CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1673

January 2020

Intellectual Property and the Organization of the Global Value Chain

Stefano Bolatto
Alireza Naghavi
Gianmarco Ottaviano
Katja Zajc Kejzar

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

UKRI Economic and Social Research Council

Abstract
This paper introduces the concept of intangible assets in a property rights model of sequential supply chains. Firms transmit knowledge to their suppliers to facilitate input customization. Yet, to avoid knowledge dissipation, they must protect the transmitted intangibles, the cost of which depends on the knowledge intensity of inputs and the quality of institutions protecting intellectual property rights (IPR) in supplier locations. When input knowledge intensity increases (decreases) downstream and suppliers' investments are complements, the probability of integrating a randomly selected input is decreasing (increasing) in IPR quality and increasing (decreasing) in the relative knowledge intensity of downstream inputs. Opposite but weaker predictions hold when suppliers' investments are substitutes. Comprehensive trade and FDI data on Slovenian firms' value chains provide evidence in support of our model's predictions. They also suggest that, in line with our model, better institutions may have very different effects on firm organization depending on whether they improve the protection of tangible or intangible assets.

Key words: sequential production, intellectual property, intangible assets, appropriability, stage complementarity, upstreamness, firm organization, outsourcing, vertical integration
JEL Codes: F12; F14; F21; F23; D23; L22; L23; L24; O34

This paper was produced as part of the Centre’s Trade Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We are grateful to Carlo Altomonte, Pol Antr_as, Andrew Bernard, Davin Chor, Paola Conconi, Bram De Lange, Wilhelm Kohler, Bohdan Kukharskyy, Ahmad Lashkaripour, Keith Maskus, Rahul Mukherjee, Stela Rubinova, Catherine Thomas, Gerald Willman, and participants of several seminars, workshops, and conferences for extremely valuable discussions. Katja Zajc Kej_zar acknowledges the financial support from the Slovenian Research Agency (research core funding No. (P5-0117)). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n 822390-MICROPROD-H2020-SC6-TRANSFORMATIONS-2018). The usual disclaimer applies.

Stefano Bolatto, University of Bologna. Alireza Naghavi, University of Bologna and Johns Hopkins University SAIS Europe. Gianmarco Ottaviano, University of Bocconi and Centre for Economic Performance, London School of Economics. Katja Zajc Ljubljana, University of Ljubljana.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Despite recent setbacks for international trade due to renewed protectionist pressures, in the last decades value chains have generally become more global in nature due to the increased participation of suppliers located across different countries. In this context, incomplete contracts and contract enforcement continue to be a central issue when studying firms' organizational choices.[1] The two canonical approaches to confronting this issue are the 'transaction costs' theory of the firm (Williamson, 1971, 1975, 1985) and the 'property rights' theory of the firm (Grossman and Hart, 1986; Hart and Moore, 1990), which have helped understanding how specific institutional features of different production locations affect firms' organizational decisions. According to the transaction cost approach, better contracting institutions reduce the hold-up problems associated with outsourcing and facilitate the exploitation of the corresponding gains from specialization. Differently, according to the property rights approach, better contracting institutions mitigate the need to create investment incentives through outsourcing and allow firms to reap a larger share of the final revenues through integration. Empirical studies, from Corcos et al. (2013) to Eppinger and Kukharskyy (2017), have found strong evidence in favor of the property rights theory: better institutional quality increases the incidence of integration. There are exceptions, such as Defever and Toubal (2013) who highlight that, in line with the transaction costs approach, outsourcing is more frequently observed for more productive firms due to its higher organizational costs.

Most existing works on international trade and firm organization have, however, focused on hold-up problems related to tangible assets, compelling Antràs and Rossi-Hansberg (2009) to underline missing research on how the non-appropriable nature of knowledge may also affect firms' organizational choices. Their comment gains particular salience in the case of sequential production along supply chains. Atalay, Hortacsu and Syverson (2014) emphasize the rationale for using vertical integration as a way to promote efficient intra-firm transfers of intangible inputs (such as marketing know-how, intellectual property or R&D capital). They show that, in line with the property rights theory, for US firms integration is not much of a tool to ensure a smooth flow of physical inputs from upstream to downstream production stages, but rather a means to secure the efficient transmission of technology along the value chain. Branstetter, Fisman and Foley (2006) provide evidence that knowledge transmission by US multinationals to their affiliates increases after IPR reforms in host countries. Canals and Şener (2014) find that US firms substantially expand their outsourcing activities in high-tech industries as a response to IPR reforms in the host countries. Naghavi, Spies and Toubal (2015) further show that, when outsourcing of complex products involves the sharing of technology with a supplier, French multinationals choose countries with better IPR enforcement. This is in line with earlier work by Yang and Maskus (2001), who argue that countries with stronger patent

[1] See the vast literature on international trade and the boundaries of firms (e.g. Antràs, 2003, 2005; Antràs and Helpman, 2004, 2008; Grossman and Helpman, 2002, 2003, 2005).

rights attract larger arm's length volumes of licensed technology. Finally, Kukharskyy (2019) shows that better IPR quality weakens a headquarter's threat of knowledge dissipation by its supplier, reducing the need to use integration to protect its knowledge against imitation.

Against this backdrop, our aim is to follow up on the foregoing comment by Antràs and Rossi-Hansberg (2009) in terms of both theory and empirical analysis. As for the former, we introduce the concept of intangible assets in a property rights model of sequential value chains *à la* Antràs and Chor (2013) and Alfaro et al. (2019). In their models, in order to produce customized inputs, suppliers along the value chain have to undertake relation-specific investments under contractual incompleteness arising from the fact that the delivered quality of an input is not verifiable by third parties (such as a court or an arbitrator) and an input of low quality cannot be used for final production. Contractual incompleteness leads to ex-post Nash bargaining on the suppliers' contributions to final revenues. Faced with the possibility of being held-up at the ex-post bargaining stage, input suppliers underinvest in the relation with the final producer. The latter can alleviate the resulting hold-up inefficiency by appropriately choosing the organization of production facing a trade-off between surplus extraction (which is better served by the vertical integration of the supplier) and supplier incentivization (which is better served by an arm's length outsourcing contract). In our model, the choice between integration and outsourcing is also affected by the fact that, in order to support input customization, firms have to transmit knowledge to their suppliers. However, to avoid knowledge dissipation, they also have to protect the transmitted intangibles, the cost of which depends on the knowledge intensity of inputs and the quality of institutions protecting intellectual property rights (IPR) in the suppliers' locations.

As in Antràs and Chor (2013) and Alfaro et al. (2019), also in our model the profit-maximizing organizational choice depends on whether suppliers' relation-specific investments are complements or substitutes along the value chain. If they are complements, when outsourcing and integration coexist, the former is chosen upstream while the latter takes place downstream. If they are substitutes, the opposite pattern holds. Once issues related to IPR are factored in, also the variation of knowledge intensity across sequential inputs affects the trade-off between surplus extraction and supplier incentivization as the firm's organizational decision about any input is not independent from its decision on how much knowledge to transmit along the entire value chain. In particular, when the knowledge intensity of inputs increases (decreases) downstream and suppliers' investments are complements, the probability of integrating a randomly selected input is decreasing (increasing) in the quality of IPR protection and increasing (decreasing) in the relative knowledge intensity of downstream inputs. Opposite but weaker predictions hold when suppliers' investments are substitutes.

Intuitively, if relatively less knowledge is transmitted upstream of a given stage z and suppliers' investments are sequential substitutes, a firm is less likely to use outsourcing at that stage, favoring rent extraction over supplier incentivization. The reason is that, with less upstream knowledge transmission, upstream suppliers contribute less to the firm's revenues and, with sequential substitutability, that raises supplier z's return on investment. Differently, if relatively less knowledge is transmitted upstream of z but suppliers' investments are sequential complements, the firm is less likely to use vertical integration at stage z, favoring supplier incentivization over rent extraction as the limited contribution of upstream suppliers to the firm's revenue reduces supplier z's return to investment. These effects associated with knowledge transmission interfere with the hold-up effects of contractual incompleteness already highlighted by Antràs and Chor (2013) and Alfaro et al. (2019) in the case of sequential production. Their relevance is, however, mitigated when the quality of IPR protection improves.

Turning to the empirical analysis, we test our model's predictions through probit regressions exploiting comprehensive data on the population of Slovenian firms from 2007 to 2010. We merge transaction-level trade data on firms with their outward cross-border direct investment and financial data. Transaction-level trade data provide us with information on the complete set of inputs imported at the firm level, while FDI data gives us the country of affiliates. The firm's decision to integrate an input is estimated at the firm-country-product level. It is measured as the probability of transacting an input in a particular source country within firm boundaries, whereby distinguishing between integration and outsourcing by exploiting information on the core activity of the firm's affiliate in a particular host country, in the wake of Alfaro et al. (2019). To locate the position of inputs along a value chain, we use industry-pair specific measures of upstreamness as in Alfaro et al. (2019) and, to determine whether inputs are sequential complements or substitutes, we use the demand elasticity of the firm's core export product as in Antràs and Chor (2013), as well as the demand elasticity of its inputs and a measure of their technological substitutability. Finally, to define the knowledge intensity of inputs, we follow the Eurostat classification based on the R&D intensity of their industry and, to measure the quality of IPR protection, we take the IPR enforcement index from Park (2008).

We find that our model's predictions hold at the most disaggregated level when controlling for unobserved firm-specific effects as in Mundlak (1978), Chamberlain (1984) and Wooldridge (2002), and for firm-country-product level unobserved heterogeneity when using a random effects probit model. Moreover, in line with the model, we find that better overall contract enforcement ('rule of law') has the opposite impact than better IPR quality, which suggests that better institutions may have very different effects on firm organization depending on whether they improve the protection of tangible or intangible assets. It also shows that our findings are specific to IPR institutions and cannot be generalized to other regulatory measures that affect contract enforcement. All our findings are robust to alternative specifications and definitions of integration/outsourcing and complements/substitutes, as well as to the inclusion of a battery of firm-level controls and additional source-country institutional variables.

The rest of the paper is organized as follows. Section 2 provides a brief overview of our theoretical framework through a simple model of a supply chain consisting of two production stages only: a downstream final stage and an upstream intermediate stage. Section 3 extends the simple model to a richer setup of sequential production, which we use to derive empirically testable predictions. Section 4 presents the data and the variables we use for our empirical analysis. Section 5 tests the predictions of the model. Section 6 concludes.

2 Intangibles and Intellectual Property Protection

To understand how knowledge transmission affects the organization of a value chain, it is useful to start by introducing a simple theory of knowledge dissipation into a property rights framework with hold-up inefficiencies where the supply chain consists of two production stages only: a final stage performed by a 'firm' and a single intermediate stage of production performed by a 'supplier'. The supplier has to solve a series of problems and come up with solutions for the provision of a fully customized ('tangible') input to the firm. The firm then uses the input to produce and sell a differentiated final product with market power that allows for the extraction of monopolistic rents from consumers. The supply contract is incomplete, giving rise to a hold-up problem that the firm deals with through an organizational choice between vertically integrating the supplier and relying on the supplier as an independent outsourced contractor.

Input customization requires the transmission of firm-specific knowledge ('intangibles') from the final producer to the supplier. The more knowledge is transmitted, the closer the input is to the firm's specifications and thus the higher is the input's productivity when used for final production by the firm. However, transmitted knowledge has to be protected by the firm to avoid the risk of 'dissipation'. This arises from the existence of a large number of potential competitors in the final market that, from any bit of unprotected knowledge, can reverse engineer all knowledge needed to reproduce the final product by themselves, thus destroying the firm's monopolistic rents. In other words, knowledge transmitted without protection by the firm becomes a public good.

In this setup, as it will be discussed in detail below, two problems affect the supplier's incentive to invest in relation-specific customization: the 'hold-up problem' due to the incompleteness of the supply contract, and the 'knowledge transmission problem' due to costly knowledge protection.

2.1 Hold-up and knowledge transmission

Consider an industry in which the final good is available in many differentiated varieties, each manufactured by a monopolistically competitive firm. Preferences are described by a standard CES utility function, thereby each firm faces the following demand for its variety:

$$q = Ap^{-\frac{1}{1-\rho}}, \tag{1}$$

where q is quantity demanded, p is price, $A > 0$ is a demand shifter that the firm treats as exogenous, and $\rho \in (0, 1)$ is a measure of the price elasticity of final demand with the elasticity of substitution between varieties equal to $1/(1 - \rho)$.

Final production of each variety requires a customized intermediate input and customization requires knowledge transmission from the firm to the input supplier. Specifically, final production obeys the linear technology

$$q = \theta\delta x, \tag{2}$$

where $q \geq 0$ is the amount of final output, $x \geq 0$ is the amount of intermediate input, $\theta > 0$ is the firm's productivity, and $\delta \in [0, 1]$ is the input's productivity as determined by the amount of knowledge transmitted by the firm to the supplier. With $\delta = 0$ no knowledge is transmitted and intermediate production cannot take place; with $\delta = 1$ all relevant knowledge is transmitted and the input's productivity is at its maximum.

In order to produce the customized input, the supplier has to undertake a relation-specific investment under contractual incompleteness. This is because the delivered quality of the input is not verifiable by third parties (such as a court or an arbitrator) and an input of low quality cannot be used for final production. Contractual incompleteness leads to ex-post Nash bargaining on the joint surplus from the relation, that is, on the revenues generated by final sales. When bargaining ex post, both parties have no outside option. For the supplier, once produced, the customized input has no value outside the relation with the firm. As for the firm, should it be unhappy with the delivered input, it would be too late to find an alternative supplier. Faced with the possibility of being held-up at the ex-post bargaining stage, the supplier underinvests in the relation.

The final producer can alleviate the resulting hold-up inefficiency by appropriately choosing the organization of production between the vertical integration of the supplier (labeled V) or an arm's length outsourcing contract (labeled O). Under vertical integration the final producer is in control of the physical assets used in intermediate production, which allows the firm to extract more surplus from the supplier when it comes to ex-post bargaining. This feature is captured by assuming that the firm's Nash bargaining weight $\beta \in (0, 1)$ is larger under vertical integration than under outsourcing ($\beta_V > \beta_O$) so that the firm appropriates a larger share of joint surplus under the former than the latter. However, foreseeing a lower return on its relation-specific investment, an integrated supplier is inevitably more prone to underinvest in its relationship with the firm than an independent supplier. Accordingly, the firm's organizational choice faces a trade-off between surplus extraction and supplier incentivization.

The final production technology (2) highlights the importance of knowledge transmission: the more knowledge is transmitted from the firm to the supplier, the higher the input's productivity. However, to avoid dissipation and rent destruction, knowledge transmission has to be protected. This is costly and the cost depends on both the characteristics of the input in terms of 'knowledge intensity' and those of the country where the input is produced in terms of IPR quality. Specifically, the cost of protecting an amount δ of transmitted knowledge is assumed to be

$$\kappa(\omega, \lambda) = \omega \ \delta^{\lambda}, \tag{3}$$

where $\omega > 0$ measures the input's knowledge intensity and $\lambda > 0$ measures the country's quality of IPR institutions.[2] The cost of protecting knowledge transmission is increasing in the amount of knowledge transmitted δ. For given δ, it is higher the larger is *input-specific* knowledge intensity (i.e. the larger is ω): more knowledge-intensive inputs are more difficult to protect from knowledge dissipation. It is also higher the worse is *country-specific* IPR quality (i.e. the smaller is λ). Given that from any bit of unprotected knowledge, potential competitors can reverse engineer all knowledge needed to reproduce the final product, all transmitted knowledge will be protected in equilibrium.

2.2 Organizational choice

The timing of events is as follows. First, the firm chooses the organizational form $\beta \in \{\beta_V, \beta_O\}$ and the amount of transmitted knowledge $\delta \in [0, 1]$. Second, the firm posts a contract for the provision of the customized input, stating the chosen organizational form and knowledge transmission. Both are verifiable by third parties and thus contractible. Third, a large number of identical potential suppliers competitively bid for the contract and the firm selects one among them. Fourth, the selected supplier decides how much to invest in the relationship with the firm, that is, how much to supply of the intermediate input x. Fifth, the firm and the supplier bargain on how to share their joint surplus consisting of revenues from final sales. Sixth and last, final production takes place, output is sold and revenues are shared according to the agreed split rule.

Given this timing, the model has to be solved backwards, characterizing first the supplier's decision on x and then the firm's decisions of β and δ. As for the former, taking $\beta \in \{\beta_V, \beta_O\}$ and $\delta \in [0, 1]$ as given, the supplier chooses x so as to maximize its profit

$$\pi_S = (1 - \beta) \ r(x) - c \ x, \tag{4}$$

where c is the marginal cost of input production, $r(x) = \theta^{\rho} A^{1-\rho} (\delta x)^{\rho}$ is revenues from final sales, and $(1-\beta)$ is the supplier's share of these revenues. The profit-maximizing amount of input supplied

[2] For example, in the case of protection through patenting, $\kappa(\omega, \lambda)$ would compound the difficulty of filing and getting a patent approved with the cost of enforcing the patent.

to the firm then evaluates to

$$x^*(\beta,\delta) = A\left(\frac{\rho\theta^\rho}{c}\right)^{\frac{1}{1-\rho}}(1-\beta)^{\frac{1}{1-\rho}}\,\delta^{\frac{\rho}{1-\rho}}, \tag{5}$$

which highlights that the supplier's relation-specific investment is increasing in its share of surplus $(1-\beta)$. This confirms that, given $\beta_V > \beta_O$, the supplier's investment is higher with outsourcing than with vertical integration.

Turning to final production, anticipating the supplier's choice (5), the firm selects $\beta \in \{\beta_V, \beta_O\}$ and $\delta \in [0,1]$ so as to maximize its own profit

$$\pi_F = A\left(\frac{\rho\theta}{c}\right)^{\frac{\rho}{1-\rho}}\beta(1-\beta)^{\frac{\rho}{1-\rho}}\delta^{\frac{\rho}{1-\rho}} - \omega\delta^\lambda, \tag{6}$$

where the cost of protected knowledge transmission (3) is subtracted from the firm's share of final revenues $\beta r(x^*)$. The optimal choice of β is thus independent from δ as it maximizes $\beta(1-\beta)^{\frac{\rho}{1-\rho}}$. Specifically, if the firm's problem were 'relaxed' so that the firm's bargaining weight β were not constrained to be either β_V or β_O but could instead take any value between 0 and 1, the firm would optimally set β at $\beta^+ \equiv 1-\rho$, as doing so would satisfy the corresponding first-order condition whatever the value of δ. This implies that three cases arise for the constrained optimization. The first two cases are unambiguous: for $\beta_O < \beta_V < \beta^+$ the firm necessarily prefers vertical integration to outsourcing, whereas for $\beta^+ < \beta_O < \beta_V$ it necessarily prefers outsourcing to vertical integration. Hence, vertical integration is the firm's optimal choice when ρ is small enough, and outsourcing is its optimal choice when ρ is large enough. This reveals that, when the demand is more elastic (larger ρ), the firm is more inclined to outsource (smaller β^+), whereas a more rigid demand (smaller ρ) increases the firm's propensity to integrate (larger β^+). Intuitively, lower vales of the demand elasticity (smaller ρ) make the firm's revenues more concave in output, hence the firm gives more weight to rent extraction through integration than to increasing scale by incentivizing the supplier through outsourcing. In the third and last case, for $\beta_O < \beta^+ < \beta_V$ the firm's choice depends on other parameter restrictions determining whether π_F is larger for β_O or β_V. Nonetheless, the general insight that more rigid demand favors vertical integration holds true.

For any given β, the first-order condition for the maximization of (6) with respect to δ implies that the optimal amount of knowledge transmission evaluates to

$$\delta^* = \left[\frac{A}{\omega\lambda}\left(\frac{\rho\theta}{c}\right)^{\frac{\rho}{1-\rho}}\frac{\rho}{1-\rho}\,\beta\,(1-\beta)^{\frac{\rho}{1-\rho}}\right]^{\frac{1}{\lambda-\frac{\rho}{1-\rho}}}, \tag{7}$$

where $\lambda > \rho/(1-\rho)$ is assumed to hold for the second-order condition to be satisfied. This reveals that, as the optimally chosen β maximizes $\beta\,(1-\beta)^{\frac{\rho}{1-\rho}}$, the preferred organizational choice is associated with more knowledge transmission than the alternative. In other words, the organizational choice that more efficiently deals with the hold-up problem happens to be also the one that maximizes knowledge transmission. In addition, expression (7) also reveals that, once the organizational form β has been chosen, larger ω leads to lower δ^* as protecting knowledge transmission is more costly. We highlight this result as:

Lemma 1 *In a value chain consisting of two stages, the final producer prefers vertical integration to outsourcing when the elasticity of final demand is low. Irrespective of the organizational form chosen, higher input knowledge intensity discourages knowledge transmission from the final producer to the input supplier, but does not affect the organizational choice.*

We now show that this independence between the parallel decisions on organization and knowledge transmission does not carry through to more complex sequential production.[3]

3 Sequential Production and Intangibles

We now assume that producing the final good requires a unit measure of inputs that have to be sequentially supplied, each of them corresponding to a different stage of a long value chain as in Antràs and Chor (2013) and Alfaro et al (2019). We index each stage by $z \in [0, 1]$ such that $z = 0$ is the first stage to be performed (i.e. the most 'upstream'), and $z = 1$ is the last one (i.e. the most 'downstream'). At the end of each stage z, a certain amount of the corresponding input $x(z)$ is delivered to the next stage of production for further reprocessing, so that any further stage brings the associated intermediate input closer to the one needed for final production (which was the only input in the previous section).

