Received: 25 June 2020 | Revised: 27 October 2021 | Accepted: 1 November 2021

DOI: 10.1111/twec.13227

ORIGINAL ARTICLE

The World Economy WILEY

Productivity and wage effects of firm-level upstreamness: Evidence from Belgian linked panel data

Benoît Mahy[1] | **François Rycx**[2] | **Guillaume Vermeylen**[3] | **Mélanie Volral**[1]

[1]DULBEA, University of Mons (Soci&ter), Mons, Belgium

[2]Soci&ter, DULBEA, CEBRIG, IRES, GLO, IZA, Université Libre de Bruxelles, Bruxelles, Belgium

[3]Soci&ter, CEBRIG, DULBEA, University of Mons (Soci&ter), Mons, Belgium

Correspondence
Guillaume Vermeylen, CEBRIG, DULBEA, University of Mons (Soci&ter), Mons, Belgium.
Email: Guillaume.Vermeylen@umons.ac.be

Funding information
National Bank of Belgium

Abstract

This paper is the first to estimate the impact of a direct measurement of firm-level upstreamness (i.e. the steps—weighted distance—before the production of a firm meets either domestic or foreign final demand) on productivity and wage costs. To this end, we merged detailed Belgian linked panel data, covering all years from 2002 to 2010, with a unique dataset containing accurate information on the yearly position of each firm in the value chain. We rely on the methodological framework pioneered by Hellerstein et al. (*Journal of Labor Economics*, 1999, 17, 409) to estimate panel data models at the firm level. Controlling for key worker and firm characteristics, our GMM-SYS and FE-IV estimates show that firms positioned more upstream (i.e. further away from the final consumer) create more value. Our results also indicate that the impact of firm-level upstreamness is stronger on productivity than on wage costs, which implies that profitability is fostered. More precisely, in line with Belgium's strong labour market institutions, our estimates suggest that the productivity gains associated with upstreamness are shared almost equally between wages and profits.

KEYWORDS

global value chains, linked employer–employee panel data, productivity, rent sharing, upstreamness

1 | INTRODUCTION

Over the last 30 years, production processes have become increasingly fragmented and divided into ever smaller parts, considered as separate activities (OECD, 2013). In order to minimise costs, the production decision process now involves the sourcing of inputs from multiple suppliers, often located in foreign countries (Antràs et al., 2012; Manello et al., 2016). This has resulted in the emergence of truly global value chains[1] (GVCs), in contrast with the integrated production processes supported by the traditional view of international trade. Baldwin (2011) explains that this radical change is due to the sharp decline in coordination costs, induced by the information and communication technology (ICT) revolution. This reduction in costs enables firms to stop bundling all major stages of their production process in the same location. In other words, it becomes easier and inevitable for firms to unbundle factories in order to achieve economies of scale and obtain comparative advantages (Baldwin, 2011). This fragmentation particularly concerns small open economies such as Belgium. If we consider, for instance, the share of imported inputs in total intermediate inputs, that is a standard measurement of offshoring (Feenstra & Hanson, 1996, 1999), Belgium is ranked 7th among 35 investigated OECD countries. This share reaches 34% in Belgium, compared to a weighted OECD average of 16% (OECD, 2010). In the same line, Dhyne and Duprez (2015) showed that, between 2002 and 2012, 82% of commercial enterprises in Belgium were producing goods and services that were directly or indirectly exported, and that 99% consumed imported goods and services over the same period.

In this context, firms' adoption of appropriate strategies for participating in and positioning themselves within GVCs is likely to be decisive for their performance and, more generally, for the competitiveness and growth of European countries that have endured a greater slowdown in productivity (total factor productivity) gains than the United States (Dhyne & Fuss, 2014). However, economic analyses have hitherto taken very little notice of the role that firms' position in GVCs plays in creating value (Dhyne & Duprez, 2015), and empirical evidence on this issue is still scarce (Amador & Cabral, 2016). At the aggregate level, the descriptive statistics presented by the OECD (2012) suggest that most of the value is created in upstream activities, that is in the initial stages of the production process (e.g. innovation, R&D, design), and downstream activities, that is those closer to the end user (e.g. marketing, branding, logistics). In contrast, the value created 'in-between', in pure manufacturing/assembling stages, appears to be limited. The OECD (2012) also stressed the need for countries to 'move up the value chain', that is to specialise in initial production stages, in order to create more value. Specialisation in more upstream activities is also likely, according to the OECD, to increase firms' control over high-value downstream stages of the production process and thus to promote economic growth.

In the economic literature, studies conducted at the firm level have also paid very little attention to the role that a firm's position in a GVC plays in value creation. To the best of our knowledge, only three papers have explicitly examined and tested whether and how that position influences the creation of value.[2] First, Rungi and Del Prete (2018) determined a firm's position in a GVC by merging its core industrial activity (at the NAICS 4-digit level) in the

[1]According to Gereffi (2014), GVCs might be defined as the full range of activities to bring a good or a service from its conception to its end use (final demand), these activities being divided among multiple firms and geographical areas.

[2]The scarcity of evidence on this issue can be explained by the relative novelty of accurate measurements of firms' position in GVCs, such as upstreamness (see Antràs et al., 2012; Fally 2011), and by the difficulty to obtain the data necessary to compute these measurements.

productive sequence with downstreamness measurements (i.e. the distance, in the GVC, between the first level of value-added creation and that of the firm) sourced from Antràs and Chor (2013). In other words, the authors used an indirect cross-sectional indicator of downstreamness, measured at the sectoral level. Based on a cross-section dataset of about 2 million firms located in the European Union for the year 2015 and controlling for country-, industry- and firm-level characteristics, their OLS and fractional probit findings suggest the existence of a 'smile curve'. According to their results, the creation of value thus appears to be the highest for activities at the top and at the bottom of the supply chain. In contrast, intermediate activities are found to bring less value and therefore to be more likely to be offshored, notably to emerging economies.

The second related study is that of Ju and Yu (2015). The authors investigated how the position of a firm in a GVC, measured through an upstreamness index (i.e. the average distance from the firm's production level to the final use), affects its productivity and profitability. Applying the methodology developed by Antràs et al. (2012) to Chinese data, they computed: (i) an industrial upstreamness index for 120 different sectors; and (ii) a firm-level upstreamness index—though only for exporting firms—in an indirect way, that is using a firm's average upstreamness in exports as a proxy for its production index. Controlling for firm characteristics, their OLS estimates suggest that upstreamness fosters productivity and profits. Ju and Yu (2015) also showed that firms in upstream industries are more capital-intensive. Accordingly, the higher cut-off productivity level required to operate in more upstream industries could explain, in turn, why firms in those industries are found to be more productive and profitable.

The third relevant study is that of de Vries et al. (2021), who analyse whether total factor productivity of Dutch manufacturing firms could depend on their functional specialisation as well as the more upstream or downstream position of their products. The authors pool cross-sectional data from firms observed in two waves of surveys conducted in 2012 and 2017. Building on Chor et al. (2021), they measure the upstream and downstream position of firms by considering the position of their exported products only. They indirectly determine the position of each firm from the weighted average of the values of its exported products in relation to its total exports, multiplied by the upstream or downstream levels calculated for 54 sectors to which these products correspond. Their results, estimated using the procedure of Wooldridge (2009), suggest that firms specialising in R&D and marketing activities are more productive than those specialising in manufacturing, and that a more upstream position has no influence on the productivity of firms as such, neither through an increase in production efficiency nor through increased market power.

Compared to previous studies and their contrasting results, this paper provides novel evidence on this important—but still under-researched—topic in three different ways. Taking advantage of a rich dataset, the first contribution of this paper comes from its ability to rely on precise information regarding each firm's productivity (i.e. on the value added per worker) and particularly on an accurate and direct indicator of upstreamness, truly computed at the firm level, in order to investigate the relationship between firms' position in GVCs and their productivity. Indeed, the merging of the National Bank of Belgium business-to-business (NBB B2B) transactions dataset with detailed linked panel data, for the purpose of our study, gives us unique longitudinal information on the number of steps (weighted distance) necessary before the production of (almost) each commercial firm[3] (operating in both manufacturing and services) meets either domestic or foreign final demand by using

[3]For instance, a few enterprises, such as sole traders and those who are not required to fill VAT declarations, are not included in the NBB B2B transactions dataset.

information on the upstreamness of each transaction in each firm, whereas the three aforementioned studies relied on indirect sectoral measurements of the upstream position of firms. Our dataset allows us not only to measure whether (and to what extent) the relationship is positive, negative or insignificant, but also to test more appropriately for possible nonlinearities in the nexus between upstreamness and productivity, for whether the impact of upstreamness on productivity tends to be linear (Ju & Yu, 2015), nonlinear (Rungi & Del Prete, 2018) or not significant at all (de Vries et al., 2021).

The ex-ante nature of the relationship between the position of firms in GVCs and the creation of value is ambiguous because of the different channels through which upstreamness can influence the productivity of firms, either positively or negatively. We identify a first set of channels suggesting a positive relationship: these focus on the productive advantages that more upstream firms might derive from their export position, interactions with more productive (downstream) partners, increased control over the value chain, greater R&D and capital intensity, or additional product market power in particular industries.

A first channel is based on the assumption that exporting firms are more likely to be higher up in the value chain (Amador & Cabral, 2016; Dhyne & Duprez, 2013, 2015) and on strong evidence that exporting firms are more productive (Baldwin & Yan, 2017; Berthou et al., 2015). The causal impact of exports (and thus of upstreamness) on productivity could be explained not only by greater specialisation and increased economies of scale/scope (OECD, 2013), but also by the larger foreign markets associated with exports, which allows exporting (i.e. more upstream) firms to benefit from learning about new technologies and products and to increase their incentives to invest and innovate (Baldwin & Yan, 2017).

Next, Serpa and Krishnan (2018) have pointed out that a more upstream position could foster firms' productivity through upstream spillovers (i.e. from large customers to suppliers), whereby a firm's vertical partner (customer) could improve the productivity of that firm (supplier). The authors' results indicate that endogenous channels, in which the firm benefits directly from the interaction with a more productive partner through knowledge transfer, tend to be the most important source of spillovers. Considering value chains led by multinational corporations, Pietrobelli and Saliola (2008) also find that the interaction between (more downstream) leading firms and their (more upstream) suppliers is also important in explaining the productivity of upstream suppliers. Their estimates suggest that more intense buyer involvement with (local) suppliers, not only in product design and quality but also in technology diffusion and R&D, is generally associated with higher upstream supplier productivity.

Although it has also been generally argued that some downstream activities (such as marketing, logistics and branding) create a lot of value (OECD, 2013), another channel could be that leading firms, who are mostly upstream because of their intensive innovation and R&D activities, tend to keep a strong control over high-value downstream activities. Indeed, upstream R&D is often seen as a key step for chain control purposes, as illustrated by the innovation that enabled Apple to create the iPod and to extract much of the value created by accompanying this innovation with appropriate managerial practices (Dedrick et al., 2010).

