composition of customer industries to which industry i delivers.

Step 3: This is an intermediate step. All concordance and IO-to-industry conversion tables have to be harmonized. Again, some of the data are not machine readable. Moreover, all codes need to be harmonized to obtain a mapping from IO codes for each period to 2002-NAICS codes. Some of the IO codes map to multiple industries. In this step, tables were created where each row indicates an IO code and all NAICS codes to which the IO maps.

Step 4: NAICS-based IO tables were harmonized and consistency checks were done. For example, I tested whether the differences in the tables e.g. regarding the sector coverage are negligible. Some normalizations of IO-flows to input (output) shares were made through division by row (column) sum. The full 6-digit list is used as row and column names.

Not every time snapshot has a full sector coverage. This is a result of reclassification issues, obsolescence and introduction of new sectors. For example, some of the finely granulate computer industries were not yet existing in 1977. For

these cases, empty vectors are included to present missing sectors to ensure that matrices have same dimensionality.

Additional steps of harmonization are done. Rows represent the range of inputs that is used, columns represent customers. After this step, NAICS 6-digit data on IO flows, sector weights (row and column sums), input shares (measured in percentage points), 6-digit distance matrix computed by the input-share dissimilarity are obtained.

Step 5: Harmonization of quadratic NAICS 2002 matrices. The matrices are 1179×1179 matrices of 6-digit industries. Empty rows and columns are included for industries that are not producing in some t , for example if an industry was not yet existing or disappeared over time.

Step 6: For each t , I create NAICS \times NAICS matrices with flows of goods $flow_{ij,t}^{\mu,in}$ as entries.

SI.1.1.3. Technical and conceptual issues

General remarks about IO codes, NAICS and SIC The original IO data in early years uses IO codes which are an internal metric of the accounting system used to construct social accounting matrices (SAM). These codes are converted into industry codes (SIC and NAICS). The classification system has changed over time. Fortunately, the IO-codes in the raw IO tables are largely consistent across time. I converted the accounting codes into SIC and from SIC into NAICS or directly into NAICS if such mapping is available.

How to make a decision about the set of economic sectors to be considered?

The accounting matrices include also dummy industries like private household industries (not same as personal consumption expenditures), government industries, and special positions (e.g. *Non-comparable imports*, *Scrap*, *Rest of world* adjustments, inventory valuation adjustments). These positions are required to ensure completeness in the calculation of GDP which is one of the original purposes of the data sets) cf. Horowitz and Planting, 2006, Chap. 4.

As a pragmatic rule, all final demand sectors were kept that can be mapped to NAICS codes. An output link to final demand as customer reveals information about the technological capabilities of the producer. As an input link, final demand is not relevant because it will not appear as a production input.

For the main analyses presented in this paper, the subset of industries is further reduced because industries were only included if having a non-zero patent stock $A_{i,t}^r$ and goods output $A_{i,t}^u$ in all periods. This excludes the majority of final demand positions.

Another concern is the consistency of the data across time. This is partly addressed by normalizations and the focus of analysis on relative industry differences instead of quantitative cross-time comparisons. In the panel data analyses, I control for FE and make checks using clustered standard errors.

Note that some of the rows and columns are empty for some periods at the 6-digit level. This is a result of the harmonization procedure to uniform NAICS codes. This may happen if an industry disappears or a new industry emerges. Often, the emergence (disappearance) of an industry is associated with a split (merge) of pre-existing industries. This problem arises more often for final demand and service industries. I cope with this problem by the use of more aggregate data and robustness checks.

How to deal with accounting codes that are mapped into multiple SIC sectors? Some of the accounting codes are associated with multiple SIC sectors, i.e. multiple industries have been aggregated into one accounting position. Information about the strengths of links to each of these subsectors is missing. For reasons of simplification, I assume that the accounting position is equally related to all of them. The strengths of single links is weighted uniformly by the number of sectors. For example, the IO code 020401 (“Fruits”) is linked to 9 SIC sectors (0171, 0172, 0174, 0175, *0179, *019, *0219, *0259, *029). The links are weighted by factor 1/9.

How to deal with inconsistencies across time in changing classification systems? The accounting codes of 1977 and 1982 data are mapped to SIC 1987. All mappings from accounting positions to SIC are based on the 1987 data after having ensured that the accounting codes are consistent across time. Also the vast majority of IO-to-SIC-mappings is consistent in 1977 and 1992 data. For 1977 some minor deviations exist but these are largely explainable by adjustments in the SIC system between 1977 to 1987. Some of the old SIC industries do not exist any longer. A reconstruction is practically not feasible with reasonable effort given that the value added of higher precision is negligibly small if existing at all. The 1977 IO-SIC mapping is only used when 1987-data is not available.

In the 2002 NAICS file, some IO codes are mapped to a very high number of subsectors. This is for example the case for aggregate positions such as retail and wholesale trade and construction. I kept them in the mapping. It should be noted that an

accounting position that has a link to more than hundred 6-digit NAICS industries is not necessarily meaningful. I cope with this problem by a series of robustness checks using only a subset of the data, higher levels of aggregation and rounding of IO links that fall below a certain threshold.

The more recent versions of the classification systems are more detailed. I used equal weights when one coarse industry mapped to several more detailed industry when using another (typically more recent) classification system. Hence, the transaction volume is equally distributed across sub-sectors.

Which NAICS version to use? I use NAICS 2002 codes. These codes have a direct mapping to SIC 1987 codes.

SI.1.2. Patent data

The raw patent data classified by CPC codes are taken from an earlier project. An extensive documentation of the data are provided along with the data which can be downloaded for re-use under a CC-BY-4.0 license (Hötte, 2021).

From the raw data, I use the CPC classification data, citations among patents with the grant number as ID, and data on the grant year of the patents. Further, to map patents classified by CPC codes to NAICS 6-digit codes, I used the concordance tables by Goldschlag et al. (2020).¹⁴ I used the *Cooperative Patent Classification (CPC) Crosswalks - Version 1603* file downloaded in October 2021.

SI.1.2.1. Processing steps in detail

To construct a patent citation networks among NAICS 6-digit industries, the data were processed in a series of steps.