Sequential supply is captured by extending the production function (2) to

$$q = \theta \left(\int_0^1 [\delta(z)\ x(z)]^\alpha I(z)\ dz \right)^{1/\alpha}, \tag{8}$$

where: $\alpha \in (0, 1)$ is the degree of substitutability between the different inputs, measuring the extent to which less processing at a given stage can be compensated by more processing at another stage; $\delta(z)$ is the productivity of input z; and $I(z)$ is an indicator function taking value 1 if stage z has been completed and 0 otherwise. This last feature is what makes the production process described by (8)) inherently 'sequential': downstream stages are useless, unless inputs from upstream stages have been delivered. To avoid unenlightening complexity, we assume that, at each stage of the production process, if the two parties cannot find an agreement, both the firm and the supplier are capable of

[3]Independence comes from our assumption that the cost of protecting knowledge transmission does not vary with the firm's organizational choice. While we make this assumption in order to highlight the distinct role of sequential production in knowledge transmission, Appendix A1 analyzes the alternative case in which the organizational choice indeed affects the cost of knowledge protection in a two-stage setup.

producing a zero-value-added input at a zero marginal cost, which simply allows for the continuation of the production process but does not contribute to increase the value of final production.

At each stage of the value chain the firm faces the same hold-up problem described in the previous section, and has to protect knowledge transmission to avoid dissipation. In particular, at generic stage z the cost of protection resembles (3):

$$\kappa(\omega(z), \lambda) = \omega(z)\delta(z)^{\lambda}, \tag{9}$$

where knowledge intensity $\omega(z)$ is now allowed to vary across inputs.

The timing of events follows the same logic as before. First, the firm chooses the organizational form $\beta \in \{\beta_V, \beta_O\}$ and the amount of transmitted knowledge $\delta(z) \in [0, 1]$ for all stages $z \in [0, 1]$. Second, the firm posts a contract for the provision of each customized input z, stating the corresponding chosen organizational form and knowledge transmission. Third, for each stage z a large number of identical potential suppliers competitively bid for the corresponding contract and the firm selects one of them. Fourth, the selected suppliers decide how much to invest in their relationships with the firm, that is, how much to supply of their intermediate input $x(z)$. Fifth, the firm and each supplier z bargain on how to split their joint surplus consisting of the corresponding stage's marginal contribution $r'(z)$ to final revenues $r(z)$. Sixth and last, final production takes place, output is sold and revenues are shared among all value chain participants according to the agreed split rules. As the simple model, also this extended model has to be solved backwards, characterizing first the suppliers' decision on $x(z)$ and then the firm's decisions of $\beta(z)$ and $\delta(z)$.

3.1 Intermediate supplies

The choice of optimal investment $x(z)$ by a supplier at stage z mimics the supplier's decision in the previous section, the only exception being that now the joint surplus consists of stage z's *incremental* contribution to final revenues,

$$r'(z) = \frac{\rho}{\alpha} \left(A^{1-\rho}\theta^{\rho}\right)^{\frac{\alpha}{\rho}} \ r(z)^{\frac{\rho-\alpha}{\rho}} \left(\delta(z) \ x(z)\right)^{\alpha}, \tag{10}$$

which is the derivative with respect to z of revenues secured up to stage z by the investments of upstream suppliers,

$$r(z) = A^{1-\rho}\theta^{\rho} \left[\int_0^z \left(\delta(s) \ x(s)\right)^{\alpha} ds\right]^{\frac{\rho}{\alpha}}. \tag{11}$$

Expression (10) shows that supplier z's contribution can be either increasing or decreasing in the revenues $r(z)$ secured up to the stage z, depending on the elasticity of final demand (ρ) and the degree of complementarity between the different inputs (α). If $\rho > \alpha$ holds, $r'(z)$ is increasing in $r(z)$ so that higher investments by upstream suppliers raise the marginal return of supplier z's own investment. Following Antràs and Chor (2013) and Alfaro et al. (2019), we will refer to this case as 'sequential complementarity' given that more investment by upstream suppliers incentivizes investment by downstream suppliers. On the contrary, if $\rho < \alpha$ holds, more upstream investment disincentivizes investment by downstream suppliers. We will therefore refer to this second case as 'sequential substitutability'.

For given $\beta(z) \in \{\beta_V, \beta_O\}$ and $\delta(z) \in [0,1]$, the supplier then chooses $x(z)$ so as to maximize

$$\pi_S(z) = (1-\beta(z))\; r'(z) - c\; x(z). \tag{12}$$

Given (10) and (11), the supplier's incentive to provide the input increases with its share of surplus $(1-\beta(z))$, the extent of upstream protection and upstream production $(\int_0^z (\delta(s)x(s))^\alpha\, ds)$, and the amount of protection specific to its stage $(\delta(z))$. The supplier's profit-maximizing provision of input then evaluates to:[4]

$$x^*(z) = \Lambda \left(\frac{1}{c}\right)^{\frac{1}{1-\rho}} (1-\beta(z))^{\frac{1}{1-\alpha}}\, \delta(z)^{\frac{\alpha}{1-\alpha}} \left[\int_0^z [(1-\beta(s))\,\delta(s)]^{\frac{\alpha}{1-\alpha}}\, ds\right]^{\frac{\rho-\alpha}{\alpha(1-\rho)}} \tag{13}$$

$$\text{with} \;\; \Lambda \equiv A\; (\rho\theta^\rho)^{\frac{\rho}{1-\rho}} \left(\frac{1-\rho}{1-\alpha}\right)^{\frac{\rho-\alpha}{\alpha(1-\rho)}}.$$

3.2 Value chain organization

Turning to the firm, for all stages, the final producer chooses $\beta(z) \in \{\beta_V, \beta_O\}$ and $\delta(z) \in [0,1]$ so as to maximize profit

$$\pi_F = \int_0^1 [\beta(z)\; r'(z) - \kappa(\omega(z), \lambda)]\, dz \tag{14}$$

anticipating the optimal input provision of all its suppliers, $x^*(z)$ for $z \in [0,1]$. Given (10), (11) and (13)), the firm's profit (14) can be rewritten as

$$\pi_F = \mathcal{L}_F - \int_0^1 \omega(z)\delta(z)^\lambda dz, \tag{15}$$

where

$$\mathcal{L}_F \equiv \Theta\; c^{\frac{\rho}{1-\rho}} \int_0^1 \beta(z)\, [(1-\beta(z))\delta(z)]^{\frac{\alpha}{1-\alpha}} \left\{\int_0^z [(1-\beta(s))\delta(s)]^{\frac{\alpha}{1-\alpha}}\;\; ds\right\}^{\frac{\rho-\alpha}{\alpha(1-\rho)}} dz,$$

$$\text{with} \;\; \Theta \equiv \frac{\rho}{\alpha} A(\rho\theta)^{\frac{\rho}{1-\rho}} \left(\frac{1-\rho}{1-\alpha}\right)^{\frac{\rho-\alpha}{\alpha(1-\rho)}}.$$

[4]To obtain $x^*(z)$, we first express the first-order condition of the supplier's maximization problem in terms of $x(z)$ as a function $r(z)$, and then we plug the resulting expression into (10). This delivers a separable differential equation that can be solved for $r(z)$. The solution, substituted in the supplier's first-order condition, delivers (13).

3.2.1 Organizational choice for given knowledge transmission

As done before for the simple model, the result of the maximization of the firm's profit (15) can be characterized by initially neglecting the constraint $\beta(z) \in \{\beta_V, \beta_O\}$. In particular, without such constraint, the first-order condition with respect to $\beta(z)$ can be used to express the firm's optimally-chosen bargaining weight at stage z as

$$\beta^{+}(z) = 1 - \alpha\left(z + \Delta(z)\right)^{\frac{\alpha-\rho}{\alpha}}, \tag{16}$$

where

$$\Delta(z) \equiv z(1-z)\left(\frac{\frac{1}{z}\int_0^z \delta(s)^{\frac{\alpha}{1-\alpha}}ds}{\int_0^1 \delta(z)^{\frac{\alpha}{1-\alpha}}dz} - \frac{\frac{1}{1-z}\int_z^1 \delta(s)^{\frac{\alpha}{1-\alpha}}ds}{\int_0^1 \delta(z)^{\frac{\alpha}{1-\alpha}}dz}\right)$$

captures the differential in (weighted) average transmitted knowledge between stages located upstream and downstream of stage z. Accordingly, $\Delta(z)$ is an index of 'upstream knowledge transmission', which is positive when more knowledge is transmitted upstream, and negative when more knowledge is transmitted downstream.

Expression (16) shows that here, differently from the previous section, the firm's organizational choice for stage z is *not independent* from its decision on how much knowledge to transmit along the value chain. In particular, given $\rho \in (0,1)$ and $\alpha \in (0,1)$, (16) implies that the more knowledge is transmitted upstream of z in relative terms, the smaller is the firm's unconstrained optimal bargaining weight at stage z whenever suppliers' investments are sequential substitutes ($\rho < \alpha$). Differently, the more knowledge is transmitted upstream of z in relative terms, the larger is the firm's unconstrained optimal bargaining weight at stage z whenever suppliers' investments are sequential complements ($\rho > \alpha$). In other words, if relatively more knowledge is transmitted upstream of z and suppliers' investments are sequential substitutes, the firm is more likely to use outsourcing (smaller $\beta^{+}(z)$) at stage z, favoring supplier incentivization over rent extraction. The reason is that, with more upstream knowledge transmission, upstream suppliers contribute more to the firm's revenues and, with sequential substitutability, that reduces supplier z's return on investment (smaller $r'(z)$). If relatively more knowledge is transmitted upstream of z but suppliers' investments are instead sequential complements, the firm is more likely to use vertical integration (larger $\beta^{+}(z)$) at stage z, favoring rent extraction over supplier incentivization, as the contribution of upstream suppliers to the firm's revenues raises supplier z's return to investment (larger $r'(z)$).

Before characterizing knowledge transmission at the different stages, it is useful to contrast our model with that of Antràs and Chor (2013) and Alfaro et al. (2019). Ours embeds theirs in the special case of complete knowledge transmission at all stages: $\delta(z) = 1$ for all $z \in [0,1]$. In this case $\Delta(z) = 0$ holds and (16) boils down to $\beta^{+}(z) = 1 - \alpha z^{\frac{\alpha-\rho}{1-\alpha}}$. Accordingly, the firm's unconstrained optimal bargaining weight $\beta^{+}(z)$ is a decreasing function of input 'upstreamness' z with sequential complements ($\rho > \alpha$), while it is an increasing function of z with sequential substitutes ($\rho < \alpha$).

To map $\beta^+(z)$ into the binary choice between β_O and β_V, one can follow the same logic we used above for the two-stage value chain: stage z is necessarily integrated if $\beta_O < \beta_V < \beta^+(z)$ and outsourced if $\beta^+(z) < \beta_O < \beta_V$. Hence, given that with $\rho < \alpha$ the function $\beta^+(z)$ decreases with z, sufficient conditions for integrated and outsourced stages to coexist along the value chain under substitutability are $\beta^+(0) > \beta_V$ and $\beta^+(1) < \beta_O$. As for $\rho < \alpha$ we have $\lim_{z\to 0} \beta^+(0) = 1$ and $\beta^+(1) = 1 - \alpha$, the exact parameter condition is $1 - \alpha < \beta_O$. Differently, given that with $\rho > \alpha$ the function $\beta^+(z)$ increases with z, sufficient conditions for integrated and outsourced stages to coexist along the value chain under complementarity are $\beta^+(0) < \beta_O$ and $\beta^+(1) > \beta_V$. As for $\rho > \alpha$ we have $\lim_{z\to 0} \beta^+(0) = -\infty$ and $\beta^+(1) = 1 - \alpha$, the exact parameter condition is $1 - \alpha > \beta_V$.

A similar logic applies to the general case of $\Delta(z) \neq 0$ given that, just like z, also $z + \Delta(z) = \left(\int_0^z \delta(s)^{\frac{\alpha}{1-\alpha}} ds\right) / \left(\int_0^1 \delta(s)^{\frac{\alpha}{1-\alpha}} ds\right)$ is an increasing function of z. The only twist here is that we have $z+\Delta(z) > z$ when more knowledge is transmitted upstream, and $z+\Delta(z) < z$ when more knowledge is transmitted downstream.

The monotonicity of $z + \Delta(z)$ ensures that, analogously to Antràs and Chor (2013) and Alfaro et al. (2019), when mapping $\beta^+(z)$ into the binary choice between β_O and β_V, expression (16) implies that the decision on which stages to integrate or outsource obeys a cutoff rule. In the case of sequential complements ($\rho > \alpha$), there is a cutoff stage $z_C^* \in [0, 1]$ at which the firm is indifferent between the two organizational forms and such that all upstream stages are outsourced, while all downstream stages are integrated: $\beta(z) = \beta_O$ for $z \in [0, z_C^*]$ and $\beta(z) = \beta_V$ for $z \in (z_C^*, 1]$. This cutoff is implicitly determined by

$$z_C^* + \Delta(z_C^*) = H_C \tag{17}$$

$$\text{with} \quad H_C \equiv \left\{ 1 + \left(\frac{1-\beta_O}{1-\beta_V}\right)^{\frac{\alpha}{1-\alpha}} \left[\left(\frac{1 - \frac{\beta_O}{\beta_V}}{1 - \left(\frac{1-\beta_O}{1-\beta_V}\right)^{-\frac{\alpha}{1-\alpha}}} \right)^{\frac{\alpha(1-\rho)}{\rho-\alpha}} - 1 \right] \right\}^{-1}.$$

Differently, in the case of sequential substitutes ($\rho < \alpha$), the cutoff stage $z_S^* \in [0, 1]$ at which the firm is indifferent between the two organizational forms is such that all upstream stages are integrated, while all downstream stages are outsourced: $\beta(z) = \beta_V$ for $z \in [0, z_S^*)$ and $\beta(z) = \beta_O$ for $z \in [z_S^*, 1]$. This threshold is implicitly determined by

$$z_S^* + \Delta(z_S^*) = H_S \tag{18}$$

$$\text{with} \quad H_S \equiv \left\{ 1 + \left(\frac{1-\beta_V}{1-\beta_O}\right)^{\frac{\alpha}{1-\alpha}} \left[\left(\frac{\frac{\beta_V}{\beta_O} - 1}{\left(\frac{1-\beta_V}{1-\beta_O}\right)^{-\frac{\alpha}{1-\alpha}} - 1} \right)^{\frac{\alpha(1-\rho)}{\rho-\alpha}} - 1 \right] \right\}^{-1}.$$

Clearly, when knowledge transmission is complete at all stages ($\Delta(z) = 0$), both (17) and (18) boil down to the corresponding expressions in Antràs and Chor (2013) and Alfaro et al. (2019).

We can summarize these cutoff results in the following propositions.

Proposition 2 *When suppliers' investments are complements ($\rho > \alpha$), there exists a cutoff stage z_C^* such that all upstream stages are outsourced and all downstream stages are integrated.*

Proposition 3 *When suppliers' investments are substitutes ($\rho < \alpha$), there exists a cutoff stage z_S^* such that all upstream stages are integrated and all downstream stages are outsourced.*

3.2.2 Knowledge transmission for chosen organization

The cutoff rule guiding the decision on which stages to integrate or outsource allows us to decompose $\mathcal{L}_F$ in the profits generated by the outsourced stages and those generated by the integrated stages. Then, depending on whether we consider sequential substitutes or complements, we can use (17) or (18) to rewrite the firm's profit 15 as

$$\pi_F = \Theta \, \frac{\alpha(1-\rho)}{\rho(1-\alpha)} \, c^{-\frac{\rho}{1-\rho}} \, \Gamma\left(\beta_V, \beta_O\right) \left[\int_0^1 \delta(z)^{\frac{\alpha}{1-\alpha}} dz\right]^{\frac{\rho(1-\alpha)}{\alpha(1-\rho)}} - \int_0^1 \omega(z) \, \delta(z)^{\lambda} dz, \tag{19}$$

with $\Gamma\left(\beta_V, \beta_O\right) \equiv \Gamma_C\left(\beta_V, \beta_O\right)$ for $\rho > \alpha$, and $\Gamma\left(\beta_V, \beta_O\right) \equiv \Gamma_S\left(\beta_V, \beta_O\right)$ for $\rho < \alpha$.[5]

Optimal knowledge transmission $\delta^*(z)$ solves the first-order condition for the maximization of (19)) with respect to $\delta(z)$, which yields

$$\delta^*(z) = \Theta(\beta_V, \beta_O)\Omega\omega(z)^{-\frac{1}{\lambda - \frac{\alpha}{1-\alpha}}} \tag{20}$$

with

$$\Theta(\beta_V, \beta_O) \equiv \left(\frac{\alpha}{1-\alpha}\frac{\Theta}{\lambda}\left(\frac{1}{c}\right)^{\frac{\rho}{1-\rho}} \Gamma(\beta_V, \beta_O)\right)^{\frac{1}{\lambda - \frac{\alpha}{1-\alpha}}\left(1 - \frac{\alpha}{1-\alpha}\frac{\alpha-\rho}{(1-\alpha)(1-\rho)\left(\lambda - \frac{\rho}{1-\rho}\right)}\right)}$$

and

$$\Omega \equiv \left[\int_0^1 \left(\frac{1}{\omega(z)}\right)^{\frac{1}{\lambda - \frac{\alpha}{1-\alpha}}\frac{\alpha}{1-\alpha}} dz\right]^{-\frac{\alpha-\rho}{(1-\alpha)(\lambda(1-\rho)-\rho)}},$$

so that the implicit definitions (17) and (18) of the cutoffs can be restated as

$$z_C^* + z_C^*(1 - z_C^*)\Omega(z_C^*) = H_C \tag{21}$$

and

$$z_S^* + z_S^*(1 - z_S^*)\Omega(z_S^*) = H_S, \tag{22}$$

[5] The expressions of the two bundling parameters $\Gamma_C\left(\beta_V, \beta_O\right)$ and $\Gamma_S\left(\beta_V, \beta_O\right)$ are reported in Appendix A2 and are such that $\Gamma_C(\beta_V, \beta_O) = \Gamma_S(\beta_O, \beta_V)$ holds.

respectively, where

$$\Omega(z) \equiv \frac{\frac{1}{1-z}\int_z^1 \omega(s)^{-\frac{\frac{\alpha}{1-\alpha}}{\lambda-\frac{\alpha}{1-\alpha}}}ds}{\int_0^1 \omega(z)^{-\frac{\frac{\alpha}{1-\alpha}}{\lambda-\frac{\alpha}{1-\alpha}}}dz} - \frac{\frac{1}{z}\int_0^z \omega(s)^{-\frac{\frac{\alpha}{1-\alpha}}{\lambda-\frac{\alpha}{1-\alpha}}}ds}{\int_0^1 \omega(z)^{-\frac{\frac{\alpha}{1-\alpha}}{\lambda-\frac{\alpha}{1-\alpha}}}dz}$$

captures the differential in (weighted) average knowledge intensity between stages located upstream and downstream of stage z. Accordingly, $\Omega(z)$ can be interpreted as an index of 'upstream knowledge intensity', which is positive when upstream stages are more knowledge intensive than downstream ones, and negative when the opposite holds. When knowledge intensity is uniform across all stages (i.e. $\omega(z) = \omega > 0$ for all $z \in [0,1]$), transmitted knowledge is also the same (i.e. $\delta(z) = \delta > 0$ for all $z \in [0,1]$) so that we have both $\Omega(z) = 0$ and $\Delta(z) = 0$, which brings us back to the cutoff expressions in Antràs and Chor (2013) and Alfaro et al. (2019).

3.3 Comparative statics and empirical predictions

The model delivers clear-cut predictions on how IPR quality affects the organization of the value chain when knowledge intensity is a monotonic function of z. For concreteness, consider the specific functional forms $\omega(s) = e^{\omega s}$ and $\omega(s) = e^{\omega(1-s)}$ such that knowledge intensity rises and falls respectively with downstreamness, at the constant rate $\omega > 0$.[6] This rate then measures the 'relative knowledge intensity' of the part of the value chain more costly to protect from knowledge dissipation: downstream inputs for $\omega(s) = e^{\omega s}$ and upstream inputs for $\omega(s) = e^{\omega(1-s)}$. If we define the bundling parameter

$$\mu \equiv \frac{\omega\frac{\alpha}{1-\alpha}}{\lambda - \frac{\alpha}{1-\alpha}}, \tag{23}$$

under rising knowledge intensity $\omega(s) = e^{\omega s}$ we obtain

$$z_r^* = -\frac{1}{\mu}\ln\left(1 - \left(1 - e^{-\mu}\right) H_r\right) \tag{24}$$

for $r = \{C, S\}$ with $H_r \in (0,1)$ ensuring $z_r^* \in (0,1)$; differently, under falling knowledge intensity $\omega(s) = e^{\omega(1-s)}$ we get

$$z_f^* = 1 + \frac{1}{\mu}\ln\left(\left(1 - e^{-\mu}\right) H_f + e^{-\mu}\right) \tag{25}$$

for $f = \{C, S\}$ with $H_f \in (0,1)$ ensuring $z_f^* \in (0,1)$.