Given that about 90% of Belgian firms operate in imperfectly competitive product markets (Dobbelaere & Vancauteren, 2014), one could also argue that a firm's more upstream position should promote labour productivity within that firm by increasing its market power and allowing it to set higher prices. Indeed, Alfaro et al. (2019) show that a firm is more likely to integrate relatively upstream steps in the value chain, while engaging in subcontracting to downstream suppliers, when the price elasticity of demand for its product is lower and contributes thereby to increasing its market power.

A final channel is related to the fact that more upstream firms are generally more R&D and capital-intensive. Indeed, as highlighted by Shen et al. (2021), this could further improve their growth prospects and give them more opportunities to outsource some of the non-core operations that offer lower value added and, thereby, to increase their supply base capabilities.

We also identify a second set of channels suggesting, contrarily, that the relationship between upstreamness and firm productivity may actually be negative. These channels focus on the productive disadvantage that more upstream firms may face when they are unable to control the entire value chain and in particular to engage in high-value-added downstream activities (which are often associated with increased market power), or that being more upstream prevents them from taking advantage of a position as both buyer and supplier.

Based on the results of de Vries et al. (2021) and Rungi and Del Prete (2018), we have already noted that engaging in downstream activities could be expected to improve productivity. Depending on the type of activity and industry, the search for higher rents in value chains may therefore involve the development of more downstream activities. For instance, in order to gain more control and market power in their value chain in the food industry, Belgian farmers have expanded their upstream activities, such as milk production, to downstream activities, such as processing the milk into cheese and selling it to final demand locally. This strategy allowed them to capture more of the value created along the chain while getting access to additional activities with higher market power. In the automotive industry, Sturgeon et al. (2008) note that for technical, political and commercial reasons, producers may prefer to keep final assembly of vehicles close to final demand, rather than exporting the assembled vehicles. Engaging in additional downstream production then still provides a source of additional value and market power that increases prices and thereby productivity. In the food industry, the wave of mega-mergers has accelerated consolidation and concentrated market power in the hands of a few dominant (downstream) firms (Woodall & Shannon, 2018). In such a case, being positioned more downstream still seems better for capturing additional value.

Giovannetti and Marvasi (2018) advance another argument for why moving further upstream may penalise labour productivity. The authors point out that labour productivity (as proxied by sales per employee) is enhanced when firms are, at the same time, buyers and suppliers in the value chain. Indeed, their OLS results suggest that Italian upstream firms that produce intermediate products for other firms are less productive than midstream firms that buy and supply intermediate products, which are in turn less productive than downstream firms that buy intermediate products and produce final consumption goods.

In addition to our ability to estimate the theoretically ambiguous relationship between the position and productivity of firms by using an accurate measurement of upstreamness, this paper's second source of added value comes from the availability of our panel data and the use of appropriate estimators they allow.

Because of this, we can better take other explanatory factors of productivity into account and estimate the role of upstreamness as such. With regard to these factors, Syverson (2011) highlights internal and external productivity drivers. The former can impact productivity by acting within the firm, such as managerial talent, information technology and R&D investments, or learning by doing. The latter refer to the operating environments of producers that induce spillover effects on productivity, such as knowledge transfer, business competition and flexible input markets.

By applying the GMM-SYS and FE-IV estimators to our panel data, we can control for the effects of unobserved (constant) heterogeneity across firms, which may result from the aforementioned productivity factors, considered separately or interdependently, assuming these sources of heterogeneity remain stable over our observation period, that is from 2002 to 2010. To that extent, our data first allow us to identify the role of (upstream) position in the overall value chain as such, when other productivity factors are taken into account. Our estimators also allow us to control for simultaneity between a firm's level of upstreamness and its productivity. Indeed, reverse causality between upstreamness and productivity/wages is likely to be encountered given that (1) the number of steps before a firm's output meets final demand is typically larger among exporting firms (OECD, 2012), and (2) huge evidence supports reverse causality between firms' export behaviour and their productivity/wages (Arnold & Hussinger, 2005; Berthou & Vicard, 2013; Eaton et al., 2007; Freund & Pierola, 2010), with more productive firms self-selecting into export markets. All in all, using our panel data also allows us to address important methodological issues, namely firm-level fixed effects (productivity drivers) and the potential endogeneity of upstreamness, which neither Ju and Yu (2015), nor Rungi and Del Prete (2018), nor de Vries et al. (2021) (with respect to firm fixed effects) control for, by using both GMM-SYS and FE-IV estimators.

Moreover, using panel data allows us to model and examine the dynamics of productivity and wages, and more specifically to account for their persistence. Our panel data also cover a large part of the private sector and allow us to control for key worker characteristics (e.g. education, age, occupation, working time) in addition to the usual firm characteristics that are considered in the few existing studies (capital stock, size, industry).

A third main interest of our paper is to add to the existing literature by focusing on distributional issues that is whether and to what extent potential productivity gains from a more upstream position are shared between extra profits and wages at the firm level. Put differently, we investigate not only whether and how upstreamness affects firm-level value added, but also how the productivity gains generated along GVCs are shared between capital and labour.

With regard to the bargaining pattern in Belgium, Dobbelaere and Vancauteren (2014) estimate that 51 per cent of Belgian firms in industry and services sectors are characterised by efficient bargaining, 22 per cent by monopsony and the remaining 27 per cent behave in perfect competitive labour markets or in line with the right-to-manage bargaining model. Labour market institutions in Belgium are notably characterised by strong centralisation/coordination of collective bargaining and high union coverage/density (Garnero et al., 2020; Kampelmann & Rycx, 2013; OECD, 2018). Belgium is also one of the few countries where wages are automatically indexed to the inflation rate. Wages are negotiated mainly at the industry level but can also be renegotiated at the company level. In this institutional setting, our objective is to estimate whether and to what extent the potential gains or losses associated with upstreamness translate into higher or lower wages and, in turn, how this affects firms' profitability. For this purpose, we estimate the impact of upstreamness not only on productivity but also on wage costs and productivity-wage gaps at the firm level. Moreover, we test for possible nonlinearities, that is, the possibility that the upstreamness-productivity-wage nexus might be U-shaped, for instance.

In order to address the three sources of contribution to the literature in this paper, in a nutshell to examine the effects of firm-level upstreamness on productivity and wages, we take advantage of our access to detailed Belgian linked employer–employee panel data, which have been merged with a unique dataset derived from the NBB B2B transactions dataset. This dataset provides a direct and accurate measurement of each firm's upstreamness (i.e. the steps—weighted distance—before a firm's production meets either domestic or foreign final demand) for each year from 2002 to 2010.[4]

[4]For a detailed description of the NBB B2B transactions dataset, see Dhyne and Duprez (2015).

Our empirical strategy relies on the estimation of a value-added function and a wage cost equation per worker at the firm level. The value-added function yields estimates for the marginal product of firms' upstreamness, whereas the wage cost equation estimates the impact of upstreamness on the wage bill paid by firms. Given that both equations are estimated on the same samples with identical control variables, we are able to compare the estimates for productivity and wage costs and to draw conclusions on how upstreamness affects firms' profits.[5] This method was developed by Hellerstein et al. (1999) and refined by van Ours and Stoeldraijer (2011), among others.

We first estimate the value-added and wage cost equations using pooled OLS, with the inclusion of a large range of covariates. Next, in order to control for unobserved time-invariant firm characteristics and the potential endogeneity of firms' upstreamness, we first rely on the dynamic system generalised method of moments (GMM-SYS) estimator, proposed by Arellano and Bover (1995) and Blundell and Bond (1998). This method boils down to simultaneously estimating a system of two equations (one in levels and one in first differences) and using internal instruments to control for endogeneity. More precisely, a firm's level of upstreamness is instrumented by its lagged levels in the differenced equation and by its lagged differences in the level equation. As a robustness test, we also adopt a two-stage least squares within estimator (FE-IV). This method addresses endogeneity on the basis of external instruments and tackles firm-level fixed unobserved heterogeneity using a within differentiated model. Our main instrumental variable (IV) is the average share of a firm's sales in total clients' purchases,[6] that is a proxy for the price elasticity of the demand for the firm's product, which should be positively correlated with the firm's upstreamness, according to Alfaro et al. (2019). Moreover, we expect our IV to fulfil the exogeneity condition as it is an imprecise measurement of a firm's market power and of its capacity to generate rents (or value added), which in turn may benefit workers' wages through rent-sharing (Matano & Natchioni, 2017). Indeed, a firm's ability to create rents (or value added) is contingent on many factors beyond our IV, including the firm's number of suppliers and clients, the ease with which clients can switch to alternative suppliers, the degree of concentration among the firm's suppliers and the overall competition on the firm's main product market. The relationship between our IV and a firm's capacity to create rents (or value added) is thus not univocal. Moreover, empirical evidence suggests that the magnitude of rent-sharing in the Belgian economy, that is the elasticity between wages and firms' rents, is not big (Rusinek & Rycx, 2013; Rycx & Tojerow, 2004). The relevance and exogeneity of our instrumentation strategy are supported by an array of diagnosis tests. Furthermore, our FE-IV estimates are found to be largely in line with those obtained by GMM-SYS.

Our main finding is that firms positioned further upstream create more value, in line with the arguments and descriptive statistics put forward by the OECD (2012). Overall, our estimates show that when upstreamness increases by one step, that is, by approximately one standard deviation, productivity rises, on average, by around 5% and 20% in the short and the long run, respectively. Testing for potential nonlinearities, we find that this relationship is monotonically increasing and even slightly convex, with considerable productivity gains in the early stages of the production chain, when upstreamness is the highest. Regarding distributional issues, our results show that the productivity gains associated with upstreamness are shared almost equally between wages and profits. This outcome is consistent with Belgium's strong labour market institutions (Garnero et al., 2020; Kampelmann & Rycx, 2013).

[5]By definition, the gap between value added and wage costs corresponds to the gross operating surplus.

[6]This variable available for each firm at each period in our data set has been provided by the NBB.

The remainder of this paper is organised as follows. Sections 2 and 3 describe our methodology, estimation strategy and dataset. Section 4 presents our main econometric results and offers details about our robustness test. Section 5 finally concludes.

2 | METHODOLOGY

2.1 | Benchmark specification

Our empirical strategy relies on the estimation of a value-added function and a wage cost equation per worker at the firm level. The former yields parameter estimates for the impact of upstreamness on firms' productivity, whereas the latter estimates the influence of upstreamness on the wage bill paid by firms. Given that both equations are estimated on the same samples with identical control variables, we can compare the parameters for productivity and wage costs and draw conclusions on how upstreamness affects firms' productivity-wage gaps. In other words, the parameters enable us to highlight whether upstreamness is beneficial or harmful to firms' productivity, and whether and how the gains or losses associated with upstreamness are shared with workers (in terms of higher or lower wages). This technique was pioneered by Hellerstein et al. (1999) and refined by van Ours and Stoeldraijer (2011), among others. It is now standard practice in the literature on the productivity and wage effects of worker and firm heterogeneity (see, e.g. Cardoso et al., 2011; Devicienti et al., 2018; Garnero et al., 2014, 2020; Giuliano et al., 2017; Göbel & Zwick, 2012; Ilmakunnas & Maliranta, 2005; Mahlberg et al., 2013; Nielen & Schiersch, 2014).