Step 1: In a first step, patents were mapped to NAICS 2002 codes to create industry-level patent stocks as 5-year aggregates covering the period 1973-2012. First, patents were sampled by time window. Then, patents in each time window were aggregated into each 4-digit CPC class ensuring uniqueness for each entry by patent grant number and CPC 4-digit code. Hence, patents that map to multiple more disaggregate CPC codes that belong to the same 4-digit aggregate were treated as unique entry. The counts at the 4-digit level were subsequently mapped to NAICS 6-digit codes taking account of the weights, i.e. the patent counts

¹⁴<https://sites.google.com/site/nikolaszolas/PatentCrosswalk> [accessed in Oct 2021]

are multiplied by the weight whenever one CPC 4-digit class maps to multiple NAICS 6-digit codes.

Step 2: During this step, the citation data is mapped from citations between individual patents by grant number to citations between NAICS 6-digit industries. The NAICS to NAICS edgelist also contains a column with the weight which indicates the number of citations that flow from one industry to another during each 5-year time window. To obtain this edgelist, both the citing and the cited patent were first mapped and aggregated into CPC 4-digit classes and then mapped to NAICS codes taking account of the weights, i.e. multiplying the number of citations between two industries by the weights.

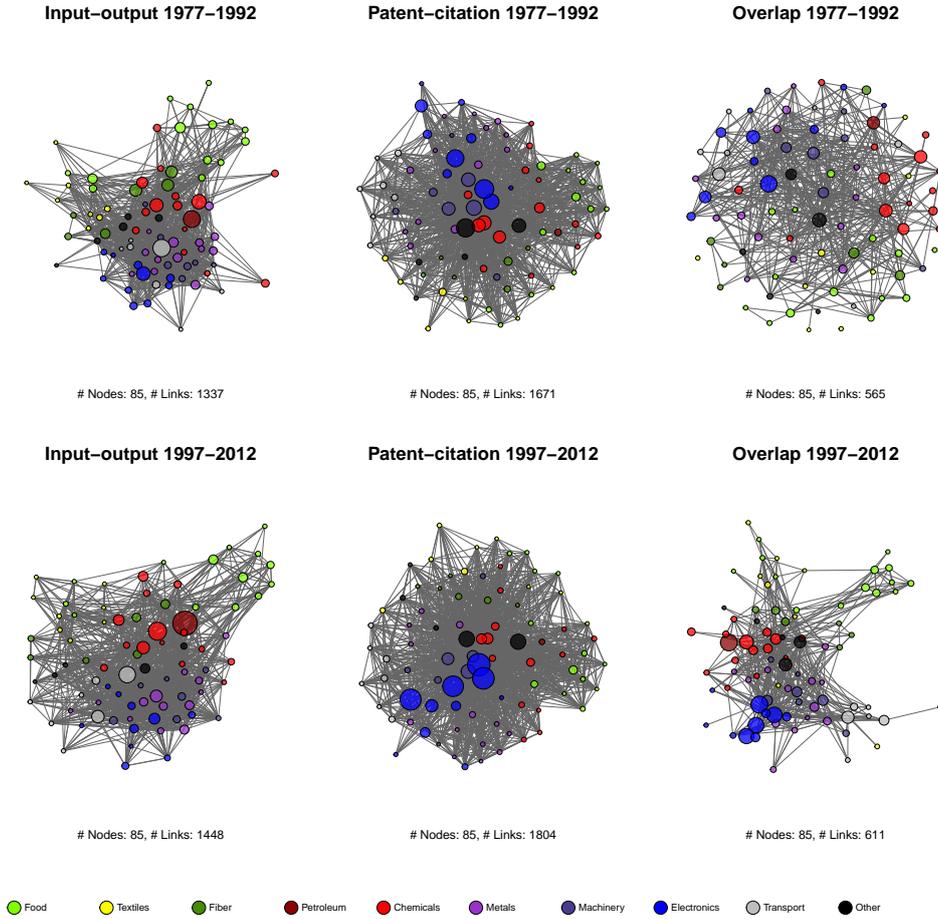
Step 3: The edgelist is used to construct an adjacency matrix with 6-digit industries as row and column names. The data are harmonized with the format of the IO adjacency matrices. Additionally, the adjacency matrices are also created for other levels of aggregation.

SI.2. Supplementary results

SI.2.1. Descriptives

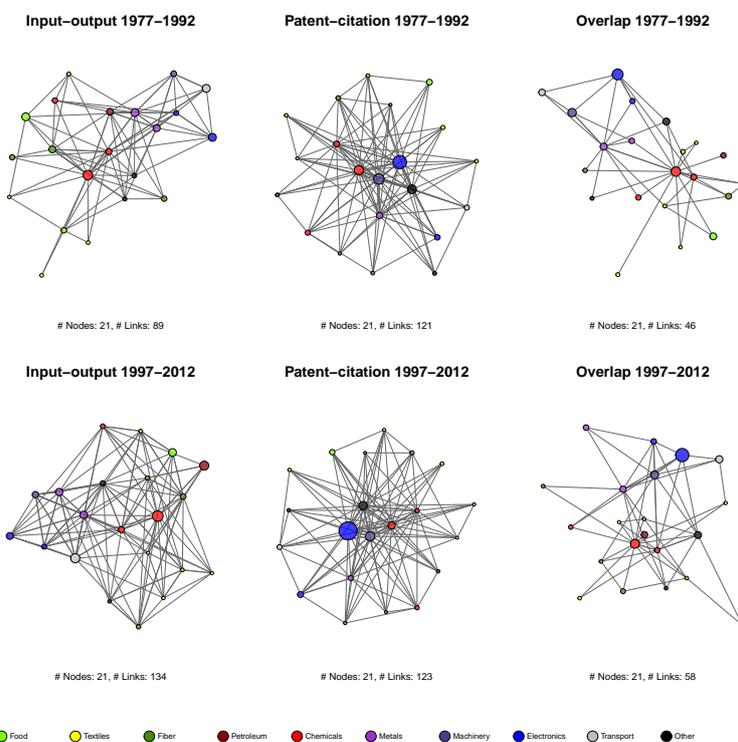
SI.2.1.1. Network plots

Fig. [SI.2](#) shows the evolution of the network plots at the 3-digit level. The change of the size of the nodes representing the Electronics sector nodes in blue color in relation to other industries illustrates well the increasing concentration in the number of patents. The size distribution in the input-output network appears more balanced, but the network becomes increasingly connected.