The cutoffs' expressions (24) and (25) are amenable to clear-cut comparative statics results that can be brought to data. In particular, with respect to IPR protection and knowledge intensity, (24) and (25) respectively imply

$$\frac{dz_r^*}{d\lambda} > 0,\ \frac{dz_r^*}{d\omega} < 0$$

[6]More details on the case $\omega(s) = e^{\omega s}$ can be found in Appendix A3, where we derive the firm's policy functions.

and

$$\frac{dz_f^*}{d\lambda} < 0, \; \frac{dz_f^*}{d\omega} > 0$$

so that we can state:

Proposition 4 *When inputs' knowledge intensity increases (decreases) downstream and suppliers' investments are complements ($\rho > \alpha$), the cutoff stage z_C^* is increasing (decreasing) in IPR quality (λ) and decreasing (increasing) in the relative knowledge intensity of downstream (upstream) inputs (ω).*

Proposition 5 *When inputs' knowledge intensity increases (decreases) downstream and suppliers' investments are substitutes ($\rho < \alpha$), the cutoff stage z_S^* is increasing (decreasing) in IPR quality (λ) and decreasing (increasing) in the relative knowledge intensity of downstream (upstream) inputs (ω).*

Moreover, given definition (23), the impact of λ on μ and therefore on the cutoffs is small for large α, which is more likely the case with substitutes than complements. Hence we can state:

Proposition 6 *The cutoff stages are less responsive to different levels of IPR quality when suppliers' investments are substitutes ($\rho < \alpha$) than when they are complements ($\rho > \alpha$).*

Together with Propositions 2 and 3, Propositions 4, 5 and 6 can be turned into empirical predictions on the probability of integrating the supply of any randomly selected input as follows. Consider some continuous distribution of inputs across stages z with c.d.f. $G(z)$ for $z \in [0, 1]$. Then, according to the model, the probability that a randomly picked input is integrated equals $1 - G(z_C^*)$ in the case of complements, and $G(z_S^*)$ in the case of substitutes. This implies that the probability of integration decreases with z_C^* in the former case, whereas it increases with z_S^* in the latter.

The empirical implications of our propositions can be thus summarized as follows.

(A) Based on Propositions 2 and 3, as in Antràs and Chor (2013) and Alfaro et al. (2019), the probability of integrating a randomly selected input increases (decreases) with its upstreamness along the value chain in the case of substitutability (complementarity).

(B) Based on Proposition 4, when inputs' knowledge intensity increases (decreases) downstream and suppliers' investments are complements, the probability of integrating a randomly selected input is decreasing (increasing) in IPR quality and increasing (decreasing) in the relative knowledge intensity of downstream inputs.

(C) Based on Proposition 5, when inputs' knowledge intensity increases (decreases) downstream and suppliers' investments are substitutes, the probability of integrating a randomly selected input is increasing (decreasing) in IPR quality and decreasing (increasing) in the relative knowledge intensity of downstream inputs.

(D) Based on Proposition 6, the impact of IPR quality on the probability of integrating a randomly selected input is stronger with sequential complementarity than substitutability.

Intuitively, when IPR quality is perfect, all transmitted knowledge is costlessly protected. Knowledge intensity is thus immaterial and our model coincides with the one by Antràs and Chor (2013): if suppliers' investments are complements, upstream stages are outsourced and downstream stages are integrated; if they are substitutes, the reverse pattern holds. When instead knowledge transmission is costly due to imperfect IPR quality, knowledge intensity matters.

Take, for instance, the case of complements when knowledge intensity increases with downstreamness. In the presence of complementarity, the chosen cutoff stage strikes the optimal balance between upstream supplier incentivization through outsourcing and downstream surplus extraction through integration. Going from perfect to imperfect IPR quality reduces the amount of knowledge transmitted, but especially downstream given that knowledge intensity increases with downstreamness. The implication is that lower IPR quality decreases the revenues generated by all stages, but especially those by the downstream ones. As it is therefore the vertically integrated part of the value chain that suffers more, the initial balance between upstream supplier incentivization and downstream surplus extraction is broken in favor of the former. Accordingly, to restore the optimal balance, the firm has to start integrating more upstream. This explains why the cutoff stage z_C^* moves towards $z = 0$, with an increased measure of integrated stages. Vice versa, an improvement in IPR quality, shifts the cutoff stage in the opposite direction, thus implying a decreased probability of integrating a randomly selected input.

The pattern is reversed in the case of substitutes when knowledge intensity again increases with downstreamness. In the presence of substitutability, the chosen cutoff stage strikes the optimal balance between upstream surplus extraction through integration and downstream supplier incentivization through outsourcing. Going from perfect to imperfect IPR quality, less knowledge is again transmitted, especially downstream. Hence, the optimal balance between upstream surplus extraction and downstream supplier incentivization is broken in favor of the former. The firm restores the optimal balance by starting to outsource more upstream. As a result, the cutoff stage z_S^* moves towards $z = 0$, with an increased measure of outsourced stages. Vice versa, an improvement in IPR quality shifts the cutoff in the opposite direction, thus implying an increased probability of integrating a randomly selected input.

Analogously, higher relative downstream knowledge intensity reinforces relative knowledge transmission upstream, resulting in a higher probability of integration of a randomly selected input when supplier investments are complements, and of outsourcing when they are substitutes.

4 Data and Key Variables

The dataset we use is composed of four distinct databases covering the population of Slovenian firms in the 2007-2010 period. Our core database includes transaction-level trade data at the 8-digit level of the European Combined Nomenclature (hereinafter CN) classification provided by the Statistical Office of the Republic of Slovenia (SURS). Using the unique firm identifiers, this transaction-level trade database is merged with (i) detailed information on the direction of firms' cross-border foreign direct investment (FDI) outflows provided by the Bank of Slovenia and (ii) firms' financial statements data from the Agency for Public Legal Records and Related Services (APLR). Hence, we have at our disposal firms' annual export and import transactions to/from partner countries as well as their outward FDI positions in the respective host partner countries. Additionally, we use a database on the performance of the foreign affiliates of Slovenian firms provided by the Bank of Slovenia, which contains further information on affiliates' performance, core industry of activity and trade flows, such as total exports and imports of affiliates, their total intra-firm trade and sales in the local (host) market. In our final sample, we have 5241 firms sourcing from 61 different partner countries.

Slovenian data are particularly well suited for studying firm organization behavior along international value chains. Slovenia is a small, highly open economy from the group of Central and Eastern European transition economies that has been heavily involved in both multilateral liberalization and regional integration processes since the mid-1990s. This involvement has been mostly related to approaching EU membership through: (i) accession to the GATT (WTO) in 1994 (1995); (ii) CEFTA membership in 1996; (iii) signing of an Association Agreement with the EU in 1996 with provisional enforcement in 1997; and (iv) EU accession negotiations between 1998-2002. In year 2004, Slovenia became a full member of the EU and adopted the Euro in 2007 as the first new EU member state. Liberalization processes contributed to the increasing involvement of Slovenian companies in global value chains (GVC). According to the WTO, Slovenia is classified among the high GVC participation economies. It recorded a GVC participation index of 58.7 in 2011 that is significantly above the average value for developed and developing countries (48.6 and 48.0 respectively). The index is high mostly on account of strong backward participation (WTO, 2016) as shown in Table 1, which is the type of participation our model is about. Figures 1 and 2 also show the value-added components of gross exports for Slovenia in 1995 and in 2011, together with the comparison between inward and outward FDIs. It is clear from Figure 2 that the strongest steady increase in Slovenian outward FDI stock has been recorded between 1999 and 2007, with the peak value in 2009, when also the gap between inward and outward FDIs has been the smallest.

4.1 Dependent variable: binary variable on the decision to integrate

Our dependent variable is a firm's propensity to transact an input in a particular source country within its boundaries. It is the outcome of the firm's binary decision on whether to integrate or outsource the supply of the input from a given country. We define inputs at the 6-digit level product groups of the CN classification, which is in full compliance with the 6-digit Harmonized system (hereinafter HS) code. Transaction-level trade data provide us with information on the complete set of inputs sourced from abroad by a firm, while FDI data give the location of its dependent establishments. However, as most related studies, we do not have information on the extent to which the firm's trade flows involve its dependent establishments ('intra-firm trade').

Antràs and Chor (2013) tackle this issue by exploiting available industry-level intra-firm trade data and using the share of intra-firm imports in total inputs as an indication of the propensity to transact a particular input within firm boundaries. The follow-up study by Alfaro et al. (2019) proposes an alternative solution based on the activities of establishments linked via ownership ties (net of subsidiaries of the 'global ultimate owner'). While the former approach lacks information on the identity (activity) of the individual buyer, the latter does not use trade data and relies instead on input-output ('I-O') tables to determine the sets of integrated and outsourced inputs without information on their source countries. We build on the latter approach in defining as traded 'intra-firm' or 'integrated' the inputs a parent firm imports from an affiliate's host country that are classified under the core activity of the affiliate, but we also exploit our detailed data to obtain the whole set of import transactions from different source countries. More specifically, inputs that a firm imports from its affiliate's host country, if classified under the core activity of the affiliate at the 4-digit industry level, are regarded as 'integrated', whereas all other inputs that the firm imports from that country are considered as 'outsourced'. Doing this also accounts for the fact that a firm may engage in both integration and outsourcing in a given country. If a firm has no FDI in a country, all imports coming from that country are regarded as 'outsourced'. This allows us to estimate the regression model at the most disaggregated firm-input-country level. As we will see, it will also allow us to consider a firm-input specific upstreamness measure for all bilateral transactions along a firm's sequential production.

We link the core activity of an affiliate and imported inputs by the parent company by first adopting the RAMON concordance from 6-digit HS 2002 to 6-digit CPA 2002 classification, and subsequently from CPA 2002 to NACE Rev. 1 at the 4-digit level based on the direct linkage in the structure of these two classifications.[7] In year 2007, the HS classification underwent a substantial revision, therefore a pairing of HS6 2007 to HS6 2002 codes is required for the purpose of linking the core activity of an affiliate with imported inputs. In converting HS 2007 to HS 2002 codes we lean on the concordance approach of Van Beveren, Bernard and Vandenbussche (2012), but assign one single code of the HS 2002 edition to each HS 2007 code. This requires certain simplifications in the event that the HS 2007 code is the result of either merging ($1 : n$ relationship) or splitting and merging ($n : n$ relationship) several codes in the previous 2002 classification. In this case, we follow the United Nations Statistics Division (2009) and give priority to the one subheading among several that has the same code as the HS 2007 subheading (if it exists). The retained code rule is based on the general practice of the World Customs Organization to maintain the existing code only if there

[7]For manufactured goods the elements of the CPA product classification are based on the HS classification.

have been no substantial changes of its scope.

We use d_integr_{ihkjt} to denote our dependent variable associated with a firm i that, in order to produce its core product k, sources input h from country j in period t. The dependent variable takes value 1 if in period t input h for product k is sourced by firm i from a country j where the firm has an affiliate whose core activity belongs to the same 4-digit industry as the input. It takes value 0 otherwise. For robustness check, we also define an alternative dependent variable based on a stricter definition of 'integrated' input that exploits information we have on whether, in a particular country-year, the affiliate reports positive intra-firm exports. This alternative dependent variable ($d\_integr\_IFEX_{ihkjt}$) takes value 1 if two conditions are fulfilled: (i) input sourced from the affiliate's host country is classified under the core activity of the affiliate at the 4-digit industry level (as for d_integr_{ihkjt}), and (ii) the affiliate reports positive intra-firm exports in a given year. It takes value 0 otherwise.

4.2 Sequential complementarity/substitutability

To distinguish between sequential substitutes and complements we first follow Antràs and Chor (2013) and Alfaro et al. (2019) and trace substitutes/complements based on low/high value of import demand elasticity faced by the buyers of a particular good. We consider the import demand elasticity of a firm i's 'core' export product, that is, the product at the 6-digit level of the HS classification that accounts for the largest share of the firm's exports. As stressed by Antràs and Chor (2013), this approach implies the assumption that any existing cross-industry variation in the degree of technological substitution across a firm's inputs (α_i) is largely uncorrelated with the demand of its core product (ρ_i). Complements ($d\_compl_{it} = 1$) are characterized by above-median import demand elasticity for a firm's core export product, whereas substitutes ($d\_compl_{it} = 0$) by below-median demand elasticity. We use the import demand elasticity estimated at the 6-digit HS product level for Slovenia by Kee, Nicita and Olarreaga (2008) following the production-based GDP function approach. This estimate is defined as the percentage change in the quantity of an imported good when its price increases by 1%, holding the prices of all other goods, as well as the productivity and the endowments of the economy, constant.

For robustness check, we complement the substitutes/complements measure of Antràs and Chor (2013) and Alfaro et al. (2019) in two ways. First, we propose a proxy for the parameter α_i, based on the presumption that the degree of technological input substitutability should be closely related to the degree of input differentiation. In particular, we presume is that inputs classified within the same industry at a given digit-level of classification exhibit higher technological substitutability compared with inputs classified in different industries at the same level of aggregation. We then compute a Herfindahl index (H_i) that measures how 6-digit imported inputs by firm i are dispersed across 3-digit industries:

$$H_i = 1 - \sum_{n=1}^{N3dig_i} \left(\frac{N6dig_{ni}}{N6dig_i} \right)^2 , \tag{26}$$

where n indexes a 3-digit HS product category while $N3dig_i$ and $N6dig_i$ refer to the numbers of 3-digit and 6-digit HS product categories involving inputs imported by firm i. When all imported inputs belong to the same 3-digit HS category n, we have $N6dig_{ni} = N6dig_i$ and thus $H_i = 0$. In contrast, when each input is classified under a different 3-digit HS category, we have $N6dig_{ni} = 1$, $N6dig_i = N3dig_i$ and thus $H_i = 1 - (1/N3dig_i)$. As H_i increases with the dispersion of 6-digit HS inputs across 3-digit product categories, we take it as an inverse measure of the technological substitutability of the firm's inputs α_i. Finally, we compute average values of the Herfindahl index across 3-digit industries to obtain industry-level inverse measures of technological substitutability. Complements and substitutes are then distinguished by considering both the estimated import demand elasticity of the core product and the industry-level average technological substitutability. Specifically, after taking the product of a firm's core import elasticity (in absolute value) and the industry-level Herfindahl index, we define a dummy variable $d\_compl_{rho \times alpha(ind.)}$ that equals 1 when the product is above the median, and zero otherwise. The underlying logic is that the higher the estimated import demand elasticity (in absolute value) and the higher the Herfindahl index, the more likely it is that $\rho > \alpha$ holds and investment along the value chain are sequential complements.

Second, building on Alfaro et al. (2019), we consider that α should be closely related to the elasticity of demand for each intermediate input in any given industry. Hence, we introduce another measure of sequential substitutability defined as the weighted average of estimated demand elasticities for a firm's intermediate and capital good imports, with weights given by the firm's import shares. We take the difference between the firm's core product import elasticity and the weighted average of its intermediate and capital good import elasticities. We then define a dummy variable $d\_compl_{rho - alpha(elast.)}$ that equals 1 when the difference is larger than 0 (sequential complementarity), and equals 0 otherwise (sequential substitutability).

In the robustness check Table 8 we will discuss later on, the columns corresponding to the three alternative dummies are labeled by rho, $rho \times alpha(ind.)$ and $rho - alpha(elast.)$ respectively.

4.3 Upstreamness/downstreamness

Since we observe import transactions at the firm-level, we are able to identify the position of each imported input h along the value chain of any given product k. This allows us to follow Alfaro et al. (2019) who define the upstreamness of input h in producing final output k as the weighted average of the number of stages it takes for h to enter (directly and indirectly) in k's final production:

$$Upstr_{hk} = \frac{d_{hk} + 2\sum_{m=1}^{M} d_{hm} d_{mk} + 3\sum_{m=1}^{M}\sum_{n=1}^{M} d_{hm} d_{mn} d_{nk} + ...}{d_{hk} + \sum_{m=1}^{M} d_{hm} d_{mk} + \sum_{m=1}^{M}\sum_{n=1}^{M} d_{hm} d_{mn} d_{nk} + ...} , \tag{27}$$

where d_{hk} denotes the direct requirement coefficient of input h in output k with $h, k = 1, ..., M$. In (27) the denominator is the sum of input h's requirement coefficients that enter product k's value chain l stages away from final production for $l = 1, 2, ..., \infty$. The numerator is also an infinite sum, but there each term is multiplied by an integer that corresponds to the number of stages upstream of k's final production at which input h enters the value chain. A larger value of $Upstr_{hk}$ (which is always greater than 1 by construction) means that a larger share of the total use value of input h is accrued further upstream in the production process of product k.[8]

Based on (27), we compute the upstreamness of each input h imported by firm i from each source country j for the final production of its core product k in year t, which we call $Upstr_{ihkjt}$. In doing so, we take the direct requirement coefficients from the 2002 US Input-Output tables provided by the Industry Benchmark Division (IBD) of the Bureau of Economic Analysis, since such detailed tables are not available for Slovenia. US SIC/NAICS product classes and industries from the US Direct Requirements matrix are matched to HS codes of firms' core export products and imported inputs based on concordance from Pierce and Schott (2009) (available at: http://www.nber.org/data-appendix/w15548/readme.txt).

4.4 Knowledge transmission

According to our model, the cost of protecting knowledge transmission is a function of two key variables: the knowledge intensity of inputs and the quality of IPR institutions in the location of production. We measure IPR quality as the logarithm of the Park (2008) index in source country j, which is a widely used proxy for patent protection in the IPR literature.[9] The variable is denoted as $lnIPR_{jt}$. In turn, we measure knowledge intensity by grouping inputs into products that are or are not knowledge intensive, based on their R&D intensity. In doing so, we adopt the Eurostat classification that, in line with the OECD, defines high-tech products as those featuring high levels of R&D expenditure over total sales.[10]

The groups classified as high-tech are aggregated on the basis of the Standard International Trade Classification (SITC) at 3-digit to 5-digit level, which we further translate to the HS classification codes that we use in our dataset. To trace the knowledge intensity of inputs along the firm's value chain, we use the upstreamness measure $Upstr_{ihkjt}$ described above. In particular, we compute the ratio $rel\_upst\_knint_k$ as the *average upstreamness of knowledge intensive inputs* relative to that of non-intensive inputs in the production of final product k, where the set of inputs used in k's production is identified based on the 2002 US Input-Output table.

[8]The fact that $Upstr_{hk}$ is specific to an input-product pair makes it different from the earlier measure of upstreamness used by Fally (2012) and Antràs et al. (2012).

[9]See e.g. Maskus (2000, 2012).

[10]A detailed list of high-tech product groups as classified by the Eurostat is provided at the following link: https://ec.europa.eu/eurostat/cache/metadata/Annexes/hte_esms_an4.pdf. Further classification details can be found at the url: https://ec.europa.eu/eurostat/ cache/metadata/en/htec_esms.htm.

The use of $rel\_upst\_knint_k$ is twofold. First, we use it to proxy the difference in knowledge intensity (or, equivalently, relative knowledge transmission) between upstream and downstream stages, which corresponds to ω in our theoretical model. In this respect, $rel\_upst\_knint_k$ measures the relative knowledge intensity of upstream stages when the upstream part of the value chain is more knowledge intensive (i.e. when knowledge intensity decreases downstream); vice versa, its inverse value $(rel\_upst\_knint_k)^{-1}$ measures the relative knowledge intensity of downstream stages when the downstream part of the value chain is more knowledge intensive (i.e. when knowledge intensity increases downstream). Second, we use $rel\_upst\_knint_k$ also to discriminate between industries with increasing and decreasing knowledge intensity as sequential production moves downstream. For this purpose, we introduce a dummy variable $d\_knint\_downstr_k$, which denotes that, along the value chain of final product k, knowledge intensive inputs tend to be located more downstream. This dummy takes the value 1 if the average upstreamness of knowledge intensive inputs is lower than the average upstreamness of inputs not intensive in knowledge. It takes value 0 otherwise.

4.5 Descriptive statistics

For concreteness, Figure 3 shows the variation of firms' core product groups in terms of their import demand elasticity rho_k and the relative upstreamness of their knowledge intensive inputs ($rel\_upst\_knint_k$). The latter is larger (smaller) than 1 when knowledge intensive (non-intensive) inputs tend to be located upstream along the value chain. The figure reveals no clear pattern of covariation between complements/substitutes and knowledge intensity increasing/decreasing downstream (or the relative knowledge intensity of upstream/downstream stages). It thus suggests that the parameters regulating sequential complementarity/substitutability and knowledge intensity along the value chain are indeed independent as assumed in the model.

Table 2 reports descriptive statistics for the pooled sample, and for the four subsamples where we distinguish between complements and substitutes based on d_compl_i, and between industries with input knowledge intensity increasing and decreasing with downstreamness based on $d\_knint\_downstr_k$.