The estimated firm-level productivity and wage cost equations are the following:

$$\ln va_{jt} = \beta_0 + \left(\beta_1 \ln va_{jt-1} + \right) \beta_2 up_{jt} + \mathbf{x_{jt}} \boldsymbol{\beta}_3 + \mathbf{z_{jt}} \boldsymbol{\beta}_4 + \left(\delta_j + \right) \partial_t + \nu_{jt} \tag{1}$$

$$\ln w_{jt} = \gamma_0 + \left(\gamma_1 \ln w_{jt-1} + \right) \gamma_2 up_{jt} + \mathbf{x_{jt}} \boldsymbol{\gamma}_3 + \mathbf{z_{jt}} \boldsymbol{\gamma}_4 + \left(\delta_j + \right) \partial_t + \omega_{jt} \tag{2}$$

The dependent variable in Equation (1) is the value added per capita of firm j, obtained by dividing the total value added (at factor costs) of firm j in period t by the total number of workers employed in firm j during the same period. The dependent variable in Equation (2) is the average wage bill of firm j (including payroll taxes and variable pay components, such as wage premia for overtime, weekend or night work, performance bonuses and other premia). It is obtained by dividing the total wage costs of firm j in period t by the total number of workers employed in firm j during the same period. In summary, the dependent variables in the estimated Equations (1) and (2) are the firm-level value added and wage cost per worker, respectively (expressed in logarithms).

The main variable of interest, up_{jt}, is the firm's level of upstreamness. It is taken from the NBB B2B transactions dataset, which provides information on the fragmentation of production chains and on the relative position of firms operating in the Belgian economy in those chains over the period 2002–2012. The level of upstreamness of firm j is measured as the steps (weighted distance) before its production at period t meets either domestic or foreign final demand.

Equations (1) and (2) also include the vector $\mathbf{x_{jt}}$, which contains a set of variables controlling for observable worker characteristics. More precisely, it includes the share of a firm's workforce that: (i) is younger than 30 and older than 49 years, respectively; (ii) has a Bachelor, a Master and a post-Master degree, respectively; (iii) is blue-collar; and (iv) works part-time. The vector $\mathbf{z_{jt}}$, also included in Equations (1) and (2), controls for firm characteristics. It includes the firm's sectoral affiliation (eight dummies), size (number of workers, in logarithm) and capital stock per worker (in logarithm), which has been estimated through the 'perpetual inventory method'.[7,8] δ_j is a dummy variable that captures, for each firm, unobserved time-invariant workplace characteristics, ∂_t is a set of year dummies (eight dummies), and ν_{jt} (ω_{jt}) is the error term.

2.2 | Estimation techniques

We first estimated Equations (1) and (2) by pooled ordinary least squares (OLS). The OLS estimator is based on the cross-section variability between firms and the longitudinal variability within firms over time. However, this estimator suffers from a potential heterogeneity bias because firm productivity and wages can be related to firm-specific time-invariant characteristics (e.g. the quality of management, an advantageous location, the ownership of a patent or other productivity drivers and firm idiosyncrasies) that are not reported in our dataset. The conventional way to remove unobserved firm characteristics that remained unchanged during the observation period is by estimating a fixed effects (FE) model. This boils down to estimating a within differentiated model, that is a model where the mean of each variable has been subtracted from the initial values. Yet, neither pooled OLS nor the FE estimator address the potential simultaneity between a firm's level of upstreamness and its productivity/wage costs. Reverse causality is likely to be an issue due to: (i) the correlation between upstreamness and the export behaviour of firms, as the number of steps before firms' production meets final demand is typically greater among exporting firms (OECD, 2012); and (ii) ample evidence supporting reverse causality between the export behaviour of firms and their productivity/wages (Arnold & Hussinger, 2005; Berthou & Vicard, 2013; Eaton et al., 2007; Freund & Pierola, 2010).[9] We performed a direct endogeneity test on the upstreamness variable in our sample and indeed reject the null hypothesis that our main variable of interest can actually be treated as exogenous.[10] To control for this endogeneity issue, in addition to the state dependence of firms' productivity/wages[11] and the presence of firm-level fixed effects, we re-estimated

[7]The 'perpetual inventory method' (PIM) incorporates the idea that capital stock results from investment flows and corrects for capital depreciation and efficiency losses. Following standard practice, we assume a 5% annual depreciation rate. See, e.g., OECD (2009) for more details.

[8]Unfortunately, our data provide no information on the export or import status of firms. Hence, future research could notably investigate whether the relative position of firms on international supply chains matters more or less than their relative position on domestic chains.

[9]As highlighted in the introduction, the traditional explanation for this phenomenon, in line with international trade models with heterogeneous firms (Melitz & Ottaviano, 2003), is related to a self-selection mechanism whereby more productive firms (also paying higher wages) are more likely to export.

[10]We performed such a test by using a two-stage least squares (2SLS) estimator on an equation in levels in which our variable of interest is instrumented by first differences. Both equations (i.e. value added and wage costs) pass standard underidentification and weak identification tests. This suggests that the endogeneity test for the upstreamness variable is valid. This test indicates that, for both equations, we have to reject the null hypothesis that upstreamness can actually be treated as exogenous. Further evidence on the endogeneity of the upstreamness variable is provided in Section 4.3.

[11]The assumption of persistent productivity, both at the industry and at the firm level, finds some support in the literature (see e.g. Bartelsman & Doms, 2000). Researchers 'documented, virtually without exception, enormous and persistent measured productivity differences across producers, even within narrowly defined industries' (Syverson, 2011: 326). Large parts of these productivity differences are still hard to explain. The persistence of wage costs is also highlighted in the literature (see e.g. Heckel et al., 2008). Wage stickiness is notably the outcome of labour market institutions, adjustment costs and efficiency wages' motives.

Equations (1) and (2) using the dynamic system generalised method of moments (GMM-SYS) proposed by Arellano and Bover (1995) and Blundell and Bond (1998). This approach is standard in the literature on firm-level determinants of productivity and wages (e.g. Buhai et al., 2017; Göbel & Zwick, 2012; van Ours & Stoeldraijer, 2011). It boils down to simultaneously estimating a system of two equations (one in levels and one in first differences) and relying on internal instruments to control for endogeneity. More precisely, a firm's level of upstreamness is instrumented by its lagged levels in the differenced equation and by its lagged differences in the level equation.[12] The implicit assumption is that differences in a firm's productivity and wages in one period, although possibly correlated with contemporaneous differences in the firm's upstreamness, are uncorrelated with its lagged levels of upstreamness. In the same line, levels of productivity and wages in one period, although possibly correlated with contemporaneous levels of the firm's upstreamness, are assumed to be uncorrelated with its lagged differences in upstreamness. Moreover, differences in the firm's upstreamness variable are assumed to be reasonably correlated with their past levels, and levels of the firm's upstreamness variable are assumed to be reasonably correlated with their past differences.

One advantage of *system* GMM is that time-invariant explanatory variables can be included among the regressors, whereas these variables typically disappear in difference GMM. Asymptotically, the inclusion of these variables does not affect the estimates of the other regressors because the instruments in the level equation (i.e. lagged differences of upstreamness) are expected to be orthogonal to all time-invariant variables (Roodman, 2009). To examine the validity of our estimates, we applied the tests proposed by Arellano and Bond (1991) and Hansen (1982). The first is a test for overidentification that allows to test the validity of the instruments. The second is a test for autocorrelation, where the null hypothesis assumes no second-order autocorrelation in the first-differenced errors. The non-rejection of the two tests is required in order to assume that our estimates are reliable.

The adoption of a *dynamic* GMM specification aims to account for the persistency in firm-level wage costs and productivity. It is also likely to improve the identification of the parameters of interest (even though the coefficients on the lagged dependent variables are not a central issue of the analysis). Indeed, as illustrated by Bond (2002), the use of a dynamic model is necessary to obtain consistent results when estimating a production function with serially correlated productivity shocks and explanatory variables that are correlated to these shocks. Although serial correlation may arise if, for instance, 'the effects from demand shocks are only partially captured by the industry-specific control variables' (Hempell, 2005), the responsiveness of input factors to productivity shocks may be explained by the above-mentioned endogeneity issue. Interestingly, the inclusion of the lagged dependent variable in the OLS, FE and GMM-SYS specifications also provides an ad hoc test for the appropriateness of the GMM-SYS method. As outlined by Roodman (2009), this test consists in checking whether the regression coefficients on the lagged dependent variables obtained with GMM-SYS fall between the OLS and FE estimates.

The GMM-SYS method has been widely used in the literature to examine the consequences of worker and firm characteristics on firm/industry-level productivity and wages (e.g. Cardoso et al., 2011; Garnero et al., 2014; Nielen & Schiersch, 2014; Szymczak et al., 2019). However, as it relies on internal instruments (i.e. lagged levels and differences of explanatory variables) to

[12]Bond and Söderbom (2005) provided a review of the literature on the identification of production functions. The authors notably highlight that the adjustment costs of labour and capital can justify the use of lagged values (of the endogenous variable) as instruments.

control for endogeneity, its accurateness could be discussed. Accordingly, as a sensitivity test, we adopted an alternative estimation method: a two-stage least squares within estimator (FE-IV). This method addresses endogeneity on the basis of external instruments and tackles firm-level fixed unobserved heterogeneity using a within differentiated model. Our main instrumental variable is the average share of a firm's sales in total clients' purchases, a proxy for the price elasticity of the demand for the firm's product. The relevance and exogeneity of our instrumentation strategy, discussed in Section 4.3, is supported by an array of diagnosis tests. Furthermore, the FE-IV estimates are found to be largely in line with those obtained by GMM-SYS.

3 | DATASET

Our empirical analysis is based on a combination of three large datasets. The first is a new and unique dataset, derived from the NBB B2B transactions dataset (hereafter NBB B2B), covering the whole Belgian private sector over the period 2002–2012.[13] It provides direct, yearly information on the upstreamness of (almost)[14] each commercial firm in Belgium, that is on the number of steps (weighted distance) before the firm's production meets either domestic or foreign final demand. More precisely, Dhyne and Duprez (2015) first built a firm-level input-output table for each year, using the values of transactions between enterprises. They then applied the methodology developed by Antràs et al. (2012), which models the upstreamness of a firm's production as the number of transactions and/or transformations (made by firms in Belgium or abroad, before importation or after exportation) that are needed, on average, for all the production of that firm to meet the final demand. The upstreamness of a firm is computed as a sum of terms: (i) the first representing the firm's share of output that reaches final demand directly, with no additional transformation; (ii) the second its share of output that reaches final demand after only one additional transformation by another firm, multiplied by 2 (i.e. the number of transactions needed to meet final demand); (iii) the third its share of output that reaches final demand after only two transformations by other firms, multiplied by 3; and so on (see Dhyne and Duprez (2015) for more details).

Our second source of data is the Structure of Earnings Survey (SES), carried out by Statistics Belgium. It is representative of firms that were operating in Belgium between 1999 and 2010, employed at least ten workers and had economic activities within sections B to N of the NACE Rev. 2[15] nomenclature. The survey contains a wealth of information, provided by the HR departments of firms, both on the characteristics of firms (e.g. sector of activity, number of

[13]We did not have access to the entire NBB B2B data set but only to a limited number of variables from this data set, including the firm-level upstreamness and the average share of the firm's sales in total clients' purchases.