Notes: These figures show the network of upstream links (suppliers) at the 3-digit level for two different time periods. A link between a pair of industries i and j is shown if j is a sufficiently important supplier to i , i.e. if the average of the weight $w_{ij,t}^{in,\alpha}$ during time periods 1977-1992 and 1997-2012 exceeds a threshold level given by the average weight over all industry pairs and all periods plus one standard deviation ($\text{mean}_{i,j,t}(w_{ij,t}^{in,\alpha}) + \text{sd}_{i,j,t}(w_{ij,t}^{in,\alpha})$). The overlap network shows nodes as being connected if they are connected on both layers, i.e. links are compiled on the basis of weights averaged across both layers. The size of the nodes is proportional to the size of an industry $A_{i,t}^\alpha$ in the respective layer, and in the overlap network to the weighted mean of both layers ($0.5 \cdot (A_{i,t}^\mu + A_{i,t}^\tau)$). Plots of the downstream network are available in 5.1. Self-citations and within-sector IO flows are not shown. The colors indicate broad industrial categories given by groups of 3-digit level industries, i.e. Food (311-312), Textiles (314-316), Fiber (321-323), Petroleum (324), Chemicals (325-327), Metals (331,332), Machinery (333), Electronics (334-335), Transport (336), Other (337-339). Data: 4-digit manufacturing industries.

Figure SI.1: Upstream networks of 4-digit level industries for different periods



Notes: These figures show the network of upstream links (suppliers) at the 3-digit level for two different time periods. A link between a pair of industries i and j is shown if j is a sufficiently important supplier to i , i.e. if the average of the weight $w_{ij,t}^{in,\alpha}$ during time periods 1977-1992 and 1997-2012 exceeds a threshold level given by the average weight over all industry pairs and all periods plus one standard deviation ($\text{mean}_{i,j,t}(w_{ij,t}^{in,\alpha}) + \text{sd}_{i,j,t}(w_{ij,t}^{in,\alpha})$). The overlap network shows nodes as being connected if they are connected on both layers, i.e. links are compiled on the basis of weights averaged across both layers. Plots of the downstream network are available in ???. Self-citations and within-sector IO flows are not shown. The colors indicate broad industrial categories given by groups of 3-digit level industries, i.e. Food (311-312), Textiles (314-316), Fiber (321-323), Petroleum (324), Chemicals (325-327), Metals (331,332), Machinery (333), Electronics (334-335), Transport (336), Other (337-339). Data: Balanced panel of 3-digit industries.

Figure SI.2: Upstream network at the 4-digit level.

SI.2.1.2. Industry rankings

<i>Top 10 industries by Aggr. output ($A_{i,t}^{\mu}$)</i>															
1977-1982				1987-1992				1997-2002				2007-2012			
1	Petroleum & Coal Prod.	3241	9.97	Motor Vhcl. Parts Mnft.	3363	5.74	Motor Vhcl. Parts Mnft.	3363	7.21	Petroleum & Coal Prod.	3241	10.55			
2	Motor Vhcl. Parts Mnft.	3363	5.31	Petroleum & Coal Prod.	3241	5.53	Semicond. & Oth. Elctr.	3344	5.02	Basic Chem. Mnft.	3251	5.26			
3	Pulp, & Paper Mills	3221	2.76	Basic Chem. Mnft.	3251	4.11	Plastics Prod. Mnft.	3261	4.53	Motor Vhcl. Parts Mnft.	3363	4.51			
4	Basic Chem. Mnft.	3251	2.56	Plastics Prod. Mnft.	3261	3.35	Petroleum & Coal Prod.	3241	4.43	Motor Vehicle Mnft.	3361	3.14			
5	Anim. Slaughter. & Prc.	3116	2.52	Semicond. & Oth. Elctr.	3344	3.14	Basic Chem. Mnft.	3251	4.25	Plastics Prod. Mnft.	3261	3.08			
6	Converted Paper Prod.	3222	2.45	Pulp, & Paper Mills	3221	3.05	Iron & Steel Mills	3311	2.95	Aerospace Products	3364	2.89			
7	Plastics Prod. Mnft.	3261	2.37	Print. & Support Act.	3231	2.87	Print. & Support Act.	3231	2.76	Anim. Slaughter. & Prc.	3116	2.54			
8	Synth. Rubber & Fibers	3252	2.34	Converted Paper Prod.	3222	2.75	Pulp, & Paper Mills	3221	2.51	Iron & Steel Mills	3311	2.36			
9	Textile & Fabric Coat.	3133	2.09	Synth. Rubber & Fibers	3252	2.48	Converted Paper Prod.	3222	2.45	Synth. Rubber & Fibers	3252	2.22			
10	Print. & Support Act.	3231	2.07	Anim. Slaughter. & Prc.	3116	2.18	Synth. Rubber & Fibers	3252	2.37	Pharm. & Medicine	3254	2.13			
Quartiles:															
0.355, 0.65, 1.205				0.37, 0.73, 1.15				0.265, 0.705, 1.165				0.25, 0.59, 1.13			

<i>Top 10 industries by Patent stock ($A_{i,t}^{\pi}$)</i>															
1977-1982				1987-1992				1997-2002				2007-2012			
1	Gen. Purpose Machines	3339	5.80	Semicond. & Oth. Elctr.	3344	5.45	Semicond. & Oth. Elctr.	3344	7.44	Semicond. & Oth. Elctr.	3344	9.71			
2	Basic Chem. Mnft.	3251	5.14	Gen. Purpose Machines	3339	5.42	Measur. & Ctrl. Instr.	3345	4.96	Measur. & Ctrl. Instr.	3345	6.45			
3	Eng. & Power Transm.	3336	4.66	Eng. & Power Transm.	3336	4.77	Oth. Misc. Mnft.	3399	4.49	Computer Equ.	3341	6.25			
4	Oth. Misc. Mnft.	3399	4.62	Oth. Misc. Mnft.	3399	4.65	Gen. Purpose Machines	3339	4.43	Communic. Equ. Mnft.	3342	5.58			
5	Measur. & Ctrl. Instr.	3345	4.10	Measur. & Ctrl. Instr.	3345	4.32	Computer Equ.	3341	4.22	Commerc. Ind. Equ.	3333	4.11			
6	Paint, Coat. & Adhesive	3255	3.95	Paint, Coat. & Adhesive	3255	3.99	Eng. & Power Transm.	3336	3.77	Eng. & Power Transm.	3336	3.76			
7	Semicond. & Oth. Elctr.	3344	3.89	Basic Chem. Mnft.	3251	3.78	Paint, Coat. & Adhesive	3255	3.63	Oth. Misc. Mnft.	3399	3.69			
8	Plastics Prod. Mnft.	3261	3.74	Chem. Prod. & Prepar.	3259	3.53	Commerc. Ind. Equ.	3333	3.34	Gen. Purpose Machines	3339	3.63			
9	Chem. Prod. & Prepar.	3259	3.60	Plastics Prod. Mnft.	3261	3.47	Chem. Prod. & Prepar.	3259	3.16	Paint, Coat. & Adhesive	3255	2.98			
10	Commerc. Ind. Equ.	3333	2.50	Computer Equ.	3341	3.03	Communic. Equ. Mnft.	3342	3.15	Magn. & Optical Media	3346	2.61			
Quartiles:															
0.17, 0.505, 1.325				0.17, 0.49, 1.315				0.14, 0.44, 1.065				0.1025, 0.36, 0.8375			