Around 18% of import transactions are carried out by firms that report outward FDI activity in at least one year, throughout the 2007-2010 period (see d_OutFDI), and about 3% of transactions by firms with outward FDI in a particular sourcing country in a given year ($d\_OutFDI\_bilateral$). Among complements, both FDI shares are higher for firms with higher relative knowledge intensity of upstream inputs. Among substitutes, we observe the opposite, with higher FDI shares recorded among the firms characterized by higher relative knowledge intensity of downstream inputs. However, less than 0.1% of import transactions are regarded as integrated when the condition of being classified under the core activity of the affiliate at the 4-digit industry level is applied (see d_integr). The percentage is slightly less when the additional condition of the existence of positive intra-firm exports by affiliates is accounted for (d_integr_IFEX).

The incidence of input integration is higher for industries characterized by higher relative knowledge intensity of downstream inputs and the more so for substitutes. There are no notable differences observed between complements and substitutes or upstream and downstream relative knowledge intensity with respect to the average upstreamness of their inputs. Yet, firms operating in industries with higher relative knowledge intensity downstream tend to source, on average, from countries with better IPR institutions and rule of law implementation, both for complements and substitutes.

The four groups of firms are as well alike in terms of inputs' demand elasticity and industry Herfindahl index, which is in agreement with the presumption that the cross-industry variation in the degree of technological substitution across firms' inputs (α) is largely uncorrelated with the elasticity of demand (ρ), again as assumed in the theoretical model. The four groups are further similar, on average, in terms of their age, export propensity and financial leverage. However, firms with their core export product characterized by sequential complementarity are, on average, smaller in terms of number of employees and feature lower average capital intensity and slightly lower labor productivity. The least capital-intensive production process with the lowest average labor productivity is evidenced for sequential complements with high relative knowledge intensity of upstream inputs.

5 Empirical Specifications and Results

5.1 Empirical model specifications and methodological issues

Our database allows us to explore integration versus outsourcing decisions made not just across different inputs at the firm level but also across different input sourcing countries. As already discussed in Section 4.1, the dependent variable in our specifications is a binary indicator (d_integr_{ihkjt}) reporting whether or not in year t firm i with core export product k imports input h from source country j within the firm's boundaries. It is this source-country dimension that will distinguish our specifications from Alfaro et al. (2019) and will allow us to test our model's predictions, summarized in Propositions 4-6, on how country-specific IPR quality affects a firm's organization decision.

Specifically, we augment the empirical model of Antràs and Chor (2013) with the knowledge intensity of inputs along the value chain and the quality of IPR institutions in the sourcing countries. We test our predictions by means of three specifications. Specification (I) reads:

$$\begin{aligned}
&Pr(d_integr_{ihkjt} = 1) = \beta_0 + \beta_1\ Upstr_{ihkjt} + \beta_2\ d_compl_i + \beta_3\ lnIPR_{jt} + \\
&+\beta_4\ Upstr_{ihkjt} * d_compl_i + \beta_5\ lnIPR_{jt} * d_compl_i + \beta_6\ lnIPR_{jt} * Upstr_{ihkjt} + \\
&+\beta_7\ lnIPR_{jt} * d_compl_i * Upstr_{ihkjt} + \beta_8\ d_knint_downstr_k + \\
&+\beta_9\ d_knint_downstr_k * d_compl_i + X'_{it}\beta_{10} + \sum \beta_{11,k}\ d_industry_k + \\
&+\sum \beta_{12,j}\ d_country_j + \sum \beta_{13,t}\ d_year_t + u_{ihkjt}\ , \qquad (28)
\end{aligned}$$

and applies to the pooled sample. In the expression above, $Pr(d\_integr_{ihkjt} = 1)$ is the probability that firm i producing product k integrates input h imported from country j in year t. Besides the explanatory variables described in the previous section, specification (I) includes a vector X_{it} of standard firm-specific controls: age, size, capital intensity of production, labor productivity, export propensity and financial leverage. Size ($size_{it}$) is measured by the number of employees. Age (age_{it}) refers to years passed since the year of foundation reported in the Business Register of the Republic of Slovenia. Capital-intensity ($Kintensity_{it}$) is measured by fixed assets per worker, which according to Olley and Pakes (1996) affect the distribution of future plant productivity and may act as a proxy for unobserved sources of efficiency. Labor productivity ($Lproductivity_{it}$) is defined as value added per employee. Export propensity ($Ex\_Propensity_{it}$) is measured by the share of exports in total sales, while financial leverage as debt-to-assets ratio ($Debt\_assets_{it}$). We also include sets of (i) annual dummy variables to control for macroeconomic shocks; (ii) partner country dummies to account for country-specific time-invariant effects; and (iii) industry-specific effects, where a firm's industry affiliation is based on its core export product at the 1-digit level of the HS classification.

For specification (II) we split our sample between sequential complements and substitutes so as to avoid the complexity of interpreting the triple interaction $lnIPR_{jt} * d\_compl_i * Upstr_{ihkjt}$. After splitting the sample based on our alternative definitions of complements/substitutes, specification (II) reads:

$$
\begin{aligned}
&Pr(d\_integr_{ihkjt} = 1) = \beta_0 + \beta_1\ Upstr_{ihkjt} + \beta_2\ lnIPR_{jt} + \beta_3\ lnIPR_{jt} * Upstr_{ihkjt} + \\
&+\beta_4\ d\_knint\_downstr_k + X'_{it}\ \beta_5 + \sum\ \beta_{6,k}\ d\_industry_k + \\
&+\sum \beta_{7,j}\ d\_country_j + \sum \beta_{8,t}\ d\_year_t + u_{ihkjt}\ .
\end{aligned}
$$

To make further comparisons between different types of industries and institutions, we zoom in specification (II) and split the complements subsample, for which intangible assets play a crucial role according to both our theory and the evidence provided. In particular, we split the final products depending on whether the knowledge intensity of inputs along their value chains increases or decreases with downstreamness. This also allows us to augment the specification with the rate at which knowledge intensity increases or decreases downstream ($rel\_upst\_knint_k$). The resulting specification (III) reads:

$$
\begin{aligned}
&Pr(d\_integr_{ihkjt} = 1) = \beta_0 + \beta_1\ Upstr_{ihkjt} + \beta_2\ lnIPR_{jt} + \beta_3\ lnIPR_{jt} * Upstr_{ihkjt} + \\
&+\beta_4\ rel\_upst\_knint_k + X'_{it}\ \beta_5 + \sum\ \beta_{6,k}\ d\_industry_k + \\
&+\sum \beta_{7,j}\ d\_country_j + \sum \beta_{8,t}\ d\_year_t + u_{ihkjt}\ .
\end{aligned}
$$

All three specifications are estimated by probit, for which some remarks are in order. First, in line with heterogeneous firm dynamics models, the variability of firm growth usually decreases with firm size, raising the concern that variance is not constant across firms. This could also hold for firms' integration decisions. Therefore, we test whether firm size affects the conditional variance of the firm's integration choice to detect potential heteroscedasticity. When Wald's test for heteroscedasticity rejects the null hypothesis of homoscedastic variance, we implement a maximum-likelihood heteroscedastic probit model that generalizes the standard probit model by allowing the scale of the inverse link function to vary from observation to observation as a function of firm size. Second, to deal with potential endogeneity caused by unobserved firm-specific effects, we employ a parameterization of unobserved firm-specific effects by firm-level means of all time-varying independent variables over the sample period, as suggested by Mundlak (1978), Chamberlain (1984) and Wooldridge (2002). Eventually, we opt for a random effects probit model in order to explicitly exploit the panel structure of our data, thereby controlling for firm-country-product fixed effects as random variables uncorrelated with the regressors.

5.2 Empirical results

5.2.1 Pooled sample results

Starting with specification (I), Table 3 depicts the results for the baseline case, where complements and substitutes are defined based on the estimated import demand elasticity (rho). Column (1) of the table shows the results of the probit model with robust standard errors adjusted for firm clusters, whereas column (2) refers to the specification that includes firm-level means of all time-varying independent variables over the sample period to control for unobserved firm-specific effects. As Wald's test fails to reject the null hypothesis of homoscedastic variance, ordinary pooled probit results are reported. Column (3) instead reports the results estimated by the random effects probit model controlling for unobserved heterogeneity for each firm-country-product that is invariant over time. Column (4) adds industry and country dummies to the random effect probit estimation. The likelihood-ratio test confirms the importance of unobserved heterogeneity ('frailty') in these specifications. For this reason we will report only the random effects probit model results in subsequent tables.

A significant negative interaction between sequential complementarity and upstreamness is present throughout all columns of Table 3. This lends support to the Antràs and Chor's (2013) prediction inherited by our model that the likelihood of integration decreases when moving upstream along the production chain for sequential complements, and downstream for sequential substitutes.

In line with our model, there are also significant differences between complements and substitutes as regards the impact of IPR quality and upstreamness. This is revealed by the significant interaction of $lnIPR$ with the dummy variable for complements *d_compl* on the one hand, and with both *d_compl* and upstreamness $Upstr$ on the other. The interaction of $lnIPR$ with complementarity is negative and highly significant, suggesting that better IPR institutions, on average, encourage outsourcing when inputs are complements compared to when they are substitutes.

The positive and significant triple interaction term, in turn, shows that this feature is less likely at the upstream stages, hence occurring more at the downstream stages of sequential production. Finally, integration is more likely when knowledge intensity is increasing with downstreamness ($d\_knint\_downstr = 1$). This is so for both complements and substitutes as indicated by the insignificant interaction term between the dummies $d\_knint\_downstr$ and d_compl.

5.2.2 Split-sample results

Table 4 reports the results for specification (II), in which the sample is split into sequential complements and substitutes. This allows us to see more directly that the coefficients associated with IPR quality are only relevant in the case of complements. This is consistent with the prediction of Proposition 6 that the impact of IPR quality should be weaker for substitutes.

The significantly negative coefficient of $lnIPR$ in columns (1-3) again suggests that IPR quality tends to reduce a firm's propensity to integrate, especially for relatively downstream stages, as denoted by the positive and significant coefficient of the interaction between $lnIPR$ and $Upstr$. As we will see below for the double-split sample, this finding can also be interpreted as the impact of IPR quality being stronger for more knowledge-intensive inputs. Recall first that, with sequential complementarity, outsourcing (integration) takes place upstream (downstream). When the downstream side of production is relatively more knowledge intensive, better IPR quality leads to more outsourcing *downstream*; when instead the upstream side of production is relatively more knowledge intensive, better IPR quality could even lead to more integration upstream. The results survive the demanding introduction of country dummies into the random effect probit model in column (3), with reduced significance of the overall effect of IPR quality, but still a strongly significant association with upstreamness, thus highlighting the importance of IPR quality for organizational decisions when production is sequential.

The impact of upstreamness on integration differs for complements and substitutes, as confirmed by the Chow test of equality of regression coefficients between the two groups. In particular, in line with the aggregate sample results, the impact of upstreamness is significantly negative for complements. Integration is more likely when knowledge intensity is increasing downstream, with the effect being more robust and of larger magnitude in the case of substitutes.

In columns (7-8) we replace IPR quality with a measure of 'rule of law' from the Worldwide Governance Indicators (2015) database as a proxy for overall contract enforcement. The results clearly show that contract enforcement has the opposite effect with respect to IPR quality on the integration decision. This is in line with with the property rights model of Antràs and Chor (2013) as the impact of better contract enforcement is significantly negative for substitutes and positive for complements with a significant negative interaction of contract enforcement and upstreamness: better contract enforcement increases the prevalence of integration over outsourcing. This divergence suggests that better institutions may have opposite effects on firm organization depending on whether they improve the protection of tangible or intangible assets. It also shows that our findings in the previous columns of Table 4 are specific to IPR institutions and cannot be generalized to other regulatory measures that directly affect contract enforcement.

As for the firm-specific controls, the results in both specifications (I) and (II) indicate that larger and older firms with higher export propensity are more likely to integrate inputs both for complements and substitutes.[11] Differently, Table 4 shows that capital intensity has opposite impact on integration for complements and substitutes: positive for former and negative for the latter. On the other hand, the impact of labor productivity is significantly negative for substitutes and mostly insignificant or weakly significant for complements. The heterogeneous effects of capital intensity and labor productivity explain why their estimated impacts are less consistent across the columns of Table 3, where complements and substitutes are pooled together. Finally, financial leverage has a negative effect on integration for both complements and substitutes.

5.2.3 Double-split sample zoom in the complements case

We saw in Table 4 that better IPR quality in the source country increases the propensity to outsource when inputs are sequential complements. However, the positive interaction between IPR quality and upstreamness conveyed the message that this is not always the case and the effect prominently takes place downstream. It may also mean that firms would become more inclined toward integration with improved IPR quality along stages on the upstream part of the value chain.

Our theory predicts that this distinction depends on whether input knowledge intensity is increasing or decreasing as sequential production moves downstream. To investigate this aspect, in specification (III) we further split the sample accordingly. In particular, we focus on the complements case as Proposition 6 and its supporting results from Table 4 suggest that cutoff stages are particularly responsive to IPR quality in that case. We also augment the specification with the relative knowledge intensity of upstream inputs (rel_upst_knint) when input knowledge intensity decreases downstream and, alternatively, with the relative knowledge intensity of downstream inputs ($(rel\_upst\_knint)^{-1}$) when input knowledge intensity increases downstream. These additional variables correspond to parameter ω in Proposition 4 and also capture the rate at which knowledge intensity is decreasing and increasing with downstreamness respectively.

Table 5 reports the double-split sample results of specification (III) for complements. In Col-

[11] The only exception appears in column (2) of Table 3, where the effect of export propensity is absorbed by period averages introduced in the wake of Mundlak (1978), Chamberlain (1984) and Wooldridge (2002).

umn (1) input knowledge intensity increases downstream ($d\_knint\_downstr = 1$) and this happens at rate $(rel\_upst\_knint)^{-1}$. In Column (2) input knowledge intensity decreases downstream ($d\_knint\_downstr = 0$) and this happens at rate (rel_upst_knint). In line with our theoretical predictions in Proposition 4, the table shows a significantly positive impact of $(rel\_upst\_knint)^{-1}$ on the likelihood of integration in the former case, and a significantly negative impact of (rel_upst_knint) in the latter case.

As for the impact of IPR quality $lnIPR$ on integration, it remains significantly negative for complements no matter whether knowledge intensity is increasing or decreasing downstream. However, the coefficient of the interaction term between $lnIPR$ and $Upstr$ is significant only when input knowledge intensity decreases downstream. In this case it is positive and the implied indirect effect of $lnIPR$ through the interaction dominates its negative direct effect for upstream stages. Accordingly, in line with Proposition 4, better IPR quality extends the set of integrated stages towards the upstream part of the value chain.

To better visualize the impact of $lnIPR$ and help us interpret its interaction with $Upstr$, Figure 4 graphically represents the estimated marginal effects reported in Table 6 relative to columns (1) and (2) of Table 5.[12] The figure plots the average marginal effects of an increase in IPR quality on the probability of integration at different stages along the supply chain when knowledge intensity increases downstream (left panel) or decreases downstream (right panel). In line with our theoretical predictions, the left panel shows a strong negative impact on the propensity to integrate for inputs that enter the value chain downstream, while the right panel shows a strong positive impact for those that enter upstream.

The double-split sample can also be exploited to compare the impact of IPR institutions with that of more general contracting institutions (rule of law). This is done in column (3) focusing on the case of sequential complements and knowledge intensity increasing downstream. The column reports an effect of rule of law that is in stark contrast to that of IPR quality in column (1) as better contract enforcement increases the propensity to integrate. Figure 5 plots the estimated average marginal effects of IPR quality versus those of rule of law reported in Table 6, based on columns (1) and (3) in Table 5 respectively. The former are both stronger and opposite to those of rule of law. This confirms the insight from specification (I) that better institutions have opposite effects on firm organization depending on whether they improve the protection of tangible or intangible assets.

5.2.4 Robustness Checks

In this section we report the results of several robustness tests of our findings to alternative variable definitions and sample restrictions. In doing so, we build on the single-split specification (II) to

[12]Regression coefficients in probit models cannot be interpreted as simple slopes as in ordinary linear regressions, but have to be interpreted in terms of Z-scores (i.e. as changes in Z-score for one unit increase in the explanatory variable).

consider all available observations and guarantee a sufficient number of observations for the different sample restrictions.

First, Table 7 reports the results obtained from modifications of specification (II) aimed at ensuring a vertical-type connection between a firm's imported inputs and its core export product. In particular, columns (1) and (2) report the results when we restrict the sample to import transactions that are classified as intermediates or capital goods according to the Broad Economic Categories classification; columns (3) and (4) show the results when we use our alternative dependent variable ($d\_integr\_IFEX_{ihjt}$), which conditions the classification of transactions as intra-firm also on the existence of a firm's affiliate in the source country declaring intra-firm export activities.

The results in Table 7 confirms those in Table 4: better IPR quality diminishes the propensity to integrate in relatively downstream stages for complements, while the impact for substitutes is not statistically significant. Moreover, the differences between complements and substitutes, in line with theoretical predictions, become more pronounced both with respect to inputs' upstreamness and relative knowledge intensity along the value chain. Specifically, the impact of *Upstr* remains significantly negative for complements, while it becomes significantly positive for substitutes in column (2); the interaction between *lnIPR* and *Upstr* becomes significantly negative in column (2); and the impact of *d_knint_downstr* turns insignificant for complements in column (3), while remaining highly significant and positive for substitutes.

Second, Table 8 presents the results obtained using two alternative indicators of sequential complements/substitutes described in Section 4.2. In particular, columns (1)-(4) use the indicator $d\_compl_{rho \times alpha(ind.)}$ based on the core product's demand elasticity *rho* (as a proxy for ρ) and the industry average of the Herfindahl index (as a proxy for (inverse) α); columns (5-8) use instead the dummy $d\_compl_{rho-alpha(elast.)}$ based on the difference between *rho* and another proxy for α based on the demand elasticity of imported intermediate and capital goods. Due to significant 'frailty' confirmed by the likelihood-ratio test, we continue to rely on a random effects probit estimator, thereby controlling for unobserved heterogeneity at detailed firm-country-product level.

The results in Table 8 show that our previous findings reported in Table 4 are robust to the alternative ways of disentangling complements from substitutes. In the case of complements, in all columns the impact of better IPR quality is again significantly negative. Moreover, the interaction term of IPR quality with upstreamness is still positive and even more significant than before. Finally, the estimates also remain positive for the coefficient on the dummy *d_knint_downstr*, which indicates when knowledge intensive inputs are located more downstream along the value chain. On the other hand, no significant effect of IPR quality is detected in the case of substitutes.

Further robustness checks can be found in Appendix B where, for both complements and substitutes, we focus on the case of higher relative knowledge transmission upstream as this case is more readily comparable with the case of higher degree of upstream contractibility of tangible investments. After presenting our baseline results for specification (III), we extend the analysis to alternative measures of sequential complements/substitutes. We then look at how our results vary across firms that differ in their reliance on inputs sourced by a single country. Finally, we control for additional source country institutional variables that could potentially influence our results in different parts of the value chain. Once more, our conclusions in Section 5.2.2 are confirmed.

6 Conclusion

We have introduced intangible assets in a property rights model of sequential supply chains. In the resulting model firms transmit knowledge to their suppliers to facilitate inputs' customization, but they must protect the transmitted intangibles to avoid knowledge dissipation. Protection is costly and depends on both inputs' knowledge intensity and the quality of institutions protecting intellectual property rights (IPR) in suppliers' locations.

Our model predicts that, when inputs' knowledge intensity increases (decreases) downstream and suppliers' investments are complements, the probability of integrating a randomly selected input is decreasing (increasing) in IPR quality and increasing (decreasing) in the relative knowledge intensity of downstream inputs. It yields opposite but weaker predictions when suppliers' investments are substitutes.

Through the analysis of comprehensive trade and FDI data covering the population of Slovenian firms from 2007 to 2010 we have found evidence in support of our theoretical predictions. Moreover, as also predicted by the model, we have shown that better overall contract enforcement ('rule of law') has the opposite impact of better IPR quality. This divergence suggests that better institutions may have very different effects on firm organization depending on whether they improve the protection of tangible or intangible assets. It also shows that our findings are specific to IPR institutions and cannot be generalized to other regulatory measures that affect contract enforcement.

References

Alfaro, L., Antràs, P., Chor, D., Conconi, P. (2019), "Internalizing Global Value Chains: A Firm-Level Analysis", *Journal of Political Economy*, 127 (2): 509–559.

Antràs, P. (2003), "Firms, Contracts, and Trade Structure", *Quarterly Journal of Economics*, 118 (4), 1375–1418.

Antràs, P. (2005), "Incomplete Contracts and the Product Cycle", *American Economic Review*, 95 (4), 1054–1073.

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.

Antràs, P., Helpman, E. (2004), "Global Sourcing", *Journal of Political Economy*, 112 (3), 552–580.

Antràs, P., Helpman, E. (2008), "Contractual Frictions and Global Sourcing", in Helpman, E., D. Marin, and T. Verdier, eds., *The Organization of Firms in a Global Economy*, Cambridge, MA: Harvard University Press.

Antràs, P., Rossi-Hansberg, E. (2009), "Organizations and Trade", *Annual Review of Economics*, 1 (1), 43–64.

Atalay, E., Hortacsu, A., Syverson, C. (2014), "Vertical Integration and Input Flows", *American Economic Review*, 104 (4), 1120–1148.