[14]See footnote 3.

[15]NACE stands for 'Nomenclature statistique des Activités économiques dans la Communauté Européenne' (Statistical classification of economic activities in the European Community), and 'Rev. 2' refers to its version.

workers) and on the individuals working there (e.g. age, education, occupation, working hours).[16]

The SES, however, provides no financial information. It has therefore been merged with a firm-level survey, called the Structure of Business Survey (SBS). Also conducted by Statistics Belgium, the SBS provides information on financial variables such as firm-level material inputs, investments, value added, wage costs and gross operating surplus per worker. In terms of coverage, the SBS differs from the SES: it does not cover the whole financial sector (NACE K), but only other financial intermediation and activities auxiliary to financial intermediation. The merger of our three datasets, that is the SES, SBS and NBB B2B datasets, was carried out by Statistics Belgium, in collaboration with the National Bank of Belgium, using firms' VAT codes.[17]

The system generalised method of moments estimator (GMM-SYS) requires firms' information for (at least) two consecutive years. Given that the sampling percentages of firms in our dataset increase with firms' size (see footnote 16), medium-sized and large firms are over-represented in our econometric investigations. Note that workers and firms for which data are missing or inaccurate have been excluded.[18] We also dropped firms with fewer than 10 observations because the use of average values of worker characteristics at the firm level requires a suitable number of observations.[19] Our final sample covering the period 2002–2010 consists of an unbalanced panel of 11,975 firm-year-observations from 3529 firms. It is representative of medium-sized and large firms in the Belgian private sector, with the exception of large parts of the financial sector (NACE K) and the electricity, gas and water supply industry

[16]The SES is a stratified sample. The stratification criteria refer to the region (NUTS-groups), the main economic activity (NACE-groups) and the size of the firm. The sample size in each stratum depends on the size of the firm. The sampling percentages of firms are equal to 10%, 50% and 100% when the number of workers is between 10 and 50, between 50 and 99, and above 100, respectively. Within a firm, the sampling percentages of employees also depend on the size of the firm. The sampling percentages of employees are equal to 100%, 50%, 25%, 14.3% and 10% when the number of workers is between 10 and 20, between 20 and 50, between 50 and 99, between 100 and 199, and between 200 and 299, respectively. Firms employing at least 300 workers are required to report information for an absolute number of employees, which ranges from 30 (for firms with 300–349 workers) to 200 (for firms with 12,000 workers or more). To guarantee that firms report information on a representative sample of their workers, they are asked to follow a specific procedure. First, they have to rank their employees in alphabetical order. Next, Statistics Belgium gives them a random letter (e.g. the letter O) from which they have to start when reporting information on their employees (following the alphabetical order of the workers' names in their list). If they reach the letter Z and still have to provide information on some of their employees, then they have to continue from the letter A in their list. Moreover, firms that employ different categories of workers, namely managers, blue- and/or white-collar workers, have to set up a separate lists in alphabetical order for each of these categories and to report information on a number of workers in these categories that is proportional to their share in total firm employment. For example, a firm with 300 employees (namely 60 managers, 180 white-collar workers and 60 blue-collar workers) will have to report information on 30 workers (namely six managers, 18 white-collar workers and six blue-collar workers). For an extended discussion, see Demunter (2000).

[17]We had access to a fully anonymized version of the merged dataset, which prevents us from directly identifying an individual firm.

[18]For instance, we eliminated a (very small) number of firms for which the recorded value added was negative.

[19]This restriction is also unlikely to affect our results as it leads to a very small drop in sample size. The average number of observations per firm in each year stands at around 35 in our final sample.

TABLE 1 Firm-level descriptive statistics of selected variables, 2002–2010

Variables	Mean	Std. dev.
Annual value added per worker (€[a])	91,809	624,962
Annual value added per worker (ln)	11.10	0.61
Annual wage cost per worker (€[a])	47,817	23,135
Annual wage cost per worker (ln)	10.69	0.43
Upstreamness (in steps)	2.34	0.93
Age (%)		
Less than 30 years	21.6	14.88
Between 30 and 49 years	60.92	14.42
50 years and more	17.50	13.02
Education (%)		
Non-tertiary education	73,84	26.34
Bachelors	15.23	16.84
Masters	10.36	15.59
Post-masters	0.56	2.99
Blue-collar workers[b] (%)	53.85	34.19
Part-time workers (%)	16.40	17.46
Sector (%)		
Mining and quarrying (B)	0.69	
Manufacturing (C)	48.74	
Electricity, gas, steam and air conditioning supply; water supply, sewerage, waste management and remediation activities (D + E)	0.59	
Construction (F)	8.27	
Wholesale and retail trade, repair of motor vehicles and motorcycles (G)	17.13	
Accommodation and food service activities (I)	1.87	
Transport and storage; information and communication (H + J)	5.84	
Financial and insurance activities (K)	2.61	
Real estate activities; professional, scientific and technical activities; administrative and support service activities (L + M + N)	14.26	
Size (number of workers)	229.69	431.64
Size (number of workers) (ln)	4.64	1.36
Capital stock per worker (€[a])	239,830	1,912,046
Capital stock per worker (ln)	10.78	1.60
Number of observations	11,975	
Number of firms	3529	

[a]At 2004 constant prices.

[b]The distinction between blue- and white-collar workers is based on the International Standard Classification of Occupations (ISCO-08). Workers belonging to groups 1–5 are considered to be white-collar workers (1: Legislators, senior officials and managers; 2: Professionals; 3: Technicians and associate professionals; 4: Clerks; 5: Service workers and shop and market sales workers) and those from groups 7–9 are considered to be blue-collar workers (7: Craft and related trades workers; 8: Plant and machine operators and assemblers; 9: Elementary occupations).

(NACE D + E) and an over-representation of firms operating in the manufacturing industry (NACE C).[20]

Table 1 presents the means and standard deviations of selected variables. It indicates that firms' mean annual value added per worker in logarithmic form (ln) stands at 11.10 (which corresponds to 91,809 EUR), whereas their mean annual wage cost per worker in ln reaches 10.69 (47,817 EUR). Regarding the upstreamness variable, we find that the mean number of steps before a firm's production meets either domestic or foreign final demand is equal to 2.34. We also observe that around 61% of workers within firms are prime-age (i.e. between 30 and 49 years old), 26% have a tertiary education degree (15% possess a Bachelor's degree, 10% a Master's degree and 1% a post-Master's degree), 54% are blue collars, and 16% are part-timers (i.e. have <30 h of paid work per week). Furthermore, we see that the mean of firm size in ln stands at 4.64 (i.e. the firms in our sample employ, on average, 230 workers) and that the mean capital stock per worker in ln is equal to 10.78 (i.e. their capital stock per worker amounts to approximately 240,000 EUR, on average). Firms are concentrated mainly in NACE sectors C (manufacturing—49%), G (wholesale and retail trade, repair of motor vehicles and motorcycles—17%) and L-M-N (real estate activities; professional, scientific and technical activities; administrative and support service activities—14%).

4 | RESULTS

4.1 | Benchmark estimates

We first estimated Equations (1) and (2) by pooled OLS, without any covariate. The results, presented in the first two columns of Table 2, point towards the existence of positive and significant relationships between upstreamness and firm productivity (coefficient = 0.104), on the one hand, and between upstreamness and wage costs (coefficient = 0.045), on the other.

After controlling for time fixed effects, worker and firm characteristics (see Columns (3) and (4)), we observe that the regression coefficients associated with the upstreamness variable somewhat decrease (to 0.075 and 0.026, respectively) but remain positive and significant. Table A2 in Appendix shows the detailed OLS estimates. We find that almost all covariates are significant with the expected sign, in both the productivity and wage regressions. Indeed, we observe that education has a positive and somewhat stronger effect on productivity than on wage costs. This finding is compatible with the 'wage compression effect' highlighted by Kampelmann et al. (2018). Not surprisingly, we also find a strong negative effect of the share of part-time workers on productivity and wages per capita. In line with the literature on inter-industry wage differentials (du Caju et al., 2011), we further observe that productivity and wages are the highest in sectors D and E (i.e. electricity, gas, steam and air conditioning supply; water supply, sewerage, waste management and remediation activities) and the lowest in sector I (accommodation and food service activities). The results also show that firm size and capital stock have a positive and significant

[20]Appendix Table A1 shows the consequences of the data cleaning process on descriptive statistics of selected variables. From this table, it can be seen that the reduction of the sample size due to the constraints of non-missing data and multiple observations over time leads to slight changes in the composition of the sample. More precisely, our final sample gathers firms that are (i) more productive and with higher labor costs, (ii) bigger, (iii) a bit more upstream and (iv) more concentrated in the manufacturing industry. In other words, our final sample therefore allows us to investigate the relation between upstreamness, productivity and labour costs among medium-sized and bigger firms operating in a large part of the private sector, within which manufacturing is somewhat over-represented.

TABLE 2 Upstreamness, productivity and wage costs firm-level OLS estimates, 2002–2010

	OLS		OLS	
	Productivity	Wage costs	Productivity	Wage costs
Dependent variable	(1)	(2)	(3)	(4)
Value added per worker (one year lagged, in ln)				
Wage costs per worker (one year lagged, in ln)				
Upstreamness[a]	0.104***	0.045***	0.075***	0.026***
	(0.006)	(0.004)	(0.005)	(0.003)
Worker characteristics[b]	No	No	Yes	Yes
Firm characteristics[c]	No	No	Yes	Yes
Year dummies (8)	No	No	Yes	Yes
Adjusted R^2	.025	.009	.473	.528
Sig. model (p-value)	.000	.000	.000	.000
Number of observations	11,975	11,975	11,975	11,975
Number of firms	3529	3529	3529	3529

Notes: ***/**/*Significant at the 1%, 5% and 10% level, respectively. Robust standard errors are shown in brackets. The dependent variable is either the value added per worker in ln ('productivity') or the wage cost per worker in ln ('wage cost') at the firm level.

[a]Steps (distance) before the production of a firm meets either domestic or foreign final demand.

[b]Share of the workforce that: (i) is younger than 30 and older than 49 years, respectively, and (ii) is highly educated (three categories). The share of blue-collar workers and the share of part-time workers are also included.