Notes: Industries are ranked by output (patent stock) $A_{i,t}^{\alpha}$ averaged across the time window indicated in the column header in decreasing order, i.e. showing the largest industries on top. The values $A_{i,t}^{\alpha}$ were normalized before through division by the economy-wide average output (patent stock) in t , i.e. the mean value for each period equals one. The bottom lines of each sub-table show the quartile values as indicators for the skewness of the distribution. Deviations of the median from the average indicate skewness. Data: Balanced panel of 4-digit manufacturing industries.

Table SI.1: Top-10 ranking of industries by output and patent stock at the 4-digit level.

SI.2.2. Regressions

SI.2.2.1. Relationships between growth rates

	$gr((VA/L)_{i,t})$		$(VA/L)_{i,t}$		$gr(TFP4_{i,t})$		$TFP4_{i,t}$		$gr(TFP5_{i,t})$		$TFP5_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$gr(A_{i,t-1}^\mu)$	0.0016 (0.0072)	-5e-04 (0.0066)	-0.0049 (0.0074)	-0.0059 (0.0062)	0.001 (6e-04)	8e-04 (6e-04)	-0.0013 (0.0036)	-0.0037 (0.004)	9e-04 (6e-04)	7e-04 (6e-04)	-0.0013 (0.0036)	-0.0037 (0.0039)
$gr(A_{i,t-1}^\tau)$	0.0033 (0.0256)	0.0034 (0.0243)	-0.0157 (0.0202)	-0.0168 (0.0188)	-0.0015 (0.0023)	-0.0013 (0.0021)	-0.0333** (0.0128)	-0.0268** (0.0097)	-0.0015 (0.0023)	-0.0013 (0.0021)	-0.0333** (0.0128)	-0.0268** (0.0097)
$gr(PR_{i,t-1}^{\mu,d})$	-7e-04 (8e-04)	4e-04 (8e-04)	-5e-04 (9e-04)	0 (9e-04)	-1e-04 (1e-04)	0 (1e-04)	-6e-04 (4e-04)	0 (8e-04)	-1e-04 (1e-04)	0 (1e-04)	-6e-04 (4e-04)	0 (8e-04)
$gr(PR_{i,t-1}^{\tau,d})$	-0.0016 (0.0012)	-0.0031* (0.0014)	0 (0.0013)	0 (0.0014)	1e-04 (1e-04)	1e-04 (1e-04)	0.0024 (0.0013)	0.0029 (0.0017)	1e-04 (1e-04)	0 (1e-04)	0.0024 (0.0013)	0.0029 (0.0017)
$Spill(gr(A)_{i,t-1}^{\mu,d})$	-0.1469 (0.6982)	1.015* (0.5051)	-1.175 (0.7361)	-0.309 (0.602)	0.0165 (0.0455)	0.0675 (0.0426)	-0.0664 (0.229)	-0.1431 (0.2235)	0.011 (0.0455)	0.0734 (0.0421)	-0.0802 (0.2291)	-0.1509 (0.2242)
$Spill(gr(A)_{i,t-1}^{\tau,d})$	-1.662 (5.08)	-0.4941 (5.037)	0.5095 (5.266)	1.678 (5.352)	-0.0313 (0.4553)	-0.198 (0.4537)	1.496 (2.487)	1.062 (3.007)	-0.095 (0.4522)	-0.2529 (0.4524)	1.427 (2.493)	1.06 (3.01)
$gr(L_{i,t-1})$	-0.134* (0.0519)	-0.1345** (0.0515)	0.0433 (0.0706)	0.0405 (0.0708)	-0.0051 (0.007)	-0.0047 (0.0071)	-0.0354 (0.0795)	-0.0312 (0.0766)	-0.0053 (0.007)	-0.0049 (0.007)	-0.0349 (0.0793)	-0.0307 (0.0764)
$gr(Wage_{i,t-1})$	0.0481 (0.1296)	0.0498 (0.1294)	-0.1169 (0.1289)	-0.1072 (0.1283)	0.0141 (0.009)	0.0132 (0.0089)	-0.0537 (0.0547)	-0.0587 (0.0565)	0.0135 (0.009)	0.0125 (0.0089)	-0.055 (0.0548)	-0.0598 (0.0566)
$gr((K/L)_{i,t-1})$	-8.895* (4.053)	-8.645* (4.019)	5.998 (4.492)	6.064 (4.483)	-0.3231 (0.399)	-0.3302 (0.4027)	-1.582 (2.116)	-1.526 (2.225)	-0.304 (0.3947)	-0.308 (0.3987)	-1.567 (2.106)	-1.508 (2.214)
$gr((Lprod/L)_{i,t-1})$	0.0245 (0.1329)	0.038 (0.1332)	0.2136 (0.1365)	0.2147 (0.1374)	0.0041 (0.0126)	0.0048 (0.0128)	0.0686 (0.0507)	0.0622 (0.0491)	0.0027 (0.0123)	0.0034 (0.0125)	0.0729 (0.0511)	0.0665 (0.0493)
$gr((I/L)_{i,t-1})$	-0.0013 (0.0197)	-0.0018 (0.0199)	-0.0401* (0.0159)	-0.0405* (0.0162)	-0.0027 (0.0016)	-0.0026 (0.0016)	-0.0068 (0.0067)	-0.0066 (0.0071)	-0.0025 (0.0016)	-0.0024 (0.0016)	-0.0069 (0.0067)	-0.0067 (0.0072)
$gr((VA/L)_{i,t-1})$	-0.266*** (0.0405)	-0.2635*** (0.0403)	0.4066*** (0.0469)	0.4065*** (0.0463)	-0.0172*** (0.0036)	-0.0168*** (0.0037)	0.0531* (0.0221)	0.0532* (0.0237)	-0.0171*** (0.0036)	-0.0167*** (0.0037)	0.0532* (0.0221)	0.0533* (0.0236)
$gr(TFP4_{i,t-1})$					-0.2033*** (0.0481)	-0.2058*** (0.0484)	1.787*** (0.379)	1.775*** (0.3685)				
$gr(TFP5_{i,t-1})$	1.825*** (0.4458)	1.789*** (0.4395)	1.336*** (0.3761)	1.354*** (0.3807)					-0.2037*** (0.0479)	-0.2062*** (0.0482)	1.794*** (0.3798)	1.782*** (0.3698)
R^2	0.0848	0.0903	0.1455	0.1432	0.0829	0.0848	0.1027	0.1044	0.083	0.0853	0.1035	0.1052
N	1661	1661	1661	1661	1661	1661	1661	1661	1661	1661	1661	1661
Average	0.2418	0.2418	4.61	4.61	0.6941	0.6941	0.6804	0.6804	0.6942	0.6942	0.6806	0.6806

