New Imported Inputs, New Domestic Products*

Italo Colantone†
Erasmus University Rotterdam
Tinbergen Institute and ERIM

Rosario Crinò‡
CEMFI

October 2013

Abstract

We study the relationship between new imported inputs and the introduction of new domestic products. To this purpose, we assemble a novel data set covering 25 European countries over 1995-2007 and containing information on domestic production and bilateral trade for the universe of goods. We develop a procedure to identify new imported inputs and new domestic products, while dealing with the complications raised by the yearly changes in the commodity classifications. We augment these data with information on prices and novel estimates of quality. We organize the empirical analysis around a version of the endogenous growth model with expanding variety, in which inputs are allowed to be heterogeneous in terms of quality. In line with this framework, we find three main results. First, new imported inputs have a strong positive effect on product creation in Europe. Second, they work through a combination of mechanisms, allowing countries to benefit from both wider and better sets of intermediate products. Finally, new imported inputs give a substantial boost to output growth in manufacturing.

JEL codes: F1.

Keywords: New Intermediates; New Goods; Output Growth; Prices; Quality; Product-Level Data.

*The paper has greatly benefited from the comments of two anonymous referees. We are especially grateful to Paolo Epifani for his precious suggestions. We also thank Carlo Altomonte, Pablo Arocena, Tibor Besedeš, Stéphane Bonhomme, Davide Castellani, Klaus Desmet, David Dorn, Jesús Fernández-Huertas Moraga, Gino Gancia, Rodolfo Helg, Amit Khandelwal, Miklós Koren, Claudio Michelacci, Samuele Murtinu, Paolo Naticchioni, Laura Ogliari, Nina Pavcnik, Enrico Pennings, Mark Roberts, Leo Sleuwaegen, Ilke Van Beveren, Maurizio Zanardi, and seminar participants at many venues for useful comments and discussions. Rosario Crinò gratefully acknowledges financial support from the Efige project (N. 225551), funded by the European Commission under the 7th Framework Programme. The usual disclaimer applies.

†Address: Department of Applied Economics. Burgemeester Oudlaan 50, 3000 DR, Rotterdam, The Netherlands. E-mail: colantone@ese.eur.nl.

‡*Corresponding author.* Address: Centro de Estudios Monetarios y Financieros (CEMFI). Casado del Alisal 5, 28014, Madrid, Spain. E-mail: crino@cemfi.es.

1 Introduction

A key message of the endogenous growth literature is that countries can sustain long-run growth by producing new and upgraded goods (see e.g. Aghion and Howitt, 2005, and Gancia and Zilibotti, 2005). It has long been argued that international trade can stimulate the introduction of new products in a country, through the efficiency gains arising when producers get access to new input varieties from abroad (Rivera-Batiz and Romer, 1991; Backus, Kehoe, and Kehoe, 1992). Despite the prominence of these models in the theoretical literature, empirical research on this issue is extremely limited. In an influential paper on India, Goldberg et al. (2010a) have provided the first, and so far only, evidence of a positive link between new imported inputs and new domestic products.

In this paper, we use unique micro data for a large group of developed countries to provide novel evidence on the effects of new imported inputs on product creation, and to answer two important questions that remain open in the empirical literature. First, what are the mechanisms through which new imported inputs operate? And second, what are the implications of new imported inputs for growth? Our results confirm that new imported inputs are a crucial determinant of product creation. More importantly, our analysis shows that new imported inputs work through the interaction of different mechanisms, allowing countries to widen the set of available intermediates and to access superior varieties. Consistent with these results, we also find new imported inputs to be an important stimulus to output growth in manufacturing. Overall, our paper portrays a complete picture of the relationship between product creation and access to new foreign inputs. The main point we make is that new imported inputs have a pervasive effect on product creation across countries, work through a complex combination of mechanisms, and constitute an important engine of growth.

Our analysis relies on a novel data set covering 25 countries of the European Union (EU) over 1995-2007. For each country, we have information on domestic production and bilateral trade for the universe of products, at the highest possible level of disaggregation (8 digits). The first task we accomplish with these data is to identify new domestic goods and new imported inputs. This task is extremely challenging, as the commodity classifications are revised every year by the European authorities. We thus develop a new procedure that keeps track of all classification changes using correspondence tables, yielding a precise indication of which products and foreign inputs are new in each country every year. Introduction of new products and imports of new intermediates are relevant phenomena in the EU countries. According to our data, new products account for 5% of all goods produced domestically each year, and their introduction is responsible for 25% of the annual growth in manufacturing output. Similarly, new foreign inputs make 13% of all input varieties imported each year, and account for 20% of the annual growth in intermediates imports.

To guide our empirical analysis and provide the key insights for interpreting our results, we start by presenting a simple theoretical framework based on the benchmark version of the endogenous growth model with expanding variety (Rivera-Batiz and Romer, 1991). In this framework, new products are invented through research and development (R&D), according to a 'lab equipment' technology which implies that all factors, including intermediate inputs, are productive in research. In the model, new imported inputs widen the set of available intermediates and thereby generate a 'scale effect' that raises productivity in research. In equilibrium, this efficiency gain leads to greater product creation and faster output growth. We then present an extension of this model in which intermediate inputs are allowed to be heterogeneous in terms of quality, similar to Aghion and Howitt (1998). This more general framework delivers the additional prediction that differences in product characteristics may amplify the scale effect of new imported inputs. Specifically, the model predicts the efficiency gain from importing new inputs to be larger the lower their quality-adjusted price. The intuition is that, in this extended model, new imported inputs not only widen the set of available intermediates, but may also change its composition toward varieties with more favorable price-quality ratios.

Having presented the theoretical framework we turn to the empirical analysis. As a starting point, we provide extensive evidence of a strong positive correlation between new imported inputs and new domestic products within countries and industries. Then, we move to instrumental variables (IV) regressions to address reverse causality. Indeed, unobserved shocks to specific industries and EU countries may lead to product creation for reasons unrelated to foreign intermediates; but once the decision to produce a new good has been made, firms may start sourcing the necessary new inputs from abroad. This would induce an upward bias in the OLS estimates. Thus, we construct an instrument capturing variation in new imported inputs not due to industry-specific shocks in the EU countries. In particular, the instrument captures variation due to changes in transportation costs, as induced by fluctuations in oil prices (Hummels, 2007; Hummels et al., 2013). The instrument turns out to be a strong predictor of new imported inputs, in the direction one would expect. At the same time, in the IV regressions, the coefficient on new imported inputs remains positive and highly significant. The size of this coefficient is roughly half the size of the baseline correlation estimated by OLS, suggesting that the instrument removes the upward bias induced by reverse causality. Overall, this first part of the analysis shows that new imported inputs are an important determinant of the introduction of new goods in the EU.

Next, we study the mechanisms through which new imported inputs operate. According to the model, new imported inputs stimulate product creation by generating efficiency gains, through two channels. First, they give rise to scale effects by expanding the range of available intermediates. Second, they may allow countries to access better varieties, i.e. varieties with lower quality-adjusted prices. To evaluate the empirical relevance of these two mechanisms, we need to measure the quality-adjusted price of each input variety imported into each EU country. While our data contain information on raw prices (the numerator of quality-adjusted prices), obviously we do not observe quality (the denominator) and must therefore estimate it. We do so using the methodology developed by Khandelwal (2010). As a result, we construct an extremely detailed and widely comprehensive data set, containing time-varying quality estimates for all input varieties imported into each EU country; to the best of our knowledge, no such data set existed before. Using these estimates, we construct quality-adjusted prices and find robust evidence in favor of both mechanisms. In particular, we show that new imported inputs boost product creation even when they have the same quality-adjusted price as the existing intermediates, consistent with a pure scale effect. We also show, however, that the effect of new imported inputs is decreasing in their quality-adjusted price, consistent with the idea that new imported inputs also work by changing the composition of the inputs set toward superior varieties. Overall, this second part of the analysis suggests that new imported inputs stimulate product creation through a combination of mechanisms, allowing countries to benefit from both wider and better sets of intermediate products.

In the final part of the analysis, we discuss the implications of new imported inputs for growth, and provide suggestive evidence on the characteristics of new goods. In endogenous growth models with expanding variety—such as the model used in this paper—the introduction of new products constitutes technical progress and thus acts as the 'engine of growth' for the economy. Our evidence that new imported inputs stimulate product creation would therefore suggest that they should also have a positive impact on growth. Indeed, we do find robust evidence of such an effect in our data. In particular, we show that new imported inputs substantially increase the growth rate of manufacturing output per worker, even after accounting for other determinants of scale effects and growth studied in the literature (see in particular Backus, Kehoe, and Kehoe, 1992). Finally, we close the paper by studying the characteristics of new goods. Consistent with existing extensions of the expanding variety model, we find new goods to be upgraded, i.e. characterized on average by higher quality and prices compared to old products.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 presents the data and some stylized facts. Section 4 illustrates the theoretical framework underlying our empirical analysis, which is performed in Section 5. Finally, Section 6 concludes.

2 Related Literature

Our paper speaks to different strands of empirical literature. In particular, as mentioned in the introduction, it is related to recent work by Goldberg et al. (2010a). Exploiting India's trade liberalization as an exogenous trade shock, the authors identify a large positive effect of new foreign inputs on the number of goods produced within firms. However, they do not study the mechanisms through which new imported inputs operate, the characteristics of new goods, or the implications of new imported inputs for growth. These are the main contributions of our paper. Moreover, we depart from Goldberg et al. (2010a) also in other important directions. First, we focus on a large group of industrialized countries, as opposed to a fast-growing developing economy. By doing so we show, for the first time, that new imported inputs are a fundamental engine of product creation and growth not only in the developing world, but also in advanced economies that lie closer to the technological frontier. Second, we focus on the introduction of new products at the economy-wide level, not at the firm level. This departure is most relevant especially in developed countries such as the ones we study, since products that are new for a firm need not be new also for the economy as a whole.

Apart from Goldberg et al. (2010a), empirical evidence on the link between new imported inputs and new domestic products has been lacking. One of the main reasons has been the unavailability of detailed data on domestic production. In different contexts, domestic production has often been proxied using data on exported goods.[1] However, proxies based on export data do not perfectly account for the introduction of new products, since some goods may not be exported, at least initially. To the best of our knowledge, we are the first to employ data on domestic production for the universe of products, in many countries and years, to study how trade affects the creation of new goods.[2]

Our paper is also related to a large empirical literature on the relationship between trade and growth. According to this literature the relationship is unclear (see Kehoe ad Ruhl, 2010, for a recent review).[3] Indeed, many cross-country studies have shown a positive correlation between trade openness and GDP growth, but the direction of causality has generally been hard to establish. Moreover, little evidence has been found on the type of scale effects predicted by endogenous growth models after opening to trade. Yet, an influential paper by Backus, Kehoe, and Kehoe (1992) has shown that scale effects do arise when focusing on manufacturing, and that proxies for intermediates trade implied by the theory of Rivera-Batiz and Romer (1991) are strongly correlated with manufacturing output growth. Our work is similar to Backus, Kehoe, and Kehoe (1992) in these two important aspects, as we also focus on manufacturing and use proxies for trade in intermediates inspired by models in the Rivera-Batiz and Romer (1991) tradition. Indeed, we do confirm the presence of scale effects in manufacturing and, notably, the robust correlation of these proxies with the growth rate of manufacturing output per worker.

Our work is also connected to the recent literature on the characteristics of new products, in particular to Broda and Weinstein (2010) and Xiang (2005). Using bar-code data on the purchases of US households, Broda and Weinstein (2010) show that inflation estimates based on the conventional Consumer Price Index are upward biased, as the CPI is computed on a fixed set of goods and is thus unable to accommodate the higher quality of new products. While the authors' data are extremely well suited to track new consumption patterns, our data are very well suited to identify the production of new goods within a country, which is at the core of our paper. Instead, Xiang (2005) shows that new products are responsible for a large fraction of the increase in US wage inequality, since their production is high-skill labor intensive compared to that of old goods. Unlike Xiang (2005) we study the determinants, not the consequences, of the

[1] Examples include Feenstra et al. (1999), Broda, Greenfield, and Weinstein (2006), Bas and Strauss-Kahn (2011), Feng, Li, and Swenson (2012), and Aristei, Castellani, and Franco (2013).

[2] A few other studies have used the same production data employed in this paper to analyze different issues. For example, Bernard et al. (2012) have documented new facts on the behavior of Belgian firms.

[3] A non-exhaustive list of contributions include Frankel and Romer (1999), Alcalá and Ciccone (2004), Dollar and Kraay (2004), Alesina, Spolaore, and Wacziarg (2005), and Spolaore and Wacziarg (2005).

introduction of new products.[4]

Finally, our paper is linked to two other streams of research. The first studies the effect of imported inputs on domestic productivity. With a few exceptions, the existing papers find such an effect to be positive and economically large.[5] In the endogenous growth literature, this productivity effect would be referred to as 'level effect' or 'static gain' (see e.g. Rivera-Batiz and Romer, 1991). Instead, our interest lies in the so called 'growth effect' or 'dynamic gain', which works through the introduction of new products and has been overlooked in the empirical literature. The second stream of research deals with the welfare effects of new imported varieties *in general*. With a few exceptions, the available studies find new foreign varieties to bring about substantial welfare gains.[6] Our analysis complements these studies by unveiling new sources of gains associated with the extensive margin of trade, namely the positive effects that new imported inputs have on the introduction of new goods.[7]

3 Data and Stylized Facts

3.1 Data

We source data on domestic production from Prodcom (PC), a database administered by Eurostat. PC covers all EU countries and contains annual information on the value and volume of sold production for the universe of products.[8] The data are based on an annual survey of firms' production activities within the territory of each reporting country. The survey covers the entire manufacturing sector and, according to the EU regulation, must encompass at least 90% of the annual production of each 4-digit industry in each country. Importantly for our purposes, the survey does not cover production activities undertaken outside the national borders, e.g. in foreign subsidiaries of domestic multinationals. As for the level of product aggregation, the PC classification contains roughly 4,500 8-digit product codes. This classification can be directly linked to the classification of industrial activities in the EU (NACE Rev. 1.1), as the first four digits of the PC code identify a 4-digit NACE industry. This feature enables us to easily map products into industries. As for the time coverage, the data are available since 1995, with some differences across countries (see Table A3). We limit the analysis to the period 1995-2007, as the PC classifi-

[4] To identify new goods, Xiang (2005) compares the 1972 and 1987 versions of the Standard Industrial Classification, and defines products as new if they are absent in the former but present in the latter version of the classification.

[5] See Trefler (2004), Amiti and Konings (2007), Kasahara and Rodrigue (2008), Kehoe and Ruhl (2008), Kugler and Verhoogen (2009, 2012), Sivadasan (2009), Halpern, Koren, and Szeidl (2011), Khandelwal and Topalova (2011), Boeler, Moxnes, and Ulltveit-Moe (2012), Ramanarayanan (2012), Saravia and Voigtländer (2012), Gopinath and Neiman (2013), and Kasahara and Lapham (2013). Muendler (2004) is a notable exception.

[6] See Kehoe and Ruhl (2010), Broda and Weinstein (2004, 2006), Broda, Greenfield, and Weinstein (2006), Feenstra (1994), and Feenstra, Markusen, and Zeile (1992). Arkolakis et al. (2008) is a notable exception.

[7] Influential studies quantifying the extensive margin of trade in different countries include Kehoe and Ruhl (2013), Hummels and Klenow (2005), Goldberg et al. (2009), and Besedeš and Prusa (2011).

[8] Data for Cyprus and Malta are confidential, so we exclude these countries from the analysis and focus on 25 rather than 27 EU Members. Belgium and Luxembourg are aggregated by Eurostat and thus constitute a single unit of analysis.

cation has been entirely restructured in 2008 and a complete mapping between the old and new version cannot be produced.[9]

A crucial task for our study is the identification of new products. We define a good as a 'new product' for a country when the first domestic firm starts producing it and thus a positive production is recorded in PC. The identification of new products is dramatically complicated by the changes that occur every year in the PC classification, following the EU legislation. These changes are of two types: (i) new products are added to the classification with new codes; (ii) some of the existing ('old') product codes are converted into new product codes. This second type of change is problematic for our purposes, as it reflects renaming of products rather than true product entry. We identify these cases using year-to-year correspondence tables provided by Eurostat. As a result, when a new code appears in the classification, we know exactly whether it represents a new product or is just a new indicator for one or more existing goods.

Taking this into account, we identify code h, produced by country c in year t, as a new product if either: (1) code h is introduced in the classification in year t and does not have any old code corresponding to it; or (2) code h is introduced in the classification in year t and has one or more old codes corresponding to it, but none of them was produced by country c in any previous year; or (3) code h is not new to the classification, but was not produced by country c in any previous year. With this identification procedure, a product can be counted as new for a country only once: if production resumes after having stopped for a while, this is not counted as entry. Hence, in our data, product entry is not spuriously driven by classification changes or by discontinuities in production over time. Examples of new products for some of the countries in our sample are as follows. Spain started producing 'flat panel video monitors, LDC or plasma' (PC 32302049) in the year 2000; in previous years, the country already produced 'color video monitors with cathode-ray tube' (PC 32302045). The Netherlands started producing 'photocopiers incorporating an optical system' (PC 30012185) in the year 2002; in previous years, the country already produced 'electrostatic photocopiers' (PC 30012170).

As for the trade data, we source them from Comext, another database administered by Eurostat. For all EU countries since 1988, Comext contains annual information on the value and volume of trade (both imports and exports) in the universe of manufacturing products with all trading partners in the world (about 200 countries). The commodity classification used by Comext is the Combined Nomenclature (CN), which contains more than 10,000 8-digit codes. This classification can be linked to the NACE classification through appropriate correspondence tables provided by Eurostat. To identify the intermediate inputs, we also map the CN classification into the Broad Economic Categories (BEC) classification.[10] We then define as inputs all CN codes that belong to the following BEC categories: 'parts and accessories' (BEC 42); 'capital goods, except transport equipment' (BEC 41); 'processed industrial supplies' (BEC 22); 'industrial transport equipment' (BEC 521); 'parts and accessories of transport equipment' (BEC 53); 'processed fuels and lubrificants' (BEC 32); 'processed food and beverages for industry' (BEC 121). This way

[9]The restructuring has followed the shift from NACE Rev. 1.1 to NACE Rev. 2.

[10]We note, instead, that a mapping between the PC and BEC classifications cannot be produced.

of defining inputs is standard, both in the empirical trade literature and in the computation of aggregate trade statistics (e.g. by Eurostat, the OECD, and the United Nations). Our definition is also similar to the one employed by Goldberg et al. (2009), who use an approach based on Input-Output tables.[11]

In most of the paper, we treat each variety (product h - partner n combination) as a different input, thereby following the standard approach in the empirical trade literature (see e.g. Broda and Weinstein, 2006, and Goldberg et al., 2009, 2010a).[12] We define a variety as a 'new imported input' for a country when the product is imported from the trading partner for the first time. The CN classification has also undergone several changes over the sample period. We keep track of all of them using year-to-year correspondence tables provided by Eurostat. We then identify variety v, imported into country c in year t, as new if either: (1) code h is introduced in the classification in year t and does not have any old code corresponding to it; or (2) code h is introduced in the classification in year t and has one or more old codes corresponding to it, but none of them was imported into country c from partner n in any previous year; or (3) code h is not new to the classification, but was not imported into country c from partner n in any previous year. Similar to domestic goods, imported varieties can be counted as new only once. Hence, the identification procedure is not affected by changes in the CN classification or by discontinuities in bilateral trade flows over time.[13]

3.2 Stylized Facts

In Table 1, we report information on the entry of domestic products and imported varieties. As for the latter, we consider the whole sample of goods as well as the subsample of intermediate inputs. All figures are in percentages, averaged across countries, NACE industries, and years. The table shows that new products account for a non-negligible share of all goods produced domestically each year (5%). Similarly, new varieties account for a substantial portion of total imported varieties in both samples (13%). The table also reports figures for the exit rates of domestic goods and imported varieties (4.8 and 10.5%, respectively), implying degrees of churning consistent with firm-level evidence for the US (Bernard et al., 2009; Bernard, Redding, and Schott, 2010).[14]

Next, we decompose the annual growth in production and import value into the contributions of new, exiting, and continuing products and varieties. To this purpose, we use the follow-

[11] In Appendix B, we perform robustness checks showing that our results are unchanged when using narrower definitions of inputs, which exclude capital goods, fuels, and lubrificants.

[12] However, in Appendix B we show that our results do not depend on this choice.

[13] We have written two Stata codes that identify new domestic products and new imported inputs, respectively. In a nutshell, the identification of new products works as follows. Consider a code h for which we observe positive production in country c at time t, but not in previous years. The program first checks for the existence of old codes corresponding to h. If there is none, code h is directly identified as a new good. If instead some old codes exist, the program verifies that country c's production was zero for each of them over all previous years. Only in that case is code h labeled as a new good. This routine runs in approximately one day on a standard computer. The program that identifies new imported inputs works similarly. However, for each code h, the above procedure is repeated across all trading partners of country c. As a consequence, the program takes on average two days for each sample country.

[14] The procedures that identify exiting goods and exiting varieties are specular to those that identify new products and new imported inputs; see the previous footnote.

Table 1: Entry and Exit Rates, %

	Entry	Exit
Domestic products	5.0	4.8
All imported varieties	13.2	10.5
Imported varieties of intermediate inputs	13.1	10.6

Notes: The *entry* rate is the number of new domestic products (new foreign varieties) divided by the total number of domestic goods (foreign varieties). The *exit* rate is the number of exiting products (exiting foreign varieties) divided by the total number of domestic goods (foreign varieties). Figures are averages across countries, NACE industries, and years. *Source*: Eurostat (Prodcom and Comext).

ing formula borrowed from Goldberg et al. (2010b):

$$\frac{X_{cit} - X_{cit-1}}{X_{cit-1}} = \frac{1}{X_{cit-1}} \cdot \left[\sum_{z \epsilon New_{cit}} X^z_{cit} - \sum_{z \epsilon Exiting_{cit}} X^z_{cit-1} + \sum_{z \epsilon Continuing_{cit}} (X^z_{cit} - X^z_{cit-1}) \right], \quad (1)$$

where c indexes countries, i denotes NACE industries, and t stands for years; depending on the specification, the superscript z indexes domestic goods or imported varieties, while X denotes production or import value. The results of these decompositions are presented in Table 2. As before, figures are in percentages, averaged across countries, industries, and years; numbers in italics are normalized by the growth rates reported in the first column. Note that new goods account for one-quarter of the average annual growth in domestic production. Goldberg et al. (2010b) and Bernard, Redding, and Schott (2010) find similar contributions using firm-level data for India and the US, respectively. At the same time, new imported varieties account for 17% of the average annual growth in total imports, and for 20% of the average annual growth in intermediates imports. All in all, these figures suggest that imports of new intermediates and introduction of new products are relevant phenomena in our sample of EU countries.

4 Conceptual Framework

In this section, we use some of the standard tools of the endogenous growth theory to formulate a simple model that clarifies the intuition for why new imported inputs affect product creation, and will help us organize our empirical analysis. To fix ideas and convey the main points in the simplest possible way, we start with a framework resting on the benchmark version of the endogenous growth model with expanding variety (Rivera-Batiz and Romer, 1991).[15] Then, to deliver richer predictions about the mechanisms, we extend this framework by allowing inputs to be differentiated in terms of quality, similar to Aghion and Howitt (1998).

[15] In this part, our exposition will closely follow chapter 13 in Acemoglu (2009).

Table 2: Decomposition of Growth Rates, %

	Growth Rate	*of which*:	New	Exiting	Continuing
Domestic production	9.4		2.3	-1.6	8.6
	100.0		*24.8*	*-16.5*	*91.7*
Overall imports	11.7		2.0	-1.3	11.0
	100.0		*17.1*	*-10.7*	*93.6*
Imports of intermediate inputs	11.6		2.3	-1.5	10.8
	100.0		*20.0*	*-12.7*	*92.7*

Notes: The table decomposes the annual growth rates of production and import value into the contributions of new, exiting, and continuing products and varieties (see equation (1) in the text). Figures are averages across countries, NACE industries, and years. Numbers in italics are normalized by the growth rates reported in the first column. *Source*: Eurostat (Prodcom and Comext).

4.1 Symmetric Inputs

Consider an economy, denoted by c, which has a single final good sector, labeled i. In each period t, competitive firms produce the final good using labor and symmetric intermediate products according to the following function:

$$Y_t = \frac{1}{1-\alpha} L^{\alpha} \left(\int_0^{N_t} x_{ht}^{1-\alpha} dh \right), \alpha \in (0,1), \tag{2}$$

where N_t is the number of intermediate goods available at time t, and x_{ht} denotes the flow output of intermediate product h; labor is inelastically supplied by a fixed number of workers, denoted by L.[16] Final good producers take both prices and N_t as given. We take the final good as our numeraire and normalize its price to 1 in every period.

New intermediate goods are invented through research and development, according to the following technology:

$$\dot{N}_t = \mu \frac{1}{1-\alpha} L^{\alpha} \left(\int_0^{N_t} x_{ht}^{1-\alpha} dh \right), \tag{3}$$

where $\dot{N}_t$ is the number of new products introduced in period t, and $\mu > 0$ is a parameter measuring productivity in research. This specification has been proposed by Rivera-Batiz and Romer (1991) under the name 'lab-equipment' and implies that all factors, including intermediate inputs, are productive in research. As standard, we assume free entry into R&D.

When a firm discovers a new blueprint, it receives a fully enforced patent and acquires perpetual monopoly power over the production of the corresponding intermediate good. Producing one unit of any intermediate product entails only a marginal cost equal to ψ.

We solve the model for a balanced growth (BG) equilibrium in which innovation occurs at a constant rate g. The derivation of the equilibrium is standard and is fully detailed in Appendix A. Here, we just explain the intuition for why new imported inputs stimulate product creation, and comment on the relevant equations for our empirical analysis. Specifically, under the standard

[16]The term $1-\alpha$ in the denominator of (2) is introduced to simplify the notation and has no consequences for the results.

assumption of a representative consumer with isoelastic preferences, the equilibrium expression for g reads as follows:

$$g = \frac{\mu\alpha L - \rho}{\theta}, \tag{4}$$

where ρ is the intertemporal discount factor and $1/\theta$ is the constant elasticity of intertemporal substitution.

Since changes in g correspond to changes in the rate of introduction of new products, i.e. our primary object of interest, we now study how g changes when the country gets access to new intermediate inputs from abroad. In particular, suppose that at time t the cost of importing inputs from some foreign country drops to zero, starting from a prohibitively high level at which imports were zero. For simplicity, assume that the number of new imported inputs is $N_t^* = N_t$. Then, as shown in Appendix A, the new equilibrium value of g is:

$$\hat{g} = \frac{2\mu\alpha L - \rho}{\theta}. \tag{5}$$

Clearly, $\hat{g} > g$, which implies that new imported inputs raise the rate of introduction of new products.

The intuition for this result is extensively discussed in the literature (see e.g. Rivera-Batiz and Romer, 1991, Backus, Kehoe, and Kehoe, 1992, and Aghion and Howitt, 1998, p. 374). In particular, the result follows from the properties of the R&D technology (3) which, as shown in Appendix A, exhibits increasing returns as a function of L and N_t (see (22)). Hence, the 'scale effect' generated by a larger number of available intermediates has the same implication for innovation as an increase in the productivity parameter μ. In equilibrium, this efficiency gain determines a higher value of g.[17] In the specific comparative-statics exercise considered above, new imported inputs double the number of available intermediates, and this has the same impact on $\dot{N}_t$ as a doubling of the productivity parameter, from μ to 2μ (see (27)). In turn, the equilibrium rate of innovation increases from g to $\hat{g}$.[18]

4.2 Heterogeneous Inputs

Assume now that the intermediate inputs are differentiated in terms of quality. In particular, each product h is characterized by a quality level $\lambda_h > 0$. As standard in the literature, we model quality as a unidimensional metric translating physical units into efficiency units: higher quality means more efficiency units per physical unit. Following Aghion and Howitt (1998, chp. 12), we assume the quality of each input to be randomly drawn from the distribution of existing intermediates,

[17] The scale effect is reminiscent of the conventional 'love-for-variety' effect that is present in static trade models with differentiated inputs, in which an increase in the number of available intermediates leads to greater specialization in the use of resources and thus to higher efficiency (Ethier, 1982).

[18] In the BG equilibrium of this model, output also grows at the same rate g (see Proposition 13.1 in Acemoglu, 2009). The intuition is that, in a model with expanding variety, the introduction of new products constitutes technological progress, and thus acts as the engine of growth for the economy. It follows that any factor changing g has analogous implications for innovation and growth. This feature of the model motivates our analysis of the effects of new imported inputs on growth; see Section 5.3.

once the corresponding blueprint has been discovered. Then, we modify (2) as follows:

$$Y_t = \frac{1}{1-\alpha} L^{\alpha} \left[\int_0^{N_t} \left(\lambda_h x_{ht} \right)^{1-\alpha} dh \right], \tag{6}$$

and accordingly rewrite (3) as:

$$\dot{N}_t = \mu \frac{1}{1-\alpha} L^{\alpha} \left[\int_0^{N_t} \left(\lambda_h x_{ht} \right)^{1-\alpha} dh \right]. \tag{7}$$

We also assume, similar to Aghion and Howitt (1998), that producing one unit of product h entails a marginal cost equal to $\psi \eta \lambda_h$, with $\eta > 0$. Countries may differ in the technology parameter η governing the quality-marginal cost relationship. As shown in Appendix A, η pins down the equilibrium value of the quality-adjusted price of the intermediate products, p_h/λ_h. The remainder of the model is exactly as above. It is easy to show (see Appendix A) that the equilibrium expression for g is now:

$$g = \frac{\mu \alpha L \eta^{(\alpha-1)/\alpha} - \rho}{\theta}, \tag{8}$$

which is equivalent to (4) except for the presence of $\eta^{(\alpha-1)/\alpha}$ in the first term.

Consider now the same comparative-statics exercise as in the previous section. In particular let $N_t^* = \xi N_t$, with $\xi > 0$, be the number of new foreign inputs and η^* their quality-adjusted price. Then, as shown in Appendix A, the new equilibrium value of g is:

$$\hat{g} = \frac{\omega \mu \alpha L \eta^{(\alpha-1)/\alpha} - \rho}{\theta}, \tag{9}$$

where

$$\omega \equiv 1 + \xi \left(\frac{\eta^*}{\eta} \right)^{(\alpha-1)/\alpha}. \tag{10}$$

Since $\omega > 1$, it follows that $\hat{g} > g$. Therefore, new imported inputs raise the rate of introduction of new products, and the increase in g is proportional to ω.

Importantly, (10) shows that the effect of new imported inputs now takes place through two channels. First, new imported inputs generate a scale effect by expanding the set of available intermediates. Note, indeed, that ω is proportional to the relative number of new imported inputs, ξ. This effect is always present, even when new imported inputs have the same characteristics as the existing intermediates, i.e. even when $\eta^* = \eta$. Second, when $\eta^* \neq \eta$, new imported inputs may generate further efficiency gains, by allowing the country to access varieties with more favorable price-quality ratios. In particular, (10) shows that a low value of η^*/η—the relative quality-adjusted price of new imported inputs—amplifies the standard scale effect. In the next sections, we provide evidence consistent with these implications.

5 Empirical Analysis

Guided by our conceptual framework, we now turn to the empirical analysis. We start by providing evidence of a positive effect of new imported inputs on product creation (Section 5.1). Then, we investigate the mechanisms underlying this effect (Section 5.2) and explore the implications of new imported inputs for growth (Section 5.3). Finally, we study the characteristics of new domestic products (Section 5.4).

5.1 New Imported Inputs and the Introduction of New Products

Our simple model implies a positive relationship between the rate of introduction of new domestic goods and the relative number of new imported inputs (see (9) and (10)). In this section, we provide evidence on this relationship. The empirical counterpart for the rate of introduction of new domestic products is NP, the share of new domestic goods in the total number of domestic products. Similarly, the empirical counterpart for the relative number of new imported inputs is NII, the share of new foreign varieties in the total number of imported input varieties.[19] We start by documenting a strong positive correlation between these two variables within countries and industries. Then, we show that this correlation is robust across a large number of sensitivity checks and extensions. Finally, we tackle reverse causality using instrumental variables.

5.1.1 Baseline Correlation

We run OLS regressions of the following form:

$$NP_{cit} = \beta_{ci} + \beta_t + \beta_1 NII_{cit-1} + \varepsilon_{cit}, \tag{11}$$

where c denotes countries, i NACE industries, and t years; β_{ci} are country-industry effects, β_t year effects, and ε_{cit} is an error term.[20] We correct the standard errors for two-way clustering by country-industry and industry-year, to accommodate autocorrelated shocks as well as correlated shocks across countries for a particular industry (Cameron, Gelbach and Miller, 2011).[21]

The results are reported in Table 3. In the first three columns, we estimate (11) at different levels of industry aggregation: 4-digit (column 1), 3-digit (column 2), and 2-digit (column 3). In

[19]We construct NP and NII using the product-level data described in Section 3. Note that, in the model laid out in Section 4, innovation leads to the introduction of new intermediate products. As discussed by Acemoglu (2009), however, the model can naturally be reinterpreted as featuring an expanding set of final goods. Consistent with these two interpretations, we use both final and intermediate products to construct our dependent variable NP. Note also that, as mentioned in Section 3, the PC and CN classifications do not have a comparable number of codes. Hence, it is natural to normalize NP by the number of domestic goods, and NII by the number of foreign input varieties, instead of using in both cases the overall number of products (i.e. domestic goods plus foreign varieties) as would be suggested by our simple model.

[20]In the interest of space, we focus on specifications including the first lag of NII, since our results are very similar across alternative lag lengths (available upon request).

[21]Similar results are obtained with one-way clustering by country-industry, two-way clustering by country-industry and country-year, and two-way clustering by country and industry (available upon request).

Table 3: New Imported Inputs and the Introduction of New Products

	(1)	(2)	(3)	(4)
NII	0.122***	0.245***	0.411***	
	[0.021]	[0.036]	[0.069]	
$NIIov$				0.582***
				[0.073]
Obs.	33521	18244	4446	4583
R^2	0.07	0.07	0.11	0.12
Industry aggregation	NACE 4	NACE 3	NACE 2	NACE 2

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. NII is the share of new imported input varieties in the total number of imported input varieties. $NIIov$ is the weighted average of NII across all industries, computed using weights from country-specific Import Matrices. All specifications are estimated by OLS and control for country-industry and year effects. Standard errors are corrected for two-way clustering at the country-industry and industry-year level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

all cases, the estimate of β_1 is positive and highly significant, with t-statistics close to 6. Next, we broaden the scope of our explanatory variable to account for the fact that industries source inputs not just from themselves (as implicitly assumed when using NII) but also from other industries in the economy. Hence, we construct a more comprehensive indicator of new imported inputs which accounts for backward linkages across industries. Using country-specific Import Matrices provided by Eurostat, for each industry i we compute the share of any industry j in its total imports of intermediates. We calculate these figures for all available years (see Table A3) and then take their average over time. Using the resulting weights (ϕ_{cij}), we construct an *overall* indicator of new imported inputs as follows:

$$NIIov_{cit} = \sum_j \phi_{cij} \cdot NII_{cjt}. \tag{12}$$

Since the Import Matrices are available only for 2-digit industries, in the rest of the paper we will work at that level of industry aggregation. In column (4), we use $NIIov$ in place of NII. The coefficient β_1 remains positive and highly significant. In terms of magnitude, an increase of 1 percentage point (1 standard deviation) in the share of new imported inputs is associated with an increase of roughly 0.6 p.p. (0.3 standard deviations) in the share of new domestic goods. Hence, our data reveal a strong positive association between new imported inputs and new domestic products. In the next section, we assess the robustness of this correlation.

5.1.2 Robustness Checks

We now present an extensive sensitivity analysis showing that the positive relationship between new imported inputs and new domestic products is remarkably robust. From here on, we focus on the baseline version of (11), presented in column (4) of Table 3.

Alternative specifications We start by showing that the correlation is robust across many alternative specifications. The results of these exercises are reported and discussed in Appendix B, while here we just highlight the main points. In particular, we find the baseline correlation to be robust to different approaches for dealing with potential outliers, as well as to alternative definitions of our explanatory variable, obtained by: (i) redefining the weights ϕ; (ii) using narrower definitions of intermediate inputs (without capital goods, fuels, and lubrificants); and (iii) restricting the analysis to entirely new *products* (as opposed to new *varieties*). Moreover, we find our results to hold under alternative ways of identifying new domestic products and new imported inputs, designed to address concerns that: (i) the commodity classifications may adjust with some delay to the invention of new goods; (ii) our procedure may overestimate the number of new products in the initial years of the sample; and (iii) the identification of new goods may also include products that remain in the sample for just a few years after entry. Finally, we prove the robustness of our main correlation to the use of alternative estimators.

In Appendix B, we also shed some light on the 'margins' underlying the positive correlation between new imported inputs and new domestic products. Indeed, any new product observed in our data could have been introduced either by an incumbent firm or by a new entrant. In other words, our data are general enough to encompass both an *intensive* and an *extensive* margin of product creation.[22] Yet, our data do not allow us to disentangle the two margins. Hence, we use information from different data sources to provide suggestive evidence that both margins are indeed at work in the EU, and thus contribute to the positive correlation between new imported inputs and new domestic goods estimated with product-level data. In particular, as for the extensive margin, we show that $NIIov$ is strongly positively correlated with the entry rate of new firms in each country and industry. Instead, as for the intensive margin, we use recent firm-level data for a cross-section of firms in seven EU countries, and show that $NIIov$ is also positively correlated with the probability that incumbent firms introduce new goods.

Related factors In this section, we augment the baseline specification with proxies for other phenomena which, although not directly relevant for our theory, may be correlated with NP and $NIIov$ in practice, and may thus influence our results. All of these proxies will be computed separately for each country, industry, and year. The results are reported in Table 4. In column (1) we control for new imports of *final goods*, as proxied by the share of new varieties in the total number of imported final products. The coefficient on this variable is very small and imprecisely estimated, while β_1 remains close to the baseline specification. In column (2), we control for new *domestic inputs*, by including their share in the total number of domestic intermediates. As expected, the coefficient on this variable is positive and significant, but β_1 is largely unaffected. Interestingly, the correlation is much stronger (by an order of magnitude) in the case of new imported inputs than for new domestic intermediates, suggesting the former to be a more relevant

[22] For comparison, note that in Goldberg et al. (2010a)—the only existing paper on new imported inputs and product creation—the analysis is restricted to the intensive margin, as the authors focus only on the introduction of new goods within existing firms.

Table 4: Related Factors

	(1)	(2)	(3)	(4)	(5)
$NIIov$	0.564***	0.665***	0.583***	0.593***	0.404***
	[0.102]	[0.110]	[0.073]	[0.074]	[0.070]
Share of new imported final goods	0.022				
	[0.083]				
Share of new domestic inputs		0.082***			
		[0.020]			
Share of exiting domestic products			-0.034		
			[0.024]		
Share of exiting foreign inputs				0.096	
				[0.108]	
ln Value added per worker					0.017
					[0.019]
ln Employment					0.012
					[0.022]
ln Capital					0.006
					[0.007]
ln Material and service inputs					-0.017
					[0.020]
Obs.	4583	4118	4554	4583	4286
R^2	0.12	0.12	0.12	0.12	0.10

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. All specifications are estimated by OLS and control for country-industry and year effects. Standard errors are corrected for two-way clustering at the country-industry and industry-year level. The level of industry aggregation is NACE2. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

driver of product creation.[23]

Next, we account for the *exit* of domestic goods and foreign inputs. While the model does not contemplate exit (as standard in models with expanding variety), in reality both NP and $NIIov$ may be correlated with measures of exit. For example, industries that undergo deep restructuring of their production activities may exhibit high rates of product creation and destruction, as well as high rates of entry and exit of foreign intermediates. In columns (3) and (4), we thus control for the shares of exiting products and foreign inputs, respectively. The first variable is defined as the number of goods that disappear from the sample in a given period, divided by the total number of domestic products. The second variable is instead defined as the number of intermediate varieties that stop being imported in a certain year, divided by the total number of imported inputs. Both variables enter with small and insignificant coefficients, and leave our main results unaffected.[24]

Finally, we consider the role of other production factors and technical change. Given data availability, we account for technical change using labor productivity (value added per worker). As for the other inputs, we control for employment, capital, and material and service inputs. All of these variables are sourced from Euklems (O'Mahony and Timmer, 2009). As shown in column

[23] For completeness, we note that the share of new domestic inputs equals 7% on average, against 13% for the share of new imported inputs.

[24] Consistent with these findings, we obtain similar results if we use as the dependent variable the net entry of domestic goods (defined as new minus exiting products as a share of total domestic goods), or if we use as the regressor the net entry of imported inputs (defined as new minus exiting varieties as a share of total imported inputs). In the former case, $NIIov$ has a coefficient (standard error) of 0.415 (0.090). In the latter case, the net entry of foreign inputs has a coefficient (standard error) of 0.424 (0.059).

Table 5: Underlying Trends

	Coeff.	Std. Err.	Obs.	R^2
(1) Initial value of new domestic products and new foreign inputs	0.620***	[0.091]	4080	0.39
(2) Initial churning of domestic products and foreign inputs	0.668***	[0.083]	4011	0.33
(3) Pre-sample output growth	0.543***	[0.075]	4583	0.14
(4) Pre-sample employment growth	0.556***	[0.074]	4581	0.13
(5) Pre-sample change in labor productivity	0.371***	[0.057]	4435	0.12
(6) Pre-sample changes in capital and material intensity	0.475***	[0.068]	4546	0.17
(7) Pre-sample changes in export intensity and import penetration	0.621***	[0.079]	4440	0.14

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. The explanatory variable is $NIIov$. In each row, the baseline specification (see column 4 of Table 3) is augmented with interaction terms between the year dummies and the indicated variables. All specifications control for country-industry and year effects. Standard errors are corrected for two-way clustering at the country-industry and industry-year level. The level of industry aggregation is NACE2. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

(5), these controls have little impact on our estimate of β_1.

Underlying trends A possible concern with our results thus far is that the correlation between NP and $NIIov$ may be spuriously driven by some underlying trend. In this section, we thus re-estimate our baseline specification allowing for differential trends across countries and industries, based on pre-existing observable characteristics. To this purpose, we augment (11) with interaction terms between the year dummies and the *pre-sample* values of specific variables in each country-industry pair. We report the results in Table 5.

The first issue we tackle is that NP and $NIIov$ may have grown faster in countries and industries with initially large shares of new products and foreign inputs. In row (1) we therefore add interactions between the year dummies and the initial values of NP and $NIIov$. If anything, the inclusion of these terms slightly increases the estimate of β_1. Another possibility is that NP and $NIIov$ have grown faster in industries and countries that were initially more dynamic, i.e. characterized by more adding and shedding of inputs and products. Hence, in row (2) we add interactions between the year dummies and the initial churning of domestic products and foreign intermediates. We define churning as the sum of new and exiting products (imported input varieties) divided by the total number of domestic products (imported input varieties). If anything, the estimate of β_1 is now even larger.

More generally, it could be the case that countries and industries that grew faster in pre-sample years also exhibited higher shares of new products and foreign inputs in subsequent periods. In row (3) we thus augment our specification with interactions between the year dummies and average output growth over the five years prior to the beginning of the sample. In row (4) we do the same, but using employment growth instead of output growth. We source output and employment data from Euklems. In both cases, β_1 is close to our baseline estimate. A related concern is that countries and industries characterized by faster technical change in pre-sample periods also had greater entry of domestic products and foreign inputs in subsequent years. In row (5) we thus extend the specification with interactions between the year dummies and the pre-sample variation in labor productivity (average over five years). Our main evidence is preserved.

It might also be the case that NP and $NIIov$ have grown faster in countries and industries characterized by deeper changes in factor intensities. In row (6) we thus add interactions between the year dummies and the pre-sample changes in capital and material intensity (averages over five years). Capital intensity is defined as capital compensation per worker, while material intensity is defined as material purchases per worker. The coefficient β_1 is little affected. Finally, NP and $NIIov$ may have grown faster in countries and industries characterized by more rapid changes in their exposure to international trade. Hence, in row (7) we add interactions between the year dummies and the pre-sample changes in export intensity and import penetration (averages over five years). Export intensity is the ratio of exports to output, while import penetration is the ratio of imports to output; these variables are constructed using trade data from Comext and output data from Euklems. After adding these interactions, β_1 is positive, highly significant, and slightly larger than our baseline estimate. Overall, we conclude that differential trends based on pre-existing characteristics have little impact on our results.

Contemporaneous shocks The positive correlation documented so far is compatible with two explanations. On the one hand, new imported inputs could stimulate the introduction of new products, as implied by our conceptual framework. On the other hand, industry-specific shocks in the EU countries may lead to product creation for reasons unrelated to the availability of foreign intermediates; but once the decision to produce a new good has been made, firms could start sourcing the necessary new inputs from abroad. This 'reverse causality' would induce an upward bias in the OLS estimate of β_1. In this section, we therefore study the role of industry-specific shocks in the EU countries over the sample period. We argue that these shocks cannot be the only explanation for the positive relationship between new imported inputs and new domestic products. Yet, they seem to induce an upward bias in the coefficient β_1 estimated by OLS. We take this evidence as a motivation for our IV analysis, which will be presented in the next section.

Our approach consists of augmenting (11) with several fixed effects, obtained by interacting the year dummies with indicators for countries and (groups of) industries. The results are reported in Table 6. In row (1) we add country-year and industry-year effects which capture, respectively, country-specific shocks common to all industries (e.g. changes in macroeconomic conditions) and industry-specific shocks common to all countries (e.g. sector-specific technical change). This specification includes more than 400 variables and is thus highly demanding. Nevertheless, β_1 remains positive and very precisely estimated. In row (2) we add a linear trend for each country-industry pair (about 500 variables). The coefficient β_1 is now identified out of deviations from these trends, and its point estimate is still positive and significant. Note, however, that in these specifications β_1 falls compared to our baseline estimate by 40% on average.

Next, we add fixed-effects for triplets of countries, *sectors*, and years. We define a sector as a small group of similar industries, where similarity is assessed in terms of a number of charac-

Table 6: Contemporaneous Shocks

	Coeff.	Std. Err.	Obs.	R^2
(1) Country-year and industry-year dummies	0.433***	[0.111]	4583	0.78
(2) Country-industry specific time trends	0.288**	[0.131]	4583	0.28
(3) Country-*sector*-year dummies: Output	0.305**	[0.145]	4583	0.69
(4) Country-*sector*-year dummies: Labor productivity	0.341**	[0.138]	4583	0.68
(5) Country-*sector*-year dummies: Material intensity	0.301**	[0.135]	4583	0.66
(6) Country-*sector*-year dummies: Capital intensity	0.237**	[0.113]	4583	0.69
(7) Country-*sector*-year dummies: Import penetration	0.327***	[0.114]	4583	0.66
(8) Country-*sector*-year dummies: Export intensity	0.347***	[0.117]	4583	0.67

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. The explanatory variable is $NIIov$. The baseline specification (see column 4 of Table 3) is augmented with country-year and industry-year effects in row (1), and with country-industry specific linear trends in row (2). In rows (3)-(8), the specification is augmented with country-*sector*-year dummies. The sectors are identified by dividing the NACE2 industries into quintiles, based on the change (over the sample period) in the variable indicated in each row. All specifications include country-industry effects. Standard errors are corrected for two-way clustering at the country-industry and industry-year level. The level of industry aggregation is NACE2. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

teristics discussed below. In these regressions, we thus control for shocks specific to a certain group of industries in a given country and time period. Our coefficient of interest is identified only from the remaining variation across the industries that fall in the same country-sector-year cell. Since we cannot condition on a full set of country-*industry*-year effects (as they would be perfectly collinear with $NIIov$), this is the most stringent test we can perform while still being able to identify β_1.

We start by allowing for the possibility that fast- and slow-growing industries experience different shocks. For each sample country, we divide industries into five bins of equal size, based on their average output growth *over the sample period*: each bin is a sector. We then add country-sector-year dummies to (11). The results are reported in row (3), showing that β_1 is positive and precisely estimated, but reduced by 47% compared to our baseline estimate. Next, we allow for different shocks across industries characterized by high and low rates of technical change. To this purpose, in row (4) we divide industries into quintiles based on the average change in labor productivity over the estimation period. Again, β_1 is positive and significant, but drops by roughly 40% compared to our baseline estimate.

In rows (5) and (6), we allow for different shocks across industries characterized by different changes in factor intensities. To this purpose, we identify the industry quintiles based on the average change in material or capital intensity over the sample period. Our main coefficient is positive and precisely estimated, but is lower than the baseline estimate by about 50%. Finally, we control for different shocks across industries experiencing different changes in their exposure to international trade. To this purpose, in rows (7) and (8) we identify the industry quintiles based on the average change in import penetration or export intensity over the estimation period. The results are in line with previous specifications. In light of the evidence reported in this section, we now turn to the IV regressions to fully account for industry-specific shocks in the EU countries.

5.1.3 Instrumental Variables Regressions

According to the previous section, the upward bias induced by industry-specific shocks in the EU countries may account for 40-50% of our baseline correlation. To identify the effect of new imported inputs, we thus need to isolate their exogenous variation, which occurs independently of these shocks. An important source of exogenous variation is given by changes in transportation costs. Intuitively, a reduction in these costs should facilitate trade and boost imports of new intermediate inputs. Following this argument, and inspired by previous work of Hummels (2007) and Hummels et al. (2013), we thus build an instrument proxying for changes in variety-specific transportation costs.[25]

Description of the instrument To explain how the instrument works, recall that we define a new input variety as an 8-digit product imported from a certain trading partner for the first time. Thus, the instrument should capture changes in transportation costs that may induce new 8-digit goods to be imported from a given foreign country. Our strategy is to compute transportation costs at the level of *6-digit* input varieties, defined as combinations of 6-digit products and trading partners. It is our contention that, when the cost of importing a 6-digit product from a certain trading partner declines, the EU countries may raise imports of all the constituent 8-digit goods, including those that were not imported before from that country.[26]

To implement this idea, we construct measures of transportation costs following Hummels (2007) and Hummels et al. (2013). In a first step, we estimate an ad-valorem cost function using data on transportation costs for the US. These data are available at the 6-digit level of the Harmonized System (HS) classification (equivalent to the 6-digit level of the CN classification) for the period 1995-2006. The specification reads as follows:

$$\begin{aligned} \ln{(f/m)_{hnt}} \quad &= \quad \beta_h + \beta_1 \ln{(k/m)_{hnt}} + \beta_2 \ln{dist_n} + \beta_3 \ln{oil_t} + \beta_4 \ln{dist_n} * \ln{oil_t} + \\ &+\beta_5 \ln{dist_n} * \ln{oil_t} * \ln{(k/m)_{hnt}} + \varepsilon_{hnt}, \end{aligned} \tag{13}$$

where β_h are product fixed-effects; m_{hnt} is the value of imports into the US of 6-digit product h from partner n in year t; f_{hnt} is the associated transportation charge; k_{hnt} is the weight of the product (in tons); $dist_n$ is distance from partner n; oil_t is oil prices; and ε_{hnt} is an error term.[27]

[25] In the theoretical section, we considered the complete elimination of the cost of importing any good from a given foreign country. In reality, transportation costs change continuously, and differ substantially across products and trading partners. Our instrument will capture these features. We are not the first to use transportation costs as an instrument for imported inputs. Most notably, Hummels et al. (2013) use changes in transportation costs to instrument for offshoring by Danish firms.

[26] At the 8-digit level, transportation costs would be observed only *after* a variety has been imported for the first time. It follows that variation in transportation costs at the *8-digit* level cannot be used to instrument for new imported inputs. Furthermore, to compute the transportation costs we employ US trade data (described below), as the European data do not contain information on transportation charges. The two data sources cannot be linked to each other at higher levels of disaggregation than six digits. For these reasons, we use *6-digit* variety-specific transportation costs in all our analysis.

[27] We source the US trade data from Feenstra, Romalis, and Schott (2002). Data on distances (number of kilometers between capital cities) are sourced from CEPII. Oil prices are Brent.

This specification allows transportation costs to depend on product characteristics (the product fixed-effects β_h and the bulk weight k/m), oil prices, and shipping distance. In a second step, we apply the estimated coefficients (including β_h) to import and distance data for the EU countries (obviously, the data on oil prices are the same as those used for estimating (13)). To avoid introducing endogeneity, we keep the bulk weight of each 6-digit variety constant at its pre-sample level.

For a given EU country c, let τ_{cvt} denote the resulting estimate of transportation costs for 6-digit variety v in year t. We aggregate these estimates at the level of each 2-digit industry in each EU country, using pre-sample information on the import share of each 6-digit variety in a given country-industry pair ($ImpSh_{civ}$). Hence, the instrument is equal to:

$$Transportation\ Costs_{cit} = \sum\nolimits_{v \in ci} ImpSh_{civ} * \tau_{cvt}. \tag{14}$$

Clearly, the instrument varies over time only due to changes in oil prices, which are determined by global factors. Instead, the cross-sectional variation of the instrument depends on (i) differences in bulk weights across varieties and (ii) differences in varieties' import shares across industries and EU countries. Both variables are measured pre-sample and maintained constant throughout. As a result, the cross-sectional variation of the instrument is unaffected by industry-specific shocks in the EU countries over the estimation period. The functioning of the instrument is intuitive: a change in oil prices has a stronger effect on transportation costs for heavier inputs traveling longer distances; in turn, this affects imports disproportionately more in the EU countries that rely more intensively on those inputs.

Summing up, to isolate the exogenous variation in new imported inputs, we use changes in variety-specific transportation costs. To construct this instrument, we take advantage of pre-sample variation in the composition of trade across industries and EU countries. Owing to this variation, changes in transportation costs (as induced by fluctuations in oil prices) will have differential effects on new imported inputs across countries and industries. We rely on this country-industry-time variation to identify the effect of new imported inputs on product creation. The results are presented in the next section.

Results and discussion The 2-Stage Least Squares (2SLS) estimates are reported in Table 7. As before, all regressions are estimated at the 2-digit industry level, and standard errors are corrected for two-way clustering by country-industry and industry-year. In the bottom of the table, we report the first-stage coefficient on our excluded instrument, together with the test statistic for weak identification (Kleibergen-Paap Wald F-statistic). The latter is corrected for error correlation within clusters.

Column (1) shows the baseline results. We start by commenting on the first-stage estimates. The coefficient on our excluded instrument is highly significant (t-statistic 17.3). The instrument explains a substantial fraction of the variation in new imported inputs, more than 30% according to Angrist and Pischke's partial R^2 (unreported). Importantly, the first-stage coefficient is neg-

Table 7: Instrumental Variables Regressions

	(1)	(2)	(3)	(4)
$NIIov$	0.325***	0.415***	0.451***	0.380***
	[0.102]	[0.103]	[0.099]	[0.138]
Obs.	4583	3676	3678	3065
R^2	0.11	0.12	0.15	0.12
	First-stage estimates and IV statistics			
$Transportation\ Costs$	-3.555***	-3.563***	-3.604***	-3.305***
	[0.206]	[0.234]	[0.240]	[0.235]
F-statistic for weak identification	298.4	232.8	225.5	197.8

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. Estimation is performed by 2-Stage Least Squares. Column (1) uses the whole sample. Columns (2)-(4) exclude, respectively, the following industries: NACE 20, 23, 26, 28, 34; NACE 21, 23, 24, 26, 27; and NACE 17-19, 26-28, 30. The F-statistic for weak identification is the Kleibergen-Paap Wald F-statistic. All specifications control for country-industry and year effects. Standard errors and IV statistics are corrected for two-way clustering at the country-industry and industry-year level. The level of industry aggregation is NACE2. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

ative and large in absolute value implying, as expected, that lower transportation costs give a substantial boost to new imported inputs in the EU. According to the point estimate, the share of new imported inputs rises by 3.5 p.p. after a 1 p.p. reduction in transportation costs. Overall, these results underscore that the instrument has strong predictive power and correlates with new imported inputs in the expected way.

Having shown that the instrument works as it should, we turn to the second-stage results. The main point to highlight is that the coefficient on $NIIov$ is positive and highly significant also in the IV regressions. Furthermore, we note that this coefficient is smaller than our baseline estimate obtained by OLS (reported in column 4 of Table 3), suggesting that the instrument is removing the upward bias induced by reverse causality. Interestingly, the IV coefficient falls within the range of estimates reported in Table 6, where we controlled for country-sector-year dummies. This further suggests that our IV strategy is isolating variation in new imported inputs not due to shocks in individual industries and EU countries. In terms of magnitude, the IV coefficient implies that a 1 p.p. increase in the share of new imported inputs raises the share of new products by 0.3 p.p., explaining half of the baseline correlation detected by OLS.

In the rest of this section, we address concerns with our identification strategy. One might worry that business cycle fluctuations may induce a correlation between the instrument and the error term of (11), as oil prices, on the one hand, and product creation, on the other, may respond to fluctuations in economic activity. We believe that this is not a major concern for us, since all our specifications include time dummies, which absorb common changes in macroeconomic conditions across countries and industries. Nevertheless, we now show that our estimates do not change when we exclude industries that are most sensitive to the business cycle.

In a first exercise, we estimate (11) after dropping the five industries characterized by the highest correlation between their own output growth and GDP growth.[28] The results, reported

[28]Excluded industries are: 'wood and wood products' (NACE 20); 'coke, petroleum products, and nuclear fuel'

in column (2), show that the coefficient on $NIIov$ actually increases slightly compared to the estimate obtained on the whole sample. In a second exercise, performed in column (3), we exclude the five most energy-intensive industries, as identified by the US Department of Energy.[29] If anything, our coefficient of interest is now slightly larger. In a third exercise, we follow Autor, Dorn, and Hanson (2013) and exclude seven industries characterized by substantial fluctuations in economic activity over the sample period, due to housing booms, sudden technological innovations, or rapid growth in emerging players like China.[30] As shown in column (4), our coefficient of interest is close to the one obtained on the whole sample.

One might also worry that the instrument affects product creation not just through new imported inputs, but also through other channels. While we cannot test for the exclusion restriction, we can augment our specification with a number of control variables, which may correlate with our instrument and affect the introduction of new goods. By showing that our estimates do not change when controlling for these factors we can be confident that, in practice, this concern is not a major threat to our findings.

The results of these exercises are reported in Table 8. For all specifications, we show both OLS estimates (panel a) and IV estimates (panel b). In the IV regressions, we instrument $NIIov$ using $Transportation\ Costs$, and the control variables using their longest available lag (see below). The first-stage coefficients on our instrument are always remarkably similar to those reported in Table 7, in terms of sign, size, and significance. In the interest of space, these estimates are not shown in the table, but are available from us upon request together with the other first-stage results.

One way in which the instrument may affect product creation is through demand. For example, a reduction in oil prices may raise the real income of consumers, increasing demand for new products. Similarly, lower trade costs may expand export market opportunities, raising foreign demand for European goods. Inspired by Hummels et al. (2013), we thus add three different proxies for demand to (11). The first proxy, used in column (1), is a time-varying price index, computed separately for each country and 2-digit industry. We source this variable from Euklems. In the IV regression, we instrument this variable with its 8th lag. The second and third proxies—used in columns (2) and (3), respectively—are total exports and the share of new varieties in total exported varieties. Both variables are constructed separately for each country, 2-digit industry, and year, using data from Comext.[31] In the IV regressions, we instrument these variables using their 9th lag. To the extent that these three demand proxies are also linked to the business cycle, their inclusion further addresses the previous concern with economic fluctuations. The results show a positive and significant coefficient on the price index, and small

(NACE 23); 'non-metallic mineral products' (NACE 26); 'fabricated metal products' (NACE 28); 'motor vehicles' (NACE 34).

[29] Excluded industries are: 'pulp and paper' (NACE 21); 'coke, petroleum products, and nuclear fuel' (NACE 23); 'chemicals' (NACE 24); 'non-metallic mineral products' (NACE 26); 'basic metals' (NACE 27).

[30] Excluded industries are: 'textiles' (NACE 17); 'apparel' (NACE 18); 'leather' (NACE 19); 'non-metallic mineral products' (NACE 26); 'basic metals' (NACE 27); 'fabricated metal products' (NACE 28); 'office machinery and computers' (NACE 30).

[31] New exported varieties are identified using the same procedure applied to imports in Section 3.

Table 8: IV Regressions, Additional Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
a) <u>OLS</u>						
$NIIov$	0.400***	0.597***	0.564***	0.552***	0.601***	0.377***
	[0.058]	[0.078]	[0.073]	[0.083]	[0.075]	[0.065]
ln Price index	0.104***					
	[0.028]					
ln Total exports		0.008				
		[0.009]				
Share of new exported varieties			0.096			
			[0.070]			
ln Imports of continuing inputs				-0.015		
				[0.012]		
ln FDI inflow					0.002	
					[0.003]	
ln FDI outflow						-0.003
						[0.003]
Obs.	4437	4582	4582	4386	4397	4277
R^2	0.11	0.12	0.12	0.12	0.12	0.10
b) <u>2SLS</u>						
$NIIov$	0.274***	0.465**	0.293**	0.328**	0.405***	0.423***
	[0.089]	[0.229]	[0.121]	[0.134]	[0.151]	[0.145]
ln Price index	0.190***					
	[0.073]					
ln Total exports		0.064				
		[0.094]				
Share of new exported varieties			0.148			
			[0.316]			
ln Imports of continuing inputs				0.004		
				[0.070]		
ln FDI inflow					0.013	
					[0.017]	
ln FDI outflow						0.015*
						[0.008]
Obs.	4437	4582	4582	4386	4397	4277
R^2	0.10	0.10	0.11	0.12	0.11	0.08
F-statistic for weak identification	94.3	10.0	21.1	15.8	30.9	136.8

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. In panel b), $NIIov$ is instrumented using $Transportation\ Costs$, and the control variables using their longest available lag. All specifications control for country-industry and year effects. Standard errors are corrected for two-way clustering at the country-industry and industry-year level. The level of industry aggregation is NACE2. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

and insignificant estimates for the export proxies. Most importantly, the coefficient on $NIIov$ remains positive, significant, and stable in size.

Another way in which the instrument may influence product creation is through imports of *existing* instead of *new* intermediates. Hence, in column (4) we add imports of continuing inputs, i.e. input varieties imported into a country also in previous years. We construct this variable using data from Comext. In the IV regression, we instrument it using its 10th lag. The coefficient on the new control is small and insignificant, while our coefficient of interest remains close to previous specifications. Finally, it might be the case that the instrument affects product creation by influencing the decision of firms to make a foreign direct investment (FDI). For example, lower trade costs may induce European firms to relocate production stages abroad ('vertical' FDI). To the extent that firms benefit from technological spillovers in foreign countries, or that offshoring induces cost savings, the outward investment may facilitate the introduction of new products. Similarly, changes in trade costs may affect the decision of foreign firms to invest in the EU countries, where they may start producing some new goods. Hence, we control for FDI inflows (column 5) and outflows (column 6) in each country and industry. We construct these variables using data from Euklems and Unctad. In the IV regressions, we instrument them using their 9th lag. Both variables enter with small and imprecisely estimated coefficients, and their inclusion leaves our main evidence unchanged.

5.2 Channels

The previous sections have shown that new imported inputs have a positive effect on the introduction of new domestic products. In this section, we discuss the mechanisms through which this effect takes place. Our conceptual framework shows that new imported inputs operate by generating efficiency gains, through two channels. First, they give rise to a scale effect by expanding the range of available intermediates. Second, they may allow countries to access varieties with more favorable price-quality ratios. According to (10), the scale effect is always present, and gets amplified when new imported inputs have relatively low quality-adjusted prices.

We can study these implications empirically by estimating the following specification:

$$NP_{cit} = \beta_{ci} + \beta_t + (\beta_1 + \beta_2 QAPNew_{cit-1}) \cdot NIIov_{cit-1} + \beta_3 QAPNew_{cit-1} + \varepsilon_{cit}, \qquad (15)$$

where c denotes countries, i 2-digit industries, and t years; $QAPNew$ is the average quality-adjusted price of new imported inputs, relative to existing intermediates (details below). The effect of new imported inputs is then given by:

$$\frac{\partial NP}{\partial NIIov} = \beta_1 + \beta_2 QAPNew. \qquad (16)$$

The sum $\beta_1 + \beta_2$ can be used to test for the first mechanism. A positive value of this sum would indeed imply that new imported inputs raise product creation even if they have the same quality-adjusted price as the existing intermediates (so $QAPNew = 1$). This would correspond to the case in which $\eta^* = \eta$ in (10). Instead, the coefficient β_2 can be used to test for the second mechanism. Indeed, if β_2 is negative, the effect of new imported inputs is decreasing in their quality-adjusted price, as implied by (10) when $\eta^* \neq \eta$.

We now explain the calculation of $QAPNew$. Our data contain information on the raw prices (c.i.f. unit values) of all imported varieties of intermediate inputs. We start by dividing each of these prices by a measure of quality (explained below). In particular, for a given variety v imported into country c at time t, let the raw price be denoted by p_{cvt} and the quality by λ_{cvt}: the quality-adjusted price is then equal to p_{cvt}/λ_{cvt}. Next, we compute the mean of these quality-adjusted prices across *new* varieties, and divide it by the mean computed across *existing* varieties; to maximize comparability across products, we compute this ratio separately within each country and year, at the finest possible level of industry disaggregation (4-digit).[32] The resulting variable measures the relative quality-adjusted price of new imported inputs, and serves as a proxy for η^*/η. To match the other variables entering (15), we then aggregate these relative prices at the 2-digit industry level, by taking their weighted average across 4-digit industries; as weights, we use the share of each 4-digit industry in the total number of new imported inputs in the corresponding 2-digit industry. Finally, we pass the resulting variable through the Import Matrices as in (12).

We obtain the quality estimates λ_{cvt} using an approach developed by Khandelwal (2010). Here, we summarize the salient aspects of this methodology, while relegating technical details and estimation results to Appendix C. In this intuitive and tractable approach, quality is the vertical component of a demand model which is devised to also accommodate differences in horizontal characteristics across products. The demand for each variety is modeled as follows: the quantity market share of the variety in the corresponding industry is a function of the variety's price and some controls for horizontal differentiation. These demand functions are estimated industry by industry, and the quality estimates are obtained by summing the variety fixed effects, the time fixed effects, and the residuals from the regressions. Intuitively, these estimates assign higher quality to varieties with greater market shares, conditional on prices and other controls.[33] Using our data on bilateral imports at the product level, we estimate separate demand functions for each country and 4-digit industry.[34] Estimation is performed by 2SLS on the subsample of imported inputs. Overall, we run 3,268 separate regressions using a total of 10 million observations. As a result of this effort, we construct an extremely detailed and widely comprehensive data set, which contains quality estimates for all imported varieties—defined at the finest level of product disaggregation—in each EU country. To the best of our knowledge, no such data set existed before.

The estimates of (15) are reported in Table 9a, where odd-numbered columns refer to OLS estimates and even-numbered columns to IV estimates. As for the latter, we instrument $NIIov$ and its interaction using $Transportation\ Costs$ and its interaction with $QAPNew$. Given that $QAPNew$ is a generated regressor, we accompany the analytical standard errors (reported in square brackets) with bootstrapped standard errors based on 100 replications (reported in round brackets). The baseline estimates are shown in columns (1) and (2). Notably, the sum $\beta_1 + \beta_2$ is positive and highly significant, as implied by the p-value reported in the bottom of panel a). This is consistent with the first mechanism, according to which new imported inputs work by widening the set of available intermediates and thereby generating a scale effect. At the same time,

[32] We do not encounter problems in comparing goods within narrowly-defined industries, as quantities in Comext are always expressed in the same unit (tons). Nevertheless, in a robustness check reported below, we will regress the individual quality-adjusted prices on product fixed-effects, so as to remove any remaining differences in product (hence industry) characteristics.

[33] This feature is in line with our simple model, where the relative demand for a variety is increasing in its relative quality—holding relative prices fixed—as can be easily seen using (29) in Appendix A.

[34] To ensure comparability across products, the demand functions must be estimated at the finest possible level of industry disaggregation, which is 4-digit NACE in the European case. See Amiti and Khandelwal (2013) for a discussion of comparability issues in quality estimation.

Table 9: Channels

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
a) Quality-adjusted prices						
$NIIov$ (β_1)	1.352***	1.844***	1.329***	1.748***	1.277***	1.756**
	[0.236]	[0.520]	[0.233]	[0.496]	[0.411]	[0.929]
	(0.249)	(0.487)	(0.245)	(0.467)	(0.413)	(0.732)
$NIIov * QAPNew$ (β_2)	-0.691***	-1.343***	-0.671***	-1.256***	-0.608*	-1.233*
	[0.190]	[0.467]	[0.188]	[0.448]	[0.367]	[0.874]
	(0.217)	(0.437)	(0.215)	(0.420)	(0.384)	(0.698)
$QAPNew$ (β_3)	0.076***	0.155***	0.075***	0.146***	0.064	0.139*
	[0.023]	[0.056]	[0.023]	[0.054]	[0.043]	[0.104]
	(0.026)	(0.052)	(0.025)	(0.050)	(0.045)	(0.080)
$\beta_1 + \beta_2$ (p-value)	0.000	0.000	0.000	0.000	0.000	0.000
Obs.	4229	4229	4229	4229	4229	4229
R^2	0.13	0.12	0.13	0.12	0.13	0.12
F-stat. for weak ident.	-	201.9	-	209.3	-	249.9
b) Raw prices and quality						
$NIIov$ (β_1)	-0.051	0.346	-0.046	0.361	-0.037	0.416
	[0.166]	[0.268]	[0.165]	[0.265]	[0.169]	[0.276]
	(0.193)	(0.326)	(0.194)	(0.322)	(0.196)	(0.345)
$NIIov * QNew$ (β_2)	1.880***	1.956***	1.821***	1.820***	1.663***	2.060**
	[0.419]	[0.690]	[0.410]	[0.651]	[0.541]	[1.114]
	(0.354)	(0.687)	(0.350)	(0.645)	(0.533)	(0.936)
$NIIov * PNew$ (β_3)	-0.936***	-1.709***	-0.895***	-1.610***	-0.750*	-1.883**
	[0.249]	[0.567]	[0.245]	[0.544]	[0.429]	[1.053]
	(0.263)	(0.571)	(0.259)	(0.549)	(0.457)	(0.894)
$QNew$ (β_4)	-0.203***	-0.200**	-0.199***	-0.188**	-0.180***	-0.216**
	[0.051]	[0.084]	[0.050]	[0.078]	[0.064]	[0.131]
	(0.044)	(0.086)	(0.043)	(0.081)	(0.061)	(0.108)
$PNew$ (β_5)	0.098***	0.196***	0.095***	0.186***	0.076	0.219**
	[0.029]	[0.068]	[0.029]	[0.066]	[0.050]	[0.126]
	(0.031)	(0.067)	(0.031)	(0.065)	(0.054)	(0.103)
$\beta_1 + \beta_2 + \beta_3$ (p-value)	0.000	0.000	0.000	0.000	0.000	0.000
Obs.	4229	4229	4229	4229	4229	4229
R^2	0.14	0.13	0.14	0.13	0.13	0.13
F-stat. for weak ident.	-	143.1	-	147.7	-	160.7
Specification	Baseline		Excluding the residuals from the quality estimates		Removing product fixed-effects from quality and prices	

Notes: The dependent variable is NP, the share of new domestic products in the total number of domestic goods. $QNew$, $PNew$, and $QAPNew$ are, respectively, the average quality, price, and quality-adjusted price of new imported inputs, relative to existing intermediates. All specifications control for country-industry and year effects. The standard errors are corrected for two-way clustering at the country-industry and industry-year level (square brackets) or bootstrapped (100 replications, round brackets). The level of industry aggregation is NACE2. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

the coefficient β_2 is negative and very precisely estimated, implying the effect of new imported inputs to be decreasing in their quality-adjusted price. This is consistent with the second mechanism, according to which new imported inputs may work also by changing the composition of the inputs set, toward varieties with more favorable price-quality ratios.[35]

In the remaining columns, we perform robustness checks using alternative variants of our

[35] By (16), the effect of new imported inputs varies across observations depending on the value of $QAPNew$. We find that in our data the effect is positive across the entire distribution of $QAPNew$. We also note that the derivative of (15) with respect to $QAPNew$, evaluated at the sample mean of $NIIov$, is negative as expected: -0.021 (s.e. 0.010).

proxy for quality-adjusted prices. Specifically, in columns (3) and (4) we compute $QAPNew$ using different quality estimates, obtained by excluding the residuals from the definition of λ_{cvt}. In columns (5) and (6), we instead regress the individual quality-adjusted prices on product fixed effects, so as to further clean up these variables from product and industry characteristics. We then use the residuals from these regressions to construct $QAPNew$. Reassuringly, our main results are robust across all specifications.

Variation in quality-adjusted prices can stem from variation in quality and/or raw prices. For completeness, we now study how each of these two components contributes to our previous results. To this purpose, we modify (15) by replacing $QAPNew$ with two new variables, $PNew$ and $QNew$, which measure, respectively, the average raw price and the average quality of new imported inputs, relative to existing intermediates. Both variables are constructed following the same steps as for $QAPNew$. The new estimating equation reads as follows:

$$NP_{cit} = \beta_{ci} + \beta_t + (\beta_1 + \beta_2 QNew_{cit-1} + \beta_3 PNew_{cit-1}) \cdot NIIov_{cit-1} + \\ + \beta_4 QNew_{cit-1} + \beta_5 PNew_{cit-1} + \varepsilon_{cit}.$$

We expect $\beta_1 + \beta_2 + \beta_3 > 0$, implying that new imported inputs boost product creation even if they have the same characteristics as the existing intermediates (so $QNew = PNew = 1$). Moreover, we expect $\beta_2 > 0$ and $\beta_3 < 0$, implying the effect of new imported inputs to be increasing in quality (conditional on prices) and decreasing in prices (conditional on quality). The results are reported in Table 9b. We consider the same specifications as in panel a). Namely, in columns (1) and (2) we use the baseline quality estimates; in columns (3) and (4) we use quality estimates obtained by excluding the residuals from the definition of λ_{cvt}; and in columns (5) and (6) we clean up prices and quality from product fixed-effects. As before, in the IV regressions we instrument $NIIov$ and its two interactions using $Transportation\ Costs$ and its interactions with $PNew$ and $QNew$. Note that the estimated coefficients conform with our expectations in all specifications.

5.3 New Imported Inputs and Growth

In the endogenous growth model with expanding variety, the introduction of new products represents technological progress and thus constitutes the 'engine of growth' for the economy. This is also true in the theoretical framework presented in Section 4: in equilibrium, output and technology grow at the same rate, so a higher rate of introduction of new products comes hand-in-hand with a higher rate of output growth (see also the discussion in footnote 18). In the previous sections, we have found robust evidence that new imported inputs stimulate product creation. Accordingly, we expect new imported inputs to play an important role also for growth. In this section, we look for evidence on this effect. To this purpose, we estimate specifications similar to (11), in which we replace the dependent variable with the change in output per worker, $\Delta \ln (Y/L)$. This approach is similar to the one employed by Backus, Kehoe, and Kehoe (1992), who regress the same proxy for growth on indicators of trade in intermediates, using

Table 10: New Imported Inputs and Growth

	a) Industry level				b) Country level			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$NIIov$	0.187***	0.398***	0.395***	0.432***	0.197**	0.466***	0.479***	0.496***
	[0.058]	[0.073]	[0.074]	[0.074]	[0.103]	[0.121]	[0.122]	[0.135]
ln Output		0.131***	0.127***	0.172***		0.148***	0.156***	0.172***
		[0.023]	[0.024]	[0.024]		[0.041]	[0.048]	[0.046]
Capital-output ratio			0.172	0.172*			-0.164	-0.062
			[0.108]	[0.097]			[0.248]	[0.229]
ln Trade openness				0.140***				0.126***
				[0.021]				[0.049]
Obs.	4908	4908	4886	4883	225	225	224	224
R^2	0.08	0.12	0.13	0.17	0.34	0.39	0.39	0.41

Notes: The dependent variable is the change in output per worker, $\Delta \ln (Y/L)$. All specifications are estimated by OLS. In panel a), the level of industry aggregation in NACE2; the specifications control for country-industry and year effects, and standard errors are corrected for two-way clustering at the country-industry and industry-year level. In panel b), the specifications control for country and year effects, and standard errors are corrected for clustering at the country level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

cross-country data for the 1970s and 1980s.

The first set of results is reported in Table 10a. In column (1), we estimate a baseline specification without controls. The coefficient on $NIIov$ is positive and highly significant, suggesting a strong positive association between new imported inputs and the growth rate of output per worker. As discussed by Backus, Kehoe, and Kehoe (1992), a number of models in the endogenous growth literature imply increases in scale to raise growth. While new imported inputs generate scale effects in our model, other theories predict different determinants of scale economies. For example, growth theories based on learning by doing (e.g. Young, 1991) suggest growth to be increasing in industry size. Accordingly, in column (2) we control for log output, a proper proxy for scale according to these models. In line with Backus, Kehoe, and Kehoe (1992), we find this variable to be strongly positively correlated with $\Delta \ln (Y/L)$. If anything, however, controlling for log output increases the size and precision of our main coefficient.[36]

Next, we add controls for other factors that may be relevant for growth and have been extensively studied in the empirical literature. In column (3) we include the capital-output ratio, a proxy for physical capital accumulation. As expected, the coefficient on this variable is positive, but our previous results are hardly affected. In column (4), we instead include the ratio of imports plus exports over output, which accounts for the effects of trade openness. This new control enters with a positive and precisely estimated coefficient, but its inclusion slightly increases the coefficient on $NIIov$. In terms of magnitude, the results in column (4) imply that a 1 p.p. increase in $NIIov$ is associated with an increase of 0.4 p.p. in $\Delta \ln (Y/L)$.

In panel b), we re-estimate the four specifications after aggregating the data at the country-level, in order to make our results more directly comparable with previous studies based on

[36]Alternative theories based on human capital accumulation suggest growth to be driven by the size and intensity of human capital. Accordingly, we have also tried to control for the number of high-skill workers and for their share in total employment. The coefficients on these variables turned out to be insignificant, and our main results were unchanged.

Table 11: New Imported Inputs and Growth, Robustness Checks

	Coeff.	Std. Err.	Obs.	R^2
a) Underlying trends				
(1) Initial value of new domestic products and new foreign inputs	0.370***	[0.074]	4265	0.19
(2) Pre-sample output growth	0.391***	[0.071]	4816	0.18
(3) Pre-sample change in labor productivity	0.369***	[0.066]	4668	0.18
b) Contemporaneous shocks				
(4) Country-industry specific time trends	0.303***	[0.093]	4883	0.26
(5) Country-*sector*-year dummies: Output	0.241**	[0.115]	4883	0.12
(6) Country-*sector*-year dummies: Labor productivity	0.308***	[0.118]	4883	0.09
c) IV regressions				
(7) Baseline	0.298***	[0.107]	4883	0.17
(8) Excl. most cyclical industries	0.265**	[0.113]	3779	0.15
(9) Excl. most energy-intensive industries	0.299***	[0.115]	3781	0.15
(10) Excl. most volatile industries (Autor, Dorn, and Hanson, 2013)	0.318**	[0.127]	3326	0.16
(11) Controlling for the price index	0.307***	[0.114]	4735	0.17
(12) Controlling for new exported varieties	0.290***	[0.106]	4883	0.17
(13) Controlling for imports of continuing inputs	0.293***	[0.110]	4669	0.19
(14) Controlling for FDI inflows	0.342***	[0.105]	4669	0.17
(15) Controlling for FDI outflows	0.271***	[0.113]	4549	0.17

Notes: The dependent variable is the change in output per worker, $\Delta \ln (Y/L)$. The level of industry aggregation is NACE2. All specifications include the same control variables as in column (4) of Table 10 (coefficients unreported), plus country-industry and year effects. Standard errors are corrected for two-way clustering at the country-industry and industry-year level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

cross-country data. Despite a dramatic reduction in sample size, our main coefficient remains positive, precisely estimated, and stable in size. At the same time, we still find a positive and highly significant coefficient on log output. Consistent with Backus, Kehoe, and Kehoe (1992), these results suggest that scale effects do arise when focusing on manufacturing, and that proxies for intermediates trade inspired by Rivera-Batiz and Romer (1991) are strongly correlated with growth in manufacturing output per worker.

In Table 11, we perform a number of robustness checks using our richest specification (see column 4 of Table 10a). In the interest of space, we only report the coefficient on $NIIov$; the remaining coefficients are similar to those shown in Table 10a and are available upon request. In panels a) and b), we control for underlying trends and contemporaneous shocks, using the most relevant specifications estimated in Tables 5 and 6, respectively. Reassuringly, the coefficient on new imported inputs remains positive and very precisely estimated, implying that our results are not driven by these factors. However, controlling for contemporaneous shocks reduces the size of the coefficient, suggesting OLS estimates to be upward biased. Hence, in panel c) we turn to IV regressions, instrumenting $NIIov$ with $Transportation\ Costs$. In the first row, we show results from a baseline specification, while in the remaining rows we perform the main robustness checks reported in Tables 7 and 8. Our coefficient of interest is positive and highly significant across all specifications, and its size generally falls within the range of estimates reported in panel b). In particular, the baseline IV regression implies that a 1 p.p. increase in $NIIov$ leads to an increase of 0.3 p.p. in $\Delta \ln (Y/L)$. All in all, these results suggest that new imported inputs are an important stimulus to output growth in manufacturing.

Table 12: Characteristics of New Products

	(1) *ln* Prices	(2) *ln* Quality
New	0.083***	0.021***
	[0.023]	[0.008]
		(0.007)
Obs.	113345	113345
R^2	0.92	0.30

Notes: The dependent variables are indicated in columns' headings. New is a dummy for new domestic products. All specifications are estimated by OLS, controlling for country-year and product-year effects. The standard errors reported in square brackets are analytical and corrected for two-way clustering at the country-product and product-year level. The standard error reported in round brackets is bootstrapped (100 replications). ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

5.4 Characteristics of New Products

In the theoretical framework presented in Section 4, new and old goods have the same characteristics, either because products are symmetric or because the quality of new goods is drawn from the distribution of existing ones. The endogenous growth literature has extended this framework for studying different issues. While the version we use is enough for our main objective (which is to discuss how new imported inputs affect product creation), in this final section we report suggestive evidence on some of the implications of richer versions for the characteristics of new goods. In Howitt (1999), for example, the framework is extended with a second R&D technology producing vertical innovations, which increase average product quality over time. The implication is that, in the data, new goods should exhibit on average higher quality and prices than old products.

Accordingly, we regress log prices and quality on a dummy for new products (labeled New and equal to 1 in the first year in which a good is produced), using the whole sample of 8-digit domestic products. The coefficient on New is expected to be positive in both specifications. We estimate the quality of domestic goods using a methodology proposed by Khandelwal, Schott, and Wei (2013), which is conceptually identical to the approach we used in Section 5.2 but solves a number of issues arising in the context of domestic products, most notably the choice of credible instruments (see Appendix C for details). In the regressions, we control for country-year effects to compare goods within the same country and time period, and for product-year effects to remove differential trends across products. We exclude extreme observations falling in the top and bottom 5% of the distribution of each characteristic. Finally, we report standard errors corrected for two-way clustering at the country-product and product-year level to accommodate, respectively, autocorrelated shocks to each country-product pair and product-specific shocks common to all countries.[37] Since quality is an estimated variable, we also show a bootstrapped standard error based on 100 replications.

[37] Our results are robust to alternative clustering schemes, including one-way clustering by country-product and two-way clustering by country and product (available upon request).

The results are reported in Table 12. In column (1) the dependent variable is log prices, while in column (2) it is log quality. In both specifications, the coefficient on New is positive and significant at the 1% level. Differences between new and old goods are not only precisely estimated but also economically meaningful. In particular, the price of new products exceeds that of old goods by 8% on average, while quality is higher by more than 2%. Overall, these results are broadly consistent with the idea that new goods are upgraded as, in the first year in which they are produced, they exhibit higher prices and quality compared to old products.

6 Conclusion

We studied the relationship between new imported inputs and product creation, using novel micro data for 25 EU countries over 1995-2007. We first showed that new imported inputs stimulate the introduction of new domestic goods. Then, we documented that this effect occurs through a combination of mechanisms, as new imported inputs allow countries to benefit from both wider and better sets of intermediate products. Consistent with these findings, we also showed that new imported inputs give a substantial boost to output growth in manufacturing. Finally, we provided suggestive evidence on the characteristics of new products, showing that the latter are upgraded compared to old goods, in terms of quality and prices. These results bear important implications. In particular, they are at odds with the widespread concern that ever increasing imports can only harm the manufacturing sector of industrialized countries. On the contrary, our findings indicate that favoring trade in intermediates may be an effective strategy to stimulate product creation and eventually boost output growth.

A Model Derivation

Here, we derive the equilibrium for the two versions of the model presented in Section 4.

A.1 Symmetric Inputs

We start by characterizing the equilibrium of the final good sector. Final good producers maximize profits by choosing the optimal quantity of labor and of each intermediate product, taking prices and the number of available intermediates N_t as given. Hence, they solve the following problem:

$$\max_{L,[x_{ht}]_{h\in[0,N_t]}} \frac{1}{1-\alpha}L^{\alpha}\left(\int_0^{N_t} x_{ht}^{1-\alpha}dh\right) - \int_0^{N_t} p_{ht}x_{ht}dh - w_tL \equiv \pi_{it},$$

where w_t denotes wages and p_{ht} the price of intermediate good h. The first-order condition with respect to x_{ht} implies the following isoelastic demand for good h:

$$x_{ht} = Lp_{ht}^{-1/\alpha}, \tag{17}$$

whereas the first-order condition with respect to L yields:

$$w_t = \frac{\alpha}{1-\alpha}L^{\alpha-1}\left(\int_0^{N_t} x_{ht}^{1-\alpha}dh\right). \tag{18}$$

Next, we consider the problem of intermediates producers. Facing the isoelastic demand given by (17), each monopolist maximizes profits by setting a price equal to a constant markup over its marginal cost:

$$p_{ht} = p = \frac{\psi}{1-\alpha} = 1, \tag{19}$$

where the last equality follows from choosing units such that $\psi = 1-\alpha$. Substituting (19) into (17) shows that each monopolist sells the same quantity in every period:

$$x_{ht} = x = L. \tag{20}$$

Monopoly profits are then also constant and equal to:

$$\pi_{ht} = \pi = \alpha L. \tag{21}$$

Substituting (20) into (18) we obtain the equilibrium wage rate:

$$w_t = \frac{\alpha}{1-\alpha}N_t,$$

while using (20) in (3) we can rewrite the R&D technology as follows:

$$\dot{N}_t = \mu\frac{1}{1-\alpha}LN_t. \tag{22}$$

Note that, as mentioned in the text, the last expression implies that there are increasing returns to scale in research.

The BG equilibrium is summarized by two equations relating the interest rate r and the rate g at which new products are introduced. The first equation comes from the demand side of the model. We assume, as in Rivera-Batiz and Romer (1991), that the representative consumer has the following isoelastic preferences:

$$U = \int_0^{\infty} e^{-\rho t} \frac{C_t^{1-\theta} - 1}{1-\theta} dt. \tag{23}$$

Maximizing (23) subject to an intertemporal budget constraint and a No-Ponzi game condition yields the familiar Euler equation:

$$\frac{\dot{C}_t}{C_t} = \frac{r_t - \rho}{\theta}. \tag{24}$$

Since in BG consumption also grows at the rate g, i.e. $\dot{C}_t/C_t$=g, (24) implies that the equilibrium interest rate must be constant:

$$r = \rho + \theta g. \tag{25}$$

Equation (25) is the first equilibrium relation between r and g. The second relation comes instead from the production side of the model and, in particular, from the free-entry condition in R&D. The latter requires that the output cost of resources needed to generate μ new blueprints be equal to the present discounted value of their profits. Using (21) and recalling that the interest rate is constant in BG, this condition can be written as follows:

$$r = \mu \alpha L. \tag{26}$$

Combining (25) and (26) finally yields the equilibrium expression for g reported in the text:[38]

$$g = \frac{\mu \alpha L - \rho}{\theta}.$$

Consider now the effect of new imported inputs. As in the main text, assume that $N_t^* = N_t$. Since inputs are symmetric, the equilibrium quantity of each foreign variety sold in country c is the same as in (20). Then, we can rewrite (22) as:

$$\dot{N}_t = 2\mu \frac{1}{1-\alpha} L N_t, \tag{27}$$

which shows that the effect of new imported inputs is equivalent to a doubling of the productivity parameter μ. It follows that the same amount of resources invested in R&D now yields twice as

[38]As in Acemoglu (2009), we assume the following two conditions to hold: (1) $\mu\alpha L > \rho$, which ensures that $g > 0$; and (2) $(1-\theta)\,\mu\alpha L < \rho$, which guarantees the utility of the representative individual to be finite and the No-Ponzi game condition to be satisfied.

many blueprints as before. Accordingly, (26) becomes:

$$r = 2\mu\alpha L, \tag{28}$$

implying that the interest rate doubles compared to the initial equilibrium in order to restore the free-entry condition. Finally, combining (28) and (25) yields the expression for $\hat{g}$ shown in the text:

$$\hat{g} = \frac{2\mu\alpha L - \rho}{\theta}.$$

A.2 Heterogeneous Inputs

Proceeding as in the previous section, the monopolist producing good h faces an isoelastic demand given by:

$$x_{ht} = Lp_{ht}^{-1/\alpha}\lambda_h^{(1-\alpha)/\alpha}, \tag{29}$$

and thus maximizes profits by setting a price equal to:

$$p_{ht} = p_h = \eta\lambda_h. \tag{30}$$

Note that, as mentioned in the text, the technology parameter η pins down the quality-adjusted price p_h/λ_h. Monopoly profits are then equal to:

$$\pi_{ht} = \pi = \alpha L\eta^{(\alpha-1)/\alpha}. \tag{31}$$

Using (29) and (30) in (7) yields:

$$\dot{N}_t = \mu\frac{1}{1-\alpha}LN_t\eta^{(\alpha-1)/\alpha}.$$

As before, there are increasing returns to scale in research. Moreover, the number of new products $\dot{N}_t$ now also depends on the quality-adjusted price η, and is decreasing in it.

To solve for the BG equilibrium, note that the first equilibrium relation—equation (25)—is unchanged. Using (31), the new free-entry condition is instead given by:

$$r = \mu\alpha L\eta^{(\alpha-1)/\alpha}. \tag{32}$$

Combining (32) and (25) finally yields the expression for g reported in the text:

$$g = \frac{\mu\alpha L\eta^{(\alpha-1)/\alpha} - \rho}{\theta}.$$

Consider now the effect of new imported inputs. Denote by $N_t^* = \xi N_t$ and η^* the number and quality-adjusted price, respectively, of the new foreign intermediates. A foreign producer faces the following isoelastic demand for its product in country c:

$$x_{ht}^{*} = L(p_{ht}^{*})^{-1/\alpha}(\lambda_h^{*})^{(1-\alpha)/\alpha}, \tag{33}$$

and thus maximizes profits by setting a price equal to:

$$p_{ht}^{*} = p_h^{*} = \eta^{*}\lambda_h^{*}. \tag{34}$$

Using (29), (30), (33), and (34) in (7) yields:

$$\dot{N}_t = \omega\mu\frac{1}{1-\alpha}LN_t\eta^{(\alpha-1)/\alpha},$$

where

$$\omega \equiv 1 + \xi\left(\frac{\eta^{*}}{\eta}\right)^{(\alpha-1)/\alpha}.$$

Accordingly, the free-entry condition (32) modifies as follows:

$$r = \omega\mu\alpha L\eta^{(\alpha-1)/\alpha}, \tag{35}$$

showing that the interest rate increases by a factor of ω compared to the initial equilibrium. Finally, combining (35) and (25) we obtain the expression for $\hat{g}$ shown in the text:

$$\hat{g} = \frac{\omega\mu\alpha L\eta^{(\alpha-1)/\alpha} - \rho}{\theta}.$$

B Robustness Checks: Alternative Specifications

In this Appendix, we show that the correlation between new imported inputs and new domestic products is robust across many alternative specifications. As in the main text, we focus on the baseline version of (11) reported in column (4) of Table 3. The results are shown in Table A1. In panel a) we deal with outliers. To this purpose, in row (1) we trim the distributions of NP and $NIIov$ at the 1st and 99th percentiles; in row (2) we replace those extreme observations with the values of the two percentiles ('winsorizing'); in row (3) we exclude industries with extreme values of NP and $NIIov$;[39] in row (4) we exclude countries with extreme values of both variables;[40] and in row (5) we estimate (11) using an outlier-robust procedure, implemented in Stata using the `rreg` routine. If anything, β_1 slightly increases, suggesting that our results are not driven by outliers.

In panel b) we use alternative definitions of the explanatory variable. First, we recompute $NIIov$ using different weights. In particular, in row (6) we calculate ϕ using only the first available Import Matrix for each country. Compared to using average weights across all years, this gives us

[39] 'Tobacco' (NACE 16), 'footwear' (NACE 19), 'coke and petroleum' (NACE 23).
[40] Germany, UK, Latvia, and Lithuania.

Table A1: New Imported Inputs and the Introduction of New Products, Alternative Specifications

	Coeff.	Std. Err.	Obs.	R^2
a) Outliers				
(1) Trimming (1%)	0.637***	[0.088]	3052	0.13
(2) Winsorizing (1%)	0.596***	[0.076]	4583	0.12
(3) Excl. industries with extreme values of dependent and explanatory variable	0.655***	[0.076]	4141	0.12
(4) Excl. countries with extreme values of dependent and explanatory variable	0.564***	[0.080]	3827	0.12
(5) Outlier-robust estimation	0.529***	[0.043]	4583	0.11
b) Alternative definitions of explanatory variable				
(6) Computing weights (ϕ) using first available Import Matrix for each country	0.585***	[0.072]	4577	0.12
(7) Using average weights calculated from Use Matrices	0.589***	[0.072]	4646	0.12
(8) Using year-specific weights calculated from Use Matrices	0.582***	[0.064]	3969	0.13
(9) Excl. capital goods	0.548***	[0.072]	4583	0.12
(10) Excl. capital goods, fuels, and lubrificants	0.537***	[0.071]	4583	0.12
(11) New imported products instead of new imported varieties	1.366***	[0.198]	4583	0.10
c) Identification of new domestic products and new imported inputs				
(12) Only codes that are always present in PC and CN classifications	0.585***	[0.078]	4550	0.15
(13) Excl. first three years of observations for each country	0.336***	[0.118]	3300	0.09
(14) Only products and foreign inputs remaining in sample for all years after entry	2.440***	[0.844]	3502	0.09
d) Alternative estimators				
(15) Pooled Tobit	0.927***	[0.073]	4583	-
(16) Fixed-effect Poisson	0.001***	[0.000]	4491	-
(17) Linear probability model	0.392**	[0.188]	4583	0.07
e) Intensive and extensive margins of product creation				
(18) Correlation between $NIIov$ and entry rate of new firms	0.064**	[0.027]	3327	0.03
(19) Correlation between $NIIov$ and intro. of new products by incumbent firms	1.481*	[0.860]	14005	0.07

Notes: Panels a)-d) report robustness checks for the baseline correlation estimated in column (4) of Table 3. All specifications control for country-industry and year effects, except for row (15) that only includes year dummies. Standard errors are corrected for two-way clustering at the country-industry and industry-year level, except for rows (15)-(17) where they are clustered by country-industry. The level of industry aggregation is NACE2. Panel e) reports results from: a regression of the entry rate of firms in each country and industry on $NIIov$, controlling for country-industry and time effects (row 18); a regression of an indicator for whether a firm has introduced a new product over 2007-2009 on the value of $NIIov$ in 2006, controlling for industry and region dummies (based on cross-sectional firm-level data) (row 19). ***,**,* = indicate significance at the 1, 5, and 10% level, respectively.

a more precise indication of the structure of backward linkages at the beginning of the sample, but may result in more noisy measures. In rows (7) and (8) we instead calculate ϕ from the Use Matrices: we use average weights across all years in row (7) and year-specific weights in row (8). The size and significance of β_1 are always unchanged. Next, we recompute $NIIov$ using different definitions of intermediate inputs. In row (9) we exclude capital goods, and in row (10) we further omit fuels and lubrificants. The results are unaffected. In row (11), instead of considering new *varieties* (combinations of products and trading partners), we restrict the analysis to entirely new *products*, i.e. 8-digit codes that were never imported before from *any* trading partner. We still obtain a strong positive correlation between new imported inputs and new domestic products: a 1 standard deviation increase in the share of new imported inputs (corresponding to 1.5 p.p.) leads to an increase of 0.2 standard deviations in NP.[41]

Panel c) addresses concerns with our identification of new domestic products and new imported inputs. The first concern is that the commodity classifications may not immediately ad-

[41] In unreported regressions, we have redefined new imported inputs (either varieties or products) using 6-digit rather than 8-digit product codes. The results are very similar and are available upon request.

just to the invention of new products. Until these goods are assigned their own codes, firms would thus report their production under existing codes (Pierce and Schott, 2012). As a consequence, we would count these products as new only with some delay. To address this issue, we reconstruct NP using only the 3,098 codes that are present in the PC classification along the entire sample period. Likewise, we reconstruct $NIIov$ using only the 6,840 codes that are always present in the CN classification. In other words, we restrict to codes that satisfy the third criterion in our definitions of new domestic products and new imported inputs (see Section 3). Reassuringly, the estimate of β_1 reported in row (12) is equal to the one obtained when considering all codes.

The second concern is that our procedure may overestimate the number of new products and imported inputs in the initial years of the sample. To see why, take a generic country and consider a good with positive production in 1996, but not in 1995. Our procedure classifies this good as new for that country in 1996. However, since we do not observe production prior to 1995, we cannot exclude that the good was already produced by the country in some previous year (e.g. in 1994 or earlier). Our procedure becomes more reliable as time passes, since we can track production (and trade) back for a longer period. In row (13) we thus estimate (11) after excluding the first three years of observations for each country. Despite the smaller sample size, our main evidence is preserved.[42]

As explained in Section 3, our procedure identifies new domestic products and new imported inputs independent of how long they remain in the sample. We view this as a strength, since our identification yields the most comprehensive definition of what counts as a new good or a new imported input. One may be concerned, however, that our measures are noisy, since they include products and imported varieties that remain in the sample for just a few years after entry. In row (14) we exclude these cases. That is, we reconstruct NP and $NIIov$ using only products and imported inputs that remain in the sample for all subsequent years after entry. The estimate of β_1 remains positive and highly significant also in this very demanding exercise: a 1 standard deviation increase in the share of new imported inputs (0.6 p.p.) is associated with an increase of 0.2 standard deviations in the share of new products.

In panel d), we consider alternative estimators. In row (15) we estimate (11) by pooled Tobit, to accommodate left censoring in NP (1,454 observations are zero in our sample). If anything, the Tobit marginal effect is larger than the OLS estimate of β_1. In row (16) we use variables in levels (i.e. *counts* of new domestic products and new foreign inputs) and estimate the resulting specification by fixed-effect Poisson with clustered standard errors. The coefficient is still positive and highly significant, implying that one additional input is associated with an increase of 0.1% in the number of new domestic products. In row (17) we change the dependent variable into a dummy equal to 1 if at least one new product is introduced in a given country, industry, and time period. The coefficient on $NIIov$ is positive and precisely estimated, confirming that new imported inputs are associated with the entry of new domestic products.

[42] Excluding the first four or five years of observations yields similar results (available upon request). In that case, however, a few countries would drop out of the sample due to data availability.

Finally, as discussed in the main text, our data encompass two margins of product creation: an *extensive margin* (entry of new firms producing new goods) and an *intensive margin* (introduction of new products within incumbent firms). While we cannot disentangle these margins using our data, we can shed some light on each of them using complementary information from other sources. We do so in panel e). In row (18), we regress the entry rate of firms in each country and industry (available from Eurostat for the period 1997-2003) on $NIIov$, controlling for country-industry and time effects. The estimated coefficient is positive and statistically significant, suggesting that new imported inputs are likely to work along the extensive margin. In row (19), we instead use data for a cross-section of firms in seven EU countries (sourced from the Efige data set), and regress an indicator for whether a firm has introduced a new product over 2007-2009 on the value of $NIIov$ in 2006, controlling for industry and region dummies.[43] The estimated coefficient is positive, suggesting that new imported inputs are also likely to work along the intensive margin.

C Quality Estimation

In this Appendix, we provide details on the quality estimates used in Sections 5.2 and 5.4. For the analysis in Section 5.2, we needed quality estimates for all input varieties imported into each EU country. We obtained these estimates using a methodology developed by Khandelwal (2010). Here, we build on his work to explain this approach.

The demand for variety v in period t is modeled as follows (we omit country and industry subscripts, since this specification is estimated separately for each 4-digit industry and country):

$$\ln s_{vt} - \ln s_{0t} = \beta_v + \beta_t + \beta_1 \ln p_{vt} + \beta_2 \ln ns_{vt} + \beta_3 \ln pop_{nt} + \varepsilon_{vt}. \tag{36}$$

In (36), s_0 is the market share of an outside variety (domestic product), which is set to 1 minus import penetration in the industry.[44] $s_v \equiv q_v/MKT$ is the market share of variety v in the corresponding 4-digit industry, where q_v denotes the quantity of v and $MKT \equiv \sum_v q_v/(1-s_0)$. p_v is the price (c.i.f. unit value) of variety v. $ns_v \equiv q_v/\sum_{v\in h} q_v$ is the share of v in the corresponding 8-digit product h ('nest share'); this variable prevents the quality estimates from being influenced by the higher substitutability of varieties within products than across products. pop_n is partner n's population, which controls for 'hidden varieties'.[45] Log quality is then given by $\ln \lambda_{vt} = \beta_v + \beta_t + \varepsilon_{vt}$, where the variety fixed effect β_v captures the time-invariant valuation of v, the year fixed effect

[43] The seven countries are Austria, France, Germany, Hungary, Italy, Spain, and the UK. The Efige data set is described at www.efige.org.

[44] We calculate import penetration in each country, 4-digit industry, and year, using import and turnover data from Eurostat.

[45] Partner n could export different subproducts of h, classified under more detailed categories than available in the trade data (e.g. different colors). These hidden varieties would increase the market share of v even if all subproducts had the same quality as the exports of h from other partners. Population size controls for hidden varieties. Together with the nest share, it thus accommodates differences in horizontal characteristics across products. We source population data from the World Development Indicators.

Table A2: Summary Statistics on the Quality Estimates

	(1)	(2)
Coefficient on price: mean (median)	-0.955 (-0.668)	-0.235 (-0.192)
Coefficient on nest share: mean (median)	0.462 (0.514)	0.850 (0.868)
Observations per estimation: mean (median)	3128 (1651)	3128 (1651)
Varieties per estimation: mean (median)	854 (459)	854 (459)
Total number of estimations	3268	3268
Total observations across all estimations	10222617	10222617
Sargan test, mean p-value	0.2	-
Estimator	2SLS	OLS

β_t captures the secular time trend common to all varieties, and the residual ε_{vt} captures shocks to the valuation of v occurring in year t.

We estimate (36) separately for each country and 4-digit industry. The estimation sample comprises all varieties of intermediate inputs. Estimation is performed by 2SLS to account for possible correlation of p_{vt} and ns_{vt} with ε_{vt}. As in Khandelwal (2010), we use the following instruments: number of varieties within product h; number of varieties exported by partner n; interactions of distance from n with oil prices and product-specific transportation costs;[46] and bilateral exchange rates.[47] We also exclude varieties with extreme unit values, falling below the 5th or above the 95th percentile of the distribution in each country and industry, in line with Khandelwal (2010). Moreover, we restrict to industries in which there are at least 20 varieties with two or more observations.

Table A2 summarizes the results. Column (1) reports summary statistics for the IV regressions. For purposes of comparison, column (2) shows the same statistics for equivalent regressions estimated by OLS. We perform 3,268 separate regressions using more than 10 million observations. The median number of observations per estimation is 1,651, and the median number of varieties per estimation is 459. As expected, the coefficient on ns is positive and the price elasticity negative. The price elasticity estimated by 2SLS is substantially lower than that estimated by OLS, suggesting that the instruments move the coefficient on p in the expected direction. This pattern of results closely matches that of Khandelwal (2010). More importantly, our estimates are also similar in size to those obtained by the author: the median 2SLS estimates reported by Khandelwal (2010) are -0.58 for the price elasticity and 0.46 for the coefficient on the nest share.

We now turn to the estimates used in Section 5.4. Applying (36) to domestic products requires valid instruments for domestic prices, which are hard to find. Hence, we borrow from

[46] To compute product-specific transportation costs, we start from the variety-specific transportation costs available for the US. These data are sourced from Feenstra, Romalis, and Schott (2002) and have been described in Section 5.1.3. We regress these costs on partner fixed-effects to remove the influence of distance from the US. Then, we take the average of the residuals within each 6-digit product, across all trading partners of the US.

[47] We source bilateral exchange rates from the International Financial Statistics of the International Monetary Fund. We have also estimated quality without including the exchange rates among the instruments, as roughly 30% of trade flows for the Euro-area countries in our sample occur at a fixed parity. Nevertheless, the quality estimates were similar to those used in the current version of the paper, and our results on the channels did not differ in any noteworthy way from those reported in Section 5.2 (see Colantone and Crinò, 2011).

Khandelwal, Schott, and Wei (2013) an alternative methodology, which does not require the use of instruments. In particular, we model the demand for variety v in period t as follows:

$$\ln q_{vt} + \sigma \ln p_{vt} = \beta_h + \beta_t + (\sigma - 1) \ln \lambda_{vt}, \tag{37}$$

where q_v is the quantity of variety v; p_v is its price; σ is the elasticity of substitution between varieties; β_h and β_t are product and year effects, respectively. After estimating (37), quality is retrieved by dividing the residuals—the third term on the right-hand side of the equation—by $\sigma - 1$. We estimate (37) by OLS, separately for each country and 4-digit industry, dropping varieties with unit values below the 5th or above the 95th percentile of the distribution in each country. We draw elasticities of substitution from Broda, Greenfield, and Weinstein (2006). These estimates are available for each country at the 3-digit level of the Harmonized System classification. Following Khandelwal, Schott, and Wei (2013), we aggregate them up at the 4-digit level of the NACE classification, by taking the median across all corresponding HS3 products.[48]

D Data Availability

Table A3: Data Availability

	Production data	Trade data	Import matrices	Use matrices
Austria	1995-2007	1995-2007	1995, 2000, 2005	1995, 1997, 1999-2006
Belgium-Luxemburg	1995-2007	1988-2007	1995, 2000, 2005	1995, 1997, 1999-2005
Bulgaria	2001-2007	1999-2007	-	2000-2004
Czech Republic	2001-2007	1999-2007	2005	1995-2007
Denmark	1995-2007	1988-2007	1995, 2000-2006	1995-2006
Estonia	2000-2007	1999-2007	1997, 2000, 2005	1997, 2000-2006
Finland	1995-2007	1995-2007	1995-2007	1995-2007
Germany	1995-2007	1988-2007	1995, 2000-2006	1995, 1997-2006
France	1995-2007	1988-2007	1995, 1997, 1999-2006	1995, 1997-2006
Greece	1995-2007	1988-2007	2000, 2005	2000-2008
Hungary	2001-2007	1999-2007	1998, 2000, 2005	1998-2006
Ireland	1995-2007	1988-2007	1998, 2000, 2005	1998, 2000-2006
Italy	1995-2007	1988-2007	1995, 2000, 2005	1995-2006
Latvia	2001-2007	1999-2007	1996, 1998	1996, 1998, 2004
Lithuania	2000-2007	1999-2007	2000, 2005	2000-2006
Netherlands	1995-2007	1988-2007	1995-2002, 2004-2006	1995-2006
Poland	2002-2007	1999-2007	2000, 2005	2000-2005
Portugal	1995-2007	1988-2007	1995, 1999, 2005	1995-2006
Romania	2000-2007	1999-2007	2000, 2003-2006	2000, 2003-2006
Slovakia	1998-2007	1999-2007	2000, 2005	1995-2006
Slovenia	2001-2007	1999-2007	1996, 2000, 2001, 2005	1996, 2000-2006
Spain	1995-2007	1988-2007	1995, 2000, 2005	1995-2006
Sweden	1995-2007	1995-2007	1995, 2000, 2005	1995-2006
United Kingdom	1995-2007	1988-2007	1995	1995-2003

[48] The elasticities are missing for Belgium, Czech Republic, Estonia, and Bulgaria. For these countries, we use the elasticities estimated on similar neighboring economies: Netherlands, Slovakia, Latvia, and Romania, respectively.

References

[1] Acemoglu, D., 2009, *Introduction to Modern Economic Growth*, Princeton: Princeton University Press

[2] Aghion, P. and Howitt, P., 1998, *Endogenous Growth Theory*, Cambridge MA: MIT Press

[3] Aghion, P. and Howitt, P., 2005, "Growth with Quality-Improving Innovations: An Integrated Framework," in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, pp. 67-110

[4] Alcalá, F. and Ciccone, A., 2004, "Trade and Productivity," *Quarterly Journal of Economics*, 119, pp. 613-646

[5] Alesina, A., Spolaore, E. and Wacziarg, D., 2005, "Trade, Growth and the Size of Countries," in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, pp. 1499-1542

[6] Amiti, M. and Khandelwal, A.K., 2013, "Import Competition and Quality Upgrading," *Review of Economics and Statistics*, 92, pp. 476-490

[7] Amiti, M. and Konings, J., 2007, "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *American Economic Review*, 97, pp. 1611-1638

[8] Aristei, D., Castellani, D. and Franco, C., 2013, "Firms' Exporting and Importing Activities: Is There a Two-Way Relationship?," *Review of World Economics*, 149, pp. 55-84

[9] Arkolakis, C., Demidova, S., Klenow, P.J. and Rodriguez-Clare, A., 2008, "Endogenous Variety and the Gains from Trade," *American Economic Review: Papers and Proceedings*, 98, pp. 444-450

[10] Autor, D.H., Dorn, D. and Hanson, G.H., 2013, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, pp. 2121-2168

[11] Backus, D.K., Kehoe, P.J. and Kehoe, T.J., 1992, "In Search of Scale Effects in Trade and Growth," *Journal of Economic Theory*, 58, pp. 377-409

[12] Bas, M. and Strauss-Kahn, V., 2011, "Does Importing more Inputs Raise Exports? Firm Level Evidence from France," *CEPII Working Paper 2011-15*

[13] Bernard, A.B., Blanchard, E.J., Van Beveren, I. and Vandenbussche, H., 2012, "Carry-Along Trade," *NBER Working Paper 18246*

[14] Bernard, A.B., Jensen, J.B., Redding, S.J. and Schott, P.K., 2009, "The Margins of US Trade," *American Economic Review: Papers and Proceedings*, 99, pp. 487-493

[15] Bernard, A.B., Redding, S.J. and Schott, P.K., 2010, "Multiple-Product Firms and Product Switching," *American Economic Review*, 100, pp. 70-97

[16] Besedeš, T. and Prusa, T.J., 2011, "The Role of Extensive and Intensive Margins and Export Growth," *Journal of Development Economics*, 96, pp. 371-379

[17] Boeler, A.E., Moxnes, A. and Ulltveit-Moe, K.E., 2012, "Technological Change, International Sourcing and the Joint Impact on Productivity," *Mimeo*, Dartmouth College

[18] Broda, C., Greenfield, J. and Weinstein, D., 2006, “From Groundnuts to Globalization: A Structural Estimate of Trade and Growth,” *NBER Working Paper 12512*

[19] Broda, C., and Weinstein, D., 2004, “Variety Growth and World Welfare,” *American Economic Review: Papers and Proceedings*, 94, pp. 139-144

[20] Broda, C., and Weinstein, D., 2006, “Globalization and the Gains from Variety,” *Quarterly Journal of Economics*, 121, pp. 541-585

[21] Broda, C., and Weinstein, D., 2010, “Product Creation and Destruction: Evidence and Price Implications,” *American Economic Review*, 100, pp. 691-723

[22] Cameron, C.A., Gelbach, J.B. and Miller, D.L., 2011, “Robust Inference With Multiway Clustering,” *Journal of Business and Economic Statistics*, 29, pp. 238-249

[23] Colantone, I., and Crinò, R., 2011, “New Imported Inputs, New Domestic Products,” *EFIGE Working Paper 34*

[24] Dollar, D., and Kraay, A., 2004, “Trade, Growth, and Poverty,” *Economic Journal*, 114, pp. F22-F49

[25] Ethier, W., 1982, “National and International Returns to Scale in the Modern Theory of International Trade,” *American Economic Review*, 72, pp. 389-405

[26] Feenstra, R.C., 1994, “New Product Varieties and the Measurement of International Prices,” *American Economic Review*, 84, pp. 157-177

[27] Feenstra, R.C., Madani, D., Yang, T. and Liang, C., 1999, “Testing Endogenous Growth in South Korea and Taiwan,” *Journal of Development Economics*, 60, pp. 317-341

[28] Feenstra, R.C., Markusen, J.R. and Zeile, W., 1992, “Accounting for Growth with New Inputs: Theory and Evidence,” *American Economic Review: Papers and Proceedings*, 82, pp. 415-421

[29] Feenstra, R.C., Romalis, J. and Schott, P.K., 2002, “U.S. Imports, Exports and Tariff Data,” *NBER Working Paper 9387*

[30] Feng, L., Li, Z. and Swenson, D.L., 2012, “The Connection between Imported Intermediate Inputs and Exports: Evidence from Chinese Firms,” *NBER Working Paper 18260*

[31] Frankel, J.A. and Romer, D.H., 1999, “Does Trade Cause Growth?,” *American Economic Review*, 89, pp. 379-399

[32] Gancia, G. and Zilibotti, F., 2005, “Horizontal Innovation in the Theory of Growth and Development,” in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, pp. 111-170

[33] Goldberg, P.K., Khandelwal, A.K., Pavcnik, N. and Topalova, P., 2009, “Trade Liberalization and New Imported Inputs,” *American Economic Review: Papers and Proceedings*, 99, pp. 494-500

[34] Goldberg, P.K., Khandelwal, A.K., Pavcnik, N. and Topalova, P., 2010a, “Imported Intermediate Inputs and Domestic Product Growth: Evidence from India,” *Quarterly Journal of Economics*, 125, pp. 1727-1767

[35] Goldberg, P.K., Khandelwal, A.K., Pavcnik, N. and Topalova, P., 2010b, "Multiproduct Firms and Product Turnover in the Developing World: Evidence from India," *Review of Economics and Statistics*, 92, pp. 1042-1049

[36] Gopinath, G., and Neiman, B., 2013, "Trade Adjustment and Productivity in Large Crises," *American Economic Review*, forthcoming

[37] Halpern, L., Koren, M. and Szeidl, A., 2011, "Imported Inputs and Productivity," *Mimeo*, University of California Berkeley

[38] Howitt, P., 1999, "Steady Endogenous Growth with Population and R&D Inputs Growing," *Journal of Political Economy*, 107, pp. 715-730

[39] Hummels, D., 2007, "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, 21, pp. 131-154

[40] Hummels, D., Jorgensen, R., Munch, J. and Xiang, C., 2013, "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data," *American Economic Review*, forthcoming

[41] Hummels, D. and Klenow, P.J., 2005, "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95, pp. 704-723

[42] Kasahara, H. and Lapham, B., 2013, "Productivity and the Decision to Import and Export: Theory and Evidence," *Journal of International Economics*, 89, pp. 297-316

[43] Kasahara, H. and Rodrigue, J., 2008, "Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence," *Journal of Development Economics*, 87, pp. 106-118

[44] Kehoe, T.J. and Ruhl, K.J., 2008, "Are Shocks to the Terms of Trade Shocks to Productivity?," *Review of Economic Dynamics*, 11, pp. 804-819

[45] Kehoe, T.J. and Ruhl, K.J., 2010, "Why Have Economic Reforms in Mexico Not Generated Growth?," *Journal of Economic Literature*, 48, pp. 1005-1027

[46] Kehoe, T.J. and Ruhl, K.J., 2013, "How Important is the New Goods Margin in International Trade?," *Journal of Political Economy*, 121, pp. 358-392

[47] Khandelwal, A.K., 2010, "The Long and Short (of) Quality Ladders," *Review of Economic Studies*, 77, pp. 1450-1476

[48] Khandelwal, A.K., Schott, P.K. and Wei, S.J., 2013, "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters," *American Economic Review*, 103, pp. 2169-2195

[49] Khandelwal, A.K. and Topalova, P., 2011, "Trade Liberalization and Firm Productivity: The Case of India," *Review of Economics and Statistics*, 93, pp. 995-1009

[50] Kugler, M. and Verhoogen, E., 2009, "Plants and Imported Inputs: New Facts and an Interpretation," *American Economic Review: Papers and Proceedings*, 99, pp. 501-507

[51] Kugler, M. and Verhoogen, E., 2012, "Prices, Plant Size, and Product Quality," *Review of Economic Studies*, 79, pp. 307-339

[52] Muendler, M.A., 2004, "Trade, Technology and Productivity: A Study of Brazilian Manufacturers 1986-1998," *CESifo Working Paper 1148*

[53] O'Mahony, M. and Timmer, M.P., 2009, "Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database," *Economic Journal*, 119, pp. F374–F403

[54] Pierce, J.R. and Schott, P.K., 2012, "Concording U.S. Harmonized System Codes Over Time," *Journal of Official Statistics*, 28, pp. 53-68

[55] Ramanarayanan, A., 2012, "Imported Inputs and the Gains from Trade," *Mimeo*, University of Western Ontario

[56] Rivera-Batiz, L.A. and Romer, P.M., 1991, "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics*, 106, pp. 531-555

[57] Saravia, D. and Voigtländer, N., 2012, "Imported Inputs, Quality Complementarity, and Skill Demand," *Mimeo*, UCLA

[58] Sivadasan, J., 2009, "Barriers to Competition and Productivity: Evidence from India," *B.E. Journal of Economic Analysis & Policy*, 9, Advances, article 42

[59] Spolaore, E. and Wacziarg, R., 2005, "Borders and Growth," *Journal of Economic Growth*, 10, pp. 331-386

[60] Trefler, D., 2004, "The Long and Short of the Canada-U.S. Free Trade Agreement," *American Economics Review*, 94, pp. 870-895

[61] Xiang, C., 2005, "New Goods and the Relative Demand for Skilled Labor," *Review of Economics and Statistics*, 87, pp. 285-298

[62] Young, A., 1991, "Learning by Doing and the Dynamic Effects of International Trade," *Quarterly Journal of Economics*, 106, pp. 369-405