[c]Sectoral affiliation (eight dummies), number of workers and capital stock.

impact on productivity and wages, which is consistent with the findings of Lallemand et al. (2006, 2007), for instance.[21,22]

However, as argued in Section 2, OLS estimates should be considered with caution due to the potential biases associated with firm-level fixed effects and endogeneity. To account for these issues, we re-estimated Equations (1) and (2) with the dynamic GMM-SYS estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998). The variables in the differenced equation have thus been instrumented by their lagged levels, and the variables in the level equation have been instrumented by their lagged differences. The time dummies have been considered as exogenous, and we use the first and second lags of other explanatory variables as instruments. The results are presented in Columns (1) and (2) of Table 3.[23] To verify their reliability, we applied the test of overidentifying restrictions proposed by Hansen (1982) and the test for second-order autocorrelation in the first-differenced errors proposed by Arellano and Bond (1991). As shown in Table 3, the two tests do not reject the null hypothesis of valid instruments and that of no autocorrelation, respectively. As expected, we also find that current productivity and wage costs are, to a significant extent, related to their past values. The coefficients associated with upstreamness in the productivity and wage cost regressions remain highly significant and are now equal to 0.045 and 0.040, respectively. These coefficients are statistically different from each other, as shown by a standard t-test ($t = 9.96$). Accordingly, our findings suggest that when a firm's upstreamness increases by one unit (i.e. when a firm's position in the value chain moves one step away from final demand), the firm's productivity and wages increase on average by 4.5% and 4%, respectively.[24]

---

[21]To examine the robustness of these results, we ran a series of additional regressions. First, we examined how the inclusion of industry dummies at the NACE 2-digit level influences our estimates in terms of the impact of upstreamness on productivity and wage costs. Second, we re-estimated our benchmark equations adding R&D expenditures at NACE 2-digit level and additional characteristics of the production network among the covariates, namely: (i) the average number of suppliers that the firm's clients have; (ii) the average number of clients that the firm's suppliers have; (iii) a concentration index of the firm's domestic suppliers; and (iv) a concentration index of the firm's domestic clients. Next, we investigated the sensitivity of the upstreamness regression coefficients to the inclusion of the above-mentioned covariates (i.e. R&D and characteristics of the production network) in addition to the NACE 2-digit industry dummies. Finally, in order to test the value added of our firm-level measurement of upstreamness compared to an industry-level one, we computed the mean value of upstreamness at the sectoral level (NACE 2-digits) and re-estimated our benchmark equations including both indicators of upstreamness, i.e. those computed at the firm- and industry-level, together in the productivity and wage cost regressions respectively. The OLS estimates of these additional regressions are reported in Table A3 in Appendix. In all specifications, we find that the regression coefficients associated with firm-level upstreamness remain significantly positive.

[22]A point raised by a referee, for which we are thankful, is that firms possessing the best technology could survive while those racing at the initial stage have to be kicked out, leading to a selection bias when we measure firm performance, especially for firms with high upstreamness indices as they are the few survivors with the best technology. We examined this potential selection bias when analysing the role played by firm survival, by estimating the relationships between upstreamness, productivity and wage costs on subsamples composed of ‘survivor’ vs. ‘no survivor’ firms. The underlying idea is to explore whether the impacts of upstreamness on both productivity and wages are greater in survivor firms. We address survival through the size (i.e. a smaller firm is expected to have a lower probability of surviving) and through the recurrence of a firm in the database (i.e. firms that are observed over more periods are expected to have survived longer). The results, available on request, do not necessarily support the hypothesis of systematic larger impacts for larger or more frequently observed firms, suggesting that productivity, wage costs and upstreamness nexus is not a stylized fact of survival firms, and in fine that our results are not driven only by a selection effect.

[23]Note that the GMM coefficients on the lagged dependent variables systematically fall between the OLS and FE estimates (available on request). As outlined by Roodman (2009), this result supports the appropriateness of our dynamic GMM-SYS specification.

[24]These estimates are computed at the mean sample value of firms’ upstreamness, which is equal to 2.34 steps. Moving one step away from this mean approximately corresponds to a one-standard-deviation change in the upstreamness variable (see Table 1).

TABLE 3 Upstreamness, productivity and wage costs firm-level GMM-SYS estimates, 2002–2010

	GMM-SYS			
	Productivity	Wage costs	Productivity	Wage costs
Dependent variable	(1)	(2)	(3)	(4)
Value added per worker (one year lagged, in ln)	0.756***		0.758***	
	(0.033)		(0.032)	
Wage costs per worker (one year lagged, in ln)		0.770***		0.775***
		(0.052)		(0.048)
Upstreamness[a]	0.045*	0.040**		
	(0.023)	(0.019)		
Upstreamness between 2.5 and 4.5[b]			0.067*	0.064**
			(0.040)	(0.029)
Upstreamness above 4.5[b]			0.214**	0.080*
			(0.094)	(0.050)
Worker characteristics[c]	Yes	Yes	Yes	Yes
Firm characteristics[d]	Yes	Yes	Yes	Yes
Year dummies (8)	Yes	Yes	Yes	Yes
Sig. model (*p*-value)	.000	.000	.000	.000
Hansen statistic	293.22	329.38	313.71	334.90
p-value	.63	.13	.73	.11
Arellano-Bond statistic (AR2)[e]	1.60	1.21	1.60	1.19
p-value	.11	.23	.11	.23
Number of observations	11,975	11,975	11,975	11,975
Number of firms	3529	3529	3529	3529

Notes: ***/**/*Significant at the 1%, 5% and 10% levels, respectively. Robust standard errors are shown in brackets. The dependent variable is either the value added per worker in ln ('productivity') or the wage cost per worker in ln ('wage cost') at the firm level. Second and third lags of all explanatory variables are used as instruments in the GMM-SYS specification, excluding time dummies, and the lag of value-added per worker in the last column.

[a]Steps (distance) before the production of a firm meets either domestic or foreign final demand.

[b]The control group is composed of firms whose upstreamness is below 2.5.

[c]Share of the workforce that: (i) is younger than 30 and older than 49 years, respectively, and (ii) is highly educated (three categories). The share of blue-collar workers and the share of part-time workers are also included.

[d]Sectoral affiliation (eight dummies), number of workers and capital stock.

[e]AR2 displays the test for second-order autocorrelation in the first-differenced errors.

What about productivity-wage gaps? Given that the mean sample values of productivity and wage costs reach 91,809 and 47,817 EUR, respectively, the GMM-SYS estimates suggest that moving up the value chain by one step increases a firm's annual productivity per worker by 4131 EUR (i.e. 0.045*91,809 EUR) and its annual wage cost per worker by 1913 EUR (0.040*47,817 EUR), on average. Put differently, we find that: (i) profitability (i.e. the productivity-wage gap) positively depends on a firm's upstreamness (i.e. profitability increases by around 5% when upstreamness increases by one step); and (ii) the productivity gains associated with upstreamness are shared almost equally between wages and profits.

4.2 | Nonlinear relationships?

In order to test for potential nonlinear relationships between upstreamness, productivity and wage costs, we estimated the following variants of Equations (1) and (2):

$$\ln va_{jt} = \beta'_0 + \left(\beta'_1 \ln va_{jt-1} +\right) \beta_{2.1} upD1_{jt} + \beta_{2.2} upD2_{jt} + \mathbf{x_{jt}\beta'_3} + \mathbf{z_{jt}\beta'_4} + \left(\delta_j +\right) \partial_t + v'_{jt} \quad (1a)$$

$$\ln w_{jt} = \gamma'_0 + \left(\gamma'_1 \ln w_{jt-1} +\right) \gamma_{2.1} upD1_{jt} + \gamma_{2.2} upD2_{jt} + \mathbf{x_{jt}\gamma'_3} + \mathbf{z_{jt}\gamma'_4} + \left(\delta_j +\right) \partial_t + \omega'_{jt} \quad (2a)$$

where $upD1_{jt}$ ($upD2_{jt}$) is a dummy variable that is equal to 1 if the level of upstreamness of firm j at time t ranges between 2.5 and 4.5 (is >4.5) steps, and equal to 0 otherwise. The reference category is thus composed of firms with a level of upstreamness below 2.5 steps. The mean value of upstreamness among firms belonging to the reference category is equal to 1.6, it stands at 3.2 among firms with an upstreamness comprised between 2.5 and 4.5, and it is equal to 4.8 among firms with an upstreamness >4.5.

The GMM-SYS estimates of Equations (1a) and (2a) are reported in Columns (3) and (4) of Table 3. Our findings support the existence of a monotonically (and slightly convex) upward-sloping profile between upstreamness and productivity. Indeed, the productivity of firms with a level of upstreamness ranging from 2.5 to 4.5 is found to be 6.7% higher, on average, than that of the reference category, and firms with a level of upstreamness above 4.5 are found to be 21.4% more productive than the reference category, on average. As for wage costs, the relationship is still found to be significantly positive, though less steep. Compared to the reference category, wage costs appear to be, on average, 6.4% higher among firms with an upstreamness index between 2.5 and 4.5, and 8% higher among firms with an upstreamness index above 4.5. Overall, this implies that the relationship between firms' upstreamness and profitability (i.e. the productivity-wage gap) is significantly positive and convex. Alternative specifications, including not only different thresholds for $upD1_{jt}$ and $upD2_{jt}$ but also polynomials of the upstreamness variable, have been tested. Overall, they confirm that productivity, wage costs and profitability rise steadily as firms move further up the value chain.

4.3 | Robustness test

The GMM-SYS estimates presented so far seem quite accurate. However, their appropriateness might be discussed because they rely on internal instruments (i.e. lagged levels and differences of explanatory variables) to control for the endogeneity of upstreamness. Accordingly, in this subsection, we test the robustness of our findings using an alternative estimation method based on external instruments. More precisely, to address endogeneity alongside firm-level fixed unobserved heterogeneity, we re-estimated our benchmark Equations (1) and (2) using a two-stage least squares within estimator (FE-IV). This approach consists in: (i) estimating a within differentiated model, that is a model where the mean of each variable has been subtracted from the initial values (FE); and (ii) finding instrumental variables (IV) that are both highly correlated with the endogenous variable (i.e. upstreamness) and uncorrelated with the dependent variable (i.e. productivity and wage costs, for Equations (1) and (2), respectively).

Our main instrumental variable (IV) is the average share of the firm's sales in total clients' purchases. This IV is used as a proxy for the price elasticity of the demand for the firm's product.

Indeed, the larger the average share of the firm's sales in total clients' purchases, the smaller this price elasticity is expected to be. The intuition for using this IV strategy is provided by Alfaro et al. (2019). The authors extended the property-rights model of the organisation of production proposed by Antràs and Chor (2013) to show, both theoretically and empirically, that a firm's decision to integrate upstream or downstream suppliers crucially depends on the elasticity of the demand for its goods. More specifically, they found that a firm is more likely to integrate relatively upstream stages of the value chain, while engaging in outsourcing to downstream suppliers, when the price elasticity of the demand for its product is more inelastic. Accordingly, we expect our main IV to have a positive effect on the value of upstreamness in the first-stage regression. Moreover, to account for potential nonlinearities, we also included the inverse value of this IV in the first step of our estimation procedure.

The estimates of our two-stage least squares within (FE-IV) regressions are reported in Columns (1) and (2) of Table 4.[25] They show that the coefficient associated with upstreamness remains significant and is now equal to 0.266 and 0.280 in the productivity and the wage cost regressions, respectively. Our findings thus suggest that when a firm's upstreamness increases by one unit, its productivity and wage costs increase, on average, by 26.6% and 28%, respectively.[26] As for the productivity-wage gaps, given that the sample mean annual values of productivity and wage costs per worker reach 91,809 EUR and 47,817 EUR, respectively, the FE-IV estimates suggest that moving up the value chain by one step increases firms' annual productivity per worker by 24,421 EUR (i.e. 0.266*91,809 EUR) and their annual wage cost per worker by 13,389 EUR (i.e. 0.280*47,817 EUR), on average. These estimates thus corroborate our conclusion that: (i) profitability (i.e. the productivity-wage gap) positively depends on a firm's upstreamness; and (ii) the productivity gains associated with upstreamness are shared almost equally between wages and profits.