Notes: The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $gr(A_{i,t}^\alpha)$, $gr(PR_{i,t}^{\alpha,d})$, $Spill(gr(A)_{i,t}^{\alpha,d})$, $gr(L_{i,t})$, $gr(Wage_{i,t})$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $(VA/L)_{i,t}$, $TFP4_{i,t}$, $TFP5_{i,t}$, $gr((VA/L)_{i,t})$, $gr(VA/L)_{i,t}$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3. The difference between four and five factor TFP is that $TFP5_{i,t}$ is a disaggregation of material inputs into energy and non-energy materials which is not done for $TFP4_{i,t}$ (see Bartlesman and Gray (1996)).

Table SI.2: Regression results: Determinants of productivity growth

	$gr(L_{i,t})$		$L_{i,t}$		$gr((K/L)_{i,t})$		$(K/L)_{i,t}$		$gr((I/L)_{i,t})$		$(I/L)_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$gr(A_{i,t-1}^\mu)$	0.0059 (0.0062)	0.0058 (0.0053)	-0.0176 (0.0107)	-0.0171 (0.0094)	0 (1e-04)	-1e-04 (1e-04)	-0.0102 (0.0053)	-0.0095* (0.0048)	0.0089 (0.0136)	0.004 (0.0121)	-0.0105 (0.0102)	-0.0137 (0.0087)
$gr(A_{i,t-1}^\tau)$	-0.0606** (0.0212)	-0.0606** (0.0193)	-0.0393 (0.0282)	-0.0426 (0.0252)	8e-04** (3e-04)	8e-04** (3e-04)	-0.0082 (0.014)	-0.0015 (0.0125)	-0.0257 (0.0489)	-0.041 (0.0445)	-0.0077 (0.035)	-0.0209 (0.0308)
$gr(PR_{i,t-1}^{\mu,d})$	-3e-04 (6e-04)	0.0022*** (6e-04)	-5e-04 (0.0011)	-9e-04 (0.0011)	0 (0)	0** (0)	1e-04 (6e-04)	-2e-04 (6e-04)	-0.002 (0.0014)	0.0023 (0.0012)	-0.0016 (0.0012)	0.0018 (0.001)
$gr(PR_{i,t-1}^{\tau,d})$	-1e-04 (0.0011)	-5e-04 (0.0011)	-0.003 (0.0019)	-0.0044* (0.0022)	0 (0)	0 (0)	0.0023** (9e-04)	0.0025* (0.0012)	-0.0039 (0.0021)	-0.0042 (0.0022)	-0.003 (0.0018)	-0.0029 (0.0019)
$Spill(gr(A))_{i,t-1}^{\mu,d}$	-1.132 (0.612)	-0.531 (0.4353)	-1.777 (1.081)	0.2639 (0.867)	0.0065 (0.0091)	0.0137 (0.0083)	-0.6032 (0.4777)	-0.5336 (0.3917)	-1.938* (0.9459)	0.4388 (0.9196)	-1.879 (1.035)	0.1101 (0.7847)
$Spill(gr(A))_{i,t-1}^{\tau,d}$	2.015 (4.723)	4.348 (4.564)	-10.83 (4.983)	-13.31 (8.983)	-0.0109 (0.0908)	-0.0483 (0.0853)	4.161 (4.702)	4.019 (4.752)	3.135 (8.712)	6.346 (8.966)	0.3162 (7.667)	0.4101 (7.534)
$gr(L_{i,t-1})$	-0.0212 (0.0608)	-0.0272 (0.0592)	0.7454*** (0.1109)	0.7398*** (0.1097)	0.0055*** (0.0015)	0.0056*** (0.0015)	-0.1753* (0.0707)	-0.1733* (0.0705)	-0.6283*** (0.1257)	-0.6397*** (0.1243)	0.0717 (0.0827)	0.062 (0.0838)
$gr(Wage_{i,t-1})$	0.013 (0.0883)	0.0379 (0.0851)	0.5527** (0.1829)	0.5735** (0.1796)	0.0015 (0.0017)	0.0011 (0.0017)	-0.2543* (0.1253)	-0.2517* (0.1267)	0.1599 (0.2229)	0.1849 (0.2224)	0.079 (0.1513)	0.1048 (0.1517)
$gr((K/L)_{i,t-1})$	-8.637 (4.509)	-7.898 (4.371)	-7.743 (8.125)	-7.824 (8.051)	0.2009 (0.1276)	0.1908 (0.1255)	18.82*** (4.099)	18.81*** (4.109)	-39.11*** (11.11)	-38.42*** (11.05)	4.998 (4.476)	5.555 (4.52)
$gr((Lprod/L)_{i,t-1})$	0.3376** (0.1109)	0.3285** (0.1117)	0.4894** (0.1796)	0.5094** (0.1798)	-0.0051** (0.0017)	-0.0049** (0.0016)	0.0231 (0.1094)	0.0145 (0.1104)	0.1948 (0.2059)	0.2056 (0.2054)	0.4135* (0.1818)	0.4205* (0.184)
$gr((I/L)_{i,t-1})$	0.0727*** (0.0141)	0.0704*** (0.0139)	0.0295 (0.0207)	0.0287 (0.0206)	8e-04** (3e-04)	8e-04** (3e-04)	0.0021 (0.0109)	0.002 (0.0107)	-0.2857*** (0.0302)	-0.2874*** (0.0306)	0.1455*** (0.0246)	0.144*** (0.0248)
$gr((VA/L)_{i,t-1})$	0.0338 (0.0476)	0.0259 (0.0457)	-0.019 (0.0546)	-0.0145 (0.055)	0.0015 (0.001)	0.0016 (0.001)	0.023 (0.0354)	0.0222 (0.0356)	-0.0126 (0.0845)	-0.0154 (0.0854)	0.131 (0.0684)	0.128 (0.0689)
$gr(TFP5_{i,t-1})$	2.783*** (0.3188)	2.781*** (0.3187)	2.546*** (0.6486)	2.596*** (0.637)	-0.0531*** (0.012)	-0.0533*** (0.0117)	-1.272** (0.4577)	-1.248** (0.4503)	1.891** (0.6463)	1.849** (0.641)	1.177** (0.4478)	1.163* (0.4512)
R^2	0.1969	0.2034	0.2387	0.2393	0.1998	0.2097	0.2737	0.2743	0.1645	0.1636	0.0824	0.0797
N	1661	1661	1661	1661	1661	1661	1661	1661	1661	1661	1661	1661
Average	-0.05	-0.05	3.137	3.137	0.0019	0.0019	0.7256	0.7256	0.1988	0.1988	1.945	1.945