Branstetter, L., Fisman, R., Foley, C.F. (2006), "Do Stronger Intellectual Property Rights Increase International Technology Transfer? Empirical Evidence from U.S. Firm-Level Data", *Quarterly Journal of Economics*, 121 (1), 321–349.

Canals, C., Şener, F. (2014), "Offshoring and Intellectual Property Rights Reform", *Journal of Development Economics*, 108, 17–31.

Chamberlain, G. (1984), "Panel Data", in *Handbook of Econometrics*, Elsevier, ed. by Z. Griliches, and M. Intriligator, 1247–1318.

Corcos, G., Irac, D.M., Mion, G., Verdier, T. (2013), "The Determinants of Intra-Firm Trade", *The Review of Economics and Statistics*, 95(3), 825–838.

Defever, F., Toubal, F. (2013), "Productivity, Relation-Specific Inputs and the Sourcing Modes of Multinationals", *Journal of Economic Behavior and Organization*, 94, 245–357.

Eppinger, P., Kukharskyy, B. (2017), "Contracting Institutions and Firm Boundaries", University of Tuebingen - Working Papers in Economics and Finance, No. 100.

Fally, T. (2012), "Production Staging: Measurement and Facts", University of Colorado-Boulder, mimeo.

Grossman, S.J., Hart, O.D. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, 94(4), 691–719.

Grossman, G.M., Helpman, E. (2002), "Integration vs. Outsourcing in Industry Equilibrium", *Quarterly Journal of Economics*, 117, 85–120.

Grossman, G.M., Helpman, E. (2003), "Outsourcing versus FDI in Industry Equilibrium", *Journal of the European Economic Association*, 1, 317–327.

Grossman, G.M., Helpman, E. (2005), "Outsourcing in a Global Economy", *Review of Economic Studies*, 72, 135–159.

Hart, O.D., Moore, J. (1990), "Property Rights and the Nature of the Firm", *Journal of Political Economy*, 98(6), 1119–1158.

Kee, H.L., Nicita, A., Olarreaga, M. (2008), "Import Demand Elasticities and Trade Distortions", *The Review of Economics and Statistics*, 90(4), 666–682.

Kukharskyy, B. (2019), "A Tale of Two Property Rights: Knowledge, Physical Assets, and Multinational Firm Boundaries"', *Journal of International Economics*, forthcoming.

Maskus, K.E. (2000), *Intellectual Property Rights in the Global Economy*, Washington, DC: Peterson Institute for International Economics.

Maskus, K.E. (2012), *Private Rights and Public Problems: The Global Economics of Intellectual Property in the 21st Century*, Washington, DC: Peterson Institute for International Economics.

Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46 (1), 69–85.

Naghavi, A., Spies, J., Toubal, F. (2015), "Intellectual Property Rights, Product Complexity and the Organization of Multinational Firms", *Canadian Journal of Economics*, 48 (3), 881–902.

Olley, G.S., Pakes, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica*, 46 (1), 1263–1297.

Park, W.G. (2008), "International Patent Protection: 1960-2005", *Research Policy*, 37, 761–766.

Pierce, J.R., Schott, P.K. (2009), "A Concordance Between Ten-Digit U.S. Harmonized System Codes and SIC/NAICS Product Classes and Industries", NBER Working Paper 15548, National Bureau of Economic Research.

United Nations Statistics Division (2009), "Note on HS 2007 Data Conversion in UN Comtrade", available at <https://unstats.un.org/unsd/trade/conversions/Note%20on%20the%20HS %202007%20data %20conversion%20in%20UN%20Comtrade.pdf>, accessed on 15 January 2016.

Van Beveren, I., Bernard, A.B., Vandenbussche, H. (2012), "Concording EU Trade and Production Data over Time", NBER Working Paper 18604, National Bureau of Economic Research.

Williamson, O. (1971), "The Vertical Integration of Production: Market Failure Considerations", *American Economic Review*, 61 (2), 112–123.

Williamson, O. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, New York,NY: The Free Press.

Williamson, O. (1985), *The Economic Institutions of Capitalism*, New York, NY: The Free Press.

Wooldridge, J.M. (2002), *Introductory Econometrics*, Mason, OH: South-Western.

Worldwide Governance Indicators (2015), The World Bank Group, available at http://data.worldbank .org/data-catalog/worldwide-governance-indicators.

WTO (2016), *World Trade Report 2016: Levelling the Trading Field for SMEs*, Washington, DC: World Trade Organization.

Yang, G., Maskus, K.E. (2001), "Intellectual Property Rights and Licensing: an Econometric Investigation", *Review of World Economics*, 137 (1), 58–79.

Appendix A: Mathematical Appendix

A1. Heterogeneous cost of protection of knowledge transmission (simple model)

In this Appendix, we build on the simple model of supply chain outlined in Section 2, where production consists of two stages only: a final stage performed by the firm, and a single intermediate stage of production performed by a supplier.

In subsection 2.1 we have introduced the problem of costly knowledge transmission, assuming that the firm faces a given cost of protecting any bit of knowledge transmitted to its supplier, unconditional from its organizational choice. One might argue that this cost may vary with the organizational mode, depending on whether the supplier is integrated within the firm boundaries or it operates as a stand-alone entity. We can accordingly adapt specification (3) as follows

$$\kappa(\omega, \lambda) = \kappa_o \omega \delta^{\lambda},$$

where $\kappa_o = \{\kappa_V, \kappa_O\}$ reflects differences in the cost (or difficulty) of protecting knowledge transmission under different organizational modes, $\beta_o = \{\beta_V, \beta_O\}$. The firm profit, in turn, becomes $\pi_F = \beta_o r(x) - \kappa(\kappa_o, \omega, \lambda)$.

Since the supplier's profit-maximizing level of investment in (5) is unaffected by this change, the firm problem can be formulated as

$$\max_{\beta_o, \varphi} \ \pi_F \ = \ \Omega \ \beta_o \ \delta^{\frac{\rho}{1-\rho}} \left(\frac{1 - \beta_o}{c_o} \right)^{\frac{\rho}{1-\rho}} - \kappa_o \omega \delta^{\lambda} \ ,$$
$$s.t. \ \ \beta_o \in \{\beta_V, \beta_O\}; \ \ \delta > 0,$$

where $c_o = \{c_V, c_O\} > 0$ is the marginal cost of input customization, that we also allow to vary with the organizational form.

The program is solved in two steps. First, we maximize π_F with respect to δ, so as to obtain $\delta^+(\beta_o)$, i.e. the optimal amount of protected knowledge to transmit for a given organizational mode. Then we solve for the optimal organizational choice β_o. In the first step, the first-order condition to satisfy is

$$\frac{\rho}{1-\rho} \ \Omega \ \beta_o \left(\frac{1 - \beta_o}{c_o} \right)^{\frac{\rho}{1-\rho}} \delta^{\frac{\rho}{1-\rho} - 1} = \kappa_o \omega \lambda \delta^{\lambda - 1},$$

which admits the following solution,

$$\delta^+(\beta_o) = \left[\left(\frac{1 - \beta_o}{c_o} \right)^{\frac{\rho}{1-\rho}} \frac{\rho \Omega \beta_o}{(1-\rho)\kappa_o \omega \lambda} \right]^{\frac{1}{\lambda - \frac{\rho}{1-\rho}}} . \tag{A1}$$

The level of firm profits implied by (A1) is then

$$\pi_F = \left(\frac{\lambda(1-\rho)}{\rho} - 1\right)\left(\frac{\rho\Omega}{\lambda(1-\rho)}\right)^{\frac{\lambda(1-\rho)}{\lambda(1-\rho)-\rho}} \left[\frac{\beta_o(1-\beta_o)^{\frac{\rho}{1-\rho}}}{(\kappa_o\omega)^{\frac{\rho}{\lambda(1-\rho)}} c_o^{\frac{\rho}{1-\rho}}}\right]^{\frac{\lambda(1-\rho)}{\lambda(1-\rho)-\rho}} , \tag{A2}$$

which is strictly positive for $\lambda > \rho/(1-\rho)$, the same restriction on parameters that applies to the baseline model with symmetric costs of knowledge protection, $\kappa_V = \kappa_O = 1$ (see Subsection 2.1).

What changes with respect to the baseline model in Section 2 is that, here, independence between the parallel decisions on organization and knowledge transmission does not hold anymore. This is evident from (A2), particularly from the ratio between square brackets, which captures the organizational trade-offs: all else being equal, firm profit is higher for the organizational mode featuring (i) lower marginal cost of input provision c_o, (ii) lower cost κ_o of protecting the transmitted amount of knowledge and, finally, (iii) firm bargaining weight closer to the relaxed optimum (still $\beta_o^+ = 1-\rho$ as in the baseline two-stage model).

Our restriction on the size of λ implies that firm profit in (A2) is negatively related to κ_o, hence the profit is lower for the organizational mode under which knowledge transmission is more costly to protect. Moreover, the gap in profits originating from the cost differential is larger, the more knowledge-intensive the input (the larger ω), with the organizational choice that becomes accordingly more relevant to the firm. These results are easily proved by observing that $\lambda > \rho/(1-\rho)$ implies

$$\frac{d\,(\kappa_o\omega)^{-\frac{\rho}{\lambda(1-\rho)-\rho}}}{d\kappa_o} = \frac{(\kappa_o\omega)^{\frac{\rho}{\rho-\lambda(1-\rho)}}}{\kappa_o\,(\rho-\lambda(1-\rho))} < 0;$$

$$\frac{d^2\,(\kappa_o\omega)^{-\frac{\rho}{\lambda(1-\rho)-\rho}}}{d\kappa_o d\omega} = \frac{d\left(\frac{(\kappa_o\omega)^{\frac{\rho}{\rho-\lambda(1-\rho)}}}{\kappa_o(\rho-\lambda(1-\rho))}\right)}{d\omega} = \frac{\rho}{(\rho-\lambda(1-\rho))^2}(\kappa_o\omega)^{\frac{\lambda(1-\rho)}{\rho-\lambda(1-\rho)}} > 0\ .$$

Traditional assumptions are that, due to gains from specialization, c_o is smaller under outsourcing (i.e. $c_O < c_V$), while κ_o is larger (i.e. $\kappa_O > \kappa_V$) as knowledge dissipation is more likely when knowledge has to be transmitted outside the firm boundaries, rendering the protection of proprietary technology a more tedious (and costly) task. Accordingly, we treat both c_o and κ_o as functions of β_o, assuming $c_o = (\beta_o)^\gamma$ and $\kappa_o = (1-\beta_o)^\eta$, where both γ and η take value in the interval $(0,1)$. The term in square brackets in (A2) then becomes

$$\frac{\beta_o\,(1-\beta_o)^{\frac{\rho}{1-\rho}}}{(c_o)^{\frac{\rho}{1-\rho}}\,(\kappa_o\omega)^{\frac{\rho}{\lambda(1-\rho)}}} = \frac{(\beta_o)^{1-\frac{\gamma\rho}{1-\rho}}\,(1-\beta_o)^{\frac{\rho(\lambda-\eta)}{\lambda(1-\rho)}}}{\omega^{\frac{\rho}{\lambda(1-\rho)}}}\ .$$

The first-order condition of the relaxed version of the firm problem (where β can take any value in $(0,1)$) yields

$$\left(1-\frac{\gamma\rho}{1-\rho}\right)(1-\beta_o) - \frac{\rho(\lambda-\eta)}{\lambda(1-\rho)}\ \beta_o = 0\ .$$

The optimal share of ownership then evaluates to

$$\beta_o^+ = \frac{\lambda\,(1-(1+\gamma)\,\rho)}{\lambda(1-\gamma\rho)-\eta\rho}\ . \tag{A3}$$

Plugging (A3) into (A1), the optimal choice of knowledge transmission is finally obtained as

$$\delta^+ = \left(\frac{\rho\;\Omega\;\beta_o^+}{(1-\rho)\,\kappa_o\omega\cdot\lambda}\cdot\left(\frac{1-\beta_o^+}{c_o}\right)^{\frac{\rho}{1-\rho}}\right)^{\frac{1}{\lambda-\frac{\rho}{1-\rho}}} = \zeta\;\omega^{-\frac{(1-\rho)}{\lambda(1-\rho)-\rho}}, \tag{A4}$$

where

$$\zeta \equiv \left[\frac{\rho\;\Omega}{(1-\rho)\lambda}\left(\frac{\lambda(1-(1+\gamma)\rho)}{\lambda(1-\gamma\rho)-\eta\rho}\right)^{1-\frac{\gamma\rho}{1-\rho}}\left(1-\frac{\lambda\,(1-(1+\gamma)\,\rho)}{\lambda(1-\gamma\rho)-\eta\rho}\right)^{\frac{\rho}{1-\rho}-\eta}\right]^{\frac{1-\rho}{\lambda(1-\rho)-\rho}}$$

is a bundling parameter. From (A4) we note that the firm's desired amount of transmitted knowledge is inversely related with the knowledge intensity of the input procured from the suppler, in tune with the evidence that stems from (7) in the baseline version of the two-stage model.

The presence of cost heterogeneity between integration and outsourcing reveals a static trade-off faced by the firm regarding its organizational decision. When the input involves no firm-specific knowledge (or little), the property rights model is fully at play, prompting the use of outsourcing to create investment incentives by offering a larger share of the surplus to the input supplier. When the input is instead knowledge-intensive, low optimal investment makes the value of supplier's efforts prone to dissipation, thereby reducing supplier returns and incentives for adequate investment in input customization. Finally, when protecting knowledge transmission is costlier under outsourcing, the differential cost of protection gets larger for more knowledge-intensive inputs. This makes the firm even more vulnerable to rent dissipation whenever outsourcing, hence the latter will represent a viable option only if (i) IPR institutions in the supplier's location are strong enough to compensate the extra costs of protecting knowledge transmission associated with outsourcing, and/or (ii) specialization gains from outsourcing are sufficiently large. We highlight this result as:

Lemma 7 (A.1) *When protecting knowledge transmission is costlier under outsourcing, higher knowledge intensity of the input disproportionately reduces knowledge transmission* (*with increased exposure to the risk of dissipation and rent destruction*) *by more under outsourcing, thereby increasing the firm's propensity towards vertical integration.*

A2. Notation in eq. (19)

We report here the analytical expressions of the bundling parameter $\Gamma(\beta_V, \beta_O)$ appearing in Subsection 3.2.2, and more precisely in (19). As explained in the main text, in the case of sequential complements $\Gamma(\beta_V, \beta_O)$ evaluates to

$$\Gamma_C(\beta_V, \beta_O) \equiv \Lambda_C (1-\beta_O)^{\frac{\rho}{1-\rho}} \left[(\beta_O - \beta_V) + \beta_V \left(\frac{1 - \frac{\beta_O}{\beta_V}}{1 - \left(\frac{1-\beta_O}{1-\beta_V} \right)^{-\frac{\alpha}{1-\alpha}}} \right)^{\frac{\rho(1-\alpha)}{\rho-\alpha}} \right],$$

with $\Lambda_C \equiv (H_C)^{\frac{\rho(1-\alpha)}{\alpha(1-\rho)}}$, where H_C corresponds to the expression in (17).

In turn, in the case of sequential substitutes $\Gamma(\beta_V, \beta_O)$ evaluates to

$$\Gamma_S(\beta_V, \beta_O) \equiv \Lambda_S (1-\beta_V)^{\frac{\rho}{1-\rho}} \left[(\beta_V - \beta_O) + \beta_O \left(\frac{1 - \frac{\beta_V}{\beta_O}}{1 - \left(\frac{1-\beta_V}{1-\beta_O} \right)^{-\frac{\alpha}{1-\alpha}}} \right)^{\frac{\rho(1-\alpha)}{\rho-\alpha}} \right]$$

with $\Lambda_S \equiv (H_S)^{\frac{\rho(1-\alpha)}{\alpha(1-\rho)}}$, where H_S corresponds to the expression in (18).

It can be easily proved that $\Gamma_C(\beta_V, \beta_O) = \Gamma_S(\beta_O, \beta_V)$ as claimed in Subsection 3.2.2.

A3. Sequential production model: a special case

In Subsection 3.3, two examples are given for concreteness when considering how IPR quality shapes the organization of a sequential supply chain in which knowledge intensity is a monotonic function of z. In particular we assume the specific functional forms $\omega(s) = e^{\omega s}$ and $\omega(s) = e^{\omega(1-s)}$ for the cases where knowledge intensity of the inputs used in production respectively rises and falls with downstreamness.

In this Appendix, we explicit solve the model for one of these two examples, namely the former. From here onwards, we therefore assume $\omega(z) = e^{\omega z}$, where $\omega > 0$ is the constant rate at which knowledge intensity rises as production moves one stage further along the value chain. Consistently, we interpret ω as a measure of the relative knowledge intensity of downstream inputs with respect to upstream ones. The model is solved starting from (19) and proceeding in two steps as before.

Optimal knowledge transmission. Given (19), the firm optimal choice of $\delta(z)$ obeys the first-order condition

$$\left(\frac{\delta(z)}{\delta(z')} \right)^{\lambda - \frac{\alpha}{1-\alpha}} = \frac{\omega(z')}{\omega(z)}, \tag{A5}$$

where $z' \in [0,1]$ denotes a generic stage of production located upstream of z (i.e., $z > z'$), while $\lambda > \alpha/(1-\alpha)$ is a necessary restriction on parameters for the second-order condition to also hold in the case of complements.

Based on (A5), the higher the knowledge intensity of downstream stages relative to upstream ones (higher $\omega(z)/\omega(z')$), the lower the knowledge transmission downstream (smaller $\delta(z)/\delta(z')$).

Accounting for the specific functional form assumed for $\omega(z)$, we can rewrite (A5) as

$$\frac{\delta(z)}{\delta(z')} = e^{-\mu\left(z-z'\right)}, \quad \text{with} \quad \mu \equiv \frac{\omega}{\lambda - \frac{\alpha}{1-\alpha}},$$

implying that the optimal choice of $\delta(z)$ decreases with z, i.e. with downstreamness, given $\delta(z) > (<)\delta(z')$ for $z < (>)z'$.

To simplify the analysis, we can pick a suitable normalization of the marginal cost of input provision c, such that the optimally-chosen amount of transmitted knowledge at stage z boils down to $\delta(z) = e^{-\mu z}$. Given our specific assumptions, the objective function in (19) can be formulated as

$$\pi_F = \Theta \ \frac{\alpha(1-\rho)}{\rho(1-\alpha)} \ \Gamma(\beta_V, \beta_O) \ c^{-\frac{\rho}{1-\rho}} \ \left[\int_0^1 \delta(z)^{\frac{\alpha}{1-\alpha}} \ dz\right]^{\frac{\rho(1-\alpha)}{\alpha(1-\rho)}} - \int_0^1 e^{\omega z}\delta(z)^{\lambda} dz.$$

Taking the first-order condition for the maximization of π_F with respect to $\varphi(z)$, then integrating, and finally setting $\delta(z) = e^{-\mu z}$, one gets

$$\Theta \ \frac{\alpha}{1-\alpha} \ \Gamma(\beta_V, \beta_O) \ c^{-\frac{1-\rho}{\rho}} \ \left[\int_0^1 e^{-\frac{\alpha\mu}{1-\alpha}z} \ dz\right]^{\frac{\alpha-\rho}{\alpha(\rho-1)}} = \lambda \int_0^1 e^{\left[\omega-\mu\left(\lambda-\frac{\alpha}{1-\alpha}\right)\right]z} \ dz.$$

Solving the two integrals, a final equation is obtained,

$$\Theta \ \frac{\alpha}{1-\alpha} \ \Gamma(\beta_V, \beta_O) \ c^{-\frac{1-\rho}{\rho}} \ \left[\left(-\frac{e^{-\frac{\alpha\mu}{1-\alpha}}}{\frac{\alpha\mu}{1-\alpha}}\right)^{\frac{\alpha-\rho}{\alpha(\rho-1)}} - \left(-\frac{1}{\frac{\alpha\mu}{1-\alpha}}\right)^{\frac{\alpha-\rho}{\alpha(\rho-1)}}\right] =$$
$$= \lambda\left[\left(\frac{e^{\left[\omega-\mu\left(\lambda-\frac{\alpha}{1-\alpha}\right)\right]}}{\omega-\mu\left(\lambda-\frac{\alpha}{1-\alpha}\right)}\right) - \left(\frac{1}{\omega-\mu\left(\lambda-\frac{\alpha}{1-\alpha}\right)}\right)\right],$$

from which a suitable normalization for c is easily derived such that the optimal policy function of the firm is indeed $\delta(z) = e^{-\mu z}$. Note that the rate at which the optimal amount of transmitted knowledge (exponentially) decreases along the value chain, namely μ, is a bundling parameter which compounds both technological variables (ω and α) and institutional ones (λ).

