Imported input varieties and product innovation : ¤ evidence from five developing countries

M. J.D. Bos, G. Vannoorenberghe

Discussion Paper 2018-9

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

Imported input varieties and product innovation: evidence from five developing countries ☆

Marijke J.D. Bos[a,*], Gonzague Vannoorenberghe[b]

[a] *Ministry of Economic Affairs, The Hague, Netherlands.*
[b] *IRES, Université catholique de Louvain*

Abstract

We examine how access to imported intermediate inputs affects firm-level product innovation in five developing counties. We combine trade data with survey data on innovation and develop a method to determine whether new inputs were essential for the product innovation. We find evidence that the number of newly imported varieties has a significant impact on product innovations that rely on new inputs and provide suggestive evidence that this effect comes from access to better quality imports. We extend our analysis to assess the consequences of the increase in the number of Chinese exporting firms on product innovation in developing countries.

Keywords: product innovation, trade, new intermediate inputs.

JEL Classification: F1

1. Introduction

Even at early stages of development, firm-level innovation is an important component of growth (OECD, 2012). Although these innovations are typically small, incremental, and far from the technology frontier, they can be substantial drivers of growth and those that specifically address local challenges may also bring important welfare improvements. Understanding the drivers of firm-level innovation in developing countries is thus of particular interest. A substantial body of literature has identified some of these drivers, from the level of human capital and financial development in the economy, to the role of sound industrial policies and institutions. A recent strand of literature has looked at the role of trade, and in particular the role of imported intermediate

☆This research was funded with support from the Department for International Development (DFID) in the framework of the research project 'Enabling Innovation and Productivity Growth in Low Income Countries (EIP-LIC)'.

*Corresponding author

Email addresses: m.j.d.bos@minez.nl (Marijke J.D. Bos), gonzague.vannoorenberghe@uclouvain.be (Gonzague Vannoorenberghe)

inputs, in promoting product innovations.

Access to foreign intermediates may be an important determinant of firm-level innovation for a variety of reasons. First, when the imported intermediate input is not available domestically, it allows for the domestic production of better or new final products. Second, the imported inputs may be of superior quality which improves the output product's quality. Finally, foreign inputs can be cheaper or more reliable than the domestic variant, leading to lower costs. Imported intermediate inputs may therefore be of particular importance to developing countries whose domestic manufacturing industries are at early stages of development. The definition of innovation we use is very broad, namely a new or significantly improved product, where new means, new to the establishment, and not necessarily, new to the market.

While existing trade and growth models link the introduction of new intermediate inputs to economic growth through firm-level product innovations, the empirical literature is scant at best. Several papers have studied the effect of intermediate input imports on productivity (Amiti and Konings (2007), Şeker (2012), Vogel and Wagner (2010)), but with the exception of Goldberg et al. (2010), there is no evidence on the link between imported inputs and innovation. In their seminal paper, Goldberg et al. (2010) find that increasing numbers of new input varieties at the industry level in India between 1989 and 1996 accounts for 31 percent of new products in that same period. We aim to provide further empirical evidence of the effect of newly imported input varieties on product innovation in five developing countries (Ghana, Tanzania, Kenya, Uganda and Bangladesh). We use firm-level data collected by the World Bank in the recently developed innovation module of the Enterprise Survey. This data contains both quantitative and qualitative questions about the innovation behavior of firms in developing countries. In the baseline regression, we look at the effect of the number of input varieties[1] imported in a firm's industry on the firm's introduction of an innovative product. We take a broad definition of product innovation, which includes incremental changes to existing products, and define the newness of the product in a local context. We propose a novel method of combining qualitative and quantitative survey data to assess the degree to which a product innovation relies on new inputs. Among all product innovations in our sample, we identify "input-using" innovations (innovations for which the firm declares that it uses new inputs) and "input-essential" innovations (of which the new input is a defining or essential feature). To illustrate this latter category, consider a firm that describes a new product as being different from

[1] In our baseline analysis, we define a variety in an Armington sense as a 6-digit HS product from a particular source country.

the most similar product because: “Now [we] use high quality copper and PVP and earlier [we] did not use improved PVP materials and copper”. Another firm, making wooden doors, mentions that: “Earlier [we] used low quality of local wood and now [we] are using high quality and imported wood”. In these cases, the new inputs are described as an important feature of the innovation, which we call “input-essential innovations”.

One of the main challenges in identifying the effect of increased varieties on innovation is the potential for reverse causality and omitted variable bias. A correlation between imported inputs and innovation can theoretically be driven by both “push” and “pull” factors. Access to previously unavailable inputs enables or inspires firms to use the inputs for a product innovation (push factor), whereas an innovation unrelated to international trade may increase the demand for imported inputs once the manufacturing of the new or improved product has begun (pull factor). By ‘access to previously unavailable inputs’, we mean an increase in the number of imported varieties, which could have been the result of push or pull events. Identifying the push effect of the access to previously unavailable inputs thus requires to rule out the pull factors, as they represent endogeneity in this case. We follow a number of routes to address this issue. First, the concern for reverse causality is mitigated by taking the number of new input varieties prior to the product innovation. Second, we control for a range of firm-level characteristics that may drive both imports and innovation, and finally we use different instrumental variable (IV) strategies.

Our analysis uses two different definitions of imported varieties. In the first part, we follow the spirit of Broda and Weinstein (2006) and define a variety as a country-product (HS 6-digits) pair. The second part uses a detailed Chinese firm-level export dataset, which allows to supplement our analysis with an investigation of the effect of imports of Chinese firm-level varieties. There are three reasons for doing so. First, recent studies have found a significant effect of Chinese import competition on productivity in the European Union (Bloom et al., 2016), employment and innovation in the United States (Autor et al., 2013, 2017), but to the best of our knowledge, we are the first to study the effect of Chinese import varieties on innovation in developing countries. The focus on Chinese firms may be particularly relevant for developing countries as Chinese exports may be of lower quality than those from Europe or the US (Schott, 2008), and thus be more suitable for firms in developing countries. Second, the surge in China’s exports represents a clear push factor: Chinese export growth to the world has been massive and exports to the five countries considered in this chapter grew even stronger. While it is plausible for China’s export growth to have primarily been driven by a reduction in global trade barriers (Autor et al., 2013), we also adopt a method that requires weaker assumptions using a gravity model of trade. Third, in contrast to the data on world exports, the Chinese data allows us to define a variety as firm-product pair, instead of as product-country pair. This firm-product definition of a variety is closer to the new trade literature following Krugman (1979).

We find that a positive and significant effect of the import of new varieties on subsequent input-essential product innovation. These results are robust to controlling for the number of new varieties at the output level, which may induce an import competition effect, as well as to instrumental variable estimations. We show suggestive evidence that this effect comes from access to better quality imports. We however find no robust evidence in favor of a causal firm-variety channel coming from China.

These insights can be used to inform innovation and trade policy, but may also inform future micro-level innovation surveys. As opposed to, for example, the roles of finance, information and markets, the role of intermediate inputs has not received sufficient attention in the WS Enterprise Survey (including the Innovation module), Community Innovation Survey (CIS) and similar firm-level innovation surveys.

The remainder of the paper is organized as follows. Section 2 provides a brief overview of the current literature on varieties, imports and innovation. In Section 3, we introduce the data and highlight the importance of new varieties, and in Section 4, we put forward the empirical model. In Sections 5 and 5.3 we report the main results and Section 6 provides a conclusion.

2. Literature

There is a large and growing body of empirical literature on imported inputs and firm-level outcomes, both in developing and developed economies. Recent studies document that lower tariffs on imported inputs raise the productivity of firms in Brazil (Schor, 2004), Indonesia (Amiti and Konings, 2007), and India (Topalova and Khandelwal, 2011). Halpern et al. (2015) show that imported inputs account for 22% of productivity growth in Hungary from 1992-2003, and that this effect is equally driven by the higher quality of imported goods and by the higher number of imported varieties, which are imperfect substitutes for domestic inputs. Another strand of the literature explores the effect of imported intermediate goods on firms' export scope (Bas, 2012; Bas and Strauss-Kahn, 2014; Aristei et al., 2013) and export quality (Fan et al., 2015).

This paper is mostly related to the influential work by Goldberg et al. (2009, 2010). Their approach is based on Romer (1994), who shows that increasing openness leads to an expansion in the number of available product varieties, thereby raising welfare. While the (static) productivity gains from increased import varieties are well-documented (e.g. Broda and Weinstein (2006), Feenstra (1994)), there is little evidence on the dynamic gains in the form of new domestic varieties (or product innovations). Exploiting exogenous variation in trade liberalization in India between 1989 and 1996, Goldberg et al. (2010) show that access to new input varieties from abroad increases the domestic product scope, defined as the number of products produced by a firm. In a related study, Colantone and Crinò (2014) show that in 25 European countries, a higher share of newly imported varieties in an industry raises the share of new domestic products in that industry. The effects appear to work through both a wider as well as a better set of intermediate inputs, and the new domestic products are an important source of growth. This research differs from these studies in four ways. First, we use a broader measure of product innovation, which includes both new and significantly improved products, thereby capturing an additional margin. Goldberg et al. (2010) count the number of products of a firm, and can therefore not identify whether a new product replaces an old one. Colantone and Crinò (2014) on the other hand identify new products as those that are in a different 8-digit category as the previous ones. Second, using qualitative survey data, we develop a novel measure of the importance of new inputs for a firm's innovation, and show that the effect of imported inputs on innovation is strongest for such input-essential innovations. To the best of our knowledge, we are the first to use qualitative data on innovation in this context. Third, we conduct our analysis for a cross-section of poor countries, giving a broader scope to our analysis. Finally, we provide suggestive evidence of a quality channel by using firm-level data on reasons for using foreign inputs from a novel World Bank survey.

We also relate to the empirical literature on trade and innovation in developing countries using data from firm-level surveys. These surveys are typically designed to measure innovation and generally include questions on product and process innovation, and spending on R&D[2]. These surveys also collect information on a range of other firm characteristics such as employment, sales, age, or export and import behavior[3]. While panel data is uncommon, these surveys often contain some retrospective questions such that information on multiple years may be available. Using innovation survey data to study the relationship between trade and innovation is not new. Alvarez and Robertson

[2]These questions are typically along the lines of: 'During the last three years, did the firm introduce any new or significantly improved product?', and 'During the last three years, did the firm conduct R&D?'.

[3]Sometimes this data is collected in a separate survey, but administered to the same set of a firms so that the information can be linked by a firm id number. This is the case with, for example, the World Bank Enterprise Survey and the follow-up Innovation Module and Innovation Capabilities Module.

(2004) use Chilean plant-level data from the First Survey of Technological Innovation and Mexican plant-level data from the National Survey of Employment, Salaries, Technology, and Training in the Manufacturing Sector. The authors find that exposure to foreign markets is positively correlated with product innovation, R&D and the use of new tools. Şeker (2012) uses data from the World Bank Enterprise Survey from firms in 43 developing countries between 2002 and 2006 to estimate the effect of trade orientation (exporter, importer or both) on innovation, employment, sales and labor productivity. His analysis, however, lacks a strong instrument and can only rely on firm-level controls correlated with both trade orientation and firm innovation for identification[4]. Almeida and Fernandes (2008) focus on the specific technology transfer channel that may affect innovation and find that process innovation (a new way in which the main product of the firm is produced) is related to a set of openness indicators. We differ from these studies by combining firm-level data from the World Bank Enterprise Survey, the subsequent World Bank Innovation Survey and product-level import data from UN Comtrade. Specifically, the Innovation Survey allows us to use novel qualitative information on innovations and substantially refine our data.

Finally, we relate to the literature using the dramatic increase in Chinese exports over the last three decades as a shock to its trade partners. Chinese trade growth has been shown to have wide economic implications, from increasing unemployment in the US and Europe (Autor et al., 2013; Bloom et al., 2016), to spurring technical change and reallocation of employment towards more innovative European firms (Bloom et al., 2016). Schott (2008) argues that developed countries compete with China by moving up the quality ladder. By assuming that observed differences in prices reflect differences in quality, he concludes that Chinese exports are of a lower quality than those from developed countries, a finding that is supported by Kneller and Yu (2008). To investigate the role of imports of Chinese varieties on product innovation, we use data on the number of Chinese firms exporting to the countries in our sample. To our knowledge, we are the first to use customs data at the firm level to measure import varieties, allowing us to stick much closer to the traditional definition of varieties in theoretical international trade literature.

[4]Although the World Bank surveys are administered to a panel, the panel is very small, rendering a fixed effect estimation problematic.

3. Data

3.1. Firm-level data

Our firm-level data comes from the Enterprise Survey (ES) of the World Bank, which covers a wide range of business-related topics and has been administered to 130,000 firms in 135 countries since 2002. The ES has two extra modules: the Innovation Follow-up Survey (IS) and Innovation Capabilities Survey (IC). The latter two follow-up surveys were administered on a subsample of the ES firms and cover the same time-span, so that the information can be merged meaningfully. At the moment, the IC module has been administered to five countries (Bangladesh, Ghana, Kenya, Uganda and Tanzania) in the latest round in 2012/2013, which covers information from financial years 2009/2010-2011/2012. Our sample contains 1898 firms, covering 105 industries (four-digit ISIC). While the ES contains mostly quantitative questions, the IS and IC surveys contain open-ended questions, in particular to describe the firm's main innovative product. As detailed in Section 4.2, this information is key to our novel and more precise measure of firm-level innovation. For these firms, the four-digit ISIC industry is recorded.

3.2. Imports and imported varieties

Product-level import data is obtained from the United Nations Commodity Trade Database (UN Comtrade). This database provides annual product-level (HS six-digit) information on trade flows between any country pair. We use the data as reported by importers on the value (in current dollars) of imports of each product.

3.2.1. Varieties as product-country pairs

We define a 'variety' as a six-digit HS commodity - (origin) country combination in a given year for a given importing country, while we refer to a six-digit HS commodity as a 'product'. In other words, if a country imports a product from four different countries, we say that it imports four varieties of that product. Table 1 below summarizes the total number of varieties per importing country per year, over the period 2005-2013.

TABLE 1 HERE

To gauge the importance of new imported varieties against a mere expansion of import volumes, we decompose total import growth between 2007 and 2009 - two years before the start of the firm's innovation period - into an intensive margin (existing product varieties), an extensive product margin (completely new products) and an extensive variety margin (new varieties of the same product). Table 2 reports the growth share of these three categories (they thus add up to one).

TABLE 2 HERE

Comparing the intensive and extensive margins, in all countries except Tanzania, import growth is largely driven by importing more of already existing varieties (the intensive margin). This is quite different from Goldberg et al. (2009, 2010), who find that about 35% of growth is due to existing varieties, and that most growth (65%) is due to new products. Given that they considered a period in which India opened up significantly, this may not be so surprising. While there is no or even slightly negative growth in the product extensive margin in our sample, there is - with the exception of Ghana - considerable growth in the variety extensive margin. Thus, over this period, more varieties of already existing products became available to the local economies. This means that a product was already imported from at least one country, and is now being imported from more countries.

3.2.2. Chinese varieties defined at the firm-level

We use Chinese firm-level export transaction data from the Chinese Customs Trade Statistics (CCTS) Database compiled by the General Administration of Customs of China, where we exclude non-production firms and middleman companies. This dataset records exports of Chinese firms to all countries in detailed (HS 8-digit) product categories[5]. When concentrating on China, we define the number of imported varieties of a product in a country as the number of Chinese firms selling that product in the country. In contrast to the "Armington" definition of varieties that we use in the rest of the analysis, using a firm as the definition of a variety is closer to the new trade literature following Krugman (1979).

4. Empirical strategy

4.1. Regression equations

We estimate the following cross-section regression:

$$INN_{ijc} = \beta_0 + \beta_1 \ln(NIV_{jc}) + \beta_2 IMG_{jc} + X_{ijc}\gamma + \varepsilon_{ijc}, \tag{1}$$

where the dependent variable is product innovation (INN) between 2009 and 2012 by firm i, in four-digit ISIC industry j in country c. The main variables of interest are the log of new input varieties (NIV) in 2009 and the log change in the value of input imports by the industry (IMG)

[5]We use the most detailed data available, which, in the case of Chinese Customs data, is eight-digit, whereas the Comtrade data is only available at the six-digit level.

as defined below. X_{ijc} is a basic set of controls including dummies for foreign-ownership and government-ownership, age of the firm, and country and industry dummies.

In a separate regression, we interact input variety with a firm-level measure of foreign input use, denoted by FI, which is the share of foreign inputs to total inputs:

$$\begin{aligned} INN_{ijc} = & \quad \beta_0 + \beta_1 \ln(NIV_{jc}) + \beta_2 FI_i + \beta_3 (\ln(NIV_{jc}) * FI_i) + \\ + & \quad \beta_4 IMG_{jc} + X_{ijc}\delta + \epsilon_{ijc}. \end{aligned} \tag{2}$$

4.2. Defining product innovation

We use three ways to measure product innovation at firm-level. The first measure is product innovation ("*Innovation*"), a dummy variable that equals one if the firm introduced any innovative product, and zero otherwise[6]. Second, to check the role of inputs in innovation, we define the variable "*input-using innovation*" which takes a value of one if the firm reports that the main innovative product uses different inputs than products it was already producing, and zero if it either did not use different inputs or did not innovate at all. Of all innovating firms, 58% report the use of different inputs for their main innovation, so new inputs appear as an important feature of innovation. Finally, we go one step beyond the self-reported use of new inputs and define a new variable that captures whether one or more new inputs are essential to the product innovation ("*input-essential innovation*"). This variable takes a value of one if using new inputs is essential to the innovation and zero if no innovative products were introduced or if new inputs were not essential. To classify an innovation as input-essential, we examine the firm's description of its main product innovation and look for a reference to the use of a particular (material) input. We find that 38% of the product innovating firms with non-missing descriptions mention the use of a (new) input for their product innovation. Consider, for example, a firm describing its main innovative product as a toothpaste that uses new chemicals compared to the previous toothpaste it produced. This answer suggests that the use of a new input is at the core of the innovation and we define it as an "input-essential innovation". Under this definition, not all input-using innovations are input-essential innovation. The underlying assumption behind this method is that if an input is (not) mentioned in the innovation's description, it is (not) an essential feature of the innovation. While this method depends on the subjective perception of the respondent, the answers from firms

[6] It is the self-reported answer to the question: 'From fiscal year 2010 to 2012, did this establishment introduce any innovative product or service?', where "innovative" is explicitly defined as "new or significantly improved", and 'new' can be new to the firm.

are the best large-scale proxy for the importance of inputs for innovations that we can obtain in developing countries. Section A in the Appendix outlines the procedure for computing this new variable in more detail.

4.3. Measuring input varieties

Using UN Comtrade import data, for each importing country (c) we calculate the number of trading partners per six-digit HS commodity code (product) (h) in a given year (t) as well as the total imports per product $M_{h,c,t}$. In our baseline estimates, we define a 'new' variety in 2009 as a variety that is imported in 2009 but was not imported in 2008. We show in the Appendix **??** that using different lags (e.g. defining varieties as imported in 2009 but not in 2007) yields similar results, as well as using 2010 or 2008 instead of 2009 as the base year[7]. We denote this number of new varieties in product code h imported by country c in year t as $V_{h,c,t}$. Given the measure of new varieties at the product-level we generate a measure of *input* varieties at the industry-level. We first aggregate from six-digit product-level to two-digit industry level (k) so that we obtain $V_{k,c,t} = \sum_{h\in k} V_{k,c,t}$ and $M_{k,c,t} = \sum_{h\in k} M_{k,c,t}$. This first aggregation is due to the higher level of aggregation of the Input-Output (IO) matrix compared to the Comtrade database of trade. Using the IO matrix, we construct the following measure of new input varieties:

$$NIV_{j,c,t} = \sum_k (\alpha_{j,k} \cdot V_{k,c,t}), \tag{3}$$

where $\alpha_{j,k}$ is the share of input k (as a fraction of total inputs) used by industry j. Similarly, we compute a measure of total imports of the industries supplying inputs to industry j as:

$$IM_{j,c,t} = \sum_k (\alpha_{j,k} \cdot M_{k,c,t}). \tag{4}$$

Since the IO matrix is not available for all countries in our sample, we use the commonly used Indian IO matrix for all countries. Taking one IO matrix for all countries also ensures that the within-industry (across country) variation in imported varieties stems from trade differences only and not from differences in IO coefficients. While large differences in the true (unknown) IO-coefficients may be a concern in theory, di Giovanni and Levchenko (2010) find reassuring evidence that the IO matrices of 55 OECD and non-OECD countries are quite similar across countries.

The growth of imported inputs in (1) is equal to $ln(IM_{j,c,t}) - ln(IM_{j,c,t-1})$. We control for the growth in imported inputs so that the number of newly imported inputs - an extensive margin

[7] The robustness of our results to using pre-crisis years only for international trade is, in that sense, reassuring.

variety effect - can be differentiated from an intensive margin effect. The full list of variables, their description and data source can be found in Table B.1 in Appendix B.

4.4. Endogeneity

A concern in our estimation of regression equation (1) is that imported varieties may be correlated with unobservables, in particular industry-specific import demand shocks. This would for example be the case if producers develop product innovations in response to a domestic demand shock and if these new or improved products require more imported varieties. The estimation may also suffer from a reverse causality bias if innovative firms are more likely to import intermediate inputs due to unobserved characteristics. While there is little empirical evidence linking innovation to importing, previous research has found that productive firms are more likely to export (see for example Wakelin (1998); Bernard and Jensen (1999); Aw et al. (2000); Alvarez and Lopez (2005); Damijan and Kostevc (2015); Şeker (2012)), and thus the decision to innovate and the decision to import may be correlated as well. In both cases, the OLS estimate of the effect of imported varieties on innovation may be biased upward.

Various features of our OLS estimate take care of some of the most obvious concerns of endogeneity. For example, we control for a number of variables that are potentially correlated with both imported varieties and innovation, including firm size (employment), sales, productivity, and the degree of competition, following Almeida and Fernandes (2008) and Alvarez and Robertson (2004) among others. The timing of our dependent and independent variables is also chosen to reduce the risk of reverse causality. Our innovation data consists of a cross-section of firms, declaring whether they introduced a product innovation between 2009 and 2012. We measure the number of newly imported varieties at the start of the period of innovation[8], making it less likely but not unthinkable that the imports are a result of product innovations.

We finally turn to an instrumental variable estimation where we explore three different instruments for new imported input varieties, that we now describe.

Instrument 1. To isolate the supply-driven component of imported varieties, we instrument for the number of new input varieties in industry k in country c using the number of new input varieties in industry k in a similar country. We define a 'similar' country as the one with the closest ranking to our country on the Global Competitiveness Index (GCI) within its geographical region (South Asia for Bangladesh; Sub-Saharan Africa for Kenya, Tanzania, Uganda and Ghana). The Global

[8]We have experimented with different timing assumptions, see Table E.1 in the appendix.

Competitiveness Report is published by the World Economic Forum every year and ranks countries based on their competitiveness, which is defined as "the set of institutions, policies, and factors that determine the level of productivity of a country"(Schwab and Porter, 2008, pp.3). The GCI is a composite measure of a large set of indicators covering 12 different topics ('pillars') that include amongst others institutions, macroeconomic stability, eduction, financial markets, and innovation. While these paired countries[9] may be different in many respects, their similarity in competitiveness as measured by the similarity in institutions and policies that affect productivity is an important reason why we expect the number of imported varieties to be similar across industries as well. The IV strategy will produce an unbiased coefficient estimate of the effect of imported varieties if the common between-industry variation of new imported varieties is driven by exogenous factors, such as falling trade costs and rising comparative advantage of the exporting countries. The exclusion restriction may however fail if industry demand shocks are correlated across similar countries or if a demand shock in country c directly affects the demand for the same industry in the similar country.

Instrument 2. The second instrument that we use is based on the costs to import at the industry-country level. Variations in import costs have been identified as a relevant source of variation in the import of new intermediate inputs. Whereas Goldberg et al. (2010) exploit exogenously imposed changes in import tariffs, Colantone and Crinò (2014) use transportation costs which vary both over time (oil prices) and across industries (weight). We use the number of days it takes to clear inputs through customs in industry j in country i (customs delay) as an instrument for new import varieties, as a speedy customs clearing process should make it easier to import new intermediate inputs[10]. Whereas some products are relatively easy to inspect, others may require more extensive laboratory testing (Fernandes et al., 2015), generating large cross-industry differences. Differences in institutions can also create substantial variations in the time that goods take to clear customs both across countries and industries. Many government agencies including health, standards and environment may also be involved in trade regulation (Choi, 2001) and the total customs clearance process, meaning the time between a good's entry into the country and when the good can be claimed by the firm, will depend on the efficiency at each of the product's relevant agencies. We use the data from the WB enterprise survey to calculate an industry-country measure of customs

[9]The paired countries are: Senegal, Ethiopia, Zambia, Cameroon and Pakistan for Kenya, Uganda, Ghana, Tanzania and Bangladesh, respectively.

[10]Hummels and Schaur (2013) estimate the effect of a day in transit to be comparable to an ad-valorem tariff of 2.1 percent.

clearance days. Specifically, question D.3 in the WB ES reads "In the last fiscal year, when this establishment imported inputs or supplies, how many days did it take on average from the time these goods arrived to their point of entry (e.g. port, airport) until the time these goods could be claimed from customs?". For each industry-country pair, we compute the average number of customs clearance days and we use the log of the number of customs delay (log customs delay) as an instrument for log new input varieties. There is considerable variation in days in customs, ranging from 2 to 42.2 days (excluding bottom and top 5%), with a mean of 16.2 days and a standard deviation of 14.8 days. An analysis of the variance indicates that about two-thirds of the variation comes from differences within countries (across industries), and the remainder stems from between-country differences (within industries). A potential concern is that the industry characteristics that are important for innovation also affect the efficiency at customs and other relevant government agencies. Speeding up customs time may be part of government policy to stimulate or maintain output in specific sectors, which could be in the form of protecting productive and innovative sectors or, alternatively, helping less productive industries. Conversely, productive and powerful sectors may effectively lobby for efficient import processes. We control for this by including an industry measure of labor productivity in 2009.

Instrument 3. To construct our last instrument, we aggregate 6-digit products to the *industry* level and compute for all exporter-importer pairs in Comtrade the log growth of trade between 2008 and 2009. We regress this log growth on a set of exporter-industry (δ_{ks}) and importer-industry (δ_{kc}) fixed effects, as well as the log of the distance between the two countries ($Log(Dist)_{sc}$) and a dummy for their contiguity ($Contig_{sc}$):

$$Log\left(M_{ksc,2009}\right) - Log\left(M_{ksc,2008}\right) = \delta_{ks} + \delta_{kc} + Log(Dist)_{sc} + Contig_{sc} + \epsilon_{ksc}, \tag{5}$$

where M_{ksct} are the imports of k by c from source country s at t. The exporter-industry fixed effect captures supply shocks specific to the industry in the exporting country, controlling for the demand conditions in importing countries. We then define ω_{ksc} as the share of imports in country c and industry k that originated from country s in 2007, i.e.:

$$\omega_{ksc,2007} = \frac{M_{ksc,2007}}{M_{kc,2007}}. \tag{6}$$

We finally compute our instrument for new input varieties in industry j and sector c as:

$$Inst_{jc} = \sum_{k} \alpha_{j,k} \left(\sum_{s} \omega_{ksc,2007} \hat{\delta}_{ks} \right). \tag{7}$$

The bracket on the right hand side captures the extent to which c typically imports from countries that have undergone a large increase in their exporter capabilities in industry k between 2008 and 2009. We then make a weighted average over all industries, where the weight reflects the importance of industry k as an input provider for industry j in the IO matrix. $Inst_{jc}$ is thus a predicted increase in the new input varieties available to industry j in country c, which is based on the change in export capability of the country's past trade partners in the industry, and is unlikely to depend on economic conditions in the importing country.

4.5. Chinese varieties

The large increase in Chinese exports in the past few decades has had a significant impact on productivity in the European Union (Bloom et al., 2016) and employment in the United States (Autor et al., 2013). Compared to 2005, the Chinese have supplied an increasing share of total import varieties and in all our sample countries, China ranks first or second as variety supplier (see Appendix D for an overview of the main import partners per country). Nevertheless, there exists little empirical evidence on the effect of Chinese exports on the performance of domestic firms in developing countries. This research aims to fill this gap.

In addition to being an interesting case to study, the recent surge in China's export is likely to represent a clear 'push' factor, allowing us to isolate the effect of trade liberalization from 'pull' factors, such as increased domestic demand. The integration of China in the global economy in the 2000's has been a striking phenomenon, with a wide array of consequences for many countries. Figure 1 shows the growth of Chinese exports to the world and to the five countries considered in this paper, where we normalize the 2004 value to 100. We also report the growth of the number of exporting firms to the world and to our five countries. While Chinese export growth to the world has been massive, exports to the 5 countries considered in this paper grew even more strongly, by a factor 10, from 2004 to 2012. This growth came hand in hand with a large expansion of the number of Chinese firms exporting to the world and to the 5 countries in our analysis[11].

As stated in Section 3.2.2, the Chinese dataset is sufficiently detailed to allow us to define the number of imported varieties of a product in a country as the number of Chinese firms selling that product in the country. We thus examine the effect of the number of Chinese firm varieties (where a variety is a firm supplying a product) on innovation.

[11]The jump in 2006 may partly be an artifact of the Chinese Customs data, which seems to undergo a structural break between 2006 and 2007. For example, the total value of exports when adding firm-level CCTS data is smaller than the value of exports in COMTRADE before 2006 but is exactly in line from 2007 onwards.

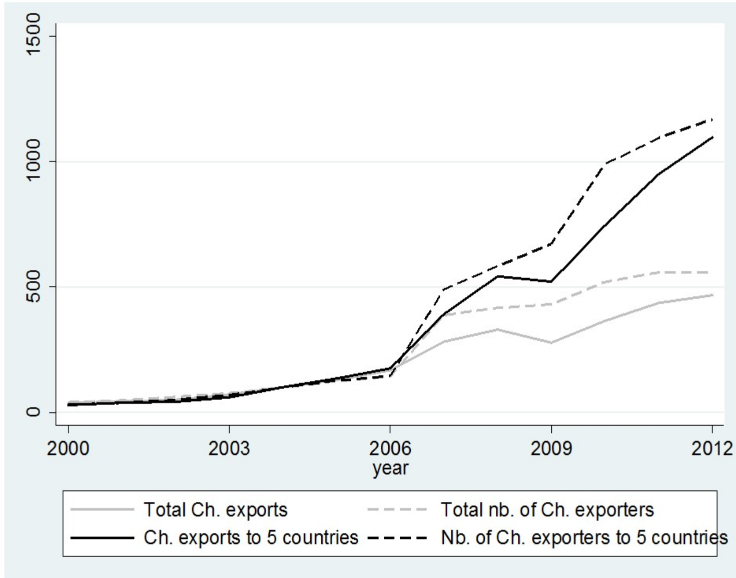

Figure 1: Expansion of Chinese trade to the world and to the 5 countries. Each series is normalized to 100 in 2004. Source: CCTS.

We conduct a similar exercise as in Section 4.3, where we now define the number of Chinese varieties $N_{k,c,t}^{CH}$ as the number of Chinese *firms* that export goods in industry k (4-digit) to importing country c in year t[12] and where $M_{k,c,t}^{CH}$ is the value of imports of input k by country c in year t from China. We then construct the number of Chinese input varieties in an industry and the value of imported inputs as:

$$N_{j,c,t}^{inp,CH} = \sum_k \left(\alpha_{j,k} \cdot N_{k,c,t}^{CH}\right), \tag{8}$$

$$M_{j,c,t}^{inp,CH} = \sum_k \left(\alpha_{j,k} \cdot M_{k,c,t}^{CH}\right). \tag{9}$$

Table 3 shows some summary statistics on the Chinese firm-level variety measure ($N_{j,c,t}^{inp,CH}$) in 2005 and 2009 in our 5 countries, as well as the same measures for French firm-level varieties, computed using data from the French customs, as a comparison. Two observations stand out. First, the number of Chinese varieties is much larger than French varieties. Given the relatively high share of Chinese compared to French imports in our sample countries, this is not surprising. Second, the number of Chinese input varieties has increased significantly between 2005 and 2009, whereas the number of French firms has remained almost the same, suggesting that the push factor is indeed

[12] Note that we here use the log number of varieties and not the number of new varieties as in the previous analysis. The reason for this is that we do not yet have data on the number of new varieties but only on the total number of varieties sold to a country in a year.

the main driver of the number of Chinese varieties.

TABLE 3 HERE

We then test the link between imports of Chinese inputs and firm-level innovation using different variants of the following equation:

$$\begin{aligned} INN_{ijc} = & \quad \beta_0 + \beta_1 \ln\left(N^{inp,CH}_{j,c,2009}\right) + \beta_2 \ln\left(M^{inp,CH}_{j,c,2009}\right) + \beta_3 \ln\left(NIV_{jc}\right) \\ + & \quad \beta_4 IMG_{jc} + \beta_5 \ln\left(N^{CH}_{j,c,2009}\right) + \beta_6 \ln\left(M^{CH}_{j,c,2009}\right) + \boldsymbol{\gamma X_{ijc}} + \varepsilon_{ijc}, \end{aligned} \tag{10}$$

where we use the same set of controls X_{ijc} as in the estimation of equation (1).

Despite the plausibility that China's export growth was primarily driven by a push factor (see Autor et al. (2013) or the comparison with French varieties above), making the problem of endogeneity less severe, the regression in (11) may still be prone to similar endogeneity concerns as in Section 4.4. We thus follow a similar strategy as above and run two regressions using instrumental variables. As a first instrument, we use the number of Chinese input varieties in industry k in a similar country, where the similar country is defined as in the description of Instrument 1 in section 4.4. As a second alternative instrument, we use the exogenous variation in China's export supply capability at the industry level, as defined in section 4.4, interacted with a country-industry measure of initial exposure to China. We thus define our second instrument similarly to equation 7 as:

$$Inst^{CH}_{jc} = \sum_k \alpha_{j,k} \left(\sum_s \omega_{k,CH,c,2002-05} * \hat{\delta}_{k,CH} \right), \tag{11}$$

where $\hat{\delta}_{k,CH}$ is the estimated change in export capability of China in equation (5), and $\omega_{k,CH,c,2002-2005}$ is the share of imports from China by country c in industry k between 2002 and 2005, which reflects the exposure to China before the large increase in Chinese varieties as depicted in Figure 1.

4.6. Summary statistics

Table 4 provides some summary statistics on the innovation variables and new input varieties in 2009. About half of the firms in our sample introduced a product innovation, which is double the average for European countries of 23.7 percent (EU-28; or 26.9% in EU-15) in 2012[13]. Finding higher propensities to innovate in developing countries is not uncommon. For example, Almeida and Fernandes (2008) find a difference of 20 percentage points between the percentage of innovative

[13] Calculated using data from the European Community Innovation Survey 2012, accessed through Eurostat.

firms in 47 developing countries (using data from the World Bank Investment Climate Surveys) and that in European countries. A possible reason for this difference is the relative size of different industries with different propensities to innovate[14], although the most likely cause is a different interpretation of what is 'new' or 'significantly' improved. There is considerable variation in input varieties in 2009, and about one-third of the firms uses at least some inputs of foreign origin in their production process, and a quarter directly imports materials or supplies.

TABLE 4 HERE

5. Empirical results

5.1. Ordinary least squares

5.1.1. Effect of input variety on innovation

This section reports the results of regressing innovation between 2009-2012 on the log of the number of new imported input varieties. A variety is defined as a country-product pair and a new variety is defined as one that is imported in the current year while not in the previous year. We take the number of new input varieties in 2009 (not imported in 2008), i.e. at the beginning of the innovation period, as the independent variable in order to reduce the potential for a reverse effect of product innovation on imported inputs. The dependent variables are innovation, new input innovation, and input-essential innovation. All regressions include four-digit industry dummies, country dummies, three size dummies based on employment (medium size is the omitted variable), a dummy for foreign-owned, a dummy for government-owned and age in years. We first report the results of the Ordinary Least Squares (OLS) regression in Table 5. The regression coefficients in the odd columns suggest that, as expected, imported varieties have a positive and significant effect on product innovations, but only for those innovations that use new inputs. The effect is significantly different from zero and not unsubstantial: a 47 percent increase in the number of input varieties from the mean, corresponding to the standard deviation of 67.39 varieties, raises the probability of an innovation by about 2.7 percentage points ($\frac{47 \cdot 0.57}{100}$). The (unreported) share of variance explained by the log of new input varieties is very small: in column 5 for example, the variable explains close to one percent of the variation. The import growth of inputs is included to control for an intensive margin effect. The number of new varieties may well be correlated with a general increase in imports, and we want to isolate the effect of variety expansion. Import growth

[14]Tables C.1 to C.3 in Appendix C show the number of innovating and non-innovating firms per ISIC sector and country for each of the three definitions of innovation.

of imports enters significantly with a negative sign in the regression with innovation and new input innovation. Out of the other control variables, ownership and age are never significant, while the coefficient for foreign firms is negative and the coefficient on government-owned is positive. The variable age enters negatively, but the effect is not significantly different from zero.

One interpretation of "input-essential innovations" is that one very specific input variety is necessary, and having access to many varieties is irrelevant since the innovation requires only that one particular input. Then it is not the number of varieties that matters, but rather the availability of a single specific input. However, having access to that necessary input variety could very well be affected by the number of varieties imported: the more varieties imported, the larger the chance that a particular variety is imported. A more likely interpretation of the positive effect of input varieties on input-essential innovation is that having more varieties to choose from induces or inspires innovation. Considering the example in the Introduction where the firm replaced a local wood type of low quality with a higher quality imported type of wood: once the foreign wood is imported and available on the local market, the firm observes this and realizes that it has the ability to improve the quality of its product (innovate) by using the better quality wood instead of the lower quality domestic wood. The fact that we observe a positive effect of new imported inputs on innovation for input-essential innovation but not for innovations in general may also suggest the existence of a substitution or crowding out effect from input-essential innovations to other types of innovations.

The even columns show the results of the regression when controlling for the number of new output varieties, as this may be correlated with the number of input varieties as well as innovation through an import-competition effect. In other words, as trade openness leads to more or better inputs for a company to use in its production, the company's output (product) may be subject to more competition now that more varieties of the output product are available on the domestic market. In this way, increasing openness to foreign trade can have an indirect effect on innovation in addition to the effect through imported intermediate inputs. While the former (indirect) effect is interesting, this paper concerns the latter (direct) effect. We control for the number of new output varieties because it might bias the coefficient of new input varieties. While the coefficient on output variety is not significantly different from zero, including it in the regression renders the effect of input varieties on new input innovation insignificant; but the effect on input-essential innovation remains significant and strong.

The results are robust to defining new varieties as those varieties that start being imported in 2008 or 2010 instead of 2009, and to defining a new variety in 2009 as a variety that was not imported in 2007 (instead of 2008) (see appendix Table E.1). We also show in appendix Table E.2 that the different coefficients obtained for the different types of innovation in Table 5 are not simply an artifact of the sample sizes for the the different types of innovation. Conditioning all regressions on the sample of firms that answers all the innovation questions does not affect our conclusions.

TABLE 5 HERE

Because we aggregate the measure of new varieties from product to industry level, the number of new varieties depends, in part, on the number of 6-digit HS products that correspond to the IO category. Compare, for example, IO category 56 ('Rubber products') which has 515 products, to IO category 59 ('Coal tar products') which has 18 products. To control for this, we construct a new measure, called 'Log weighted new input varieties' which divides the number of new varieties per IO by the number of 6-digit HS products (N_j) in that IO before the input variety measure is constructed:

$$NIV^W_{j,c,t} = \sum_k \left(\alpha_{j,k} \cdot \frac{V_{k,c,t}}{N_j} \right). \tag{12}$$

Table 6 reports the results using this measure. The effect of new input varieties on input-essential innovation remains significant and strong.

TABLE 6 HERE

5.1.2. Interactions

Table 7 reports the regression results of Eq. 2, which includes an interaction term of input variety and a measure of access to foreign inputs. Foreign input share is the share of foreign inputs in the firm's total inputs. The same controls as in Tables 5 and 6 are included, but not reported for the sake of brevity, given that their coefficients are of similar size and sign as in the previous Tables. The effect of new input varieties remains significant and the interaction term is positive for input-essential innovation, but not significantly different from zero. We thus find no evidence that firms using foreign inputs innovate more because of access to new input varieties, but rather that all firms benefit from increased foreign input variety. One potential explanation may be that firms do not know the origin of their inputs and therefore misreport the use of foreign inputs, causing a negative bias due to measurement error. It is not unlikely that firms buy their foreign inputs on the domestic market from an importer, making is difficult for the input-using firm to know the origin of the input. Moreover, the share of foreign inputs may not capture the importance of the input for the innovation if it represents only a small fraction of all inputs used. This may hold especially true for firms with multiple products. Another potential explanation could be that the effect of foreign input varieties on innovation runs mainly through an effect on domestic inputs caused by increased competition from foreign input suppliers, inducing domestic input producers to produce better of different intermediate inputs. While the output variety coefficient (which could drive this effect) was positive but insignificant in Table 6, the growth of output enters significantly in the fourth column, providing some evidence in favor of this channel. Unfortunately, the data does not allow us to further investigate what explains the insignificant interaction effect.

The interaction term remains insignificant if the measure of foreign exposure is instead a dummy for direct exporter, or if four-digit industry-country dummies are included, in which case the input variety effect itself cannot be estimated due to collinearity, but the interaction term remains insignificant[15].

TABLE 7 HERE

5.1.3. Channel

To get a better understanding of the channel through which new varieties may positively affect product innovation, we use data from the World Bank Enterprise Innovation Capabilities survey, which asks firms that use foreign inputs why these inputs were sourced abroad rather than domestically. Based on this information we create four dummy variables that equal one if the firm finds the following reasons important; in their respective order, these are: (1) there are no domestic suppliers, (2) similar domestic inputs are more expensive, (3) similar domestic inputs are of poor quality, (4) similar domestic inputs are too unreliable. These reasons are not mutually exclusive. The variables equal zero if the reason was deemed moderately important or not important. Summary statistics on these four dummy variables are reported in Table 8 below. Almost half of the firms indicate availability as an important reason; and one third of the firms deems the other three reasons important. Note that the capability survey is administered to a subset of the WB Innovation Survey sample (which itself is a subset of the World Bank Enterprise Survey), and that this question is only answered by firms that use raw materials of foreign origin (71% of the sample, 821 firms). Moreover, the reasons for importing are not mutually exclusive: firms can report more than one reason[16].

TABLE 8 HERE

[15] For the sake of brevity, these results are not reported but are available upon request.

[16] Even the category 'not available' is not mutually exclusive, because a firm can import more than one input.

Table 9 reports the results of a regression with innovation (yes/no) as the dependent variable and the four reasons for importing on the right-hand side. We use the same controls and fixed effects as in the previous regressions. We find that, for new input innovation and input-essential innovation, the quality reason is significant. These findings are in line with the observed increase in the variety extensive margin in Table 2 in Section 3.2, where most of the increase in new varieties was found to stem from importing more varieties of already existing products.

TABLE 9 HERE

5.1.4. Additional controls

Next, we control for a number of variables that are potentially correlated with both imported varieties and innovation, including firm size (employment), sales, productivity, and degree of competition. The level of competition is captured by three dummies (weak, medium and strong, for 0-5, 6-20 or more than 20 competitors, respectively) and the missing category is medium. Labor productivity is the log of real sales over employment in 2009, and mean labor productivity is industry-country mean labor productivity in 2009. As shown in Table 10, the coefficient for Log new input varieties remains significant and stable around 0.5-0.6.

TABLE 10 HERE

5.1.5. Probit

Our dependent variable in all specifications is a dummy variable for some type of product innovation. While all our previous estimates rely on a linear probability model, we also experiment a probit specification, which allows for a non-linear marginal effect of our explanatory variables. We replicate Table 6 using a probit estimation and report our results in Table 11. The results are very similar to the ones of the linear probability model, and the marginal effects of our main covariates of interest, when estimated at the mean of all independent variables, are very close to the ones of the linear probability model.

TABLE 11 HERE

5.2. Instrumental variables

5.2.1. Baseline IV

While using a lagged measure of input varieties and controlling for a number of observed firm and industry characteristics may alleviate some endogeneity concerns, the OLS coefficient estimates for imported input varieties may still be biased for reasons outlined in section 4.4. Table 12 reports the results of the IV estimation instrumenting for new input varieties using the three instruments described in section 4.4.

The first three columns of Table 12 report the results when using new imported inputs in a similar country as an instrument. The instrument appears strong, and the results are in line with the OLS results in Table 5, with a particularly significant effect of new imported inputs on input-essential innovations. Columns 4-6 of Table 12 show the results using customs delay as an instrument. The sample size is reduced by the fact that some firms do not report customs delays, which are thus missing for some country-industry pairs. The instrument appears a bit weaker, with an F-stat below 10, but the results confirm the positive effect of new input varieties on innovation, in particular for innovations that are related to inputs (columns 5 and 6). Finally, columns 7 to 9 report the results using our measure based on the change in export capability of partners ($Inst_{jc}$) as an instrument. Again, the coefficient on new imported inputs is positive and significant for input-essential innovations. On the whole, our three instrumentation strategies all point to a positive and significant effect of new imported inputs on those innovations that we call "input-essential", while the picture for other types of product innovation is more mixed and less consistent.

TABLE 12 HERE

5.2.2. Robustness

Table E.3 in the appendix presents some extensions of our IV analysis. It reports the results of instrumenting new imported input varieties by our three instruments at the same time. The results confirm the positive and significant effect of new input varieties on input-essential innovations, and a positive but insignificant effect for innovations and new input innovations, while the Hansen test cannot reject the validity of our instruments. In another extension presented in Table E.3, we instrument both the new input varieties and the import growth of inputs. We may be concerned about the endogeneity of the import growth of inputs for the same reasons as we are concerned about the endogeneity of the new imported input varieties. We use our two strongest instruments[17] - newly imported inputs in a similar country and $Inst_{jc}$ as defined in (7) - to instrument the two potentially endogenous variables. The results again confirm the positive and significant (at the 10%) effect of new input varieties and the insignificant effect on the import growth of inputs for input-essential innovations.

[17] Using all three instruments results in a very low F-stat of excluded instruments and in insignificant results across the board

5.3. A variety as a firm-product: China

We now turn to studying how the emergence of China impacted firm-level innovation in our five developing countries. As explained in Section 4.5, in addition to being an interesting case to study, the Chinese data provides a number of benefits over the previous analyses using world-level trade data supplied by UN Comtrade. First, the large increase in Chinese exports is likely to represent a considerable push factor, thus reducing the concern for reverse endogeneity through pull factors. Second, unlike the UN Comtrade dataset, the Chinese data is recorded at the firm-product level.

The OLS estimation of equation (11) is reported in Table 13, where we add controls step by step. In columns 1, 4 and 7, we only use the two measures of Chinese imported inputs (the number of varieties $N^{inp,CH}_{j,c,2009}$ and the value of imported inputs $M^{inp,CH}_{j,c,2009}$) as our main regressors. While these are jointly significant for all types of innovations, the number of imported inputs only appears significantly positive when using the input-essential innovation as our measure. In columns 2, 5 and 8, we add the controls for input imports that we used in Table 5 and show that, once more in the case of input-essential innovations, both the number of imported inputs from China and the new imported varieties defined as country-product pairs appear significant. Finally, in columns 3, 6 and 9, we show that these patterns are robust to controlling for a potential import competition effect measured by the number of varieties and the value of imports in the firms' output industry.

The positive link between Chinese exports and product-innovation in developing countries balances empirical studies that find a negative impacts of China's exports on the exports of other Asian and African countries (Giovannetti and Sanfilippo, 2009; Eichengreen et al., 2004)[18]. It may seem counter-intuitive that intermediate inputs from China, a country that has a low position on the quality ladder (as suggested by Schott (2008) and Kneller and Yu (2008)), can have a substantial contribution to innovation. However, our sample consists of developing countries whose domestic intermediate goods are likely to be of the same or even lower quality. Moreover, while inputs from high income countries may carry the best available technology, they may be less appropriate for developing countries due to the gap in technology and the resulting low absorptive capacity. For developing countries, Chinese imported inputs may be of better quality without being too far away (or too expensive) in terms of technology. Moreover, our definition of innovation includes incremental changes that are new to the firm only, which is less likely to require high technology.

[18]Athukorala (2009) warns, however, that although some crowding-out effects are present, these effects are vastly overstated in the current policy debate.

TABLE 13 HERE

To address issues of endogeneity, we run two IV regressions in a similar spirit as section 5.2.1. We first instrument the number of imported input varieties from China in 2009 by the number of Chinese imported inputs in a similar country, where the similar country is as defined in section 4.4. The results, shown in columns 1 to 3 of Table 14, broadly confirm the results of the OLS regressions and show that the number of Chinese varieties plays a positive and weakly significant role only for input-essential innovations. We should however be very careful in interpreting these results as the instrument is too weak to draw reliable conclusions.

As an alternative instrument, we use the exogenous variation in China's export supply capability at the industry level, to construct an instrument for Chinese input varieties ($Inst_{jc}^{CH}$, as defined in (11). The estimates, reported in columns 4-6 in Table 14, show that the instrument is strong but that there is no clear evidence in favor of a causal impact of Chinese imports of inputs on any type of product innovations that we consider. Unreported results show that adding the other variables used in Table 13 as uninstrumented controls does not affect any of the IV results reported in Table 14, or that using various periods for the change in export capabilities (e.g. 2007-2009 instead of 2008-2009) do not affect the results.

The results for the analysis using Chinese firm-level varieties is thus more mixed. While the OLS estimation confirms a positive association between the number of new Chinese varieties and input-essential innovation, the IV regressions do not give much backing for a strong causal effect.

TABLE 14 HERE

6. Conclusion

Innovation is considered central to growth in developing countries. Even when innovations are incremental and new only to the firm, they can bring important changes that improve welfare. Understanding the determinants of innovation is therefore of great importance to policy makers. This research contributes to a recent and growing body of literature on the effect of intermediate inputs on innovation, productivity and growth. Combining quantitative trade data with survey data from five developing countries - including a novel detailed survey on innovation - we showed that the number of new intermediate input varieties has a significant positive and sizable impact on product innovations that use new inputs and in particular innovations for which a new input is an essential feature. Making use of quantitative data on the product innovation, this finding is supported by establishing a link between imported varieties and innovations for which new inputs are an essential feature. The results are robust to controlling for the number of new varieties at the output level, which may induce an import competition effect, and are robust to instrumental variables estimations. We provide suggestive evidence that the intermediate input effect comes from access to better quality imports, but are unable to confirm that foreign-input using firms benefit more from increased varieties. Our research thus indicates that openness to trade is an important contributor to input-essential innovations in developing countries through its effect on the availability of new input varieties. Policies to increase openness may therefore have a positive effect on the economy through increased innovation, although there seems to be a substitution effect from non-input using innovations to innovations that use new inputs. Despite the importance of Chinese imports, we find no robust evidence in support of a causal innovation effect from Chinese firm varieties. Further research on the origin effect of imported intermediate inputs is warranted to base thorough conclusions on this finding. Innovation has gained a more important role in firm-level surveys, but there is a need for more detailed questions on the role of imports in innovation to better understand this effect.

References

Almeida, R. and A. M. Fernandes (2008). Openness and technological innovations in developing countries: evidence from firm-level surveys. *The Journal of Development Studies 44*(5), 701–727.

Alvarez, R. and R. A. Lopez (2005). Exporting and performance: evidence from Chilean plants. *Canadian Journal of Economics/Revue canadienne d'économique 38*(4), 1384–1400.

Alvarez, R. and R. Robertson (2004). Exposure to foreign markets and plant-level innovation: evidence from Chile and Mexico. *The Journal of International Trade & Economic Development 13*(1), 57–87.

Amiti, M. and J. Konings (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review 97*, 1611–1638.

Aristei, D., D. Castellani, and C. Franco (2013). Firms exporting and importing activities: is there a two-way relationship? *Review of World Economics 149*(1), 55–84.

Athukorala, P.-c. (2009). The Rise of China and East Asian Export Performance: Is the Crowding-Out Fear Warranted? *The World Economy 32*(2), 234–266.

Autor, D., D. Dorn, and G. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *The American Economic Review 103*(6), 2121–2168.

Autor, D., D. Dorn, G. Hanson, G. Pisano, and P. Shu (2017). Foreign competition and domestic innovation: evidence from u.s. patents.

Aw, B. Y., S. Chung, and M. J. Roberts (2000). Productivity and turnover in the export market: micro-level evidence from the Republic of Korea and Taiwan (China). *The World Bank Economic Review 14*(1), 65–90.

Bas, M. (2012). Input-trade liberalization and firm export decisions: Evidence from Argentina. *Journal of Development Economics 97*(2), 481–493.

Bas, M. and V. Strauss-Kahn (2014). Does importing more inputs raise exports? Firm-level evidence from France. *Review of World Economics 150*(2), 241–275.

Bernard, A. B. and J. B. Jensen (1999). Exceptional exporter performance: cause, effect, or both? *Journal of international economics 47*(1), 1–25.

Bloom, N., M. Draca, and J. Van Reenen (2016). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *The Review of Economic Studies 83*(1), 87–117.

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains From Variety. *The Quarterly Journal of Economics 121*(2), 541–585.

Choi, J. Y. (2001). A survey of single window implementation. Technical report, WCO Research Paper.

Colantone, I. and R. Crinò (2014). New imported inputs, new domestic products. *Journal of International Economics 92*(1), 147–165.

Damijan, J. P. and Č. Kostevc (2015). Learning from trade through innovation. *Oxford Bulletin of Economics and Statistics 77*(3), 408–436.

di Giovanni, J. and A. A. Levchenko (2010). Putting the parts together: trade, vertical linkages, and business cycle comovement. *American Economic Journal: Macroeconomics 2*(2), 95–124.

Eichengreen, B., Y. Rhee, and H. Tong (2004). The impact of China on the exports of other Asian countries. Technical report, National Bureau of Economic Research.

Fan, H., Y. A. Li, and S. R. Yeaple (2015). Trade liberalization, quality, and export prices. *Review of Economics and Statistics 97*(5), 1033–1051.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review 84*(1), 157–177.

Fernandes, A. M., R. H. Hillberry, and A. Mendoza Alcantara (2015). Trade effects of customs reform: evidence from Albania. Technical report, World Bank Policy Research Working Paper.

Giovannetti, G. and M. Sanfilippo (2009). Do Chinese exports crowd-out African goods? An econometric analysis by country and sector. *The European Journal of Development Research 21*(4), 506–530.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2009). Trade Liberalization and New Imported Inputs. *The American Economic Review 99*, 494–500.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010). Imported intermediate inputs and domestic product growth: Evidence from India. *The Quarterly journal of economics 125*(4), 1727–1767.

Halpern, L., M. Koren, and A. Szeidl (2015). Imported inputs and Productivity. *The American Economic Review 105*(12), 3660–3703.

Hummels, D. L. and G. Schaur (2013). Time as a trade barrier. *The American Economic Review 103*(7), 2935–2959.

Jaramillo, H., G. Lugones, and M. Salazar (2001). Bogota Manual. Technical report, RICYT.

Kneller, R. and Z. Yu (2008). Quality selection, Chinese exports and theories of heterogeneous firm trade. Technical report, University of Nottingham, GEP Research Paper.

Krugman, P. (1979). Increasing returns, monopolistic competition and international trade. *Journal of International Economics 9*(4), 469–479.

OECD (2012). Innovation for development. Technical report, OECD.

OECD and Eurostat (2005). *Oslo manual: Guidelines for collecting and interpreting innovation data*. Paris: OECD.

Romer, P. (1994). New goods, old theory, and the welfare costs of trade restrictions. *Journal of development Economics 43*(1), 5–38.

Schor, A. (2004). Heterogeneous productivity response to tariff reduction. Evidence from Brazilian manufacturing firms. *Journal of Development Economics 75*(2), 373–396.

Schott, P. K. (2008). The relative sophistication of Chinese exports. *Economic policy 23*(53), 6–49.

Schwab, K. and M. Porter (2008). The global competitiveness report 2008-2009. Technical report, World Economic Forum.

Şeker, M. (2012). Importing, exporting, and innovation in developing countries. *Review of International Economics 20*(2), 299–314.

Topalova, P. and A. Khandelwal (2011). Trade liberalization and firm productivity: The case of India. *The Review of Economics and Statistics 93*(3), 995–1009.

Vogel, A. and J. Wagner (2010). Higher productivity in importing German manufacturing firms: self-selection, learning from importing, or both? *Review of World Economics 145*(4), 641–665.

Wakelin, K. (1998). Innovation and export behaviour at the firm level. *Research policy 26*(7), 829–841.

Tables

Table 1: Total number of varieties (HS6 - origin country) per year

Country	2005	2006	2007	2008	2009	2010	2011	2012	2013	% growth*
Bangladesh	36117	33904	33915	35755	35413	37081	38342	.	.	6.16
Ghana	45637	46612	50005	48042	52946	47161	53916	54313	53520	17.27
Kenya	39396	37719	37867	36654	39277	44888	.	.	49153	24.77
Uganda	26075	26218	27921	29600	31320	32486	32381	33595	32290	23.84
Tanzania	37479	37209	36206	39530	41321	41883	47285	47939	47544	26.86

*The growth rate is the total growth over the period 2005-2013, except for Bangladesh where growth is computed over the period 2005-2011 due to missing data from 2012 and 2013.

Table 2: Share of growth (2007-2009) due to intensive and extensive margin

	Bangladesh	Ghana	Kenya	Uganda	Tanzania
(1) Intensive margin	0.80	1.06	0.80	0.70	0.43
(2) Product ext. margin	-0.03	-0.04	-0.06	0.002	-0.009
(3) Variety ext. margin	0.23	-0.02	0.26	0.30	0.58

Table decomposes total import growth into the extensive and intensive margins between 2007 and 2009. Intensive margin is the contribution to growth due to importing more of already existing varieties, product extensive margin is the share of total growth due to importing completely new products and the variety extensive margin is the share due to importing a product from a new source country. Values are in constant US dollars and are deflated using US wholesale price indices.

Table 3: Chinese and French input varieties: 2005-2009

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Chinese input varieties 2005	60.55	52.46	0.6	190.4	1893
Chinese input varieties 2009	250.52	178.18	3.23	683.85	1893
French input varieties 2009	1.33	1.54	0.01	11.52	1893
French input varieties 2005	1.22	1.86	0.01	15.06	1893

Table 4: Summary statistics

Variable	Mean	St.Dev.	Min	Max	N
Product innovation	0.48	0.5	0	1	1895
New input product innovation	0.28	0.45	0	1	1888
Input-essential product innovation	0.13	0.34	0	1	1537
New input varieties 2009	143.67	67.39	41.67	401.43	1893
Import growth of inputs 2009	-0.13	0.17	-1.52	0.26	1893
Import growth of output 2009	-0.1	0.51	-2.63	2.25	1862
Customs delay	16.23	14.8	1	120	1467
Share of inputs of foreign origin	0.31	0.37	0	1	1813
Direct importer	0.25	0.43	0	1	1868

Details on the variable description and data sources are in Appendix B.

Table 5: Estimation results: Product innovation between 2009-2012 (I)

	Innovation		New input innovation		Input-essential innovation	
	(1)	(2)	(3)	(4)	(5)	(6)
Log new input varieties	0.28	0.29	0.57**	0.47*	0.57***	0.49***
	(0.19)	(0.25)	(0.23)	(0.28)	(0.14)	(0.15)
Import growth of inputs	-0.19*	-0.22*	-0.32**	-0.45***	0.041	0.030
	(0.10)	(0.11)	(0.13)	(0.13)	(0.080)	(0.092)
Log new output varieties		0.0040		0.040		0.038
		(0.052)		(0.053)		(0.030)
Import growth of output		0.030		0.055**		0.0056
		(0.025)		(0.028)		(0.020)
Small	-0.017	-0.016	-0.024	-0.022	-0.038*	-0.042**
	(0.027)	(0.028)	(0.025)	(0.025)	(0.021)	(0.020)
Large	0.058	0.056	0.073*	0.077*	-0.020	-0.023
	(0.036)	(0.037)	(0.040)	(0.042)	(0.033)	(0.033)
Foreign owned	-0.028	-0.024	-0.050	-0.050	-0.011	-0.012
	(0.036)	(0.036)	(0.034)	(0.034)	(0.033)	(0.033)
Government owned	0.043	0.052	0.21	0.23	0.13	0.14
	(0.16)	(0.16)	(0.15)	(0.15)	(0.14)	(0.14)
Age	-0.001	-0.0008	-0.0005	-0.0003	-0.0004	-0.0003
	(0.00085)	(0.00086)	(0.00074)	(0.00074)	(0.00066)	(0.00067)
N	1837	1806	1830	1799	1485	1461

The table reports OLS regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on log new input varieties, import growth of inputs, log new output varieties and import growth of output in 2009. All regressions include country dummies and four-digit industry dummies. Small is a dummy that equals one if the firm has between 5 and 19 employees, large is a dummy that equals one if the firm has more than 100 employees. The omitted category is medium, a dummy that equals one if the firm has between 20 and 99 employees. The sample does not contain micro firms (less than 5 employees). Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 6: Estimation results: Product innovation between 2009-2012 (II)

	Innovation		New input innovation		Input-essential innovation	
	(1)	(2)	(3)	(4)	(5)	(6)
Log new input varieties weighted by HS products	0.031 (0.18)	-0.015 (0.21)	0.26 (0.23)	0.057 (0.23)	0.40*** (0.14)	0.27* (0.16)
Import growth of inputs weighted by HS products	0.029 (0.042)	0.033 (0.045)	-0.0085 (0.047)	-0.013 (0.051)	0.055* (0.031)	0.057* (0.033)
Log new output varieties weighted by HS products		0.047 (0.051)		0.098* (0.056)		0.064* (0.035)
Import growth of output weighted by HS products		-0.0035 (0.023)		0.0030 (0.026)		-0.0052 (0.019)
Small	-0.021 (0.028)	-0.019 (0.028)	-0.028 (0.025)	-0.026 (0.025)	-0.038* (0.021)	-0.042** (0.020)
Large	0.057 (0.036)	0.059 (0.037)	0.071* (0.041)	0.083** (0.041)	-0.025 (0.033)	-0.025 (0.033)
Foreign owned	-0.024 (0.035)	-0.020 (0.036)	-0.045 (0.035)	-0.045 (0.035)	-0.013 (0.033)	-0.014 (0.033)
Government owned	0.028 (0.16)	0.041 (0.16)	0.19 (0.15)	0.21 (0.15)	0.12 (0.13)	0.14 (0.14)
Age	-0.001 (0.00085)	-0.001 (0.00085)	-0.0007 (0.00073)	-0.0005 (0.00073)	-0.0004 (0.00066)	-0.0003 (0.00067)
N	1837	1806	1830	1799	1485	1461

The table reports OLS regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on new input varieties in 2009, import growth of inputs, log new output varieties and import growth of output, where all variables are weighted by HS products. All regressions include country dummies and four-digit industry dummies. Small is a dummy that equals one if the firm has between 5 and 19 employees, large is a dummy that equals one if the firm has more than 100 employees. The omitted category is medium, a dummy that equals one if the firm has between 20 and 99 employees. The sample does not contain micro firms (less than 5 employees). Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 7: Estimation results - Interacting new varieties and access to foreign inputs

	Innovation	New input innovation	Input-essential innovation
Log new input varieties	0.31	0.70***	0.56***
	(0.21)	(0.25)	(0.15)
Import growth of inputs	-0.19	-0.31**	0.053
	(0.12)	(0.14)	(0.088)
Foreign input share	0.18	0.33	-0.20
	(0.44)	(0.38)	(0.32)
(Log new input varieties * Foreign input share)	-0.042	-0.060	0.037
	(0.090)	(0.078)	(0.063)
(Import growth of inputs * Foreign input share)	-0.027	-0.10	-0.079
	(0.25)	(0.21)	(0.15)
N	1770	1763	1427

The table reports OLS regressions of innovation (innovation, new input innovation or input-essential innovation) on log new input varieties, and log new input varieties interacted with foreign input share. All regressions include country dummies, four-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 8: Reasons for importing inputs (not mutually exclusive)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Domestic input not available	0.47	0.5	0	1	820
Domestic input more expensive	0.35	0.48	0	1	821
Domestic input of poor quality	0.33	0.47	0	1	820
Domestic input unreliable	0.3	0.46	0	1	820

Table 9: Estimation results - Reasons for using foreign inputs

	(1) Innovation	(2) New input innovation	(3) Input-essential innovation
Poor quality domestically	0.028	0.069*	0.090**
	(0.040)	(0.041)	(0.035)
Not available domestically	-0.050	0.0093	-0.021
	(0.033)	(0.034)	(0.027)
More expensive domestically	-0.025	-0.037	-0.035
	(0.040)	(0.036)	(0.026)
Unreliable domestically	0.0021	-0.012	-0.013
	(0.036)	(0.036)	(0.030)
N	788	786	622

The table reports OLS regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on reason to use foreign rather than domestic inputs. All regressions include country dummies, four-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 10: Estimation results - Additional controls

	(1) New input innovation	(2) Input-essential innovation	(3) New input innovation	(4) Input-essential innovation	(5) New input innovation	(6) Input-essential innovation
Log new input varieties	0.68^{**}	0.64^{***}	0.54^{*}	0.66^{***}	0.52^{*}	0.61^{***}
	(0.28)	(0.18)	(0.31)	(0.18)	(0.31)	(0.18)
Import growth of inputs	-0.42^{**}	-0.041	-0.30	-0.025	-0.30	-0.046
	(0.18)	(0.11)	(0.20)	(0.11)	(0.20)	(0.10)
Weak competition	-0.016	-0.00066	0.042	0.014	0.040	0.010
	(0.049)	(0.040)	(0.058)	(0.043)	(0.058)	(0.044)
Strong competition	-0.015	0.021	0.017	0.055	0.016	0.051
	(0.040)	(0.032)	(0.047)	(0.034)	(0.047)	(0.034)
Labor productivity			-0.0063	0.0067	-0.0036	0.012
			(0.0093)	(0.0081)	(0.011)	(0.0091)
Mean labor productivity					-0.013	-0.029^{*}
					(0.022)	(0.015)
N	1372	1101	1082	843	1082	843

The table reports OLS regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2010 on log new input varieties in 2009 and additional controls competition and labor productivity in 2009. All regressions include country dummies, four-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 11: Probit estimation results: Product innovation between 2009-2012

	Innovation		New input innovation		Input-essential innovation	
	(1)	(2)	(3)	(4)	(5)	(6)
Log new input varieties	0.95	1.03	1.86**	1.52	2.64***	1.90**
	(0.59)	(0.85)	(0.82)	(1.11)	(0.80)	(0.96)
Marginal effect	*0.38*	*0.41*	*0.61*	*0.50*	*0.53*	*0.38*
Import growth of inputs	-0.50	-0.59*	-1.02**	-1.39***	0.47	0.66
	(0.31)	(0.35)	(0.45)	(0.46)	(0.54)	(0.71)
Marginal effect	*-0.20*	*-0.24*	*-0.33*	*-0.46*	*0.09*	*0.13*
Log new output varieties		-0.0063		0.11		0.31
		(0.17)		(0.22)		(0.21)
Import growth of output		0.077		0.16		-0.059
		(0.076)		(0.10)		(0.13)
Small	-0.034	-0.030	-0.070	-0.064	-0.17	-0.19*
	(0.085)	(0.086)	(0.089)	(0.090)	(0.11)	(0.11)
Large	0.16	0.16	0.23*	0.25*	-0.055	-0.054
	(0.11)	(0.12)	(0.13)	(0.13)	(0.15)	(0.15)
Foreign owned	-0.091	-0.082	-0.20	-0.21	-0.017	-0.027
	(0.11)	(0.11)	(0.13)	(0.13)	(0.19)	(0.19)
Government owned	0.14	0.16	0.66	0.70*	0.52	0.61
	(0.44)	(0.44)	(0.40)	(0.40)	(0.47)	(0.49)
Age	-0.0030	-0.0024	-0.0015	-0.00087	-0.0018	-0.0012
	(0.0026)	(0.0026)	(0.0026)	(0.0026)	(0.0038)	(0.0038)
N	1779	1752	1740	1715	1322	1302

The table reports the coefficients of a Probit regression of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on log new input varieties, import growth of inputs, log new output varieties and import growth of output in 2009. The line "marginal effect" for the first two variables computes their marginal effects estimated at the mean of all variables. All regressions include country dummies and four-digit industry dummies. Small is a dummy that equals one if the firm has between 5 and 19 employees, large is a dummy that equals one if the firm has more than 100 employees. The omitted category is medium, a dummy that equals one if the firm has between 20 and 99 employees. The sample does not contain micro firms (less than 5 employees). Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%. The number of observations differs from the OLS estimates as observations that are perfectly predicted (e.g. if there is only one firm in a sector) are dropped by Stata in the Probit estimation.

Table 12: Estimation results - Instrumental variables estimation

	Inputs in similar country			Customs delay			$Inst_{jc}$		
	(1) Innovation	(2) New input innovation	(3) Input-essential innovation	(4) Innovation	(5) New input innovation	(6) Input-essential innovation	(7) Innovation	(8) New input innovation	(9) Input-essential innovation
Panel A: Second stage									
Log new input varieties	0.74*	0.63	0.53**	0.57	1.62**	0.80*	-0.56	0.15	0.81**
	(0.41)	(0.42)	(0.22)	(0.79)	(0.69)	(0.44)	(0.53)	(0.57)	(0.35)
Import growth of inputs	-0.11	-0.22	0.060	-0.13	-0.25	0.013	-0.078	-0.21	0.050
	(0.11)	(0.13)	(0.083)	(0.11)	(0.15)	(0.080)	(0.13)	(0.15)	(0.083)
Small	-0.012	-0.024	-0.040*	-0.020	0.0079	-0.020	-0.022	-0.027	-0.038*
	(0.028)	(0.025)	(0.021)	(0.034)	(0.031)	(0.026)	(0.029)	(0.025)	(0.021)
Large	0.062*	0.079*	-0.014	0.044	0.065	-0.019	0.057	0.078*	-0.013
	(0.037)	(0.042)	(0.034)	(0.042)	(0.047)	(0.039)	(0.038)	(0.042)	(0.033)
Foreign owned	-0.021	-0.045	-0.00049	-0.028	-0.048	-0.012	-0.028	-0.047	0.0016
	(0.036)	(0.035)	(0.033)	(0.040)	(0.040)	(0.040)	(0.036)	(0.036)	(0.033)
Age	-0.00091	-0.00047	-0.00046	-0.0011	0.000092	-0.00030	-0.0011	-0.00055	-0.00040
	(0.00087)	(0.00075)	(0.00066)	(0.00098)	(0.00086)	(0.00076)	(0.00087)	(0.00076)	(0.00067)
Government owned	0.064	0.21	0.11	0.12	0.31*	0.12	0.0073	0.18	0.12
	(0.16)	(0.15)	(0.14)	(0.18)	(0.16)	(0.18)	(0.17)	(0.15)	(0.14)
Panel B: First stage (main coefficients)									
Dependent variable: Log new input varieties									
Log new input varieties	0.47***	0.46***	0.46***						
in similar country	(0.070)	(0.069)	(0.070)						
Log customs delay				-0.031***	-0.031***	-0.028***			
				(0.0099)	(0.0098)	(0.010)			
$Inst_{jc}$							0.55***	0.54***	0.52***
							(0.12)	(0.12)	(0.11)
N	1787	1780	1442	1418	1412	1136	1787	1780	1442
F-stat	44.4	44.9	42.0	9.75	9.75	7.84	20.5	20.5	21.7

The table reports IV regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on log new input varieties in 2009. In columns 1-3, the instrument for log new input varieties in 2009 is $Inst_{jc}$ as defined in equation 7. In columns 4-6, we instrument both log new input varieties and the import growth of inputs by $Inst_{jc}$ and by the . All regressions include unreported country dummies, four-digit industry dummies and the industry's mean labor productivity in 2009. Small is a dummy that equals one if the firm has between 5 and 19 employees, large is a dummy that equals one if the firm has more than 100 employees. The omitted category is medium, a dummy that equals one if the firm has between 20 and 99 employees. The sample does not contain micro firms (less than 5 employees). The coefficients of the other controls in the first-stage regression are not reported for conciseness. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 13: Estimation results - Log Chinese input (firm) varieties and product innovation

	Innovation			New input innovation			Input-essential innovation		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Ch. input varieties	0.073	0.056	0.076	0.20	0.13	0.25^{*}	0.26^{***}	0.18^{**}	0.24^{***}
	(0.13)	(0.14)	(0.16)	(0.13)	(0.14)	(0.15)	(0.091)	(0.085)	(0.087)
Log Ch. imports of inputs	0.056	0.035	0.019	0.036	-0.021	-0.034	-0.089^{*}	-0.090	-0.093
	(0.086)	(0.100)	(0.11)	(0.082)	(0.095)	(0.095)	(0.053)	(0.057)	(0.058)
Log new input varieties		0.11	0.26		0.42	0.32		0.51^{***}	0.46^{**}
		(0.21)	(0.29)		(0.26)	(0.34)		(0.15)	(0.18)
Import growth of inputs		-0.10	-0.16		-0.26	-0.35^{**}		0.041	0.042
		(0.14)	(0.14)		(0.17)	(0.17)		(0.087)	(0.085)
Log Ch. output varieties			0.022			-0.012			-0.051^{**}
			(0.038)			(0.040)			(0.020)
Log Ch. imports of output			-0.0014			0.011			-0.0012
			(0.015)			(0.015)			(0.0085)
Log new output varieties			-0.042			-0.00013			0.086^{*}
			(0.091)			(0.079)			(0.047)
Import growth of output			0.083^{**}			0.11^{***}			-0.0062
			(0.034)			(0.032)			(0.024)
N	1837	1837	1738	1830	1830	1731	1485	1485	1403

The table reports OLS regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on log Chinese (firm) varieties in 2009. All regressions include (unreported) country dummies, four-digit industry dummies, three size dummies, dummies as well as controls for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table 14: Estimation Results - Chinese varieties: IV estimation

	Inputs in similar country			Export-supply capability		
	(1) Innovation	(2) New input innovation	(3) Input-essential innovation	(4) Innovation	(5) New input innovation	(6) Input-essential innovation
Panel A: Second stage						
Log Ch. input varieties	-0.87	0.45	0.99*	0.21	0.46	0.11
	(0.85)	(0.66)	(0.58)	(0.31)	(0.34)	(0.20)
Log Ch. imports of inputs	0.53	-0.090	-0.45	-0.011	-0.096	-0.011
	(0.43)	(0.33)	(0.30)	(0.17)	(0.18)	(0.10)
Panel B: First stage						
Log Ch. input var. similar country	0.12**	0.12**	0.13**			
	(0.060)	(0.060)	(0.063)			
$Inst_{jc}^{CH}$				6.75***	6.77***	7.53***
				(1.16)	(1.17)	(1.29)
N	1837	1830	1485	1837	1830	1485
F-stat	3.99	3.91	3.96	33.9	33.7	34.1

The table reports IV regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on log Chinese (firm) varieties in 2009. In columns 1-3, the instrument is log Chinese varieties in the same industry in a similar country (see Section 4.4 for the similar countries) and in columns 4-6, the instrument is change in export capability interacted with initial exposure to Chinese imports as defined by $Inst_{jc}^{CH}$ in equation (11). All regressions include country dummies, four-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Appendix A Defining input-essential innovation

The Innovation Follow-up Survey (IM) of the World Bank Enterprise Survey (ES) contains two open questions that are used for the construction of the variable input-essential innovation. Specifically, HB5x describes the main innovative product and HB7x describes how the main innovative product is different from the most similar product already produced by the firm. Using this descriptive information, we aim to capture product innovations for which new inputs played an essential role, as opposed to product innovations that did not require new inputs.

The variable input-essential innovation is coded either 1 (yes), 0 (no) or . (missing) based on specific words or combinations of words that occur in the innovation's descriptions. We follow definitions from the Oslo Manual, which contains guidelines for collecting and interpreting technological innovation data (OECD and Eurostat, 2005) - including the annex based on the Bogotá Manual (Jaramillo et al., 2001) aimed at less-developed and non-OECD countries. We start with 908 product innovations which may or may not be input-essential innovations. First, we try to reduce some of the measurement error stemming from the fact that respondents may not fully understand the survey question on product innovation. A commonly made mistake is to confuse product with process innovation or marketing innovation. For this purpose, we set the variable for product innovation equal to missing (and thus new input innovation to missing) if HB5x or HB7x contains one of the following words: machine, manual, technology, logo, design, package, packaging. This is the case for 128 observations. Second, items in the remaining list of 780 product innovations are classified as input-essential innovations if the descriptions of the innovative product contain one of the following words: using, use, used, uses, ingredient, ingredients, input, inputs, recipe, made from, made of, material, materials. We rule out the word combinations 'industrial use', 'purpose use', 'used to', 'used as', 'used on', and 'used not'. Third, we classify the innovation as input-essential if the respondent self-reports having used a new input for the innovation (question Hb9b in the WB ES)[19] *and* the descriptions mentions a specific input. We require both to rule out the use of new inputs that were less than essential, based on the premise that if they were essential, they would be mentioned in the description. An example of such a description for which the innovation is classified as input-essential innovation is "earlier [we] made toothpaste with normal chemicals and now [we] make toothpaste with improved chemicals". After having a look at the descriptions, the following

[19]Without requiring a positive answer to HB9b, we might define innovations as input-driven, while the input is not new to the firm.

inputs were found: aluminum, polythene, carbonate, carbonate, chemicals, cloth , concrete, cotton, flavor, metal, paper, plastic, polyester, rubber, silk, soya, timber, tin, wood, wooden, yarn, leather. Finally, the variable input-essential innovation is set to missing if there is no answer to HB5x and HB75x.

Table A.1 below reports the cross-tabulation between input-essential innovation and new input innovation. Note that the sample of 780 is innovators only, excluding innovations that were classified as process, organizational or marketing innovation.

Table A.1: Cross-tabulation of input essential and new input innovation

	New input innovation (HB9b)			
Input essential innovation	No	Yes	Missing	Total
No	**220**	24	0	344
	63.95	36.05	0.00	100.00
Yes	33	**173**	0	206
	16.02	83.98	0.00	100.00
Missing	69	156	5	230
	30.00	67.83	2.17	100.00
Total	322	453	5	780
	41.28	58.08	0.64	100.00

First, we note that the number of missing descriptions is fairly large, which is a downside of using quantitative data. Second, for the non-missing data, we find that when an innovation is classified as input-essential, 173 firms (84%) have self-reported use of new inputs. The other 33 firms do not report using new inputs, but in 31 cases the description in HB7x clearly mentions using a different input. Our procedure thus reduces measurement error in the variable product innovation (HB9b). The other two may be misclassifications, as one description mentions what the product is used for, and one mentions using the same input as for its other (old) products. Third, for the non-missing data, we find that, when an innovation is classified as not input-essential, 124 firms (64%) have self-reported use of new inputs. This may be driven by exactly what we are aiming to capture, namely that, in a subset of the new input using innovations, these new inputs were not essential to the innovation. On the other hand, it may also be that a new input is used, but its use is too obvious and therefore not mentioned in the innovation's description. For example, if a firm used to make cement and its new product is soap, it is probably too obvious for the respondent to mention that this new product required new or different inputs. While the former will decrease measurement error, the latter may actually increase the error as we wrongly classify the innovation as non-input-related, while it actually is input-related. In the regression analysis, this effect will bias the estimated coefficient downwards, which means we have to interpret the coefficient estimate as a lower-bound estimate for the true effect.

Appendix B List of variables, descriptions and data sources

Table B.1: List of variables, descriptions and data sources

Variable	Description	Source
Innovation outcomes		IM
Innovation	1 if the firm introduced an innovative product, 0 otherwise	IM
New input innovation	1 if the firm introduced an innovative product using new inputs, 0 otherwise	IM
Input-essential innovation	1 if the firm introduced an innovative product for which a new input was essential, 0 otherwise	IM
Input non-essential innovation	1 if the firm introduced an innovative product for which a new input was not essential, 0 otherwise	IM
Number of innovations	number of innovative product	IM
Varieties		
Total input varieties	The number of product varieties in each input-supplying industry, weighted by the input share	Comtrade
New input varieties	The number of new product varieties in each input-supplying industry, weighted by the input share	Comtrade
New output varieties	The number of new product varieties in the firm's output product industry	Comtrade
Trade		
Growth of inputs	Total import growth in each input-supplying industry, weighted by the input share	Comtrade
Growth of output	Total import growth in the firm's output product industry	Comtrade
Chinese firm varieties		
Chinese input varieties	The number of firm varieties in each input-supplying industry, weighted by the input share	CCTS
Controls and instrument		
Size dummies	Three dummies for size (5-19, 20-99, or more than 100 employees)	ES
Foreign Owned	1 if the firm is (partly) foreign-owned, 0 otherwise	ES
Government Owned	1 if the firm is (partly) state-owned, 0 otherwise	ES
Age	Years since establishment started operations	ES
Competition	Three dummies (low, medium, and strong) for competition (0-5, 6-20, or more than 20 competitors, respectively)	ES
Labor productivity	Real sales divided by the number of full-time employees	ES
Customs delay	Number of days from the time inputs arrive at their point of entry (e.g. port, airport) until the time these goods can be claimed from customs	ES
Reasons for importing		
Poor quality	1 if firm imported because domestic input is of poor quality	IC
Not available	1 if firm imported because domestic input is not available	IC
More expensive	1 if firm imported because domestic input is more expensive	IC
Unreliable	1 if firm imported because domestic input is unreliable	IC

IM= Innovation Module, ES= Enterprise Survey, IC= Innovation Capabilities Module), CCTS = Chinese Customs Trade Statistics Database. Input shares are taken from the Indian 1993 Input-Output matrix.

Appendix C Sample details

Table C.1: Innovation by country and two-digit sector

	Country and Did the firm introduce a product innovation?									
	Ghana		Bangladesh		Tanzania		Uganda		Kenya	
Industry	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Food	34	10	49	75	38	8	27	33	78	43
Textiles	2	2	35	91	19	3	16	17	12	15
Garments	14	5	33	96	36	7	4	6		
Leather	2	2	59	40			1		4	2
Wood	13	1	8	7	14	1	3	4	4	1
Paper	1	2	11	15			1			2
Publishing, printing, and Recorded media	38	10	36	14	12		2	4	4	4
Refined petroleum product	2	1							2	1
Chemicals	6	11	20	69	6	1	1	2	10	18
Plastics & rubber	12	5	11	19	6	2	1	2	2	6
Non metallic mineral products	13	3	7	5	7	3	1	6	4	2
Basic metals	5	3	10	14	4		1	2	1	4
Fabricated metal products	45	6	7	18	20	2	15	16	6	4
Machinery and equipment			9	8	2		2	3	12	8
Electronics (31 & 32)	3		2	5	5	2	1	3	3	4
Precision instruments										1
Transport machines (34&35)		2	5	19	2		1	1	6	10
Furniture	16	11	14	41	51	21	13	18	5	3
Recycling		2						1		

Table C.2: New input innovation by country and two-digit sector

	Country and Did the firm introduce a product innovation that uses new inputs?									
	Ghana		Bangladesh		Tanzania		Uganda		Kenya	
Industry	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Food	38	6	78	46	41	5	44	16	95	26
Textiles	3	1	62	64	21	1	27	4	15	11
Garments	15	4	73	56	40	3	7	3		
Leather	3		77	22			1		4	2
Wood	13	1	11	4	14	1	5	2	4	1
Paper	1	2	19	6			1			2
Publishing, printing, and Recorded media	42	4	41	9	12		4	2	7	1
Refined petroleum product	2	1							2	1
Chemicals	10	7	51	38	6	1	1	2	17	11
Plastics & rubber	13	4	17	13	6	2	1	2	4	4
Non metallic mineral products	13	3	9	3	7	3	7		5	1
Basic metals	6	2	19	5	4		3		2	3
Fabricated metal products	47	4	15	10	21	1	25	6	6	4
Machinery and equipment			12	5	2		3	2	14	6
Electronics (31 & 32)	3		3	4	6	1	3	1	5	2
Precision instruments										1
Transport machines (34&35)	1	1	14	10	2		2		10	6
Furniture	21	6	41	14	56	16	20	11	6	2
Recycling	1	1						1		

Table C.3: Input-essential innovation by country and two-digit sector

	Country and Did the firm introduce a product innovation for which new inputs were essential?									
	Ghana		Bangladesh		Tanzania		Uganda		Kenya	
Industry	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Food	39	1	85	13	43	2	41	6	98	7
Textiles	2		57	38	21		22	3	15	1
Garments	15	2	74	21	39	1	7	1		
Leather	2		72	17			1		4	
Wood	13		9	3	15		4	2	4	
Paper	2	1	18	5			1		1	
Publishing, printing, and Recorded media	40	1	39	10	12		4		7	
Refined petroleum product	3								3	
Chemicals	12	1	53	12	7		2		16	2
Plastics & rubber	12	3	18	7	7		2	1	2	2
Non metallic mineral products	13	1	10	1	8		4		5	1
Basic metals	6		18	3	4		1		2	1
Fabricated metal products	47		11	6	20	2	19	1	7	1
Machinery and equipment			9		2		4		13	3
Electronics (31 & 32)	3		4	2	6	1	3	1	5	
Transport machines (34&35)			19		2		1		11	
Furniture	17	3	32	8	58	6	17	2	6	
Recycling	1									

Appendix D Main import trading partners

Table D.1 reports the list of main import partners in terms of total import values and Table D.2 reports the list in terms of number of varieties. In both cases, China's position has increased from 2005 to the top 3 in 2010 in all sample countries.

Table D.1: Top 5 import countries measured by share of value of imports

	2005			**2010**	
Position	**Country of origin**	**Share**	**Position**	**Country of origin**	**Share**
			Bangladesh		
1	**China**	0.16	1	**China**	0.17
2	India	0.11	2	India	0.12
3	Kuwait	0.07	3	Thailand	0.06
4	Japan	0.06	4	Singapore	0.05
5	Rep. of Korea	0.05	5	China, Hong Kong SAR	0.05
			Ghana		
1	Nigeria	0.12	1	USA	0.14
2	**China**	0.08	2	**China**	0.13
3	United Kingdom	0.08	3	France	0.06
4	USA	0.07	4	Belgium	0.06
5	Belgium	0.06	5	United Kingdom	0.05
			Kenya		
1	United Arab Emirates	0.14	1	**China**	0.13
2	South Africa	0.10	2	United Arab Emirates	0.12
3	USA	0.10	3	India	0.11
4	Saudi Arabia	0.06	4	South Africa	0.06
5	United Kingdom	0.06	5	Japan	0.06
			Uganda		
1	Kenya	0.25	1	India	0.15
2	Japan	0.07	2	Kenya	0.11
3	South Africa	0.07	3	**China**	0.09
4	United Arab Emirates	0.07	4	United Arab Emirates	0.08
5	India	0.06	5	Japan	0.07
			Tanzania		
1	Bahrain	0.16	1	India	0.11
2	South Africa	0.12	2	**China**	0.11
3	**China**	0.07	3	South Africa	0.10
4	Japan	0.06	4	United Arab Emirates	0.08
5	United Arab Emirates	0.06	5	Japan	0.07

Table D.2: Top 5 import countries measured by share of number of varieties

	2005			2010	
Position	**Country of origin**	**Share**	**Position**	**Country of origin**	**Share**
			Bangladesh		
1	**China**	0.21	1	**China**	0.21
2	India	0.17	2	India	0.17
3	Singapore	0.06	3	China, Hong Kong SAR	0.08
4	Other Asia, nes	0.06	4	Singapore	0.08
5	Rep. of Korea	0.05	5	Other Asia, nes	0.06
			Ghana		
1	United Kingdom	0.13	1	United Kingdom	0.12
2	**China**	0.09	2	**China**	0.12
3	Germany	0.08	3	USA	0.10
4	South Africa	0.08	4	South Africa	0.09
5	USA	0.08	5	India	0.07
			Kenya		
1	India	0.14	1	**China**	0.15
2	United Kingdom	0.14	2	India	0.14
3	**China**	0.11	3	United Kingdom	0.11
4	South Africa	0.10	4	South Africa	0.08
5	United Arab Emirates	0.09	5	USA	0.07
			Uganda		
1	United Arab Emirates	0.19	1	**China**	0.16
2	Kenya	0.18	2	United Arab Emirates	0.15
3	South Africa	0.13	3	Kenya	0.14
4	India	0.11	4	India	0.13
5	United Kingdom	0.11	5	South Africa	0.09
			Tanzania		
1	South Africa	0.15	1	**China**	0.14
2	United Arab Emirates	0.13	2	South Africa	0.13
3	**China**	0.12	3	United Arab Emirates	0.13
4	India	0.11	4	India	0.11
5	United Kingdom	0.08	5	United Kingdom	0.07

Appendix E Robustness

Table E.1: Robustness - different years for the definition of innovation

	Innovation			New input innovation			Input-essential innovation		
Log new input varieties 2008 w.r.t. 2007	0.10			0.47***			0.45***		
	(0.17)			(0.17)			(0.11)		
Import growth of inputs from 2007 to 2008	0.44***			0.15			0.0040		
	(0.17)			(0.19)			(0.11)		
Log new input varieties 2009 w.r.t 2007		0.32**			0.61***			0.46***	
		(0.15)			(0.20)			(0.12)	
Import growth of inputs from 2007 to 2009		0.061			-0.19			0.13*	
		(0.11)			(0.14)			(0.079)	
Log new input varieties 2010 w.r.t 2009			0.11			0.50**			0.49***
			(0.21)			(0.25)			(0.15)
Import growth of inputs from 2009 to 2010			-0.069			-0.096			-0.29**
			(0.16)			(0.17)			(0.12)
Observations	1837	1837	1837	1830	1830	1830	1485	1485	1485

The table reports OLS regressions of innovation between 2009-2012 on log new input varieties in different years and the growth of import value of inputs between different years. Import growth of inputs from 2007 to 2008 computes the growth of the value of imports of the inputs of the industry between 2007 and 2008 (Table 5 defined growth as between 2008 and 2009). Log new input varieties 2008 w.r.t 2007 is the log number of input varieties that were not exported in 2007 and are in 2008 (Table 5 defined new input varieties as those not imported in 2008 and imported in 2009). All other year combinations are defined accordingly. All regressions include country dummies, 4-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table E.2: Robustness - Subsample: product innovation between 2009-2012

	Innovation		New input innovation		Input-essential innovation	
	(1)	(2)	(3)	(4)	(5)	(6)
Log new input varieties	0.33*	0.37*	0.60***	0.51***	0.57***	0.49***
	(0.19)	(0.22)	(0.16)	(0.19)	(0.14)	(0.15)
Import growth of inputs	-0.018	-0.097	-0.14	-0.22**	0.041	0.030
	(0.11)	(0.12)	(0.10)	(0.11)	(0.080)	(0.092)
Log new output varieties		-0.020		0.037		0.038
		(0.047)		(0.034)		(0.030)
Import growth of output		0.033		0.032		0.0056
		(0.025)		(0.024)		(0.020)
Small	-0.014	-0.014	-0.029	-0.027	-0.038*	-0.042**
	(0.026)	(0.026)	(0.025)	(0.025)	(0.021)	(0.020)
Large	0.034	0.029	0.044	0.048	-0.020	-0.023
	(0.039)	(0.040)	(0.043)	(0.045)	(0.033)	(0.033)
Foreign owned	0.00014	0.0020	0.00011	0.00059	-0.011	-0.012
	(0.039)	(0.039)	(0.036)	(0.036)	(0.033)	(0.033)
Age	-0.0015	-0.0013	-0.0013*	-0.0012	-0.00042	-0.00031
	(0.00094)	(0.00095)	(0.00076)	(0.00076)	(0.00066)	(0.00067)
Government owned	-0.082	-0.079	0.075	0.091	0.13	0.14
	(0.15)	(0.16)	(0.15)	(0.15)	(0.14)	(0.14)
Observations	1485	1461	1485	1461	1485	1461

The table replicates the OLS estimates of Table 5 in Section 4, restricting all regressions to the subsample of firms for which the variable input-essential innovation is not missing. All regressions include country dummies, 4-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Table E.3: IV estimation: extensions

	Instrument for new input varieties			Instrument for new input varieties & import growth of inputs		
	Innovation	New input innovation	Input-essential innovation	Innovation	New input innovation	Input-essential innovation
Panel A: Second stage						
Log new input varieties	0.43	0.71	0.65^{***}	1.11^{**}	0.77	0.44^{*}
	(0.47)	(0.50)	(0.20)	(0.54)	(0.50)	(0.26)
Import growth of inputs	-0.13	-0.24	0.017	-0.68^{**}	-0.43	0.19
	(0.12)	(0.17)	(0.083)	(0.30)	(0.28)	(0.15)
Panel B: First stage						
Dependent variable: Log new input varieties						
Log new input similar country	0.43^{***}	0.43^{***}	0.47^{***}	0.41^{***}	0.41^{***}	0.40^{***}
	(0.080)	(0.080)	(0.078)	(0.066)	(0.065)	(0.067)
$Inst_{jc}$	0.27^{**}	0.27^{**}	0.22^{**}	0.17^{***}	0.17^{***}	0.17^{***}
	(0.11)	(0.11)	(0.099)	(0.061)	(0.060)	(0.063)
Customs delay	-0.021^{**}	-0.020^{**}	-0.018^{**}			
	(0.0084)	(0.0083)	(0.0081)			
Dependent variable: Import growth of inputs						
Log new input similar country				-0.0067	-0.0061	0.0095
				(0.11)	(0.11)	(0.11)
$Inst_{jc}$				1.15^{***}	1.15^{***}	1.11^{***}
				(0.22)	(0.22)	(0.23)
N	1418	1412	1136	1787	1780	1442
F-stat	21.5	21.6	23.2	14.9	15.1	12.0
p-value Hansen test	0.17	0.36	0.19			

The table reports IV regressions of innovation (innovation, new input innovation or input-essential innovation) between 2009-2012 on log new input varieties in 2009. In columns 1-3, the instrument is log new input varieties in the same industry in a similar country (see Section 4.4 for the similar countries) measured at the two-digit industry-country level. In columns 4-5 customs delay is measured at the 4-digit industry-country level, but the regression is run on a sub-sample of non-importing firms. All regressions include country dummies, 4-digit industry dummies, three size dummies, dummies for government and foreign ownership and age. Robust standard errors (clustered by 4-digit-industry-country) are reported in parentheses. Significance: *10%, **5%, ***1%.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2018/3082/9