Notes: The regressions aim to identify the drivers of technological change reflected in changing input factor use in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $gr(A_{i,t}^\alpha)$, $gr(PR_{i,t}^{\alpha,d})$, $Spill(gr(A))_{i,t}^{\alpha,d}$, $gr(L_{i,t})$, $gr(Wage_{i,t})$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $L_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $gr((VA/L)_{i,t})$, $gr((VA/L)_{i,t})$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table SI.3: Regression results: Patterns of changing input factor use

SI.2.2.2. Robustness check using non-citation weighted patents

	$gr(A_{i,t}^\mu)$		$A_{i,t}^\mu$		$gr(A_{i,t}^\tau)$		$A_{i,t}^\tau$	
	up	down	up	down	up	down	up	down
$A_{i,t-1}^\mu$	-0.4962*** (0.0294)	-0.4943*** (0.0255)	0.4368*** (0.0334)	0.4179*** (0.0298)	-5e-04 (0.0066)	0 (0.0063)	0.0017 (0.0068)	0.0019 (0.0065)
$A_{i,t-1}^\tau$	-0.0909 (0.0748)	-0.0131 (0.0734)	-0.0932 (0.0876)	-0.002 (0.0823)	-0.1831*** (0.0261)	-0.2154*** (0.0258)	0.8069*** (0.0287)	0.7759*** (0.0274)
$PR_{i,t-1}^{\mu,d}$	-0.0426 (0.0351)	0.1529*** (0.0348)	-0.0971* (0.0429)	0.2115*** (0.0419)	0.0133 (0.0104)	-0.0154 (0.0096)	0.0113 (0.0103)	-0.0152 (0.0098)
$PR_{i,t-1}^{\tau,d}$	0.4725** (0.1713)	0.2797 (0.1898)	0.4894* (0.1945)	0.2889 (0.2161)	0.0544 (0.0726)	0.1314* (0.0641)	0.0401 (0.0786)	0.114 (0.068)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0036 (0.0338)	-0.017 (0.0318)	0.0103 (0.0397)	0.0151 (0.0366)	0.0089 (0.0109)	0.003 (0.0091)	0.0093 (0.0117)	0.0014 (0.0094)
$Spill(A)_{i,t-1}^{\tau,d}$	-0.0377 (0.1533)	-0.1128 (0.1504)	-0.0234 (0.1759)	-0.1755 (0.1632)	0.2288*** (0.0574)	0.1751*** (0.049)	0.2108*** (0.0506)	0.1579*** (0.0459)
$L_{i,t-1}$	-0.1855 (0.0968)	-0.2125* (0.1009)	-0.2092 (0.1192)	-0.247* (0.1242)	0.005 (0.0226)	0.0121 (0.0227)	0.0036 (0.0242)	0.01 (0.0241)
$Wage_{i,t-1}$	0.3852 (0.4128)	0.454 (0.4066)	0.7406 (0.4601)	0.8078 (0.4521)	-0.0454 (0.104)	-0.0575 (0.1059)	-0.0464 (0.1114)	-0.058 (0.1127)
$(K/L)_{i,t-1}$	-0.6886** (0.2219)	-0.7053** (0.2215)	-0.7485** (0.25)	-0.768** (0.2493)	-0.0081 (0.0582)	0.0076 (0.0561)	-0.0162 (0.0624)	-0.0017 (0.0606)
$(Lprod/L)_{i,t-1}$	-0.5226 (0.5864)	-0.5062 (0.5836)	-1.124 (0.6656)	-1.111 (0.667)	-0.0894 (0.1706)	-0.1508 (0.1735)	-0.0955 (0.1813)	-0.1521 (0.1853)
$(I/L)_{i,t-1}$	-0.0742 (0.0881)	-0.0651 (0.0894)	-0.0305 (0.101)	-0.0115 (0.1008)	0.0334 (0.0235)	0.0364 (0.0234)	0.0317 (0.0243)	0.0344 (0.0242)
$(VA/L)_{i,t-1}$	0.2935 (0.158)	0.2383 (0.1521)	0.1523 (0.1871)	0.0819 (0.1805)	0.127*** (0.0377)	0.1303** (0.0397)	0.1254** (0.041)	0.1287** (0.0428)
$TFP5_{i,t-1}$	-0.2974 (0.1868)	-0.2658 (0.1809)	-0.4206 (0.2256)	-0.3983 (0.2135)	0.06 (0.0712)	0.0562 (0.0698)	0.0863 (0.0734)	0.0825 (0.0706)
R^2	0.2806	0.2882	0.1762	0.187	0.1112	0.1027	0.6429	0.6401
N	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.3638	0.3638	7.683	7.683	0.0966	0.0966	5.85	5.85

Notes: The regressions aim to explain the factors that influence the evolution of industry sizes in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^\alpha$, $gr(A_{i,t}^\tau)$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table SI.4: Regression results: The evolution of industry sizes