Organizational choices. We can now derive the optimal share of ownership for any stage z. Given $\omega(z) = e^{\omega z}$ and $\delta(z) = e^{-\mu z}$, the objective function in (19) becomes

$$\pi_F = \Phi \int_0^1 \beta(z)\left[e^{-\mu z}\left(1-\beta(z)\right)\right]^{\frac{\alpha}{1-\alpha}} \ \left[\int_0^z \left[e^{-\mu s}(1-\beta(s))\right]^{\frac{\alpha}{1-\alpha}} ds\right]^{\frac{\rho-\alpha}{\alpha(1-\rho)}} dz, \tag{A6}$$

where

$$\Phi \equiv A \ \rho^{\frac{\rho}{1-\rho}} \left(\frac{1-\rho}{1-\alpha}\right)^{\frac{\rho-\alpha}{\alpha(1-\rho)}}.$$

Following Antràs and Chor (2013), we introduce a real-valued function of z,

$$\upsilon(z) \equiv \int_0^1 \left(e^{-\mu z} \ [1-\beta(z)]\right)^{\frac{\alpha}{1-\alpha}} \ dz,$$

such that the firm problem can be reformulated as follows,

$$\max_{\upsilon(z),u(z)} \Phi \int_0^1 \left[1 - e^{\mu z} \ u(z)^{\frac{1-\alpha}{\alpha}}\right] \ u(z) \ \upsilon(z)^{\frac{\rho-\alpha}{\alpha(1-\rho)}} d\upsilon \ ,$$

with $u(z) = \upsilon'(z) = [e^{-\mu z}(1-\beta(z))]^{\frac{\alpha}{1-\alpha}}$.

The associated Euler-Lagrange equation yields

$$\frac{1}{\alpha} \ e^{\mu z} \ u^{\frac{1-\alpha}{\alpha}} \ \upsilon^{\frac{\rho-\alpha}{\alpha(1-\rho)}} \left[\frac{(\rho-\alpha)(1-\alpha)}{\alpha(1-\rho)} \ \frac{u}{\upsilon} + \mu + \frac{1-\alpha}{\alpha} \ \frac{u'}{u}\right] = 0 \ , \tag{A7}$$

with $\upsilon = \upsilon(z)$, $u = u(z) = \upsilon'$, and $u' = \upsilon''$. Out of the three admissible solutions for (A7), only one generates strictly positive profits,

$$\frac{(\rho-\alpha)(1-\alpha)}{\alpha(1-\rho)} \ \frac{u}{\upsilon} + \mu + \frac{1-\alpha}{\alpha} \ \frac{u'}{u} = 0. \tag{A8}$$

The optimal share of ownership for each stage z can be retrieved by solving the second-order differential equation implied by (A8), in light of the transversality condition $e^{\mu} \ \upsilon'(1)^{\frac{1-\alpha}{\alpha}} = \alpha$, and the initial condition $\upsilon(0) = 0$.

The solution that we obtain is

$$\beta^*(z) = 1 - \alpha \left(\frac{1-e^{-\frac{\alpha}{1-\alpha}\mu z}}{1-e^{-\frac{\alpha}{1-\alpha}\mu}}\right)^{\frac{\alpha-\rho}{\alpha}}, \tag{A9}$$

which can be proved to satisfy a sufficient condition for the maximum and thus qualifies as the solution of the firm problem in its relaxed version, where $\beta(z)$ is not restricted to either β_V or β_O but can be chosen from the whole set of piece-wise continuously differentiable real-valued functions.

The policy function $\beta^*(z)$ in (A9) does not violate the constraint $0 \leq \beta(z) \leq 1$, for all $\rho \in (0,1)$ and $\alpha \in (0,1)$ such that $\rho < \alpha$. Hence, in the case of *substitutes*, the function above is admitted as the solution to the unconstrained problem, which necessarily corresponds to the one which yields the maximum for the constrained version of the same problem, where restriction $\beta(z) \in \{\beta_V, \beta_O\}$ applies.

If $\rho > \alpha$ holds, the optimal share $\beta^*(z)$ instead violates the constraint at least for some values of $z \in [0,1]$. In the case of *complements*, the solution to program (A6) must therefore be obtained by solving the following constrained problem,

$$\max_{\upsilon(z),u(z)} \quad \pi_F = \Phi \int_0^1 \left[1 - e^{\mu z}\ u(z)^{\frac{1-\alpha}{\alpha}}\right]\ u(z)\ \upsilon(z)^{\frac{\rho-\alpha}{\alpha(1-\rho)}}\, d\upsilon \tag{A10}$$

$$s.t.\ \ 0 < u(z)\ e^{\frac{\alpha}{1-\alpha}\mu z} < 1$$

$$\upsilon(0) = 0 \quad \text{(initial condition)}.$$

The associated Hamiltonian function

$$H(\upsilon, u, z, \ell) = \left[1 - e^{\mu z}\ u^{\frac{1-\alpha}{\alpha}}\right]\ u\ \upsilon^{\frac{\rho-\alpha}{\alpha(1-\rho)}} + \ell\ u + \vartheta\ (1 - e^{\frac{\alpha}{1-\alpha}\mu z}\ u)$$

implies the costate equation

$$\ell' = -\frac{\partial H}{\partial \upsilon} = -\frac{\rho-\alpha}{\alpha(1-\rho)}\ \upsilon^{\frac{\rho-\alpha}{\alpha(1-\rho)}}\ \left[1 - e^{\mu z} u^{\frac{1-\alpha}{\alpha}}\right]\ \frac{u}{\upsilon} \tag{A11}$$

Solving the first-order condition ($\partial H/\partial u = 0$) for ℓ and then taking the total derivative, a second expression for ℓ' is obtained. The latter, combined with the costate equation, delivers

$$\frac{1-\alpha}{\alpha^2}\ e^{\mu z}\ u^{\frac{1-\alpha}{\alpha}}\ \upsilon^{\frac{\rho-\alpha}{\alpha(1-\rho)}}\ \left[\frac{\rho-\alpha}{1-\rho}\ \frac{u}{\upsilon} + \frac{u'}{u} + \frac{\alpha}{1-\alpha}\ \mu\right] + F(z, \vartheta', \vartheta) = 0,$$

which coincides with (A8) insofar as the constraint $u \leq 1$ (i.e. $\beta(z) \geq 0$) does not bind, implying $\vartheta' = \vartheta = 0$.

Nevertheless, for $\rho > \alpha$, the solution in (A9) is renown to violate the above constraint, which is proved to happen in the neighborhood of $z = 0$, when $\upsilon(z)$ gets sufficiently small. This implies $\vartheta > 0$. If the constraint binds at some point $\widehat{z} \in (0,1)$, then it necessarily binds (i.e. $\theta > 0$) for any $z < \widehat{z}$. As a result, we pose $\beta(z) = 0$ for all $z \in [0, \widehat{z}]$, from which the boundary condition $e^{\frac{\alpha}{1-\alpha}\mu\widehat{z}}\ \upsilon'(\widehat{z}) = 1$ is easily derived. Then, we look for a solution of the first-order differential equation that solves (A11) only limited to $z > \widehat{z}$. In our search, we take advantage of two pieces of additional information: the first is the transversality condition ($e^{\mu}\upsilon'(1)^{\frac{1-\alpha}{\alpha}} = \alpha$), while the second is the fact that, at point $\widehat{z}$, we necessarily have

$$\upsilon(\widehat{z}) = \int_0^{\widehat{z}} \upsilon'(z)\, dz = \int_0^{\widehat{z}} u(z)\, dz$$

from which we obtain $\upsilon(\widehat{z}) = \frac{1-\alpha}{\alpha\mu}\ \left[1 - e^{-\frac{\alpha}{1-\alpha}\mu\widehat{z}}\right]$. After a few manipulations, we pin down stage $\widehat{z}$ implicitly defined by the following condition

$$e^{-\frac{\alpha}{1-\alpha}\mu\widehat{z}} = \frac{e^{-\frac{\alpha}{1-\alpha}\mu} - (1 - \alpha^{-\frac{\alpha}{\rho-\alpha}})^{\frac{1-\alpha}{1-\rho}}}{1 - (1 - \alpha^{-\frac{\alpha}{\rho-\alpha}})^{\frac{1-\alpha}{1-\rho}}}. \tag{A12}$$

The policy function that applies to all $z > \widehat{z}$ can be finally proved to be

$$\beta^*(z) = 1 - \alpha \left[1 + \chi \, \frac{e^{-\frac{\alpha}{1-\alpha}\mu} - e^{-\frac{\alpha}{1-\alpha}\mu z}}{e^{-\frac{\alpha}{1-\alpha}\mu} - 1} \right]^{\frac{\alpha-\rho}{\alpha}} \quad \text{with} \quad \chi \equiv \frac{(1-\rho)(1-\alpha^{-\frac{\alpha}{\rho-\alpha}} - (1-\alpha)}{(1-\rho)\alpha^{-\frac{\alpha}{\rho-\alpha}}}.$$

Hence, the solution to the constrained version of the firms' problem, which solves the relaxed program in (A6) in the case of complements ($\rho > \alpha$), can be characterized as

$$\beta^{**}(z) = \max \left\{ 0, 1 - \alpha \left[1 + \chi \, \frac{e^{-\frac{\alpha}{1-\alpha}\mu} - e^{-\frac{\alpha}{1-\alpha}\mu z}}{e^{-\frac{\alpha}{1-\alpha}\mu} - 1} \right]^{\frac{\alpha-\rho}{\alpha}} \right\}, \tag{A13}$$

where the double asterisk differentiates the solution above from the one relative to the unconstrained problem, i.e. $\beta^*(z)$ in (A9).

In Figure 6, the policy function $\beta^{**}(z)$ in (A13), which solves the constrained problem in (A10) for the case of *complements* ($\rho > \alpha$), is represented with a solid line, upward-sloping for all $z > \hat{z}$. It is plotted together with the solution to the unconstrained problem in (A9) for the cases where $\rho > \alpha$ (dotted line) and $\rho < \alpha$ (solid line, downward-sloping). As in Antràs and Chor (2013), the optimal share of ownership turns out to be decreasing with z in the case of substitutes ($\rho < \alpha$) and increasing for complements ($\rho > \alpha$). In this second case (complements), at all stages $z > \widehat{z}$ the share is higher in the unconstrained problem than in the constrained one: $\beta^*(z) > \beta^{**}(z)$. Moreover, when upstream suppliers cannot be incentivized by being offered a payoff exceeding their marginal contribution (as it would be optimal in the lack of the restriction $0 < \beta(z) < 1$), then the firm optimally offers "their full marginal contribution to a larger measure of suppliers, and a higher share of their marginal contribution to the remaining suppliers"(Antràs and Chor, 2013).

Appendix B: Sensitivity Analysis

To test the sensitivity of our results to different specifications, we first replicate the double-split sample results obtained in column (1) of Table 5 for both complements and substitutes, adding here industry fixed effects. To simplify the comparison and better grasp the intuition behind our results, we take a subsample of industries in which relatively more knowledge transmission takes place upstream. Results are depicted in Table 9.

We observe a significantly negative impact of the ratio $rel\_upst\_knint_k$ on the likelihood of vertical integration in the case of complements. This indicates that, at least for complements, when knowledge intensity of inputs increases downstream, the probability of integration is increasing in the relative knowledge intensity of downstream inputs. Recall that here the sign of $rel\_upst\_knint_k$ is negative since this is an inverse measure of relative knowledge intensity of the downstream stages. This finding supports Proposition 4. On the other hand, the impact of relative knowledge intensity of downstream inputs on integration of a certain input within the firm boundary is negative for substitutes (in line with Proposition 5) yet not significant, which conforms with Proposition 6.

As regard to the impact of IPR institutions on our dependent variable, it remains significantly negative for complements also when limiting the sample to those industries where knowledge intensity is increasing with downstreamness (in line with Proposition 4). Again, a negative impact on the likelihood of integration tends to be most pronounced for relatively downstream stages, as denoted by the positive and significant coefficient of the interaction between $lnIPR$ and $Upstr$. As downstream stages are the knowledge intensive ones in the sample, this can also be interpreted as the impact of IPR quality being stronger for more knowledge-intensive inputs. The optimal organizational choice is far less responsive to the quality of IPR institutions in the sourcing partner-country when considering substitutes.

Next, we extend our results to the alternative methods of categorizing complements versus substitutes. Table 10 illustrates the results. In columns (1)-(4) we consider both *rho* (the estimated import demand elasticity) and industry averages of the Herfindahl index, thereby distinguishing between complements and substitutes based on $d\_compl_{rho \times alpha\ (ind.)}$. In columns (5)-(8), the specifications are based on the difference between *rho* and the measure *alpha* (*elast.*) obtained from estimated demand elasticities of the intermediate and capital goods imported: the distinction therefore hinges on $d\_compl_{rho-alpha\ (elast.)}$.

The results on the effect of IPR institutions on the propensity towards integration are robust to the baseline specification with the *rho* measure and in line with Propositions 4 and 5. We can instead observe a change as regard to the effect of relative knowledge intensity of downstream inputs. The impact of the ratio $rel\_upst\_knint_k$ on the likelihood of vertical integration shifts from significant to insignificant in the case of complements, while it becomes significantly positive in the case of substitutes, as stated in Proposition 5. In other words, provided that input knowledge intensity increases with downstreamness and supplier investments are sequential substitutes, the probability of integration is decreasing in the relative knowledge intensity of downstream inputs.

Despite this switch in the level of significance, under all alternatives measures of complementarity/substitutability, the response of our dependent variable to relative knowledge intensity of downstream inputs significantly differs between complements and substitutes, in a manner that does not contradict and even reinforces our theoretical predictions. An additional result obtained empirically is that the strength of IPR institutions is more relevant when inputs are complements, whereas knowledge intensity plays a larger role for organizational decisions when inputs are substitutes.

The vast majority of firms in our sample (and basically all firms reporting outward FDIs) source their inputs from more than one partner country; therefore, we are not able to replicate exactly the scope of the one-partner country model with our empirical setting. Instead, we test the robustness of our results by gradually restricting the baseline database to firms which import a certain proportion of their inputs from a single country.

We start with a sub-sample of firms with at least 10% share of inputs being sourced from one-county (columns (1) and (2) in Table 11), and further increase the threshold concentration level to 20% and 30% of inputs obtained from a single country in columns (3)-(4) and (5)-(6), respectively. The results in terms of the impact of relative knowledge intensity of downstream inputs and of IPR enforcement on the integration decision (and other regressors) are very stable and fully robust when pushing the threshold from 10% to 20% and further to a 30% share within a single (primary) source country. The magnitude of coefficients for relative knowledge intensity of downstream inputs even slightly increases and becomes more significant with higher thresholds.

Finally, we control for the possibility that the interaction term $lnIPR * Upstr$ may pick up the effect of upstreamness with other time-varying effects in the partner country, since there is limited variation in the quality of IPR institutions over time. We therefore adjust the empirical model specification by including additional partner-country institutional variables that are likely to be correlated with $lnIPR$, i.e., rule of law, government effectiveness, and control of corruption obtained from Worldwide Governance Indicators (2015). We then interact upstreamness with these institutional variables simultaneously. Results presented in Table 12 are obtained with the rho-based categorization of complements and substitutes and show that the coefficient on $lnIPR*Upstr$ remains significantly positive after adding other institutional variables and their interactions with upstreamness. The impact of other regressors is fully robust to the baseline results.

Tables and Figures

Table 1: The GVC participation index, Slovenia 2011 (% share in total gross exports).

	Slovenia	Developing countries	Developed countries
Total GVC participation	58.7	48.6	48.0
Forward participation	22.6	23.1	24.2
Backward participation	36.1	25.5	23.8

Source: WTO.

Table 2: Descriptive statistics

	Pooled sample	Complements with IP intensity downstream[1]	Complements with IP intensity upstream[2]	Substitutes with IP intensity downstream[3]	Substitutes with IP intensity upstream[4]
	mean (std dev.)	mean (std dev.)	mean (std dev.)	mean (std dev.)	mean (std dev.)
d_OutFDI	0.184	0.165	0.218	0.197	0.161
	(0.388)	(0.371)	(0.413)	(0.398)	(0.367)
d_OutFDI_bilateral	0.031	0.026	0.042	0.040	0.015
	(0.173)	(0.159)	(0.201)	(0.197)	(0.120)
d_integr	0.0004	0.0004	0.0003	0.0007	0.00003
	(0.019)	(0.019)	(0.017)	(0.026)	(0.006)
d_integr_IFEX	0.0003	0.0002	0.0002	0.0006	0.00001
	(0.017)	(0.015)	(0.015)	(0.024)	(0.003)
Upstreamness	2.523	2.523	2.503	2.531	2.530
	(1.072)	(1.033)	(1.115)	(1.045)	(1.105)
IMP demand elasticity (abs.)	1.167	1.725	1.357	0.892	0.848
	(2.391)	(4.707)	(1.406)	(0.167)	(0.219)
Inputs' demand elasticity	1.150	1.185	1.196	1.108	1.134
	(0.903)	(0.817)	(0.647)	(1.172)	(0.748)
Industry Herfindahl index ($\bar{H}_{jt}$)	0.718	0.720	0.694	0.737	0.711
	(0.082)	(0.086)	(0.095)	(0.066)	(0.079)
rel_upst_knint$_k$	0.994	0.951	1.054	0.937	1.058
	(0.072)	(0.045)	(0.034)	(0.056)	(0.035)
IPR index	4.525	4.530	4.515	4.534	4.517
	(0.241)	(0.221)	(0.253)	(0.234)	(0.258)
Rule of law index	1.300	1.320	1.273	1.350	1.241
	(0.649)	(0.643)	(0.660)	(0.618)	(0.678)
Age	16.808	16.721	16.767	17.029	16.647
	(8.011)	(7.985)	(8.363)	(8.112)	(7.620)
Employment	361.775	136.495	316.481	435.193	512.303
	(1,336.96)	(306.311)	(743.912)	(1,466.0)	(1,939.6)
Ex_propensity	0.313	0.297	0.290	0.354	0.295
	(0.336)	(0.331)	(0.326)	(0.349)	(0.329)
Kintensity	86,064.2	72,283.4	64,065.8	91,761.8	108,545.4
	(576,600)	(177,074)	(208,802)	(488,467)	(971,779)
Lproductivity	46,252.9	43,827.6	37,954.3	56,949.5	41,666.5
	(112,858)	(45,776.8)	(47,796.3)	(184,371.5)	(64,002.0)
Debt_assets ratio	0.610	0.608	0.638	0.576	0.631
	(0.242)	(0.241)	(0,245)	(0.244)	(0.233)
No of observations	791,911	185,156	155,278	249,187	202,290

Note: Labour productivity (L_productivity) and capital intensity (K_intensity) are expressed in EUR.
[1] $d\_comp = 1$ & $d\_knint\_downstr = 1$; [2] $d\_comp = 1$ & $d\_knint\_downstr = 0$
[3] $d\_comp = 0$ & $d\_knint\_downstr = 1$; [4] $d\_comp = 0$ & $d\_knint\_downstr = 0$

Table 3: Probit and random effects probit model of integration at firm-market-product level for pooled sample/triple interaction specification, *rho*

	(1) Probit	(2) Probit Chamberlain -Mundlak	(3) RE Probit	(4) RE probit
d_comp	12.11**	12.44**	30.51***	35.41***
1	(5.245)	(5.309)	(10.67)	(11.41)
lnIPR	0.478	0.372	-3.085	0.604
	(2.336)	(2.431)	(5.435)	(9.320)
d_comp ∗ lnIPR	-7.673**	-7.888**	-19.00***	-22.54***
1	(3.528)	(3.568)	(7.166)	(7.706)
Upstr	0.490	0.541	-0.687	-0.388
	(1.068)	(1.050)	(3.003)	(4.568)
d_comp ∗ Upstr	-6.164*	-6.347*	-13.13**	-16.55**
1	(3.228)	(3.297)	(6.163)	(6.826)
lnIPR ∗ Upstr	-0.500	-0.533	0.0520	-0.348
	(0.786)	(0.773)	(2.050)	(3.094)
d_comp ∗ lnIPR ∗ Upstr	4.029*	4.147*	8.530**	10.98**
1	(2.147)	(2.191)	(4.116)	(4.564)
d_knint_downstr	0.791***	0.803***	3.023***	2.474***
1	(0.267)	(0.271)	(0.772)	(0.650)
d_knint_downstr ∗ d_comp	-0.454	-0.449	-1.038	-1.036
1 1	(0.289)	(0.296)	(0.938)	(0.852)
lnSize(-1)	0.104	0.338	0.424***	0.569***
	(0.0641)	(0.246)	(0.137)	(0.130)
Age	0.0397***	0.0409***	0.215***	0.183***
	(0.0138)	(0.0142)	(0.0278)	(0.0254)
Ex_prop(-1)	1.369***	-0.373	5.537***	4.696***
	(0.358)	(1.155)	(0.748)	(0.787)
ln Kintensity(-1)	0.159	0.466*	0.231	0.618***
	(0.179)	(0.269)	(0.195)	(0.203)
ln Lproductivity(-1)	-0.289**	-0.307**	-1.051***	-1.018***
	(0.116)	(0.134)	(0.313)	(0.314)
Debt_assets(-1)	-0.847**	-0.886	-0.472	-1.985**
	(0.427)	(0.568)	(0.704)	(0.792)
Constant	-5.407	-5.452	-12.54	-23.98
	(3.485)	(3.663)	(8.220)	(14.61)
Country dummies	yes	yes	no	yes
Time dummies	yes	yes	yes	yes
Industry dummies	yes	yes	no	yes
Firm-level means	no	yes	no	no
Log (pse.)likelihood	-1424.0673	-1416.8795	-988.77034	-878.37635
Wald test	chi2(42)=	chi2(47)=	chi2(21)=	chi2(42)=
	3802.43***	4881.52***	297.96***	336.83***
Wald test for heteroscedasticity (H0: lnsigma2=0)				
lnsigma2	0.016	0.014		
lempllag	(0.053)	(0.058)	/	/
chi2(1)	0.09	0.06		
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)				
	/	/	1458.58***	1091.38***
Observations	615,847	611,495	791,911	615,847
No. of firm_market_product			445,249	347,470