One may wonder how the magnitude of the FE-IV estimates can be reconciled with those obtained by GMM-SYS. Following standard practice and in order to account for serial correlation, the GMM-SYS regression coefficients were obtained through the estimation of *dynamic* productivity and wage cost equations (i.e. specifications including the lagged dependent variable among covariates). Accordingly, the GMM-SYS estimates correspond to short-run effects. To compute the long-run impact of upstreamness on productivity and wage costs, Koyck (1954) has shown that GMM-SYS estimates (such as those reported in Table 3) should be divided by 1 minus the coefficient associated with the lagged dependent variable. Applying this transformation to our GMM-SYS results, we find that semi-elasticities associated with upstreamness are more than four times bigger in the long run than in the short run, increasing from around 4% to circa 18%.[27] The long-run effects of upstreamness on productivity and wage costs, as estimated by GMM-SYS, are thus not very different in magnitude from those obtained by FE-IV.[28]

[25]To examine the existence of firm-level fixed unobserved heterogeneity, we applied a Breusch-Pagan LM test. This test clearly supports the existence of firm-level fixed effects.

[26]Note that moving up the value chain by one step is quite substantial, as it corresponds to a one-standard-deviation change in the upstreamness variable (see Table 1).

[27]To obtain the long-run effect of upstreamness on productivity, we divided the GMM-SYS coefficient of upstreamness (0.045) by 1 minus the coefficient of the lagged dependent variable (0.756), yielding a result of 0.184 (i.e. 0.045/(1–0.756)). Applying the same calculus to the GMM-SYS estimates of the wage equation (see Table 3), we find that the long-run effect of upstreamness on wage costs is equal to 0.174 (i.e. 0.040/(1–0.770)).

[28]Given that the FE-IV estimator has been applied to static productivity and wage cost equations, the FE-IV regression coefficients directly estimate the long-run effects of upstreamness on productivity and wage costs.

TABLE 4 Upstreamness, productivity and wage costs firm-level FE-IV estimates, 2002–2010

	FE-IV	
	Productivity	Wage costs
Dependent variable	(1)	(2)
Upstreamness[a]	0.266*	0.280*
	(0.141)	(0.159)
Worker characteristics[b]	Yes	Yes
Firm characteristics[c]	Yes	Yes
Year dummies (8)	Yes	Yes
Model significance		
p-value of *F* test	.000	.000
Underidentification test[d]		
p-value Kleibergen-Paap rk LM statistic	.000	.000
Weak identification test[e]		
Kleibergen-Paap rk Wald *F* statistic	24.42	24.42
Overidentification test[f]		
p-value of Sargan-Hansen *J* statistic	.832	.309
Endogeneity test[g]		
p-value associated with Chi-squared statistic	.081	.098
First-stage regression estimates (Dependent variable: upstreamness[a])		
Average share of firm' sales in total clients' purchases	1.662***	1.663***
	(0.257)	(0.257)
Inverse of average share of firm sales in total clients' purchases	−0.000**	−0.000**
	(0.000)	(0.000)
Worker characteristics[b]	Yes	Yes
Firm characteristics[c]	Yes	Yes
Year dummies (8)	Yes	Yes
Model significance		
p-value of *F* test	.000	.000
Number of observations	29,507	29,507
Number of firms	7070	7070

Notes: ***/**/*Significant at the 1%, 5% and 10% level. Robust standard errors are shown in brackets. The dependent variable is either the value added per worker in ln ('productivity') or the wage cost per worker in ln ('wage cost') at the firm level.

[a]Steps (distance) before the production of a firm meets either domestic or foreign final demand.

[b]Share of the workforce that: (i) is younger than 30 and older than 49 years, respectively, and (ii) is highly educated (three categories). The share of blue-collar workers and the share of part-time workers are also included.

[c]Sectoral affiliation (eight dummies), number of workers.

[d]The Kleibergen-Paap rk LM statistic for under-identification tests whether the equation is identified, that is whether the excluded instruments are all relevant. The null hypothesis in this test is that the equation is underidentified.

[e]Kleibergen-Paap rk statistic for weak identification is a Wald *F* statistic testing whether the excluded instruments are sufficiently correlated with the endogenous regressor. The null hypothesis is that the instruments are weak. According to the standard 'rule of thumb', weak identification is problematic for *F* statistics smaller than 10 (as suggested by van Ours and Stoeldraijer (2011)).

[f]The Sargan-Hansen *J* statistic tests the null hypothesis that the instruments are valid, that is uncorrelated with the error term.

[g]The Durbin–Wu–Hausman endogeneity test is based on the difference of two Sargan-Hansen statistics: one for the equation in which firm-level upstreamness is treated as endogenous, and one in which it is treated as exogenous. If the null hypothesis of this test cannot be rejected, then instrumentation is actually not necessary, that is upstreamness can actually be considered as exogenous.

To assess the soundness of the FE-IV approach, we also performed an array of diagnosis tests. The results of these tests are reported at the bottom of Table 4. The first-stage estimates indicate that both instrumental variables (IV) are significant with the expected sign. More precisely, they show that the average share of a firm's sales in total clients' purchases has a positive, though slightly concave, effect on upstreamness. The first-stage estimates thus suggest that our IVs are not weak, which is also corroborated by the Kleibergen-Paap rk Wald *F* statistic for weak identification. This *F* statistic is indeed much higher than 10.[29] Moreover, we can reject the null hypothesis that our first-stage equation is underidentified as the Kleibergen-Paap rk *LM* statistic is found to be highly significant. Concerning the quality of our instruments, we further find that the *p*-values associated with the Sargan-Hansen's *J* overidentification test are equal to 0.832 and 0.309 in the productivity and wage cost regressions, respectively (see Columns (1) and (2) of Table 4), which suggests that our instruments are valid. For information, we also computed bivariate correlations between our IVs, productivity and wage costs. Our findings show that all correlation coefficients are very small, fluctuating between 0.005 and 0.068. In line with the Sargan-Hansen's *J* overidentification test, they thus support the assumption that our IVs are fairly exogenous with respect to both productivity and wages. This outcome is not surprising. Indeed, our main IV is a proxy of the price elasticity of the demand for the firm's product. Put differently, it is an imprecise measurement of a firm's market power and of its capacity to generate rents (or value added), which in turn may benefit workers' wages through rent sharing (Dobbelaere & Mairesse, 2018; Matano & Naticchioni, 2017). The ability of a firm to create rents (or value added) is contingent on many factors beyond our IVs. These factors notably include the firm's number of clients, the ease with which clients can switch to alternative suppliers, the degree of concentration among the firm's suppliers and the overall competition on the firm's main product market. The relationship between our IVs and firms' capacity to create rents (or value added) is thus not univocal. Moreover, empirical evidence suggests that the magnitude of rent-sharing in the Belgian economy, that is the elasticity between wages and firms' rents, is not big (Rusinek & Rycx, 2013; Rycx & Tojerow, 2004). This outcome is consistent with studies showing that the dispersion of inter-industry wage differentials in Belgium is quite limited compared to other advanced economies, a finding that is notably attributed to the strong centralisation of the Belgian collective bargaining system (du Caju et al., 2011; Rycx, 2002). Finally, as regards the Durbin-Wu-Hausman endogeneity test, the *p*-values associated with the Chi-squared statistics are equal to 0.081 and 0.098.[30] These results suggest that the null hypothesis of no endogeneity should be rejected. The estimates thus indicate that our main explanatory variable, that is firm-level upstreamness, is endogenous and that our instrumentation strategy is warranted.

[29]As suggested by van Ours and Stoeldraijer (2011), we rely on the standard 'rule of thumb' that weak identification is problematic for *F* statistics smaller than 10.

[30]The Durbin-Wu-Hausman test is based on the difference of two Sargan-Hansen statistics: one for the equation in which the upstreamness variable is treated as endogenous, and one in which it is treated as exogenous. If the null hypothesis of this test cannot be rejected, then instrumentation is actually not necessary.

5 | CONCLUSION

This paper is the first to estimate the impact of a direct measurement of firm-level upstreamness (i.e. the steps—weighted distance—before a firm's production meets either domestic or foreign final demand) on productivity, wage costs and profits (i.e. productivity-wage gaps). To this end, we take advantage of our access to detailed Belgian linked employer–employee panel data, which has been merged with the NBB B2B transactions dataset, a unique dataset containing accurate information on the yearly position of virtually all commercial firms in the value chain. Moreover, relying on the methodological framework pioneered by Hellerstein et al. (1999), we estimate panel data models at the firm level. Our main finding, based on the generalised method of moments (GMM) and a two-stage least squares within estimator (FE-IV), is that firms positioned more upstream (i.e. further away from the final consumer) create more value. Overall, our estimates show that if a firm's upstreamness increases by one step—that is, by approximately one standard deviation—its productivity rises on average by around 5% and 20% in the short and long run, respectively. Testing for potential nonlinearities, we find that this relationship is monotonically increasing and even slightly convex, with considerable productivity gains in the early stages of the production chain, when upstreamness is the highest. Regarding distributional issues, our results show that the productivity gains associated with upstreamness are shared almost equally between wages and profits.

How can these results be interpreted? Overall, our findings are in line with the hypothesis that upstream activities (such as innovation, R&D and design) create a lot of value added,[31] whereas pure manufacturing/assembly stages located closer to the final consumers in the production chain tend to add less value (OECD, 2012). It is also generally thought that some downstream activities (such as marketing, branding and logistics) create a lot of value (OECD, 2013), but our findings do not support this assumption.[32] However, leading firms,[33]whose activities are mostly upstream due to intensive upstream activities (e.g. innovation and R&D), generally keep a strong control over high-value-added downstream activities (e.g. marketing). Accordingly, our estimates are compatible with the assertion that more upstream firms are

[31]Our data provide no information on the type of activities that are performed by firms. So, we cannot directly test the hypothesis that more upstream firms would be more intensively involved in activities such as innovation, R&D and design. However, to overcome this limitation, we gathered information from Eurostat regarding R&D expenditures at the industry level (NACE 2 digits) and regressed firms' R&D expenditures on upstreamness using OLS. The corresponding estimates, available on request, support the hypothesis that more upstream sectors are spending significantly more on R&D. More precisely, we find that if upstreamness increases by one step, R&D expenditures at the industry level increase, on average, by 26.8 million EUR. Though, R&D is not the whole story. Indeed, the relationship between R&D expenditures and upstreamness is found to be nonlinear. In particular, results show that firms presenting the largest values of upstreamness (i.e. higher than 4.5 steps) do not spend significantly more on R&D than firms presenting the lowest values of upstreamness (i.e. below 1.5). Furthermore, as highlighted in footnote 21, upstreamness remains positive and significant, even when controlling for R&D both in the productivity and wage equations.

[32]In contrast with Rungi and Del Prete (2018), our estimates indeed provide no direct evidence of a productivity premium in firms whose main activity is close to final demand. Yet, it should be recalled that our study differs from theirs in several key dimensions, including the measurement of firms' position in the value chain (direct longitudinal indicator at the firm level vs. indirect cross-sectional indicator at the sectoral level), the estimation method (GMM-SYS and FE-IV estimators vs. OLS/Probit/Logit estimators), the data coverage (Belgium *vs*. EU countries) and the number of covariates. For more details, see the discussion in the introduction.

[33]Within the Belgian production network, one could notably think of international companies such as ABInBev, Bekaert, Engie, GSK, Ontex, Pfizer, Proximus, Solvay or UCB.

more productive and profitable, but also that being higher in the value chain is likely to facilitate firms' control over strategic downstream activities that bring extra value.[34] Finally, our estimates showing that the productivity gains associated with upstreamness are quite equally split between capital and labour seem quite sound in the light of the Belgian industrial relations system, which is notably characterised by strong collective bargaining centralisation/coordination and high trade union coverage/density (Garnero et al., 2020; Kampelmann & Rycx, 2013; OECD, 2018).

Our findings reinforce and extend those reported for China by Ju and Yu (2015), who also suggested a positive relationship between upstreamness and corporate performance (measured through both productivity and profitability). Yet, our study provides distinctive improvements compared to theirs. Indeed, we consider a direct and accurate measurement of upstreamness for (almost) all commercial firms (operating both in manufacturing and services) for each year from 2002 to 2010. In addition, we rely on panel data estimation techniques (GMM-SYS and FE-IV) that account for firm-level fixed effects and the endogeneity of firms' upstreamness. Moreover, our data enable us to control not only for firm characteristics (i.e. industry, size and capital stock) but also for key variables reflecting the composition of the workforce within those firms (i.e. education, age, occupation, working time). Finally, we pay attention to distributional issues, that is the way the productivity gains associated with upstreamness are shared between capital and labour. Be that as it may, both our studies suggest that firms' upstreamness fosters their productivity and profitability.

ACKNOWLEDGEMENTS

We are most grateful to Statistics Belgium and the National Bank of Belgium for giving us access to the data. Funding for this research has been provided by the National Bank of Belgium. The usual disclaimer applies.

ORCID

Guillaume Vermeylen https://orcid.org/0000-0001-9641-060X

REFERENCES

Alfaro, L., Antràs, P., Chor, D., & Conconi, P. (2019). Internalizing global value chains: A firm-level analysis. *Journal of Political Economy*, *127*(2), 508–559. https://doi.org/10.1086/700935

Amador, J., & Cabral, S. (2016). Global value chains: A survey of drivers and measures. *Journal of Economic Surveys*, *30*(2), 278–301.

Antràs, P., & Chor, D. (2013). Organizing the global value chain. *Econometrica*, *81*(6), 2127–2204.

Antràs, P., Chor, D., Fally, T., & Hillberry, R. (2012). Measuring the upstreamness of production and trade flows. *American Economic Review: Papers and Proceedings*, *102*(3), 412–416. https://doi.org/10.1257/aer.102.3.412

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, *58*(2), 277–297. https://doi.org/10.2307/2297968

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-component models. *Journal of Econometrics*, *68*(1), 28–51.

[34]Buciuni et al. (2014) highlight, on the basis of a case study relative to the furniture industry, that control over operations is key for product innovation and competitiveness of firms participating in a GVC. Along the same lines, Dedrick et al. (2010) showed that Apple has captured a great deal of value from the innovation embodied in the iPod.

Arnold, J., & Hussinger, K. (2005). Export behavior and firm productivity in German manufacturing: A firm-level analysis. *Review of World Economics*, *141*(2), 219–243. https://doi.org/10.1007/s10290-005-0026-8

Baldwin, R. (2011). *Trade and industrialisation after globalisation's 2nd unbundling: How building and joining a supply chain are different and why it matters* (NBER Working Paper, No. 17716). Cambridge, MA.

Baldwin, R., & Yan, B. (2017). Global value chain participation and the productivity of Canadian manufacturing firms. In A. van Assche, R. Wolfe, & S. Tapp (Eds.), *Redesigning Canadian trade policies for new global realities*, (93–117). Institute for Research on Public Policy.

Bartelsman, E., & Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic Literature*, *38*(3), 569–594. https://doi.org/10.1257/jel.38.3.569

Berthou, A., Dhyne, E., Bugamelli, M., Cazacu, A.-M., Demian, C.-V., Harasztosi, P., Lalinsky, T., Merikül, J., Oropallo, F., & Soares, A. (2015). *Assessing European firms' exports and productivity distributions: The CompNet trade module* (NBB Working Paper, No. 282). Brussels.

Berthou, A., & Vicard, V. (2013). *Firms' export dynamics: Experience vs. size* (ECB Working Paper, No. 1616). Frankfurt.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143. https://doi.org/10.1016/S0304-4076(98)00009-8

Bond, S. (2002). Dynamic panel data models: A guide to micro data methods and practice. *Portuguese Economic Journal*, *1*(2), 141–162. https://doi.org/10.1007/s10258-002-0009-9

Bond, S., & Söderbom, M. (2005). *Adjustment costs and the identification of Cobb Douglas production functions* (IFS Working Paper, No. 05/04). London.

Buciuni, G., Coro, G., & Micelli, S. (2014). Rethinking the role of manufacturing in global value chains: An international comparative study in the furniture industry. *Industrial and Corporate Change*, *23*(4), 967–996. https://doi.org/10.1093/icc/dtt048

Buhai, S., Cottini, E., & Westergaard-Nielsen, N. (2017). How productive is workplace health and safety? *Scandinavian Journal of Economics*, *119*(4), 1086–1104. https://doi.org/10.1111/sjoe.12184

Cardoso, A., Guimaraes, P., & Varejao, J. (2011). Are older workers worthy of their pay? An empirical investigation of age-productivity and age-wage nexuses. *De Economist*, *159*, 95–111. https://doi.org/10.1007/s10645-011-9163-8

Chor, D., Manova, K., & Yu, Z. (2021). Growing like China: Firm performance and global production line position. *Journal of International Economics*, *130*, 25. https://doi.org/10.1016/j.jinteco.2021.103445

de Vries, G., Jiang, A., Lemmers, O., & Wei, S.-J. (2021). Firm productivity and functional specialisation. *The World Economy*, *44*(5), 1232–1260.

Dedrick, J., Kreamer, K., & Linden, G. (2010). Who profits from innovation in global value chains? A study of the iPod and notebook PCs. *Industrial and Corporate Change*, *19*(1), 81–116. https://doi.org/10.1093/icc/dtp032

Demunter, C. (2000). *Structure and distribution of earnings survey: Analysis 1995* (Statistics Belgium Working Paper). Brussels.

Devicienti, F., Grinza, E., & Vannoni, D. (2018). The impact of part-time work on firm total factor productivity: Evidence from Italy. *Industrial and Corporate Change*, *27*(2), 321–347.

Dhyne, E., & Duprez, C. (2013). Structural dynamics of Belgium's foreign trade. *NBB Economic Review*, June 2013, 27–38.

Dhyne, E., & Duprez, C. (2015). Has the crisis altered the Belgian economy's DNA? *NBB Economic Review*, Sept. 2015, 29–40.

Dhyne, E., & Fuss, C. (2014). Main lessons of the NBB's 2014 conference 'Total factor productivity: Measurement, determinants and effects'. *NBB Economic Review*, Dec. 2014, 63–76.

Dobbelaere, S., & Mairesse, J. (2018). Comparing micro-evidence on rent sharing from two different econometric models. *Labour Economics*, *52*, 18–26. https://doi.org/10.1016/j.labeco.2018.02.009

Dobbelaere, S., & Vancauteren, M. (2014). *Market imperfections, skills and total factor productivity: Firm-level evidence on Belgium and the Netherlands* (NBB working paper, No. 267).

du Caju, P., Rycx, F., & Tojerow, I. (2011). Inter-industry wage differentials: How much does rent sharing matter? *The Manchester School*, *79*(4), 691–717. https://doi.org/10.1111/j.1467-9957.2010.02173.x

Eaton, J., Eslava, M., Kugler, M., & Tybout, J. (2007). *Export dynamics in Colombia: Firm-level evidence* (NBER Working Paper, No.13531). Cambridge, MA.

Fally, T. (2011). *On the fragmentation of production in the US*. Mimeo, University of Colorado-Boulder.

Feenstra, R., & Hanson, G. (1996). Globalisation, outsourcing, and wage inequality. *American Economic Review*, *86*(2), 240–245.

Feenstra, R., & Hanson, G. (1999). The impact of outsourcing and high-technology capital on wages: Estimates for the United States, 1979–1990. *Quarterly Journal of Economics*, *114*(3), 907–940. https://doi.org/10.1162/003355399556179

Freund, C., & Pierola, M.-D. (2010). *Export entrepreneurs: Evidence from Peru* (World Bank Policy Research Working Paper, No. 5407). Washington.

Garnero, A., Kampelmann, S., & Rycx, F. (2014). Part-time work, wages and productivity: Evidence from Belgian matched panel data. *Industrial and Labor Relations Review*, *67*(3), 926–954. https://doi.org/10.1177/0019793914537456

Garnero, A., Rycx, F., & Terraz, I. (2020). Productivity and wage effects of firm-level collective agreements: Evidence from Belgian linked panel data. *British Journal of Industrial Relations*, *58*(4), 936–972. https://doi.org/10.1111/bjir.12525

Gereffi, G. (2014). Global value chains in a post-Washington consensus world. *Review of International Political Economy*, *21*(1), 9–37.

Giovannetti, G., & Marvasi, E. (2018). Governance, value chain positioning and firms' heterogeneous performance: The case of Tuscany, international. *Economics*, *154*, 86–107. https://doi.org/10.1016/j.inteco.2017.11.001

Giuliano, R., Kampelmann, S., Mahy, B., & Rycx, F. (2017). Short notice, big difference? The effect of temporary employment on firm competitiveness across sectors. *British Journal of Industrial Relations*, *55*, 421–449. https://doi.org/10.1111/bjir.12236

Göbel, C., & Zwick, T. (2012). Age and productivity: Sector differences. *De Economist*, *160*(1), 35–57. https://doi.org/10.1007/s10645-011-9173-6

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*, 1029–1054. https://doi.org/10.2307/1912775

Heckel, T., Le Bihan, H., & Montornès, J. (2008). *Sticky wages: Evidence from quarterly microeconomic data* (ECB Working Paper, No. 893). Frankfurt.