	$gr((VA/L)_{i,t})$		$(VA/L)_{i,t}$		$gr(TFP4_{i,t})$		$TFP4_{i,t}$		$gr(TFP5_{i,t})$		$TFP5_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$A_{i,t-1}^{\mu}$	0.0164** (0.0063)	0.0053 (0.0057)	0.0173** (0.0065)	0.0059 (0.0058)	4e-04 (5e-04)	3e-04 (5e-04)	0.004 (0.0026)	0.0039 (0.002)	3e-04 (5e-04)	2e-04 (5e-04)	0.0041 (0.0026)	0.0039 (0.0021)
$A_{i,t-1}^{\tau}$	-0.0358 (0.0272)	-0.0206 (0.0273)	-0.0364 (0.0273)	-0.0205 (0.0275)	2e-04 (0.0016)	1e-04 (0.0016)	0.0014 (0.0065)	0.0057 (0.0068)	2e-04 (0.0016)	1e-04 (0.0016)	0.0013 (0.0065)	0.0056 (0.0068)
$PR_{i,t-1}^{\mu,d}$	-0.0361** (0.0124)	0.0285** (0.0103)	-0.0377** (0.0128)	0.0294** (0.0106)	-0.0011 (0.001)	-4e-04 (7e-04)	-0.0017 (0.006)	0.0066 (0.0037)	-0.001 (0.001)	-4e-04 (7e-04)	-0.0018 (0.006)	0.0067 (0.0037)
$PR_{i,t-1}^{\tau,d}$	0.0496 (0.0498)	-0.0122 (0.0626)	0.0434 (0.0521)	-0.0214 (0.0664)	0.003 (0.0042)	0.0032 (0.0048)	0.0111 (0.0168)	0.0033 (0.0196)	0.003 (0.0042)	0.0033 (0.0048)	0.0112 (0.0168)	0.0034 (0.0196)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.004 (0.0089)	0.008 (0.0094)	-0.0039 (0.0091)	0.0076 (0.0096)	-0.0011 (9e-04)	-0.0015 (8e-04)	-2e-04 (0.0041)	7e-04 (0.0036)	-9e-04 (9e-04)	-0.0014 (8e-04)	-3e-04 (0.0041)	8e-04 (0.0036)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0256 (0.0362)	0.0236 (0.0383)	0.0214 (0.0373)	0.0185 (0.0391)	-4e-04 (0.003)	-0.0046 (0.0032)	-0.0122 (0.0172)	-0.0211 (0.018)	-6e-04 (0.003)	-0.005 (0.0032)	-0.0123 (0.0172)	-0.021 (0.018)
$L_{i,t-1}$	-0.0166 (0.025)	-0.0182 (0.0257)	-0.0165 (0.0252)	-0.0184 (0.026)	2e-04 (0.0021)	2e-04 (0.0021)	-0.0039 (0.0104)	-0.0052 (0.0101)	1e-04 (0.002)	2e-04 (0.0021)	-0.0039 (0.0104)	-0.0053 (0.0101)
$Wage_{i,t-1}$	0.0953 (0.1369)	0.105 (0.1387)	0.1105 (0.1392)	0.1208 (0.1412)	-0.0096 (0.0099)	-0.0104 (0.0098)	-0.0156 (0.0372)	-0.0146 (0.0365)	-0.0104 (0.0098)	-0.0113 (0.0098)	-0.0158 (0.0372)	-0.0148 (0.0365)
$(K/L)_{i,t-1}$	0.1139 (0.0859)	0.1159 (0.087)	0.1068 (0.0865)	0.1086 (0.0878)	0.0204*** (0.0053)	0.0214*** (0.0053)	0.1492*** (0.0272)	0.1491*** (0.026)	0.0202*** (0.0053)	0.0212*** (0.0052)	0.1495*** (0.0271)	0.1494*** (0.0259)
$(Lprod/L)_{i,t-1}$	0.1205 (0.1652)	0.1042 (0.1661)	0.1507 (0.1769)	0.1347 (0.1773)	-0.0047 (0.0169)	-0.0063 (0.0168)	0.0472 (0.0811)	0.0481 (0.0803)	-0.0064 (0.0167)	-0.0079 (0.0166)	0.0448 (0.0815)	0.0458 (0.0808)
$(I/L)_{i,t-1}$	-0.016 (0.0252)	-0.0137 (0.0246)	-0.0185 (0.0255)	-0.0163 (0.0249)	-6e-04 (0.0024)	-9e-04 (0.0024)	-0.029** (0.0099)	-0.0286** (0.0098)	-4e-04 (0.0024)	-7e-04 (0.0024)	-0.0289** (0.0099)	-0.0285** (0.0098)
$(VA/L)_{i,t-1}$	-0.4393*** (0.0511)	-0.443*** (0.0516)	0.5579*** (0.0525)	0.554*** (0.0531)	-0.0148*** (0.004)	-0.0142*** (0.0039)	-0.0547* (0.0213)	-0.0568** (0.0204)	-0.0147*** (0.004)	-0.0141*** (0.0039)	-0.055** (0.0213)	-0.0571** (0.0204)
$TFP4_{i,t-1}$					-0.0575*** (0.009)	-0.058*** (0.0087)	0.9738*** (0.0827)	0.9739*** (0.0829)				
$TFP5_{i,t-1}$	0.0982 (0.0719)	0.0997 (0.0736)	0.0985 (0.0714)	0.1001 (0.0733)					-0.0574*** (0.009)	-0.0579*** (0.0086)	0.9736*** (0.0825)	0.9737*** (0.0827)
R^2	0.1791	0.1767	0.3319	0.3299	0.181	0.183	0.7083	0.7095	0.1806	0.1829	0.7083	0.7095
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	0.2664	0.2664	4.48	4.48	0.6937	0.6937	0.6754	0.6754	0.6938	0.6938	0.6756	0.6756

Notes: The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $Wage_{i,t}$, $(K/L)_{i,t}$, $(I/L)_{i,t}$, $(VA/L)_{i,t}$, $TFP4_{i,t}$, $TFP5_{i,t}$, $gr((VA/L)_{i,t})$, $gr(TFP4_{i,t})$, $gr(TFP5_{i,t})$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3. The difference between four and five factor TFP is that $TFP5_{i,t}$ is a disaggregation of material inputs into energy and non-energy materials which is not done for $TFP4_{i,t}$ (see Bartlesman and Gray (1996)).

Table SI.5: Regression results: Determinants of productivity growth

	$gr(L_{i,t})$		$L_{i,t}$		$gr((K/L)_{i,t})$		$(K/L)_{i,t}$		$gr((I/L)_{i,t})$		$(I/L)_{i,t}$	
	up	down	up	down	up	down	up	down	up	down	up	down
$A_{i,t-1}^{\mu}$	0.0026 (0.0066)	-0.0012 (0.0057)	0.0052 (0.0086)	8e-04 (0.0076)	-1e-04 (1e-04)	-1e-04 (1e-04)	-0.0027 (0.0045)	-0.002 (0.0038)	0.0094 (0.0117)	-0.0057 (0.0107)	0.0089 (0.01)	-0.0043 (0.0093)
$A_{i,t-1}^{\tau}$	-0.0177 (0.0158)	-0.0275 (0.0153)	-0.021 (0.0198)	-0.0315 (0.0188)	1e-04 (3e-04)	2e-04 (3e-04)	0.0081 (0.0111)	0.0069 (0.0097)	-0.0401 (0.0351)	-0.0475 (0.0362)	-0.0441 (0.0315)	-0.0466 (0.0318)
$PR_{i,t-1}^{\mu,d}$	-0.009 (0.0087)	0.0144 (0.008)	-0.0122 (0.0106)	0.0209 (0.0108)	1e-04 (1e-04)	-1e-04 (2e-04)	0.0049 (0.0056)	-0.0056 (0.0053)	-0.0328* (0.0163)	0.034* (0.0132)	-0.028 (0.0147)	0.0323** (0.0119)
$PR_{i,t-1}^{\tau,d}$	-0.0217 (0.0379)	0.0111 (0.0461)	-0.0367 (0.0476)	-0.0049 (0.0578)	8e-04 (6e-04)	6e-04 (8e-04)	0.0247 (0.0244)	0.0404 (0.0306)	0.1035 (0.0764)	0.1773 (0.0907)	0.1116 (0.0718)	0.1612 (0.0838)
$Spill(A)_{i,t-1}^{\mu,d}$	-0.0202* (0.0092)	0.0103 (0.008)	-0.0202 (0.0128)	0.011 (0.0105)	0 (2e-04)	1e-04 (2e-04)	0.0108 (0.0063)	0 (0.0053)	-0.02 (0.0157)	0.032* (0.0139)	-0.0136 (0.0144)	0.0295* (0.0128)
$Spill(A)_{i,t-1}^{\tau,d}$	0.0934** (0.0322)	0.0959** (0.0313)	0.107** (0.0399)	0.1164** (0.0399)	-9e-04 (6e-04)	-0.0011* (6e-04)	-0.0189 (0.0214)	-0.0478* (0.0228)	0.0908 (0.0626)	0.058 (0.0602)	0.0808 (0.056)	0.049 (0.0562)
$L_{i,t-1}$	-0.1919*** (0.023)	-0.1926*** (0.0236)	0.7955*** (0.0304)	0.7935*** (0.031)	0.001 (5e-04)	0.001* (5e-04)	0.0203 (0.017)	0.0201 (0.0172)	-0.0839* (0.0392)	-0.0886* (0.0392)	-0.0942** (0.035)	-0.098** (0.0349)
$Wage_{i,t-1}$	-0.0508 (0.0774)	-0.0468 (0.0769)	-0.0259 (0.0995)	-0.0162 (0.0978)	3e-04 (0.0015)	1e-04 (0.0014)	0.0532 (0.0595)	0.0436 (0.0577)	0.2709 (0.1435)	0.2588 (0.1458)	0.3466* (0.1366)	0.3368* (0.1392)
$(K/L)_{i,t-1}$	-0.1993*** (0.0501)	-0.1957*** (0.0474)	-0.1746** (0.0651)	-0.1737** (0.0617)	-0.0064*** (0.0011)	-0.0063*** (0.0011)	0.7915*** (0.0394)	0.7928*** (0.0381)	-0.173 (0.0977)	-0.1701 (0.096)	-0.2314* (0.0904)	-0.2283* (0.0889)
$(Lprod/L)_{i,t-1}$	0.8272*** (0.1627)	0.811*** (0.1609)	0.9709*** (0.2121)	0.9567*** (0.2091)	-0.0106** (0.0035)	-0.0106** (0.0035)	-0.5231*** (0.1211)	-0.5213*** (0.1201)	0.5792* (0.2591)	0.575* (0.2562)	0.6188* (0.2514)	0.6107* (0.2491)
$(I/L)_{i,t-1}$	0.0371 (0.0201)	0.0378 (0.02)	0.0502* (0.024)	0.0518* (0.0241)	0.0022*** (4e-04)	0.0022*** (4e-04)	0.1199*** (0.0171)	0.1208*** (0.0171)	-0.7633*** (0.0421)	-0.7586*** (0.0413)	0.3046*** (0.0434)	0.3095*** (0.0428)
$(VA/L)_{i,t-1}$	0.0809* (0.0376)	0.0791* (0.0358)	0.0573 (0.0485)	0.0527 (0.0469)	6e-04 (9e-04)	6e-04 (9e-04)	-0.0067 (0.0281)	-0.0063 (0.0273)	0.1048 (0.0663)	0.1009 (0.0661)	0.113 (0.0609)	0.1084 (0.0604)
$TFP5_{i,t-1}$	-0.1332* (0.0621)	-0.1435* (0.0602)	-0.1925* (0.0758)	-0.2018** (0.0722)	0.0031** (0.0011)	0.0031** (0.001)	0.1337** (0.0438)	0.1334** (0.0417)	0.1512 (0.1153)	0.1286 (0.1125)	0.1133 (0.1092)	0.0963 (0.1064)
R^2	0.1541	0.1517	0.683	0.6832	0.2155	0.2161	0.6538	0.6548	0.3219	0.3235	0.1485	0.1514
N	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950	1950
Average	-0.0354	-0.0354	3.166	3.166	0.0018	0.0018	0.6928	0.6928	0.2482	0.2482	1.865	1.865

Notes: The regressions aim to identify the drivers of technological change reflected in changing input factor use in a balanced panel of 6-digit level NAICS manufacturing industries. The regression analyses include industry and time fixed effects. To cope with skewness and to obtain tractable coefficients, most variables pre-processed (taking logs, removing outliers, scaling). Data in logs are $A_{i,t}^{\alpha}$, $PR_{i,t}^{\alpha,d}$, $Spill(A)_{i,t}^{\alpha,d}$, $L_{i,t}$, $gr(L_{i,t})$, $Wage_{i,t}$, $(K/L)_{i,t}$, $gr((K/L)_{i,t})$, $gr((I/L)_{i,t})$, $(VA/L)_{i,t}$, $TFP5_{i,t}$ with $\alpha = \mu, \tau$ and $d = in, out$. A detailed description of the transformation steps and descriptive statistics of the regression data before and after the transformations are provided in A.3.

Table SI.6: Regression results: Patterns of changing input factor use