Note: Robust Std. Err. in round brackets, adjusted for firm clusters in (heteroskedastic) probit models; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: Random effects probit model of integration at firm-market-product level, *rho*

	(1) RE probit d_integr	(2) RE probit d_integr	(3) RE probit d_integr	(4) RE probit d_integr	(5) RE probit d_integr	(6) RE probit d_integr	(7) RE probit d_integr	(8) RE probit d_integr
	Comp	Comp	Comp	Subst	Subst	Subst	Comp	Subst
lnIPR	-15.519***	-38.072***	-23.50*	-1.264	-2.847	-24.21		
	(2.344)	(6.285)	(13.37)	(3.679)	(8.869)	(19.11)		
Upstr		-20.879***	-21.90***		-2.073	-3.605	-0.128	-1.129***
		(6.657)	(6.833)		(4.763)	(5.810)	(0.162)	(0.217)
lnIPR ∗ Upstr		12.851***	13.70***		0.866	1.699		
		(4.413)	(4.511)		(3.241)	(3.891)		
wgi_rule_law							5.379***	-6.386***
							(1.992)	(1.925)
upst ∗ wgi_rule_law							-0.590***	0.186
							(0.200)	(0.238)
d_knint_downstr	0.878	2.085**	1.401**	3.028**	2.577***	3.122***	1.472***	3.941***
	(0.663)	(0.891)	(0.639)	(1.193)	(0.782)	(0.784)	(0.429)	(0.660)
lnSize(-1)	1.114***	1.863***	1.329***	1.012***	0.899***	0.992***	0.873***	0.934***
	(0.353)	(0.422)	(0.320)	(0.228)	(0.169)	(0.188)	(0.168)	(0.139)
Age	0.243***	0.202***	0.176***	0.210***	0.174***	0.177***	0.0831***	0.155***
	(0.057)	(0.069)	(0.0523)	(0.052)	(0.037)	(0.0403)	(0.0259)	(0.0269)
Ex_prop(-1)	6.583***	9.232***	6.036***	3.264***	2.925***	3.217***	6.731***	3.650***
	(1.196)	(1.856)	(1.395)	(1.157)	(1.040)	(1.194)	(0.669)	(0.753)
ln Kintensity(-1)	1.279***	2.693***	2.015***	-0.856***	-0.713***	-0.628*	1.357***	-0.390
	(0.406)	(0.586)	(0.370)	(0.3032)	(0.277)	(0.335)	(0.240)	(0.240)
ln Lproductivity(-1)	-0.785	-1.365*	-1.027*	-1.420***	-1.103**	-1.115**	-0.132	-1.509***
	(0.615)	(0.774)	(0.612)	(0.504)	(0.466)	(0.517)	(0.299)	(0.359)
Debt_assets(-1)	-5.702***	-7.327***	-5.393***	-4.431***	-2.863**	-3.042**	-2.877***	-2.130**
	(1.894)	(2.590)	(1.676)	(1.677)	(1.300)	(1.430)	(0.911)	(0.865)
lnDist	0.021	0.072		-0.143	-0.125			
	(0.354)	(0.433)		(0.282)	(0.283)			
lnGDP	0.567*	0.717*		-0.389**	-0.447**			
	(0.302)	(0.378)		(0.191)	(0.194)			
lnGDPpc	-0.309	-1.174*		-0.975*	-1.038*			
	(0.549)	(0.697)		(0.539)	(0.562)			
Constant	-25.118***	-0.451	-0.0927	7.474	14.651	22.23	-43.96***	-7.668
	(9.454)	(17.799)	(21.44)	(7.488)	(12.457)	(30.21)	(5.445)	(5.288)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
Country dummies	no	no	yes	no	no	yes	yes	yes
Log likelihood	-395.2353	-375.8171	-335.4060	-513.4540	-509.6724	-445.5180	-887.1148	-891.4427
Wald test	chi2(19)=	chi2(21)=	chi2(33)=	chi2(18)=	chi2(20)=	chi2(29)=	chi2(40)=	chi2(34)=
	199.31***	231.80***	141.44***	271.77***	340.92***	379.20***	324.0***	425.32***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)								
	368.95***	311.41	246.84***	915.31***	872.81***	692.60***	745.82***	1303.24***
Observations	308,518	308,518	246,902	390,751	390,751	312,789	277,561	362,193
No. of firm_market_prod	197,751	197,751	155,372	243,737	243,737	192,766	175,414	221,836

Note: Standard errors in in round brackets; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 5: Random effects probit model of integration at firm-market-product level for sequential complements (i.e., double-split subsample), *rho*

	(1) RE probit	(2) RE probit	(3) RE probit
	Comp d_knint_downstr=1	Comp d_knint_downstr=0	Comp d_knint_downstr=1
rel_upst_knint_k		-111.6*	
		(59.70)	
$(\text{rel\_upst\_knint}_k)^{-1}$	23.60***		18.72***
	(7.861)		(5.320)
Upstr	-12.00*	-97.61***	-0.111
	(6.688)	(32.24)	(0.256)
lnIPR	-24.08***	-92.37***	
	(6.850)	(26.38)	
lnIPR ∗ Upstr	6.831	62.81***	
	(4.508)	(20.71)	
wgi_rule_law			1.434**
			(0.730)
upst ∗ wgi_rule_law			-0.972***
			(0.290)
lnSize(-1)	2.138***	7.724**	1.608***
	(0.589)	(3.121)	(0.260)
Age	0.229**	0.211	0.0424
	(0.0909)	(0.144)	(0.0400)
Ex_prop(-1)	8.881***	32.32***	6.946***
	(1.591)	(11.81)	(0.840)
ln Kintensity(-1)	2.829***	4.924***	2.185***
	(0.696)	(1.860)	(0.392)
ln Lproductivity(-1)	-3.071***	4.628**	-0.652
	(1.083)	(2.256)	(0.558)
Debt_assets(-1)	-11.57***	1.303	-5.183***
	(3.266)	(6.132)	(1.477)
lnDist	0.154	1.242**	-1.959***
	(0.410)	(0.548)	(0.298)
lnGDPpc	-0.661	-2.174**	-3.803***
	(0.836)	(1.028)	(0.662)
Constant	-18.90	68.20	-14.86
	(19.17)	(60.47)	(10.80)
Time dummies	yes	yes	yes
Log likelihood	-280.249	-77.091049	-552.01198
Wald test	chi2(15)=	chi2(15)=	chi2(15)=
	93.73***	79.25***	251.15***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)			
	201.70***	92.79***	497.89***
Observations	185,156	155,278	192,940
No. of firm_market_prod	123,964	102,680	129,438

Note: Standard errors in in round brackets; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 6: Marginal effect of IPR and rule of law, *rho*

	(1) IPR Comp d_knint_downstr=1	(2) IPR Comp d_knint_downstr=0	(3) Rule of Law Comp d_knint_downstr=1
Upstr=1	-0.0044***	-0.0070***	0.0002
	(0.0014)	(0.0024)	(0.0003)
Upstr=2	-0.0013**	0.0010*	-0.0002
	(0.0006)	(0.0006)	(0.0002)
Upstr=3	-0.0002	0.0025***	-0.0003**
	(0.0005)	(0.0009)	(0.0001)
Upstr=4	0.0001	0.0045	-0.0005***
	(0.0004)	(0.0069)	(0.0002)
Upstr=5	0.0002	0.0129**	-0.0006***
	(0.0004)	(0.0059)	(0.0002)
Upstr=6	0.0002	0.0103	-0.0007**
	(0.0003)	(0.0163)	(0.0003)

Note: Standard errors in in round brackets; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 7: Random effects probit model of integration at firm-market-product level on subsample of intermediate and capital goods and with intra-firm corrected dependent variable, *rho*

	(1) d_integr intermediate & capital goods	(2) d_integr intermediate & capital goods	(3) d_integr_IFEX full sample	(4) d_integr_IFEX full sample
	Comp	Subst	Comp	Subst
lnIPR	-43.10***	8.015	-33.52***	-0.383
	(6.291)	(6.494)	(6.321)	(6.963)
Upstr	-24.86***	10.86***	-17.21**	-0.529
	(7.097)	(3.357)	(6.734)	(4.004)
lnIPR $*$ Upstr	15.59***	-8.576***	10.51**	0.0834
	(4.695)	(2.375)	(4.446)	(2.731)
d_knint_downstr	1.365*	3.983***	-0.303	3.841***
	(0.795)	(1.321)	(0.800)	(1.147)
lnSize(-1)	1.702***	0.661**	1.315***	1.263***
	(0.471)	(0.261)	(0.405)	(0.234)
Age	0.186***	0.320***	0.248***	0.185***
	(0.0632)	(0.0661)	(0.0745)	(0.0674)
Ex_prop(-1)	9.323***	9.572***	7.093***	2.598**
	(2.150)	(2.698)	(2.451)	(1.220)
ln Kintensity(-1)	2.790***	-1.138**	2.691***	-1.153***
	(0.487)	(0.474)	(0.662)	(0.339)
ln Lproductivity(-1)	-1.213	-1.752**	-1.303*	-1.576***
	(0.744)	(0.861)	(0.772)	(0.562)
Debt_assets(-1)	-6.258***	-3.213	-3.802*	-4.153**
	(2.241)	(2.273)	(2.033)	(1.967)
lnDist	0.0123	-0.642	-0.226	-0.0433
	(0.440)	(0.527)	(0.418)	(0.277)
lnGDP 0.881**	0.0119	1.350***	-0.532***	
	(0.364)	(0.303)	(0.470)	(0.195)
lnGDPpc	-1.039	-1.764**	-0.575	-0.626
	(0.654)	(0.896)	(0.699)	(0.544)
Constant	1.200	-7.079	-22.06	12.34
	(15.86)	(15.06)	(17.08)	(12.16)
Time dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
Log likelihood	-354.6491	-359.3736	-230.43421	-393.6205
Wald test	chi2(20)= 271.16***	chi2(19)= 140.90***	chi2(18)= 113.05***	chi2(20)= 192.00***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)				
	300.22***	751.71***	216.87***	725.37***
Observations	218,495	246,591	208,942	390,751
No. of firm_market_prod	141,696	154,273	137,891	243,737

Note: Standard errors in in round brackets; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 8: Random effects probit model of integration at firm-market-product level, alternative combined $rho \times alpha$ (*ind.*) and $rho - alpha$ (*elast.*) measures

	(1) rho × alpha (ind.)	(2) rho × alpha (ind.)	(3) rho × alpha (ind.)	(4) rho × alpha (ind.)	(5) rho-alpha (elast.)	(6) rho-alpha (elast.)	(7) rho-alpha (elast.)	(8) rho-alpha (elast.)
	Comp	Comp	Subst	Subst	Comp	Comp	Subst	Subst
lnIPR	-6.967***	-14.966***	-2.781	-1.904	-8.413***	-16.557***	-1.391	-2.415
	(1.361)	(2.808)	(3.038)	(5.118)	(1.562)	(3.209)	(3.325)	(5.447)
Upstr		-7.811***		-0.220		-8.446***		-1.679
		(2.348)		(3.149)		(2.493)		(3.443)
lnIPR ∗ Upstr		4.863***		-0.699		5.287***		0.589
		(1.580)		(2.142)		(1.684)		(2.333)
d_knint_downstr	0.740**	1.099**	1.427***	1.536***	1.368***	1.786***	2.221***	2.332***
	(0.369)	(0.442)	(0.459)	(0.539)	(0.398)	(0.480)	(0.576)	(0.572)
lnSize(-1)	0.100	0.186*	0.565***	0.688***	0.171	0.298**	0.358***	0.486***
	(0.088)	(0.103)	(0.165)	(0.207)	(0.118)	(0.137)	(0.110)	(0.123)
Age	0.202***	0.218***	0.044***	0.048**	0.233***	0.216***	0.147***	0.142***
	(0.027)	(0.030)	(0.021)	(0.025)	(0.033)	(0.035)	(0.027)	(0.028)
Ex_prop(-1)	5.469***	5.967***	1.598**	1.284*	6.089***	6.382***	3.912***	3.865***
	(1.018)	(1.126)	(0.648)	(0.702)	(1.285)	(1.357)	(0.839)	(0.904)
ln Kintensity(-1)	0.490***	0.621***	0.018	0.091	0.440**	0.586***	0.516**	0.563**
	(0.183)	(0.209)	(0.233)	(0.265)	(0.189)	(0.209)	(0.234)	(0.246)
ln Lproductivity(-1)	-0.730**	-0.884**	-0.552*	-0.574	-0.681*	-0.828**	-1.290***	-1.336***
	(0.341)	(0.393)	(0.313)	(0.355)	(0.358)	(0.368)	(0.439)	(0.437)
Debt_assets(-1)	-2.301***	-2.793***	-0.202	-0.268	-1.775**	-2.113**	-4.143***	-4.658***
	(0.756)	(0.833)	(0.812)	(0.919)	(0.873)	(0.943)	(0.937)	(1.006)
lnDist	0.008	0.062	-0.446	-0.537	-0.118	-0.101	0.036	0.060
	(0.153)	(0.174)	(0.365)	(0.391)	(0.192)	(0.205)	(0.217)	(0.231)
lnGDP	-0.045	-0.125	0.205	0.265	-0.066	-0.161	-0.218	-0.249
	(0.112)	(0.124)	(0.222)	(0.236)	(0.131)	(0.136)	(0.161)	(0.167)
lnGDPpc	-0.317	-0.464	-1.126**	-1.092*	-0.456	-0.692**	-0.679*	-0.643
	(0.256)	(0.314)	(0.560)	(0.581)	(0.317)	(0.341)	(0.404)	(0.420)
Constant	-2.085	12.353*	0.641	-0.809	-1.055	16.923**	-2.112	0.771
	(4.316)	(6.657)	(6.008)	(8.332)	(4.934)	(7.018)	(5.741)	(9.046)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
Log likelihood	-781.7466	-758.7575	-216.9733	-203.2858	-639.0314	-623.4490	-553.2302	-541.8337
Wald test	chi2(18)= 193.35***	chi2(20)= 186.66***	chi2(19)= 46.03***	chi2(21)= 32.58***	chi2(18)= 177.15***	chi2(20)= 138.54***	chi2(18)= 161.41***	chi2(20)= 147.99***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)	1102.3***	1061.28	139.98***	110.18***	581.49***	525.49	288.90***	267.81***
Observations	336,484	336,484	371,962	371,962	265,050	265,050	396,920	396,920
No. of firm_market_prod	216,899	216,899	239,516	239,516	176,958	176,958	255,152	255,152

Note: Standard errors in in round brackets; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Random effects probit model of integration at firm-market-product level for input knowledge intensity downstream (i.e., double-split subsample), *rho*

	(1) RE probit	(2) RE probit	(3) RE probit	(4) RE probit
	Comp	Comp	Subst	Subst
rel_upst_knint$_k$	-14.724**	-19.229*	9.028	2.656
	(6.715)	(10.265)	(6.682)	(5.878)
lnIPR	-12.718***	-24.956***	-2.519	-2.830
	(3.794)	(8.578)	(4.071)	(6.077)
Upstr		-11.422**		-1.741
		(5.846)		(4.099)
lnIPR ∗ Upstr		6.785*		0.601
		(3.921)		(2.808)
lnSize(-1)	1.372***	1.649***	0.798***	0.895***
	(0.386)	(0.527)	(0.209)	(0.221)
Age	0.206***	0.178**	0.201***	0.227***
	(0.064)	(0.073)	(0.032)	(0.039)
Ex_prop(-1)	6.042***	6.848***	4.518***	3.947***
	(1.174)	(1.899)	(1.143)	(1.206)
ln Kintensity(-1)	1.994***	2.939***	-0.972***	-1.082***
	(0.478)	(0.594)	(0.271)	(0.297)
ln Lproductivity(-1)	-1.724**	-1.969**	-1.208**	-0.971*
	(0.744)	(1.009)	(0.506)	(0.525)
Debt_assets(-1)	-10.170***	-8.125***	-4.070***	-2.957**
	(2.370)	(2.716)	(1.339)	(1.354)
lnDist	-0.116	-0.099	-0.297	-0.322
	(0.435)	(0.505)	(0.320)	(0.339)
lnGDP	0.280	0.235	-0.352*	-0.393**
	(0.324)	(0.407)	(0.190)	(0.192)
lnGDPpc	-0.466	-0.786	-1.096*	-1.220*
	(0.643)	(0.836)	(0.599)	(0.629)
Constant	0.952	20.036	9.223	18.942
	(11.297)	(17.671)	(10.527)	(12.823)
Time dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
Log likelihood	-278.9142	-268.8527	-468.8406	-461.2941
Wald test	chi2(17)=	chi2(19)=	chi2(17)=	chi2(19)=
	115.95***	89.97***	436.70***	227.04***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)				
	212.83***	181.28	826.32***	787.54***
Observations	155,087	155,087	200,575	200,575
No. of firm_market_prod	104,585	104,585	126,215	126,215

Note: Standard errors in in round brackets; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 10: Random effects probit model of integration at firm-market-product level for knowledge intensity downstream (i.e., double-split subsample); alternative combined $rho \times alpha$ and $rho - alpha$ measures

	(1) $rho \times alpha$ (ind.)	(2) $rho \times alpha$ (ind.)	(3) $rho \times alpha$ (ind.)	(4) $rho \times alpha$ (ind.)	(5) rho-alpha (elast.)	(6) rho-alpha (elast.)	(7) rho-alpha (elast.)	(8) rho-alpha (elast.)
	Comp	Comp	Subst	Subst	Comp	Comp	Subst	Subst
rel_upst_knint$_k$	-5.005	-5.377	13.60*	12.866*	-5.238	-5.274	10.54**	9.126**
	(3.564)	(3.666)	(7.047)	(7.467)	(3.764)	(3.738)	(4.233)	(3.996)
lnIPR	-3.685	-8.351**	-4.432	-0.416	-5.045**	-10.49***	0.467	1.309
	(2.288)	(3.413)	(3.537)	(7.383)	(2.151)	(3.769)	(3.047)	(5.154)
Upstr		-4.501*		2.087		-4.904*		-0.244
		(2.393)		(3.543)		(2.566)		(2.957)
lnIPR $*$ Upstr		2.695*		-3.822		2.958*		-0.331
		(1.620)		(2.803)		(1.747)		(1.998)
lnSize(-1)	0.0849	0.159	0.620***	0.800***	0.184	0.260*	0.365***	0.459***
	(0.107)	(0.115)	(0.212)	(0.281)	(0.154)	(0.155)	(0.110)	(0.114)
Age	0.269***	0.264***	0.0257	0.035**	0.240***	0.233***	0.102***	0.099***
	(0.0383)	(0.039)	(0.0289)	(0.033)	(0.0425)	(0.043)	(0.0266)	(0.026)
Ex_prop(-1)	5.365***	5.458***	2.098***	1.708**	5.780***	5.959***	3.872***	3.700***
	(1.029)	(1.078)	(0.751)	(0.871)	(1.404)	(1.470)	(0.827)	(0.867)
ln Kintensity(-1)	0.771***	0.844***	0.0383	0.111	0.936***	0.976***	0.297	0.322
	(0.241)	(0.244)	(0.307)	(0.373)	(0.245)	(0.243)	(0.234)	(0.237)
ln Lproductivity(-1)	-1.329***	-1.388***	0.0876	0.091	-1.443***	-1.445**	-1.029**	-1.035**
	(0.462)	(0.456)	(0.511)	(0.602)	(0.397)	(0.395)	(0.442)	(0.437)
Debt_assets(-1)	-3.220***	-3.455***	0.886	0.809	-3.038***	-3.136***	-4.082***	-4.172***
	(0.962)	(0.979)	(1.011)	(1.235)	(1.087)	(1.079)	(0.924)	(0.932)
lnDist	-0.0935	-0.089	-0.560	-0.908	-0.242	-0.260	0.00715	0.010
	(0.181)	(0.192)	(0.476)	(0.596)	(0.216)	(0.222)	(0.206)	(0.217)
lnGDP	-0.233*	-0.281**	0.255	0.398	-0.319**	-0.348**	-0.236	-0.253*
	(0.131)	(0.133)	(0.274)	(0.331)	(0.149)	(0.149)	(0.146)	(0.149)
lnGDPpc	-0.554*	-0.625*	-0.884	-0.960	-0.625*	-0.671*	-0.765**	-0.717*
	(0.315)	(0.350)	(0.751)	(0.854)	(0.366)	(0.385)	(0.389)	(0.403)
Constant	8.064	18.052**	-18.02	-20.600	14.14**	24.189***	-5.045	-3.779
	(5.947)	(7.865)	(11.13)	(14.516)	(6.364)	(8.173)	(6.888)	(8.812)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
Log likelihood	-656.8959	-648.3597	-154.9976	-268.8527	-522.7978	-514.4387	-493.4107	-481.2305
Wald test	chi2(17)=	chi2(19)=	chi2(18)=	chi2(20)=	chi2(17)=	chi2(19)=	chi2(17)=	chi2(19)=
	98.50***	106.32***	23.35	25.18	90.48***	93.79***	130.06***	120.57***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)	896.09***	862.98***	103.50***	64.78***	386.85***	356.24***	213.54***	194.69***
Observations	179,011	179,011	154,932	155,087	149,175	149,175	197,972	197,972
No. of firm_market_prod.	117,954	117,954	104,408	104,585	103,857	103,857	129,282	129,282