Hellerstein, J., Neumark, D., & Troske, K. (1999). Wages, productivity and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labor Economics*, *17*(2), 409–446. https://doi.org/10.1086/209926

Hempell, T. (2005). What's spurious? What's real? Measuring the productivity impacts of ICT at the firm level. *Empirical Economics*, *30*(2), 427–464. https://doi.org/10.1007/s00181-005-0248-6

Ilmakunnas, P., & Maliranta, M. (2005). Technology, labour characteristics and wage-productivity gaps. *Oxford Bulletin of Economics and Statistics*, *67*, 623–645. https://doi.org/10.1111/j.1468-0084.2005.00134.x

Ju, J., & Yu, Y. (2015). Productivity, profitability, production and export structures along the value chain in China. *Journal of Comparative Economics*, *43*(1), 33–54. https://doi.org/10.1016/j.jce.2014.11.008

Kampelmann, S., & Rycx, F. (2013). Does institutional diversity account for pay rules in Germany and Belgium? *Socio-Economic Review*, *11*(1), 131–157. https://doi.org/10.1093/ser/mws009

Kampelmann, S., Rycx, F., Saks, Y., & Tojerow, I. (2018). Does education raise productivity and wages equally? The moderating roles of age and gender and industry. *IZA Journal of Labor Economics*, *7*(1), 1–37.

Koyck, L. (1954). *Distributed lags and investment analysis*. North Holland.

Lallemand, T., Plasman, R., & Rycx, F. (2006). Establishment size and the dispersion of wages: Evidence from European countries. *Applied Economics Quarterly*, *52*(4), 309–336.

Lallemand, T., Plasman, R., & Rycx, F. (2007). The establishment-size wage premium: Evidence from European countries. *Empirica*, *34*(5), 427–451. https://doi.org/10.1007/s10663-007-9042-3

Mahlberg, B., Freund, I., Cuaresma, J., & Prskawets, A. (2013). Ageing, productivity and wages in Austria. *Labour Economics*, *22*, 5–15. https://doi.org/10.1016/j.labeco.2012.09.005

Manello, A., Calabrese, G., & Frigero, P. (2016). Technical efficiency and productivity growth along the automotive value chain: Evidence from Italy. *Industrial and Corporate Change*, *25*(2), 245–259. https://doi.org/10.1093/icc/dtv021

Matano, A., & Naticchioni, P. (2017). The extent of rent sharing along the wage distribution. *British Journal of Industrial Relations*, *55*(4), 751–777. https://doi.org/10.1111/bjir.12234

Melitz, M., & Ottaviano, G. (2003). *Market size, trade, and productivity*. Mimeo, Harvard University.

Nielen, S., & Schiersch, A. (2014). Temporary agency work and firm competitiveness: Evidence from German manufacturing firms. *Industrial Relations*, *53*, 365–393. https://doi.org/10.1111/irel.12062

OECD (2009). *Measuring capital*. OECD Publishing.
OECD (2010). *OECD economic globalisation indicators 2010*. OECD Publishing.
OECD (2012). *Mapping global value chains*. OECD Publishing.
OECD (2013). *Interconnected economies: Benefiting from global value chains*. OECD Publishing.
OECD (2018). *OECD employment outlook 2018*. OECD Publishing.
Pietrobelli, C., & Saliola, F. (2008). Power relationships along the value chain: Multinational firms, global buyers and performance of local suppliers. *Cambridge Journal of Economics*, *32*(6), 947–962. https://doi.org/10.1093/cje/ben016
Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, *9*(1), 86–136. https://doi.org/10.1177/1536867X0900900106
Rungi, A., & Del Prete, D. (2018). The smile curve at the firm level: Where value is added along supply chains. *Economics Letters*, *164*, 38–42. https://doi.org/10.1016/j.econlet.2017.12.038
Rusinek, M., & Rycx, F. (2013). Rent sharing under different bargaining regimes: Evidence from linked employer-employee data. *British Journal of Industrial Relations*, *51*(1), 28–58. https://doi.org/10.1111/j.1467-8543.2011.00877.x
Rycx, F. (2002). Inter-industry wage differentials: Evidence from Belgium in a cross-national perspective. *De Economist*, *150*(5), 555–568.
Rycx, F., & Tojerow, I. (2004). Rent sharing and the gender wage gap in Belgium. *International Journal of Manpower*, *25*(3–4), 279–299. https://doi.org/10.1108/01437720410541407
Serpa, J., & Krishnan, H. (2018). The impact of supply chains on firm-level productivity. *Management Science*, *64*(2), 511–532. https://doi.org/10.1287/mnsc.2016.2632
Shen, J.-H., Zhang, L., Lee, C.-C., Zhang, J., & Shen, L. (2021). Towards a dynamic model of the industrial upgrading with global value chains. *The World Economy*, *44*(9), 2683–2702.
Sturgeon, T., Van Biesebroeck, J., & Gereffi, G. (2008). Value chains, networks and clusters: Reframing the global automotive industry. *Journal of Economic Geography*, *8*(3), 297–321. https://doi.org/10.1093/jeg/lbn007
Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, *49*(2), 326–365. https://doi.org/10.1257/jel.49.2.326
Szymczak, S., Parteka, A., & Wolszczak-Derlacz, J. (2019). *Position in global value chains: The impact on wages in Central and Eastern European countries* (Working Paper, No. 53). Faculty of Management and Economics, Gdansk University of Technology, Gdansk.
van Ours, J., & Stoeldraijer, L. (2011). Age, wage and productivity in Dutch manufacturing. *De Economist*, *159*(2), 113–137. https://doi.org/10.1007/s10645-011-9159-4
Woodall, P., & Shannon, T. (2018). Monopoly power corrodes choice and resiliency in the food system. *Antitrust Bulletin*, *63*(2), 198–221. https://doi.org/10.1177/0003603X18770063
Wooldridge, J. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*(3), 112–114. https://doi.org/10.1016/j.econlet.2009.04.026

How to cite this article: Mahy, B., Rycx, F., Vermeylen, G., & Volral, M. (2021). Productivity and wage effects of firm-level upstreamness: Evidence from Belgian linked panel data. *The World Economy*, 00, 1–29. https://doi.org/10.1111/twec.13227

APPENDIX

TABLE A1 Means and standard deviations of selected variables in different sub-samples (2002–2010)

Sample	Initial sample	Non-missing information	GMM-SYS sample
Annual value added per worker (€)[a]	78,805	91,291	91,809
	(438,598)	(593,423)	(624,962)
Annual wage cost per worker (€)[a]	35,623	43,437	47,817
	(436,001)	(591,126)	(23,135)
Upstreamness (in steps)	2.14	2.30	2.34
	(0.80)	(0.94)	(0.93)
Size (number of workers)	130.15	216.01	229.69
Sector (%)			
Mining and quarrying (B)	0.38	0.65	0.69
Manufacturing (C)	38.77	48.16	48.74
Electricity, gas, steam and air conditioning supply; water supply, sewerage, waste management and remediation activities (D + E)	0.51	0.65	0.59
Construction (F)	9.53	8.14	8.27
Wholesale and retail trade, repair of motor vehicles and motorcycles (G)	21.46	17.01	17.13
Accommodation and food service activities (I)	2.54	1.95	1.87
Transport and storage; information and communication (H + J)	7.22	5.89	5.84
Financial and insurance activities (K)	2.48	2.75	2.61
Real estate activities; Professional, scientific and technical activities; Administrative and support service activities (L + M + N)	17.11	14.80	14.26
Number of firm-year observations	27,226	13,905	11,975
Number of firms	10,847	4177	3529

[a]At 2004 constant prices.

TABLE A2 Upstreamness, productivity and wage costs detailed static OLS estimates, 2002–2010

	OLS	
	Productivity	Wage costs
Dependent variable	(1)	(2)
Upstreamness[a]	0.075***	0.026***
	(0.005)	(0.003)
% Of workers younger than 30	−0.405***	−0.440***
	(0.033)	(0.025)

(Continues)

TABLE A2 (Continued)

Dependent variable	OLS Productivity (1)	Wage costs (2)
% Of workers older than 49	0.154***	0.151***
	(0.043)	(0.026)
Highly educated – short (Bachelor)	0.506***	0.410***
	(0.034)	(0.023)
Highly educated – long (Master)	0.972***	0.834***
	(0.040)	(0.027)
Highly educated – third level (post-Master)	1.974***	1.536***
	(0.209)	(0.141)
Blue collars	−0.129***	−0.067***
	(0.019)	(0.014)
Part-timers	−0.759***	−0.741***
	(0.029)	(0.027)
Nace C	0.148	0.030
	(0.096)	(0.046)
Nace D + E	0.825**	0.133*
	(0.353)	(0.073)
Nace F	−0.025	0.028
	(0.028)	(0.021)
Nace G	0.010	0.031**
	(0.026)	(0.016)
Nace I	−0.241**	−0.243***
	(0.099)	(0.084)
Nace H + J	−0.019	0.025
	(0.038)	(0.024)
Nace K	0.342**	0.031
	(0.163)	(0.140)
Nace L + M + N	−0.010	−0.018
	(0.033)	(0.019)
Size (in ln)	0.052***	0.072***
	(0.004)	(0.002)
Capital stock per worker (in ln)	0.139***	0.043***
	(0.004)	(0.003)
Year dummies (8)	YES	YES
Adjusted R^2	.473	.528
Sig. model (p-value)	.000	.000
Number of observations	11,975	11,975
Number of firms	3529	3529

Notes: ***/**/*Significant at the 1%, 5% and 10% level, respectively. Robust standard errors are shown in brackets.
[a]Steps (distance) before the production of a firm meets either domestic or foreign final demand.

TABLE A3 Upstreamness, productivity and wage costs—sensitivity tests firm-level OLS estimates, 2002–2010

	Productivity					Wage costs				
Dependent variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Upstreamness at firm level[a]	0.075***	0.044***	0.083***	0.055***	0.043***	0.026***	0.012***	0.031***	0.017**	0.012***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Upstreamness at industry level (NACE 2 digits)					0.107***					0.085***
					(0.010)					(0.007)
R&D expenditures at NACE 2-digit level (in ln)			0.009***	0.017				0.022***	−0.015	
			(0.003)	(0.021)				(0.002)	(0.012)	
Average number of suppliers of the firm's clients			0.000***	0.000***				0.000***	0.000***	
			(0.000)	(0.000)				(0.000)	(0.000)	
Average number of clients of the firm's suppliers			−0.000***	−0.000***				−0.000***	−0.000***	
			(0.000)	(0.000)				(0.000)	(0.000)	
Concentration index of domestic suppliers of the firm			0.169***	0.149***				0.066***	0.064***	
			(0.032)	(0.030)				(0.018)	(0.018)	
Concentration index of domestic clients of the firm			0.077***	0.072***				0.036***	0.030**	
			(0.018)	(0.017)				(0.012)	(0.013)	
NACE 2-digit industry dummies	No	Yes	No	Yes	No	No	Yes	No	Yes	No
Individual and job characteristics[b]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm characteristics[c]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	.473	.514	.418	.451	.477	.528	.509	.509	.513	.457
Model significance										
p-Value of *F* test	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Number of firm-year observations	11,975	11,975	8061	8061	11,975	11,973	11,973	8060	8060	11,973

Notes: ***/**/*Significant at the 1%, 5% and 10% level, respectively. Robust standard errors are shown in brackets.

[a]Steps (distance) before the production of a firm meets either domestic or foreign final demand.

[b]Share of the workforce that: (i) is younger than 30 and older than 49 years, respectively, and (ii) is highly educated (three categories). The share of blue-collar workers and the share of part-time workers are also included.

[c]Sectoral affiliation (eight dummies), number of workers and capital stock.