Note: Standard errors in in round brackets; $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 11: Random effects probit model of integration at firm-market-product level for knowledge intensity downstream (i.e., double-split subsample) on subsample of firms with increasing concentration of sourcing from one country, *rho*-based

	(1) above 10%	(2) above 10%	(3) above 20%	(4) above 20%	(5) above 30%	(6) above 30%
	Comp	Subst	Comp	Subst	Comp	Subst
rel_upst_knint$_k$	-19.23*	2.656	-20.04*	0.542	-30.18***	-1.048
	(10.27)	(5.878)	(10.84)	(5.476)	(9.490)	(7.864)
lnIPR	-24.96***	-2.830	-25.54***	-1.113	-23.64***	4.430
	(8.578)	(6.077)	(8.310)	(9.252)	(7.715)	(7.136)
Upstr	-11.42*	-1.741	-11.77**	-0.722	-11.16*	1.746
	(5.846)	(4.099)	(5.889)	(5.010)	(6.563)	(4.370)
lnIPR * Upstr	6.785*	0.601	7.013*	-0.0884	6.006	-1.742
	(3.921)	(2.808)	(3.955)	(3.430)	(4.375)	(2.979)
lnSize(-1)	1.649***	0.895***	1.698***	0.829***	1.994***	1.067***
	(0.527)	(0.221)	(0.569)	(0.200)	(0.571)	(0.333)
Age	0.178**	0.227***	0.181**	0.202***	0.0554	0.356***
	(0.0733)	(0.0388)	(0.0753)	(0.0395)	(0.0679)	(0.0548)
Ex_prop(-1)	6.848***	3.947***	7.089***	3.818***	5.679***	11.50***
	(1.899)	(1.206)	(2.010)	(1.148)	(2.035)	(3.666)
ln Kintensity(-1)	2.939***	-1.082***	3.034***	-0.975***	2.897***	-1.683**
	(0.594)	(0.297)	(0.586)	(0.275)	(0.632)	(0.711)
ln Lproductivity(-1)	-1.969*	-0.971*	-1.998*	-0.877*	-1.152	-1.687
	(1.009)	(0.525)	(1.062)	(0.493)	(0.939)	(1.415)
Debt_assets(-1)	-8.125***	-2.957**	-8.318***	-2.597**	-9.874***	-3.007
	(2.716)	(1.354)	(2.772)	(1.281)	(2.686)	(2.141)
lnDist	-0.0988	-0.322	-0.107	-0.303	-0.193	-0.354
	(0.505)	(0.339)	(0.513)	(0.325)	(0.446)	(0.417)
lnGDP	0.235	-0.393**	0.244	-0.387*	0.349	-0.406*
	(0.407)	(0.192)	(0.405)	(0.205)	(0.348)	(0.217)
lnGDPpc	-0.786	-1.220*	-0.825	-1.176*	-0.0730	-1.363
	(0.836)	(0.629)	(0.855)	(0.656)	(0.735)	(0.841)
Constant	20.04	18.94	20.31	18.35	17.84	13.50
	(17.67)	(12.82)	(18.15)	(14.31)	(17.29)	(18.52)
Time dummies	yes	yes	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes	yes	yes
Log likelihood	-268.8527	-461.2941	-268.7725	-462.5582	-196.2061	-387.2041
Wald test	chi2(19)=	chi2(19)=	chi2(19)=	chi2(19)=	chi2(17)=	chi2(18)=
	89.97***	227.04***	99.05***	197.13***	74.40***	73.26***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)						
	181.28***	787.54***	181.34***	781.45***	158.96***	634.44***
Observations	155,087	200,575	154,321	195,135	109,507	126,557
No. of firm_market_prod	104,585	126,215	104,174	124,701	75,496	83,042

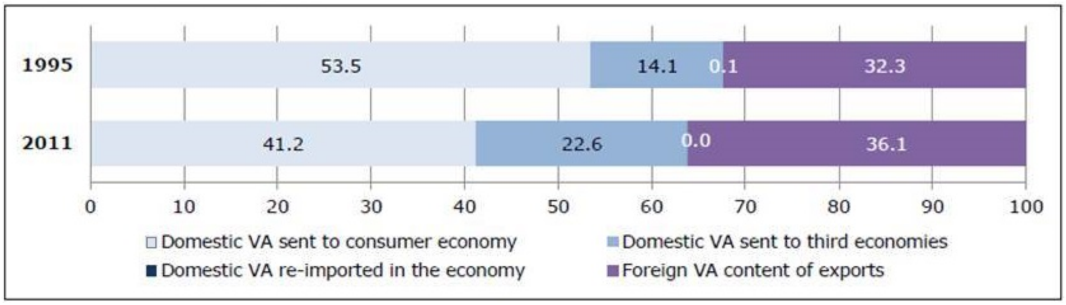
Note: Standard errors in in round brackets; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 12: Random effects probit model of integration at firm-market-product level for knowledge intensity downstream (i.e., double-split subsample) augmented with WGI interactions, *rho*-based

	(1)	(2)	(3)	(4)	(5)	(6)
	Rule of law		Govern effectiveness		Control corruption	
	Comp	Subst	Comp	Subst	Comp	Subst
rel_upst_knint$_k$	-16.72*	1.347	-19.45**	2.314	-19.24*	3.414
	(9.562)	(5.557)	(9.650)	(5.774)	(10.45)	(6.244)
lnIPR	-24.49***	-1.529	-26.13***	-1.688	-25.96***	-4.872
	(7.740)	(7.058)	(8.371)	(7.166)	(8.315)	(10.18)
Upstr	-11.54**	-0.535	-11.91**	-0.515	-12.19**	-2.494
	(5.487)	(4.341)	(5.935)	(4.303)	(5.950)	(5.269)
lnIPR ∗ Upstr	7.289*	-0.291	7.333*	-0.342	7.505*	1.171
	(3.834)	(3.061)	(4.060)	(3.045)	(4.126)	(3.665)
WGI	1.244	-0.719	0.323	-0.835	0.596	0.119
	(1.217)	(0.799)	(1.259)	(0.879)	(1.107)	(0.716)
WGI ∗ Upstr	-0.454	0.103	-0.276	0.174	-0.245	-0.112
	(0.521)	(0.343)	(0.566)	(0.384)	(0.491)	(0.308)
lnSize(-1)	1.491***	0.832***	1.632***	0.865***	1.641***	0.942***
	(0.477)	(0.211)	(0.505)	(0.220)	(0.550)	(0.230)
Age	0.157**	0.210***	0.181**	0.219***	0.174**	0.229***
	(0.0652)	(0.0383)	(0.0725)	(0.0369)	(0.0729)	(0.0376)
Ex_prop(-1)	6.019***	3.892***	6.822***	3.952***	6.764***	4.006***
	(1.557)	(1.180)	(1.805)	(1.203)	(1.799)	(1.223)
ln Kintensity(-1)	2.530***	-1.008***	2.928***	-1.047***	2.881***	-1.107***
	(0.587)	(0.289)	(0.583)	(0.297)	(0.576)	(0.301)
ln Lproductivity(-1)	-1.603*	-0.935*	-2.006**	-0.954*	-1.859*	-0.975*
	(0.896)	(0.507)	(0.986)	(0.520)	(1.026)	(0.540)
Debt_assets(-1)	-7.255***	-2.683**	-8.121***	-2.922**	-7.965***	-3.254**
	(2.431)	(1.303)	(2.700)	(1.329)	(2.659)	(1.396)
lnDist	-0.0859	-0.268	-0.0754	-0.244	-0.122	-0.332
	(0.486)	(0.315)	(0.526)	(0.326)	(0.537)	(0.363)
lnGDP	0.201	-0.401**	0.232	-0.397**	0.236	-0.367*
	(0.371)	(0.190)	(0.410)	(0.196)	(0.404)	(0.215)
lnGDPpc	-1.064	-0.476	-0.565	-0.549	-0.960	-1.017
	(0.958)	(0.853)	(1.043)	(0.830)	(1.099)	(1.094)
Constant	22.87	12.17	19.83	11.50	22.54	17.71
	(16.48)	(13.75)	(18.35)	(13.73)	(18.49)	(16.06)
Time dummies	yes	yes	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes	yes	yes
Log likelihood	-269.1421	-462.5164	-268.7403	-461.6240	-268.9165	-460.5996
Wald test	chi2(21)=	chi2(21)=	chi2(21)=	chi2(21)=	chi2(21)=	chi2(21)=
	77.74***	194.50***	90.01***	227.25***	93.25***	280.49***
Likelihood-ratio test; rho=0: chi2(1) ($Prob > chi2$)						
	179.55***	783.60***	180.36***	785.71***	180.75***	783.93***
Observations	155,087	200,575	155,087	200,575	155,087	200,575
No. of firm_market_prod	104,585	126,215	104,585	126,215	104,585	126,215

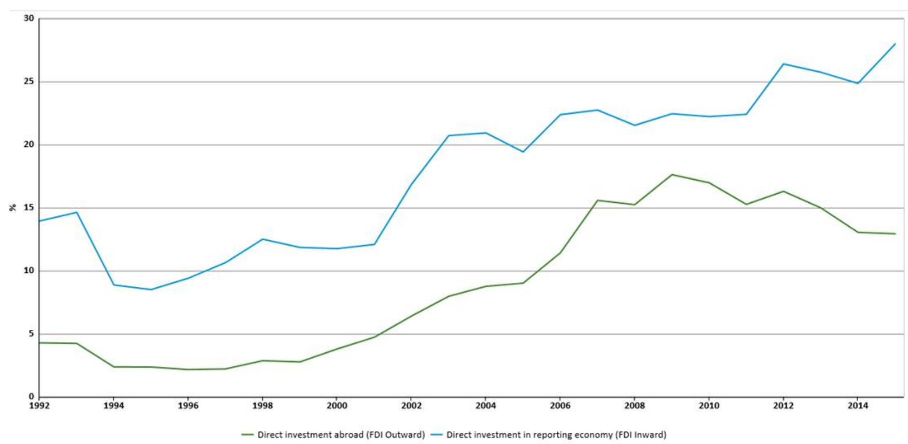
Note: Standard errors in in round brackets; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Figure 1: The value-added (VA) components of gross exports, Slovenia 1995 and 2011.
(% share in total gross export)



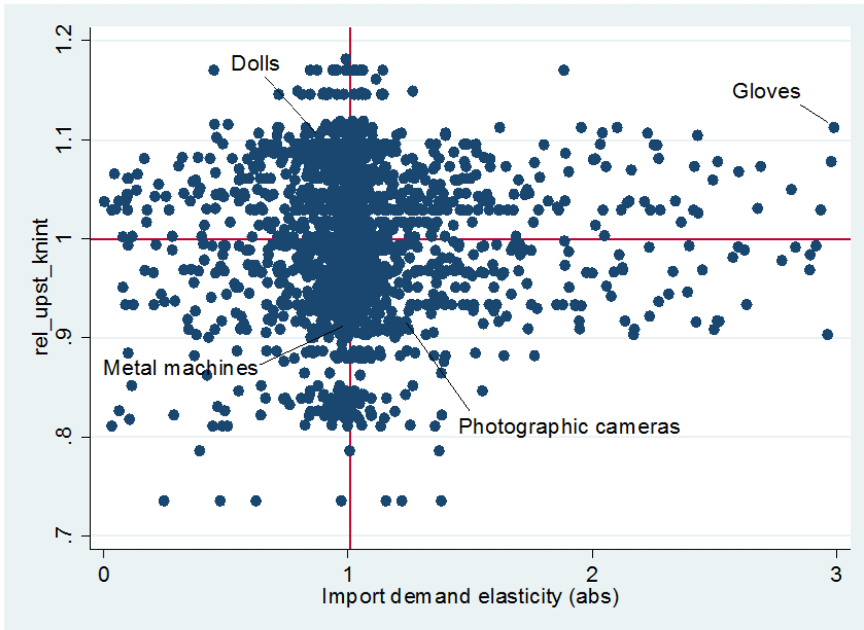
Source: WTO.

Figure 2: Slovenian FDI stock.
(% of GDP)



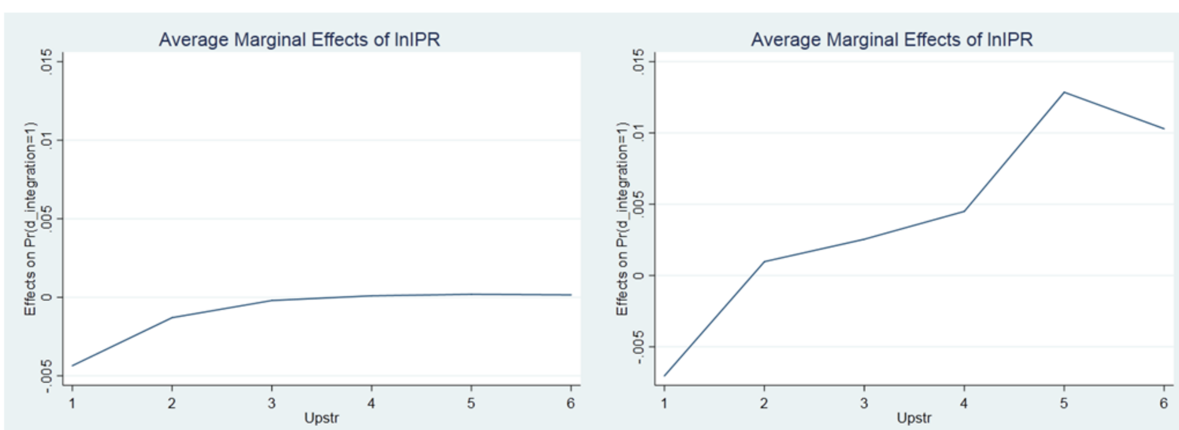
Source: WTO.

Figure 3: Relative upstreamness of knowledge intensive inputs (rel_upst_knint) and import demand elasticity (ρ) of product groups



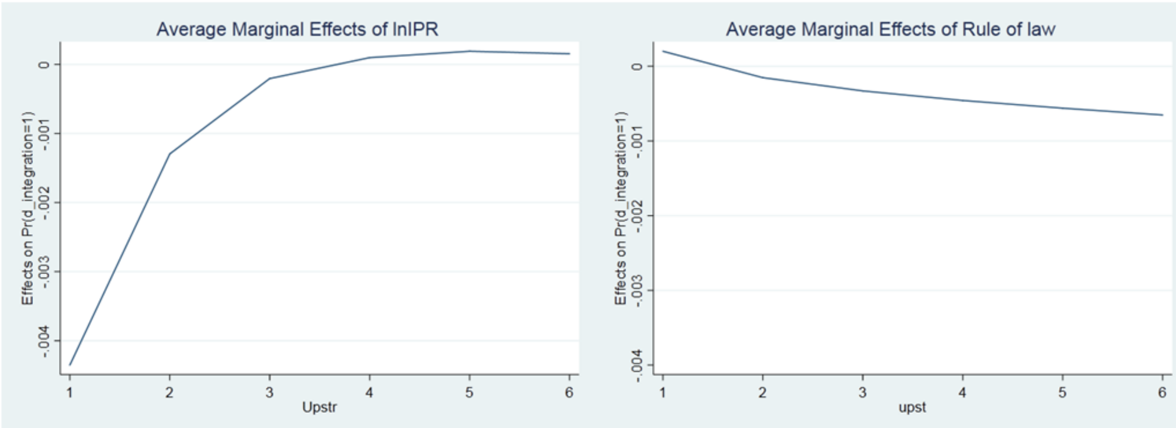
Notes: 950210 - Dolls; representing only human beings; 611692 - Gloves, mittens and mitts; of cotton, knitted or crocheted, (other than impregnated, coated or covered with plastics or rubber); 900610 - Cameras, photographic (excluding cinematographic); of a kind used for preparing printing plates or cylinders; 845730 - Metal machines; multi-station transfer machines, for working metal.

Figure 4: (Average) Marginal effects of the quality of IPR institutions for complements: Knowledge intensity increasing downstream (left) and decreasing downstream (right)



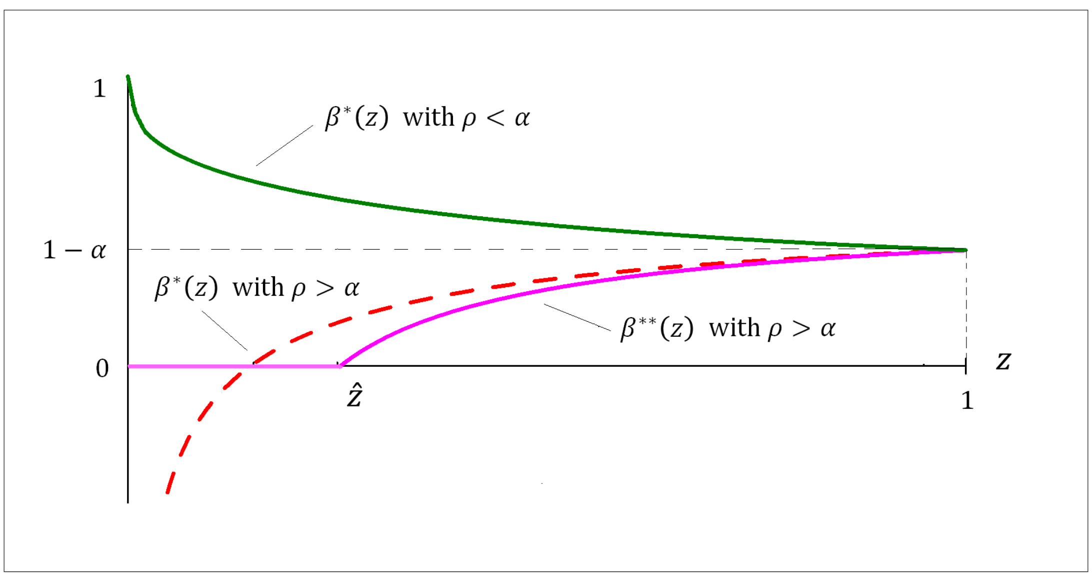
Notes: Based on regression from Table 5, columns 1-2.

Figure 5: (Average) Marginal effects of the quality of institutions for complements with relatively more knowledge transmission upstream: IPRs (left) versus Rule of Law (right)



Notes: Based on regression from Table 5, columns 1 and 3.

Figure 6: Profit-maximizing division of surplus along the supply chain (relaxed problem).



CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1672	Timothy Besley Isabel Roland John Van Reenen	The Aggregate Consequences of Default Risk: Evidence from Firm-Level Data
1671	Jan-Emmanuel De Neve Daisy Fancourt Christian Krekel Richard Layard	A Local Community Course That Raises Mental Wellbeing and Pro-Sociality
1670	Tommaso Sonno	Globalization and Conflicts: The Good, the Bad and the Ugly of Corporations in Africa
1669	Michael Amior	Immigration, Local Crowd-Out and Undercoverage Bias
1668	Antoine Berthou John Jong-Hyun Chung Kalina Manova Charlotte Sandoz Dit Bragard	Trade, Productivity and (Mis)allocation
1667	Holger Breinlich Elsa Leromain Dennis Novy Thomas Sampson	Exchange Rates and Consumer Prices: Evidence from Brexit
1666	Fabrice Defever Michele Imbruno Richard Kneller	Trade Liberalization, Input Intermediaries and Firm Productivity: Evidence from China
1665	Philippe Aghion Antonin Bergeaud Richard Blundell Rachel Griffith	The Innovation Premium to Soft Skills in Low-Skilled Occupations
1664	Filip Gesiarz Jan-Emmanuel De Neve Tali Sharot	The Motivational Cost of Inequality: Pay Gaps Reduce the Willingness to Pursue Rewards

1663	Felix Koenig	Technical Change and Superstar Effects: Evidence From the Roll-Out of Television
1662	Enrico Moretti Claudia Steinwender John Van Reenen	The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers
1661	Decio Coviello Andrea Ichino Nicola Persico	Measuring the Gains from Labor Specialization
1660	Nicolás González-Pampillón	Spillover Effects from New Housing Supply
1659	Walter Bossert Andrew E. Clark Conchita D'Ambrosio Anthony Lepinteur	Economic Insecurity and the Rise of the Right
1658	Paul Frijters Andrew E. Clark Christian Krekel Richard Layard	A Happy Choice: Wellbeing as the Goal of Government
1657	Philippe Aghion Antonin Bergeaud Matthieu Lequien Marc Melitz	The Heterogeneous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports
1656	Clare Leaver Renata Lemos Daniela Scur	Measuring and Explaining Management in Schools: New Approaches Using Public Data
1655	Clément S. Bellet Jan-Emmanuel De Neve George Ward	Does Employee Happiness Have an Impact on Productivity?

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE