BANCO CENTRAL DE COSTA RICA

DOCUMENTO DE INVESTIGACIÓN

N. 002 | 2018

Costa Rican Production Network: Stylized Facts

Alonso Alfaro Ureña
Mariany Fuentes Fuentes
Isabela Manelici
José P. Vásquez

December, 2018

La serie Documentos de Investigación del Banco Central de Costa Rica en versión PDF se encuentra en http://www.bccr.fi.cr

Fotografía de portada: "Presentes", conjunto escultórico en bronce del artista costarricense Fernando Calvo Sánchez, año 1983. Colección del Banco Central de Costa Rica.

The series of Research Documents of the Central Bank of Costa Rica can be found in PDF format at http://www.bccr.fi.cr

Photography on the title page: "Presentes" (*Present*), a set of sculptures from the Costa Rican artist Fernando Calvo Sánchez, 1983. Collection of the Central Bank of Costa Rica.

Costa Rican Production Network: Stylized Facts*

* Alonso Alfaro Ureña† Mariany Fuentes Fuentes‡ Isabela Manelici§ José P. Vásquez¶

Abstract

This paper characterizes the Costa Rican production network through a set of stylized facts. We use a firm-to-firm transactions dataset, in combination with other administrative data sources to describe for the first time the domestic production network of a developing country. In particular, we show that (i) The distributions of the number of connections resembles a Pareto distribution, (ii) Firm size and productivity are positively correlated to the number of connections to other firms, (iii) There is negative degree of assortativity between sellers and buyers, (iv) Top clients account for a high percentage of a firm's sales, and (v) Firm's distributions of sales and purchases do not vary much with the number of connections. In general, the Costa Rican production network shares many features in common to the ones recently studied in Japan and Belgium. This implies that our stylized facts reflect intrinsic characteristics of firms' behavior that appear to be constant across different levels of countries' development.

Key words: Production networks, productivity.

Resumen

Este documento de investigación realiza una caracterización de la red de producción costarricense mediante una serie de hechos estilizados. El documento usa una base de datos de transacciones entre empresas, en combinación con otras bases de datos administrativosm para describir por primera vez la red de producción doméstica de un país en desarrollo. En particular, se muestra que: (i) El número de conexiones entre empresas es similar a una distribución Pareto, (ii) El tamaño y la productividad de las empresas está positivamente correlacionada con el número de conexiones con otras empresas, (iii) Existe asortatividad negativa entre vendedores y compradores, (iv) Los clientes más grandes representan un porcentaje alto en las ventas de las empresas y (v) La distrubuciones de compras y ventas no varían con respecto al número de conexiones. En general, la red doméstica de producción comparte muchas características en común con las de Japón y Bélgica, que han sido estudiadas recientemente. Esto implica que nuestros hechos estilizados reflejan características propias del comportamiento de las empresas que parecen ser constantes a lo largo de diferentes niveles de desarrollo económico de los países.

Palabras clave: Redes de producción, productividad.

*The views expressed herein are those of the authors and do not necessarily represent the views of BCCR. All results have been reviewed by BCCR to ensure no condential information is disclosed.

†Departamento de Investigación Económica. División Económica, BCCR. alfaroua@bccr.fi.cr

‡Departamento de Investigación Económica. División Económica, BCCR. fuentesfm@bccr.fi.cr

§University of California, Berkeley. imanelici@berkeley.edu

¶University of California, Berkeley. jpvasquez@econ.berkeley.edu

Costa Rican Production Network: Stylized Facts

Contents

List of Tables

List of Figures

Costa Rican Production Network: Stylized Facts

1. Introduction

The mapping of production networks has opened new doors for economic research in recent years, posing many questions that have yet to be answered. This paper discusses general features of the data that have been just been studied recently, and characterize the behavior of the Costa Rican domestic production network with a set of stylized facts. The results documented in this paper provide valuable input for the development of models that will help to determine the origin (at the micro level) and quantify the magnitude (at the aggregate level) of various economic phenomena that occur within and outside of the production network of a developing country.

To our knowledge, this is the first paper that documents production networks in a developing country. Similar work has been done for developed countries such as Japan, Belgium, and Norway, in the works of Dhyne, Magerman and Rubínová (2015), Bernard, Moxnes and Saito (2018), Bernard, Dhyne, Magerman, Manova and Moxnes (2017), and Bernard, Moxnes and Ulltveit-Moe (2018). Research thus far has found interesting similarities between domestic (within-country) relations and international trade relations between buyers and sellers in those countries. This paper finds that the Costa Rican production network shares these similarities, which suggests that such characteristics may not be specific to developed countries, but rather that there may exist general patterns regarding the manner in which firms interact and are positioned in a production network.

A better understanding of the domestic production network is crucial for macroeconomic stability. Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012) state that when a network is highly asymmetrical, shocks that occur at the level of a firm or sector could have a much greater aggregate propagation. In addition, even though not all firms trade directly with each other, the structure of the production network could serve as an amplification mechanisms of economic shocks (Carvalho, 2014). Finally, in the presence of different production network structures, the largest firms might not necessarily be the most central in terms of contributing to aggregate volatility of the economy.

Not long ago economists spoke of economies with aggregate exchanges, where industries traded

among themselves, without much information on each firm's share in production or input purchases, and where all distortions occurring in the aggregate were due to exogenous (and sometimes inexplicable) phenomena. Most of the current literature has focused on trying to explain how micro-economic shocks can have significant effects on aggregate outcomes. Firm-level shocks might reach other firms connected to the first firm, and, through ripple effect, these shocks could add up to affect aggregate production.

Such a perspective is contrary to that of Lucas (1977), who argued that aggregate output volatility is due to exogenous macro-economic shocks, and that micro-economic shocks are so small and random that they would average out, and would have insignificant effects at the aggregate level. Apparently, Lucas did not consider an economy with production networks, where propagation effects can occur.

The studies of Carvalho, Nirei and Saito (2014), Magerman, De Bruyne, Dhyne and Van Hove (2016), and Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012), contain early empirical evidence regarding the topic of micro-economic shocks and their importance. Carvalho, Nirei and Saito (2014) studied the aggregate-level effects of the 2011 earthquake in Japan. The results showed a significant decrease in sales for those firms that had a direct or indirect relation (maximum third degree) with a firm affected by the earthquake. In addition, these firms were also more likely to create new buyer-seller connections.

Carvalho (2014) argued that aggregate product volatility can be predicted by knowing the growth rate of the total factor productivity for the largest and most connected firms, as there exists a high correlation between aggregate product growth and the production growth of the ten most connected sectors. Magerman, De Bruyne, Dhyne and Van Hove (2016) corroborated this by analyzing the economic relations among firms in Belgium, where they found that 58% of GNP growth volatility could be explained by shocks that affected firms, but that the 100 firms that contributed most to aggregate product influenced 90% of that volatility.

Additionally, Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012), using data on transactions among U.S. sectors between 1972 and 2000, detected a great deal of heterogeneity and asymmetric, heavy-tailed supplier sector sales distributions (in other words, few large sectors selling inputs to a large number of other sectors). When one of these large sectors suffers an adverse shock, the lower productivity propagates to all of its direct clients and, because there is a sequence of sectors inter-connected with its clients, the original shock creates a cascade effect that decreases production for all sectors involved.

The recent consensus from the literature is that microeconomic shocks can significantly affect aggregate outcomes, if considerable asymmetries exist in the production network. On the other hand, in a symmetrical network where the relations between suppliers and buyers are homogeneous, Lucas' argument would hold true (Carvalho, 2014). One question still remains: What determines the structure of the production network, and how does the production network structure evolve over time?

This paper makes an initial inquiry into this topic by showing stylized facts that characterize the Costa Rican production network. Our analysis uses information from the Ministry of Finance, Social Security and customs. Our final database contains records of firm-to-firm transactions for the universe of Costa Rican formal firms, their number of employees, costs, total income, exports, imports, economic sector, and domicile province, among others.

We find a positive correlation between the number of connections and the size (productivity) of a firm: the firms with the highest sales (productivity) also have more connections. In general, firms have higher sales (purchases) because they have more clients (suppliers), not because they sell more to each client. This striking pattern seems not to be unique to the Costa Rican context and was also found in Bernard, Dhyne, Magerman, Manova and Moxnes (2017) for the Belgian production network.

We also find a negative degree of assortativity among firms, meaning that small firms buy from (sell to) few suppliers (clients), but these suppliers (clients) are considerably large, with their median number of clients (suppliers) being high. This results in an one-way dependence: the small firms depend on the large firms. This pattern is also confirmed for other countries. For example, Bernard, Moxnes and Ulltveit-Moe (2018) also found a negative degree of assortativity with data on exports and imports of Norwegian firms.

The remainder of this paper is structured as follows: Section 2 contains a description of the data; Section 3 presents a brief description of the Costa Rican Production Network, Section 3.1 a set of stylized facts on the role of firms in trade and the heterogeneity of buyers and sellers; and Section 4 sets forth the conclusions.

2. Data

This paper is based on information gathered by the Ministry of Finance of Costa Rica (Ministerio de Hacienda). The first data set used comes from Form D-101 ("Declaración Jurada del Impuesto sobre la Renta" or the "Affidavit of Income Tax"). This form reports earnings, income, costs, and assets, among other items, that all active firms must declare annually. It also contains

variables such as the number of employees and the yearly firm-level total wage bill reported to the Costa Rican Social Security Fund ("Caja Costarricense del Seguro Social"). The database covers the period between 2005 and 2015.

Although the Form D-101 indicates the Corporate ID Number of active firms, it does not indicate whether the owner(s) of a given firm hold shares in other firms. For this reason, it was necessary to identify this shared ownership, in order to group the related firms together as an unique firm or as part of the same corporate group. The National Company Registry kept by the Central Bank was used for this purpose. A total of 3,947 corporate groups and 1,881 firm groups were identified, with the latter being treated as unique firms.

The second data set comes from Form D-151, the "Declaración annual resumen de clientes, proveedores y gastos específicos" (Declaration of the yearly summary of clients, suppliers and specific expenses), with data from 2008 to 2015. With this form each firm must report annually on all transactions with their clients or suppliers, in order to allow cross-checking between buyers and sellers to ensure that firms do not declare less income or higher costs in order to pay less corporate income tax. All economic entities (whether private or public) that buy or sell goods or services are required to file this form, even if they pay no income tax, otherwise they are subject to fines for non-compliance.

Since all the transactions made by firms are recorded, this allows each firm's relations with other firms to be tracked, as well as the transactions' magnitude and frequency. In other words, the Form D-151 is a double-entry record. As an example, consider an economy where only Firm A and Firm B exist. Firm A is a client of Firm B. When Firm A records a transaction as a purchase on Form D-151, Firm B must record the same transaction as a sale, since Firm B is a supplier of Firm A. At the end of the year, when all the transactions made by Firm A and Firm B are added up separately, Firm A's total purchases will be equal to Firm B's sales. If the figures are not equal, then it is likely that one of the firms is not recording all of its transactions. This is very important, as it allows the Ministry of Finance to perform cross-checks between existing transactions and ensure corporate compliance regarding the payment of the corporate income tax.

The names assigned to the variables in the original files obtained from the Ministry of Finance changed from one year to the next, a problem that had to be solved and standardized. In order to process and unify all the D-151 data, as to obtain a data panel for all firms in the country between 2008 and 2015 and their respective transactions, a multi-phase process was applied to obtain a properly filtered database.

The first phase consisted of identifying the definitive Corporate ID Number (CIDN) of all the corporate entities that filed returns. Some numbers contained errors, as the Ministry of Finance occasionally assigns additional digits to a firm's CIDN. This resulted in cases in which a firm with a given CIDN in Year X showed up with a different CIDN in Year Y, which caused it to be treated as two separate firms. Another problem had to do with foreign CIDNs, because a firm can appear with a foreign CIDN for Year X and then appear with a Costa Rican CIDN in subsequent years. The Central Bank of Costa Rica (BCCR) tracks such changes in order to avoid inconsistencies.

The second phase consisted of defining preliminary error cases. Here the amounts were totalled to determine whether total sales equalled total purchases for a given period. In doing this, it was noted that there existed a gap between sales and purchases (with sales being higher than purchases). Once this was discovered, a search was carried out to determine the origin of said gap and find solutions to correct this.

Errors such as the following were detected: 1) the seller's CIDN was the same as the buyer's CIDN, 2) the CIDN was the same as the transaction amount, 3) duplicated transactions, 4) transaction entry pairs with differing amounts because one of the firms recorded the operation as a multiple of ten, and 5) one firm recorded an operation as zero while the other firm did otherwise.

The third phase consisted of rectifying the errors identified in the second phase. To do this, income and sales data were obtained from Forms D-101 and D-104 in order to compare them with the transaction amount reported on the Form D-151. Inconsistencies were identified in this manner. In addition, transaction histories between firms were obtained in order to identify inconsistencies in the amounts thereof.

The errors found were rectified as follows. For cases 1) and 2): Firm A declared the amount of a transaction with Firm B as the latter's CIDN, but Firm B declared the correct transaction amount. For case 3), all transactions identified as duplicated were eliminated. In such a case, Firm A's transaction amount (incorrect figure) is substituted for the amount recorded by Firm B (correct figure). For case 4): historical data was used to identify the correct value. For case 5): as in the previous cases, based on an individual analysis it was determined that the correct transaction amount was the higher figure.

In addition, firms are not required to declare individual purchases/sales made from/to the general public, and so about one-fifth of the income reported by firms has no matching entries, but can not be classified as erroneous figures.

The fourth and final phase consisted of defining what were the final cases, to be kept for analysis. This final dataset incorporates the corrections just described. The two tables below (Table 1 and 2) show the number of transactions and the total value of the transactions analyzed. The data deemed fit for use were classified as: 1) Data in pairs, and 2) No partner. The data deemed unfit for use were classified as: 1) Unsolved, 2) No partner and rejected by the Ministry of Finance based on a review of the income tax returns, 3) Duplicate data already considered, and 4) Data reported erroneously.

For the purposes of this paper, we can use at least 80% of the transactions and cases (over 90% for some years). Only 12.3% of the transactions could not be used, with those categorized as unsolved representing less than 10% of the total for most periods. It is important to mention that the filtering process gave priority to high-amount transactions.

Table 1: Number of cases, B2B micro-data

	2008		2012		2015	
Type of Case	Count	%	Count	%	Count	%
Data in pairs	535,863	41.9	998,355	40.5	1,383,820	42.2
No partner	493,769	38.7	1,256,978	51	1,626,907	49.6
Subtotal of data used	1,029,632	80.6	2,255,333	91.5	3,010,727	91.9
Unsolved	128,599	10.1	202,710	8.2	251,499	7.7
No partner and rejected	108,969	8.5	-	0	-	0
Duplicated	4,904	0.4	5,936	0.2	14,652	0.4
Excluded	5,414	0.4	34	0	32	0
Total	1,277,518	100	2,464,013	100	3,276,910	100

Source: Own elaboration.

Table 2: Value of transactions (millions of dollars), B2B micro-data

	2008		2012		2015	
Type of Case	Value	%	Value	%	Value	%
Data in pairs	45,812	63.6	55,489	67.5	69,450	69.1
No partner	11,808	16.4	16,637	20.2	18,496	18.4
Subtotal of data used	57,620	80	72,126	87.7	87,946	87.6
Unsolved	7,766	10.8	10,002	12.2	12,324	12.3
No partner and rejected	6,145	8.5	-	0	-	0
Duplicated	170	0.5	71	0.1	172	0.2
Excluded	359	0.5	1	0	2	0
Total	72,060	100	82,200	100	100,444	100

Source: Own elaboration.

Upon the completion of the four filtering phases, the result is a database with data from both

Forms D-101 and Form D-151, along with CCSS data for 2008-2015. Since the goal was a database comprised of all domestic non-financial private firms, observations regarding non-governmental organizations (NGOs), public entities (including utility companies), and financial entities were excluded, along with observations regarding household consumption. A total of 75,707 firms were identified.

A minimum size restriction was also imposed, to eliminate firms with annual income below US$50,000 (based on 2013 PPI), leaving 60,479 firms and 1,995,970 observations. These restrictions do not affect the coverage of the variables, as the remaining firms represent a majority of the workforce, sales, exports, income, costs, and assets.

3. The Costa Rican Production Network

3.1. Descriptive Statistics

This section starts to characterize the Costa Rican domestic production network. Previously it was mentioned that the filtered database contained a total of 1,995,970 supplier-buyer connections and 60,478 firms for the 2008-2015 period. 59.9% are sellers and buyers at the same time, 9.7% are only sellers, and 30.4% are only buyers. Of these 60,478 firms, 42,323 sell to other firms in the network and 54,637 buy from other firms.

In other words, 9.66% do not buy input from domestic firms and 30.01% of the firms sell only to final demand (consumers) or sell less than US$4,800 to other firms. It should be noted that the Form D-151 only records transactions greater than US$4,800, meaning that those 30.01% of firms, now appearing to have had no sales to other firms, might have actually had them, just that said sales were not recorded due to their being below the reporting threshold.

Table 3 shows the percentage of total sales that represent the ten most important sectors in 2014. It can be noted that these ten sectors represent almost all sales, with Wholesale and Retail Trade being the largest sector among the total sales recorded for said year.

Note that with respect to the stylized facts, some of the figures will be estimated with controls by sector, province, and year. For this reason, the three sectors with highest sales volumes were selected (Wholesale and Retail Trade; Manufacturing; and Agriculture, Forestry, and Fishing), which together represented 68.71% of total sales.

Tables 4 to 7 provide summary statistics on the number of clients and suppliers each firm had each year. This data is also separated by sector for the year 2014. It should be noted that only the

Table 3: Percentage of total sales by sector, 2014

Industry	%
Wholesale and Retail Trade	38.39
Manufacturing	24.60
Administrative and Support Services	6.44
Agriculture, Forestry, and Fishing	5.72
Construction	5.34
Transportation and Storage	4.86
Professional, Scientific, and Technical	3.50
Lodging and Food Services	3.35
Information and Communication	3.15
Education	1.14
Total	96.48

Source: Own elaboration.

Table 4: Number of buyers per firm by year

Year	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
2008	19,530	18.22	111.58	1	2	4	10	30	60	229
2009	20,395	17.14	88.56	1	2	4	10	29	58	229
2010	21,721	16.88	89.36	1	2	4	10	28	58	227
2011	22,869	17.88	106.64	1	2	4	10	29	59	239
2012	23,901	18.27	117.76	1	2	4	10	29	60	256
2013	24,418	18.76	120.96	1	2	4	10	30	60	260
2014	24,640	19.81	125.58	1	2	4	11	32	63	266
2015	24,741	20.99	140.59	1	2	4	11	33	67	290
All	182215	18.57	114.84	1	2	4	10	30	61	250

Source: Own elaboration.

ten aforementioned sectors are used, as they are the most representative.

Table 5: Number of buyers per firm by sector, 2014

Industry	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
Wholesale and Retail Trade	7,354	36.4	202.6	1	3	7	24	67	139	426
Manufacturing	2,880	28.1	128.5	1	3	6	15	45	98.5	438
Administrative and Support Service	1,603	15.3	68.4	1	2	4	10	28	50	150
Agriculture, forestry and fishing	2,590	4.4	9.0	1	1	2	4	8	13	33
Construction	2,269	6.2	12.0	1	2	3	6	12	19	45
Transportation and Storage	2,211	10.1	31.6	1	2	3	8	20	36	122
Professional, Scientific and Technical	2,124	10.4	65.7	1	2	4	9	18	28	87
Accommodation and Food Services	975	7.3	18.3	1	1	2	6	15	28	71
Information and Communication	612	18.5	79.8	1	2	6	12	28	52	256
Education	180	48.9	151.4	1	1	3	15.5	111	335	809
Total	22,798	20.6	129.3	1	2	4	11	33	67	282

Source: Own elaboration.

Table 6: Number of suppliers per firm by year

Year	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
2008	26,246	11.40	32.18	1	2	4	10	24	40	112
2009	26,386	11.80	31.69	1	2	5	11	24	41	113
2010	27,736	11.79	30.88	1	2	5	11	24	41	111
2011	29,711	12.17	37.42	1	2	5	11	25	42	114
2012	30,761	12.67	39.53	1	2	5	11	26	43	117
2013	31,475	12.86	36.12	1	3	5	12	27	44	117
2014	31,541	13.53	36.61	1	3	6	13	28	46	127
2015	31,841	14.05	39.30	2	3	6	13	29	48	132
All	235,697	12.59	35.82	1	2	5	11	26	43	118

Source: Own elaboration.

According to these tables, among seller firms, the average number of clients was 18.57 for the entire period and 20.6 for 2014, with standard deviations of 114.84 and 129.3, respectively. The average number of suppliers per buyer firm for 2008-2015 was 12.59, and 14.05 for 2014, with standard deviation of 35.82 and 37.89, respectively.

This means that, on average, a firm has more clients than suppliers. In addition, based on the standard deviations and percentile distributions, the distribution of clients per seller is much more dispersed than the distribution of suppliers per buyer.

The distributions of clients and suppliers are right skewed. Note that in 2014 the median number of clients was 4, while the 99^{th} percentile firm had 282 clients. The median number of suppliers was 6, while the 99^{th} percentile firm had 133 suppliers. There is a high concentration of firms with less than 6 connections, which means that there is a high asymmetry in the distribution of

Table 7: Number of suppliers per firm by sector, 2014

Industry	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
Wholesale and Retail Trade	11,052	14.5	29.8	2	3	7	15	32	49	117
Manufacturing	3,527	23.3	81.0	2	3	7	16	42	86	292
Administrative and Support Service	1,358	16.0	29.6	1	3	6	15	39	66	150
Agriculture, forestry and fishing	2,496	10.3	25.1	1	2	4	9	22	38	103
Construction	2,610	12.6	26.2	1	2	5	12	29	47	116
Transportation and Storage	2,049	10.7	19.2	1	2	4	11	26	40	92
Professional, Scientific and Technical	1,921	8.7	13.5	1	3	5	9	18	29	77
Accommodation and Food Services	2,774	10.4	21.1	2	3	5	10	20	32	96
Information and Communication	618	13.6	37.4	1	3	6	12	25	44	120
Education	429	15.1	25.1	2	3	7	16	37	57	97
Total	28,834	14.05	37.89	1	3	6	13	29	48	133

Source: Own elaboration.

connections.

The distribution by sector is even more asymmetric. For seller firms, the "Manufacturing" and "Education" sectors have the highest extreme values, with the 99^{th} percentile having 438 and 809 clients, respectively, and 50% of the data being concentrated in firms with less than 6 and 3 clients.

In the case of buyer firms, the "Manufacturing" and "Administrative and Support Services" sectors present the starkest asymmetries, with the top percentile having 292 and 150 suppliers, respectively. Once again, the same pattern of concentrated values is found, with 50% of the firms having less than 7 suppliers.

The same relationships were found in Belgium (also for 2014) by Bernard, Dhyne, Magerman, Manova and Moxnes (2017), with the distribution of buyers per seller being more dispersed than that of suppliers per buyer. The majority of firms had less than nine connections, with a strong heterogeneity among sectors.

Figure 1 displays this asymmetry in the out-degree of firms. It shows the distribution of customers per firm for the three largest sectors. It can be noted that, in fact, more than 50% of the firms in those three sectors have less than ten customers.

Table 8 and 9 report summary statistics on the total income of firms in Costa Rica per year (and by sector for 2014). These Tables also showcase a strong dispersion in sales across firms.

The average sales volume for all active firms during 2008-2015 was US$1.64 million (real 2013 USD), with a standard deviation of 10.5 millions. For 2014, firms in the 90^{th} percentile generated 60% of the total transactions.

Figure 1: Firm sales distribution by sector

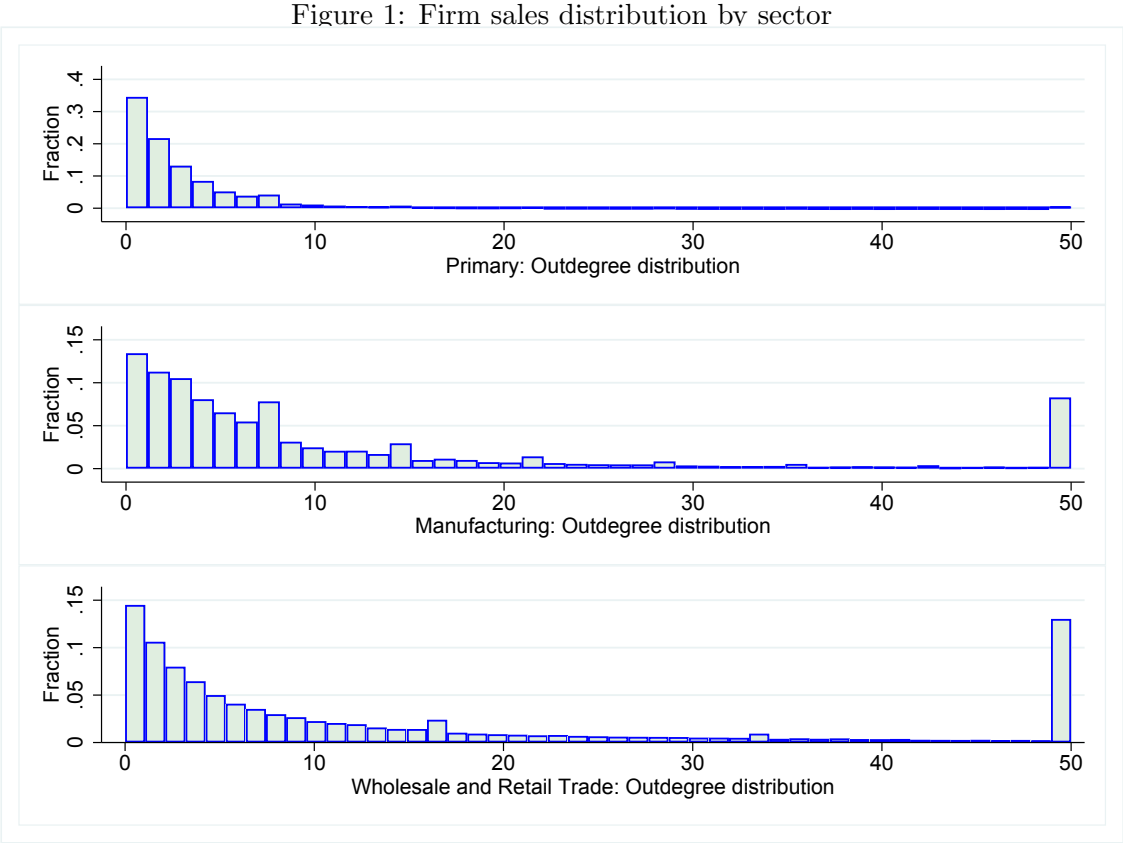


Table 8: Firm sales by year
(Million dollars, PPI-deflacted to 2013 US$)

Year	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
2008	30,153	1.701066	11.20	0.072	0.115	0.24	0.66	2.09	4.86	25.70
2009	30,708	1.541638	9.94	0.070	0.110	0.23	0.61	1.93	4.44	22.60
2010	32,583	1.614353	10.60	0.072	0.113	0.23	0.62	2.02	4.56	24.00
2011	34,690	1.635914	10.70	0.073	0.114	0.24	0.63	2.02	4.71	24.20
2012	36,009	1.638618	10.70	0.073	0.115	0.24	0.64	2.01	4.70	24.20
2013	36,702	1.647876	10.20	0.074	0.117	0.24	0.64	2.06	4.80	24.90
2014	36,572	1.648487	10.10	0.073	0.115	0.24	0.64	2.06	4.83	25.00
2015	36,869	1.713172	10.70	0.074	0.121	0.25	0.66	2.16	4.96	26.50
All	274,286	1.643977	10.50	0.073	0.115	0.24	0.64	2.05	4.74	24.70

Table 9: Firm sales by sector, 2014
(Million dollars, PPI-deflacted to 2013 US$)

Industry	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
Wholesale and Retail Trade	11849	1.953	11.00	0.084	0.15	0.33	0.90	3.00	6.41	25.80
Manufacturing	3820	3.882	19.40	0.078	0.13	0.30	0.87	4.12	13.60	81.90
Administrative and Support Service	2003	1.937	9.97	0.074	0.12	0.28	0.84	2.58	6.30	34.80
Agriculture, forestry and fishing	2928	1.177	8.68	0.068	0.10	0.20	0.45	1.35	2.99	16.20
Construction	3150	1.023	4.64	0.069	0.10	0.21	0.55	1.64	3.28	14.60
Transportation and Storage	2657	1.102	4.82	0.071	0.10	0.20	0.55	1.83	3.86	17.20
Professional, Scientific and Technical	2687	0.784	5.61	0.066	0.09	0.18	0.40	1.02	1.83	10.20
Accommodation and Food Services	2917	0.693	3.94	0.074	0.11	0.19	0.39	0.87	1.62	8.25
Information and Communication	811	2.344	11.70	0.078	0.13	0.33	1.03	3.06	6.99	36.20
Education	473	1.454	3.83	0.103	0.20	0.45	1.22	3.85	5.75	14.60
Total	33295	1.747	10.50	0.075	0.12	0.25	0.68	2.22	5.15	26.80

Source: Own elaboration.

The distributions of sales across firms appear to differ strongly by sector. The largest number of firms (11,849) was concentrated in the Wholesale and Retail Trade sector, with the smallest number (811) in the Information and Communication sector. On average, however, sales of firms in the Information and Communication sector were 20% higher than those of firms in the Wholesale and Retail Trade sector.

Table 10 shows the distribution of transactions between clients-suppliers for 2014. The average value of the 265,451 transactions recorded that year was US$67,000, with a standard deviation of US$621,000. The median sales transaction was for US$19,000, with the 90^{th} percentile representing 56% of the total number of transactions for 2014. This transaction value asymmetry was also found by Bernard, Dhyne, Magerman, Manova and Moxnes (2017) in Belgium, with an even higher magnitude than in Costa Rica.

As in the previous cases, sectors differ strongly among each other. "Wholesale and Retail Trade" accounts for the highest number of transactions, whereas "Agriculture, Forestry, and Fishing" shows the highest average transaction value (US$97,000) and a total of only 16,950 transactions.

Finally, Table 11 shows the number of employees per sector for 2014. The mean number of employees per firm among all active firms that year was 18.81, with a standard deviation of 114.15. Note once again the level of heterogeneity among sectors. The difference between the highest mean and the lowest mean is considerable (43.98 employees). The "Administrative and Support Services" sector had the highest mean (53.69 employees), while the "Professional, Scientific, and Technical" sector had the lowest (only 9.70).

Table 10: Firm-to-firm transaction values, 2014
(Million dollars, PPI-deflacted to 2013 US$)

Industry	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
Wholesale and Retail Trade	113,850	0.068	0.63	0.005	0.01	0.01	0.03	0.09	0.18	0.86
Manufacturing	50,577	0.081	0.70	0.006	0.01	0.01	0.04	0.13	0.26	1.05
Administrative and Support Service	13,424	0.057	0.79	0.005	0.01	0.01	0.03	0.07	0.15	0.63
Agriculture, forestry and fishing	16,950	0.097	0.73	0.006	0.01	0.02	0.04	0.14	0.28	1.41
Construction	22,776	0.054	0.24	0.005	0.01	0.01	0.03	0.10	0.19	0.68
Transportation and Storage	13,224	0.062	0.35	0.005	0.01	0.01	0.03	0.09	0.20	0.86
Professional, Scientific and Technical	9,582	0.039	0.93	0.005	0.01	0.01	0.02	0.06	0.10	0.31
Accommodation and Food Services	17,397	0.037	0.34	0.005	0.01	0.01	0.02	0.05	0.10	0.41
Information and Communication	4,238	0.071	0.54	0.005	0.01	0.01	0.03	0.09	0.20	0.76
Education	3,433	0.030	0.12	0.005	0.01	0.01	0.02	0.05	0.09	0.27
Total	265,451	0.067	0.62	0.005	0.00	0.01	0.01	0.01	0.03	0.09

Source: Own elaboration.

Table 11: Number of firm workers per industry, 2014

Industry	N	Mean	St Dev	10th	25th	50th	75th	90th	95th	99th
Wholesale and Retail Trade	11,849	11.72	53.40	0	1	3	8	20	39	147
Manufacturing	3,820	35.60	154.46	1	2	5	15	53	130.5	679
Administrative and Support Service	2,003	53.70	230.87	1	2	7	25	93	206	979
Agriculture, forestry and fishing	2,928	23.47	182.50	0	1	3	8	27	62	399
Construction	3,150	13.98	51.59	0	1	3	10	27	52	190
Transportation and Storage	2,657	12.75	43.84	0	1	3	7	25	54	202
Professional, Scientific and Technical	2,687	9.71	46.80	0	1	3	6	16	35	126
Accommodation and Food Services	2,917	15.57	72.73	1	2	4	10	25	43	187
Information and Communication	811	26.92	98.78	1	2	6	16	50	88	471
Education	473	51.13	168.70	2	5	15	42	114	176	582
Total	36,572	18.81	114.15	0	1	3	9	28	58	286

Source: Own elaboration.

3.2. Stylized Facts

This section presents stylized facts that represent the first characterization of the Costa Rican productive network and, more broadly, of the production network of a developing country. These facts help the reader understand how firms are distributed in the network, their patterns of connections, productivity, and asymmetry, thus establishing a basis for future theoretical and empirical research efforts.

Most of these stylized facts have already been illustrated by Bernard, Dhyne, Magerman, Manova and Moxnes (2017) and Bernard, Moxnes and Saito (2018) for the cases of Belgium and Japan, respectively. Our results show that these stylized facts represent an intrinsic characteristic of firms' interactions and are not specific to developed economies.

We have relied on the aforementioned database of 60,478 active firms during the 2008-2015 period. We document two types of stylized facts: 1) those that calculate statistics by firm, and, then,

2) those that use aggregates that group together firms with the same characteristic. For the first group, fixed effects of location, year, and sector were used to prevent that any of these variables drive the results. We decided to use only the three sectors with highest sales volumes ("Wholesale and Retail Trade," "Manufacturing," and "Agriculture, Forestry, and Fishing"). Together, these three sectors represented 68.71% of total sales in 2014 (see Table 3).

These fixed effects could not be used for the second type of stylized facts. Aggregate graphs were calculated for all sectors. Finally, starting with Figure 10, we report statistics for firms that are either exporters or importers.

Fact 1. The distributions of the number of customers and suppliers resembles a Pareto distribution.

As can be noted in Figure 2, the number of clients and suppliers have a heavy-tailed distribution. Most firms have a low number of connections, while a minority are very well connected. Only 1% of firms have more than 150 connections (either buyers or sellers) and only one tenth of 1% have more than 1,000 connections.

Around 50% of firms have less than five connections. The parameters estimated for the Pareto distributions of per-firm suppliers and per-firm customers are -0.58, and -0.73, respectively. The estimated slope of the CDF is slightly stepper for Japan, with parameters of -1.50 for customers and -1.32 for suppliers (Bernard, Moxnes and Saito, 2018). However, the qualitative findings are very similar.

Fact 2. Firm size is positively correlated to the number of connections to other firms.

From Figure 3, we learn that there is a positive correlation between a firm's sales and the number of customers or suppliers that this firm has, as in Bernard, Moxnes and Saito (2018). The green (solid) line plots the relationship between the number of suppliers of firms and their total sales. The slope is 0.89, meaning that a 10% increase in the number of suppliers is associated to a 8.9% increase in sales. The black (dashed) line plots the relationship between the number of customers of firms and their sales. The slope is 1.2, meaning that a 10% increase in the number of suppliers is associated to a 12% increase in sales.

Figure 2: Number of suppliers and buyers (CDFs)

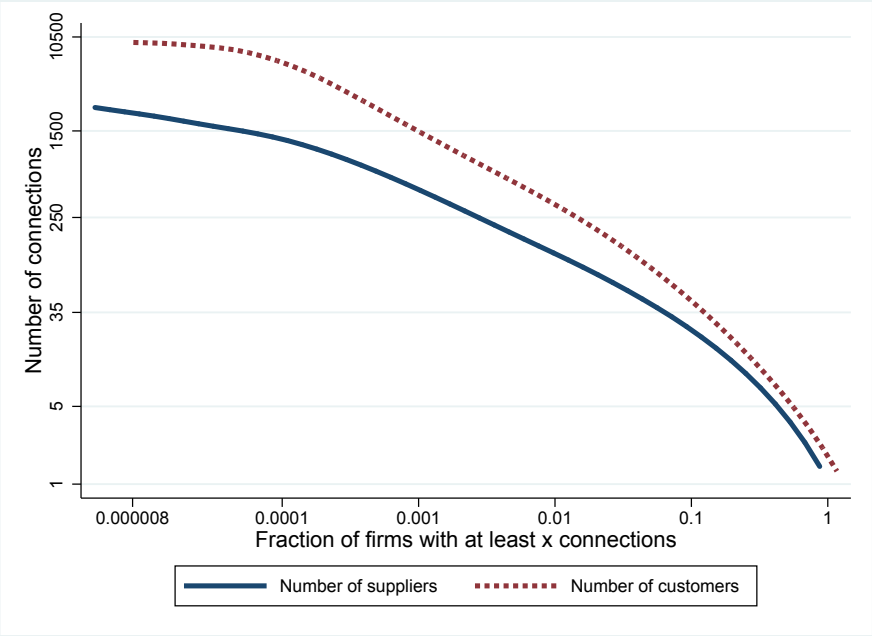
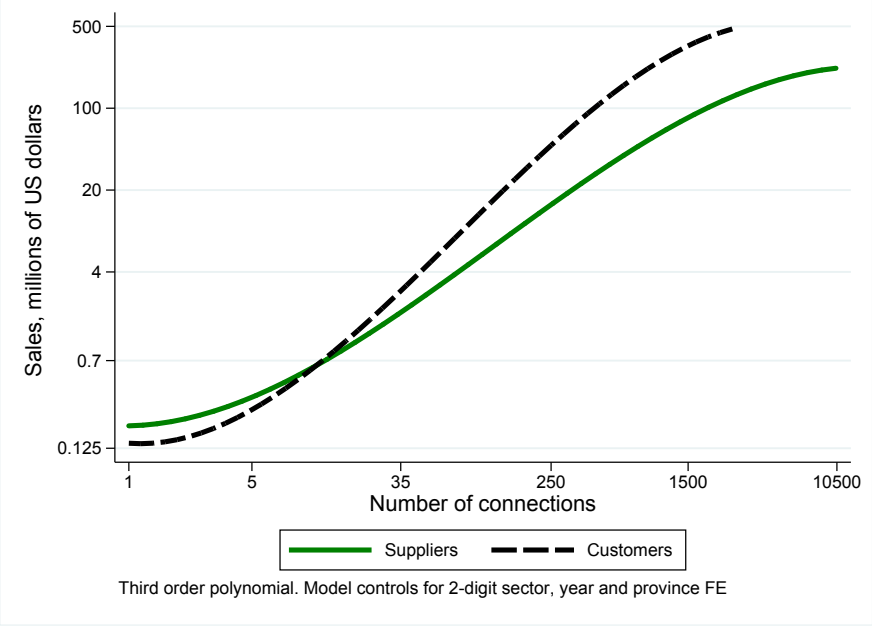


Figure 3: Size of sellers and buyers, related to firm number of connections



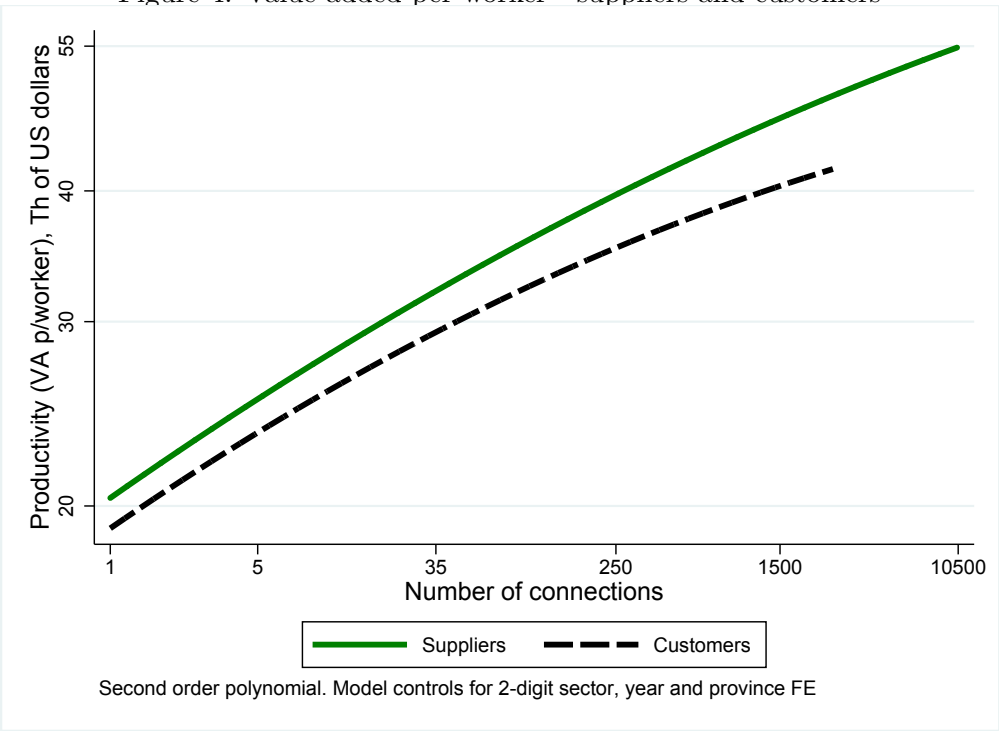
Notes: The figure shows the local polynomial regression of log total firm income (y axis) on log number of connections (x axis). The green (solid) line relates to supplying connections, the black (dashed) line relates to buying connections. We controlled for the economic sector, year, and province, along with a third-degree polynomial with interaction effects for the x variable.

Fact 3. Firm productivity is positively correlated with the number of firm connections.

Productivity is a measure of efficiency in the use of productive inputs. We study the relationship between a firm's productivity and its number of connections. We measure productivity as valued added per worker. According to Figure 4, the correlation is positive for both customer and supplier firms. The slope is 0.11 for seller firms and 0.10 for buyer firms. In other words, a 10% increase in connections is correlated with a 10% higher value-added per-worker.

Figures 3 and 4 are consistent with each other. A more productive firm is likely to also be a large firm, both in terms of its sales and in terms of its number of connections. Large firms have access to better technology, more productive workers, and better cost-minimization practices. The situation is different for small firms, due to their low income level and higher fixed costs that prevent them from competing with larger firms.

Figure 4: Value added per worker - suppliers and customers



Notes: The figure shows the local polynomial regression of log value added per worker (y axis) on the number of connections (x axis). The green (solid) line relates to supplying relationships, the black (dashed) line relates to buying relationships. We control for the economic sector, year, and province, along with a second-degree polynomial with interaction effects for the x variable. We measure productivity as value added (total income less total costs) per worker.

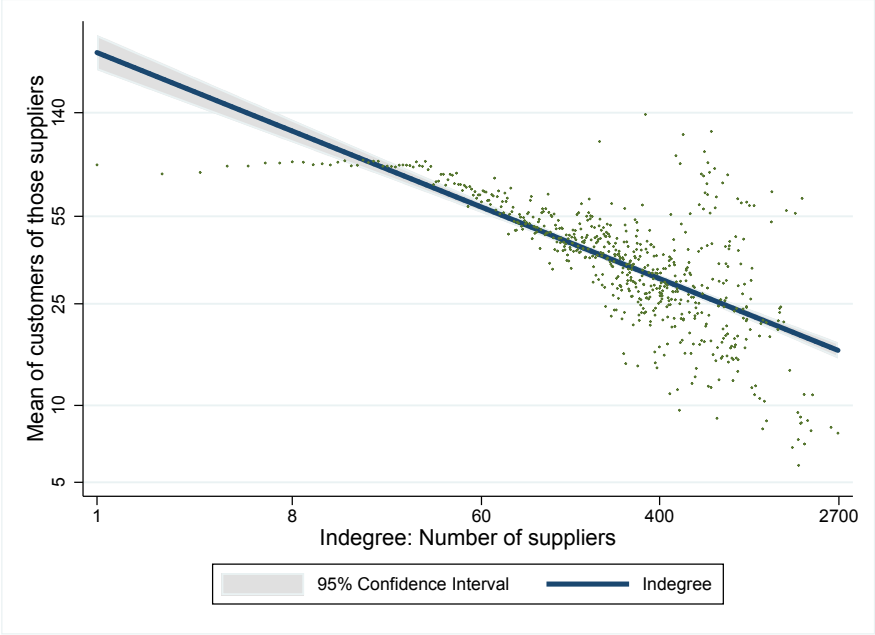
Fact 4. There is a pattern of negative degree of assortativity between sellers and buyers.

Positive assortativity can be defined as an entity's preference to connect with others that have similar characteristics. In our case, this would mean that a firm that is connected with a large number of companies would want the latter to also be connected to a large number of firms.

Figure 5 shows a negative degree of assortativity for buyer firms. The interpretation in this case would be that when a buyer firm is connected with a large number of firms, the latter would, on average, have a low number of connections. The slope is -0.34, meaning that a 10% increase in a firm's number of suppliers is associated to a 3.4% reduction in suppliers' average number of customers.

In order to provide a more precise explanation, let us focus on coordinates (400,50) of Figure 5. This ordered pair represents a firm with 400 suppliers, each of which sells to an average of 50 customers. As the number of suppliers decreases, the average number of customers of these suppliers increases. This means that a one-way dependence exists, with small suppliers depending on large clients.

Figure 5: Degree assortativity - Suppliers and customers of suppliers

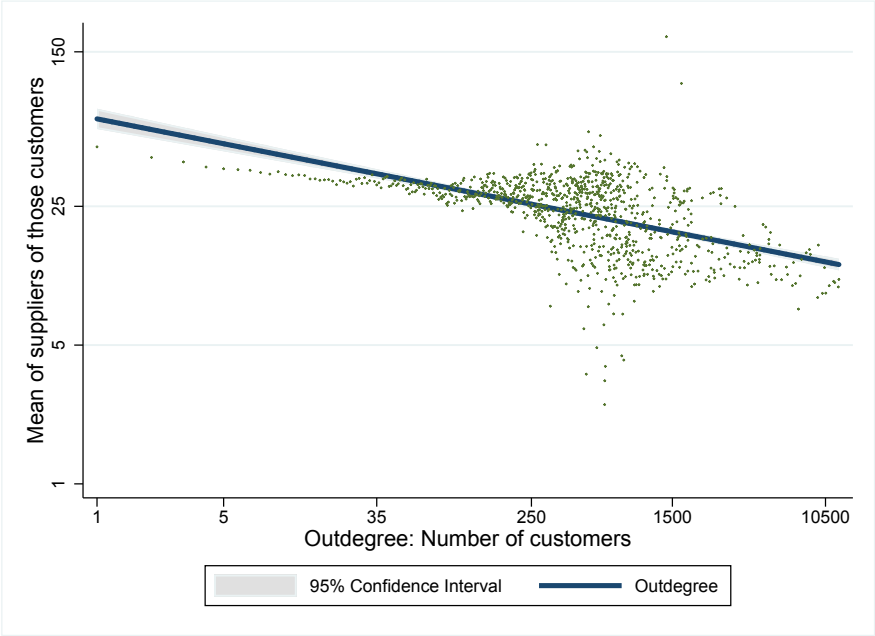


Notes: The figure shows the local polynomial regression of the supplier number log (x axis) and the log for the average number of customers for said suppliers (y axis). To calculate the y axis, the firms with the same number of suppliers are grouped together, and then their average number of customers is calculated.

Figure 6 shows the degree of assortativity for seller firms, with a slope of -0.18 [1] As in the preceding case, this assortativity is negative. When firms have many customers, the latter buy from a low number of suppliers. Now the dependence goes the other way, with these customers depending on few suppliers.

Considering that 50% of the firms have less than seven connections (Tables 4 to 7), this means that the Costa Rican production network could be sufficiently asymmetric, with a few large companies buying from or selling to the other firms, and thus sustaining the network in general.

Figure 6: Degree assortativity - Suppliers and customers of suppliers



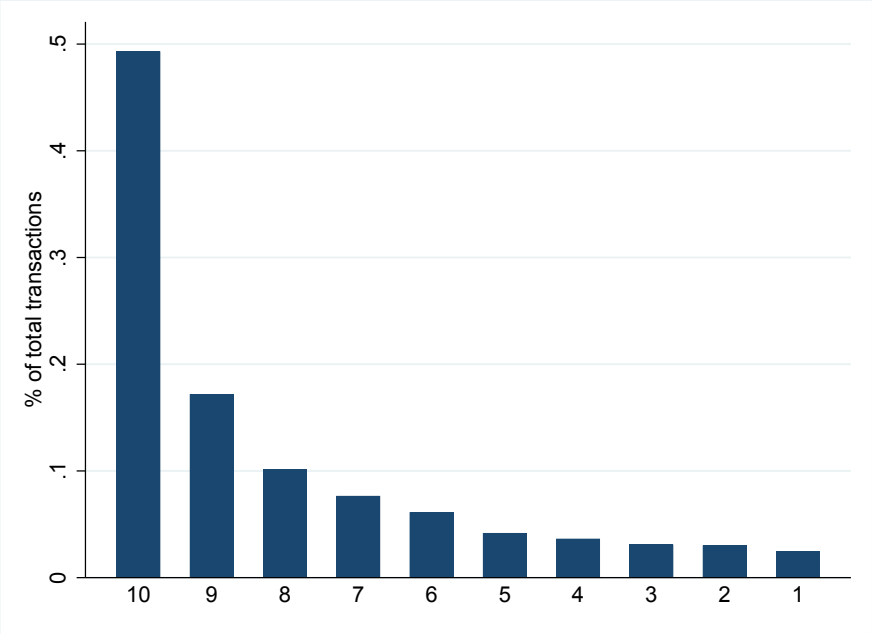
Notes: The figure shows the local polynomial regression of the customer number log (x axis) and the log for the average number of suppliers for said customers (y axis). To calculate the y axis, the firms with the same number of customers are grouped together, and then their average number of suppliers is calculated.

Fact 5. Top clients account for a high percentage of a firm's sales.

Figures 7 and 8 provide an alternative visualization of the heterogeneity among firms. Figure 7 shows the distribution of transactions, ordered by deciles, for seller firms. The figure shows the share that each decile of clients represents in each firm's total sales (on average). By construction, clients in the 10^{th} decile account for the same or more sales than those in the 9^{th} decile, and so on. However, it is still striking to learn how important, overall, the top decile of clients are. According

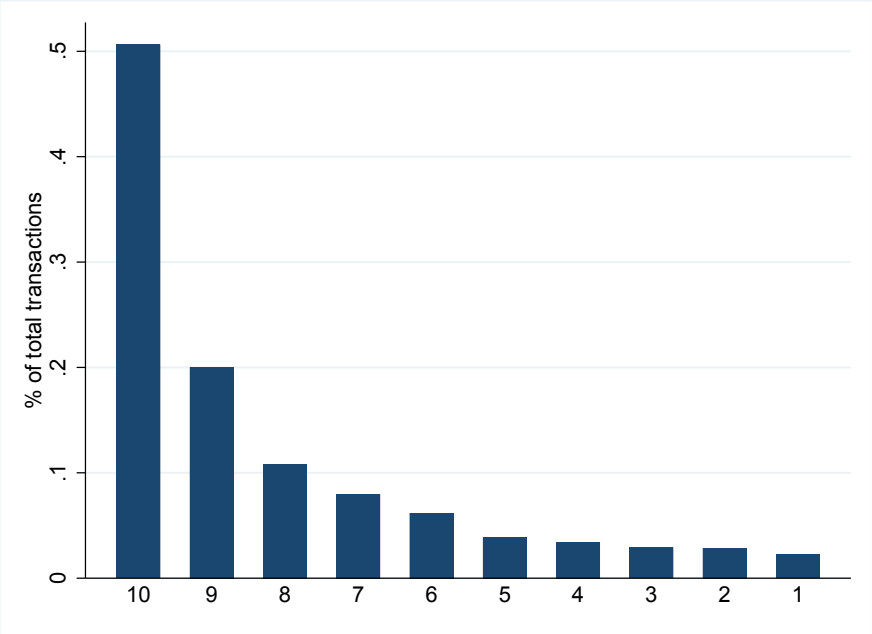
[1] Bernard, Moxnes and Saito (2018) found the same negative correlation for Japanese firms, with a slope of -0.2, which is close to the one for Costa Rica.

Figure 7: Sellers, deciles and share of transactions



Notes: The x axis shows each decile and the y axis indicates the percentage of sales that each decile accounts for in the total transactions. To calculate this share, the transactions of each firm are separated into deciles, and then the transactions for each decile are summed up in order to calculate the ratio between deciles' totals and the total sales of each firm. Finally, each firm's percentage is used to calculate an average value for each decile.

Figure 8: Buyers, deciles and share of transactions



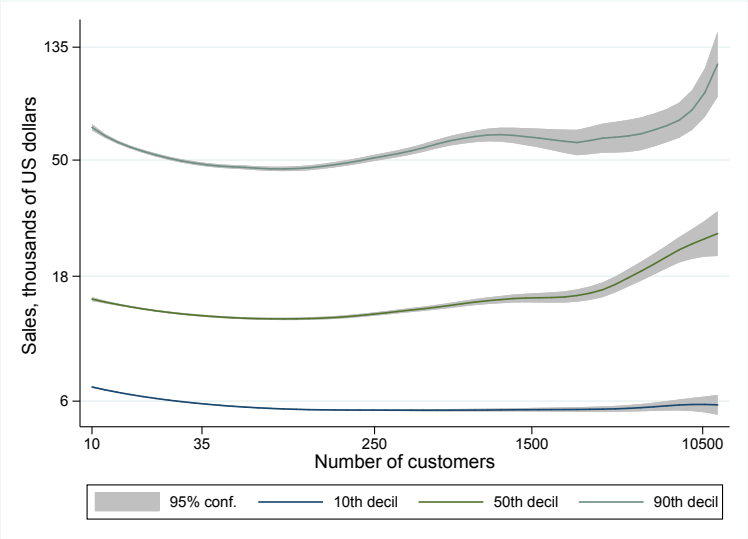
to the graph, the top 10% of clients represent, on average, just under 50% of firms' total sales, while the lowest decile represents approximately 2.5%. This suggests that, on average, sellers rely heavily on a small share of their top clients. This finding does not depend on the size of the seller: both large and small sellers depend on the top 10% of their customers for almost 50% of their total sales.

Figure 8 shows the same stylized fact for buying firms. The same behavior is observed as in the preceding graph, with the only difference being that the highest decile represents, on average, just over 50% of the total sales volume, with the lowest decile representing less than 2%. This pattern is again present irrespective of the size of the buyer.

Fact 6. The distribution of sales (purchases) among customers (suppliers) does not vary much with the number of customers (suppliers).

Figure 9 shows the average across sellers of their 10th, 50th and 90th percentile values of its bilateral sales for firms with more than 10 customores. It can be noted that the three lines are close to parallel. This means that, within a decile, the number of customers that each firm has does not matter, as sales to each are approximately the same whether the firm has ten customers or a hundred customers. In other words, a firm is large because it has many customers, not because it sells more to each of its customers. This finding is consistent with what Bernard, Dhyne, Magerman, Manova and Moxnes (2017) find for Belgium.

Figure 9: Number of buyers and bilateral sales



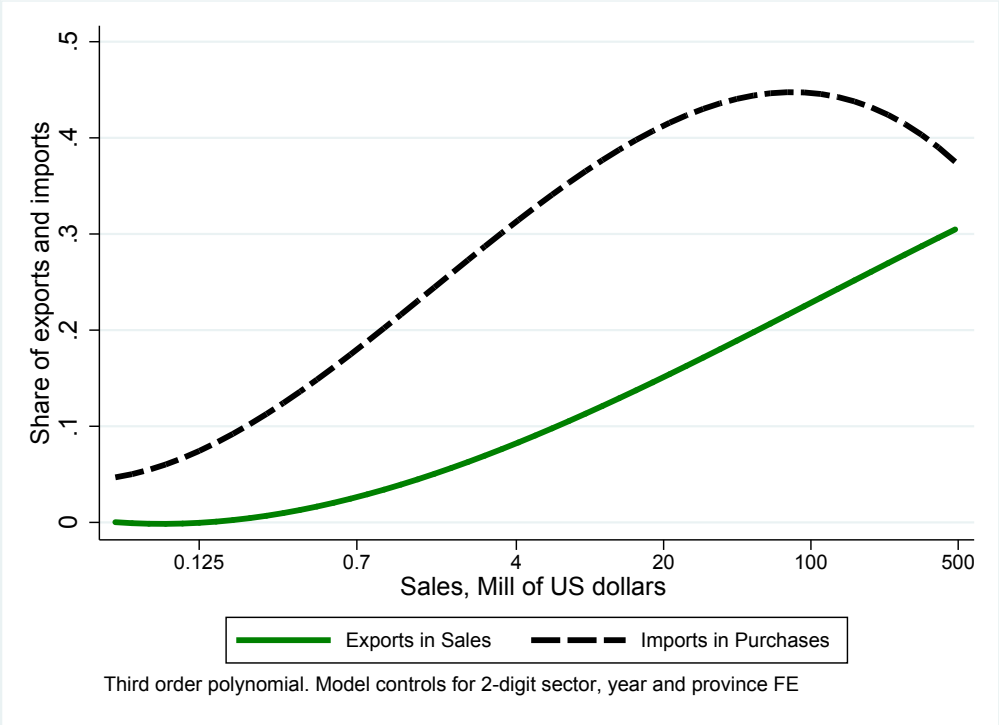
Notes: The figure shows polynomial regressions of the value of firm-to-firm sales at the 10th, 50th and 90th percentile of the distribution of sales on log total number of customers. The regressions only include firms with more than ten customers.

Fact 7. Larger firms have a higher share of imports and exports.

Figure 10 presents the share that exports represent in total sales, the share that imports represent in total purchases, and their relationship with a company's sales. It can be noted that both the import and export shares increase as sales volume increases. The correlation between exports and sales is positive for all values, but begins to decrease when sales exceed US$65 million. Even if a decrease might seem counterintuitive, the confidence intervals (not shown in the figure) are wide enough after US$65 million that we cannot reject the curve to remain constant after that point.

From Figures 3 and 4 and, it is known that a large firm has more connections and is more productive, but now it can also be known that if that firm is an importer or exporter, it will have a high share of imports or exports, and thus it is highly likely that the firm will have a large number of foreign connections as well. In fact, Bernard, Moxnes and Ulltveit-Moe (2018) found a positive correlation between large firms and a high number of foreign connections in the case of Norway.

Figure 10: Share of exports and imports

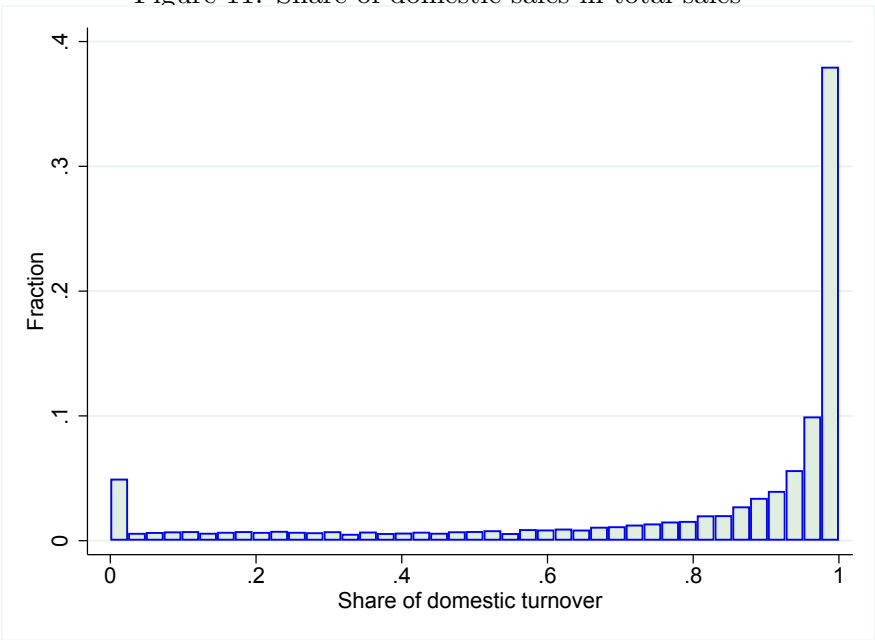


Notes: The figure shows a local polynomial regression of log total sales (x axis) on the share of export (solid line) or import (dashed line) on the y axis. We control for the economic sector, year, and province were controlled, along with a third-degree polynomial with interaction effects for variable x. The export share is calculated as the ratio between total exports and total income, while the import share is calculated as the ratio between total imports and total purchases.

Fact 8. Most sales of exporter firms are domestic sales. Most purchases of importer companies are domestic purchases.

Figure 11 shows the density of the domestic sales share for exporter firms. Worth highlighting is the fact that most exporter firms obtain the majority of their income from sales in the local market. 70% of exporter firms sell more than 80% of their products in the domestic market.

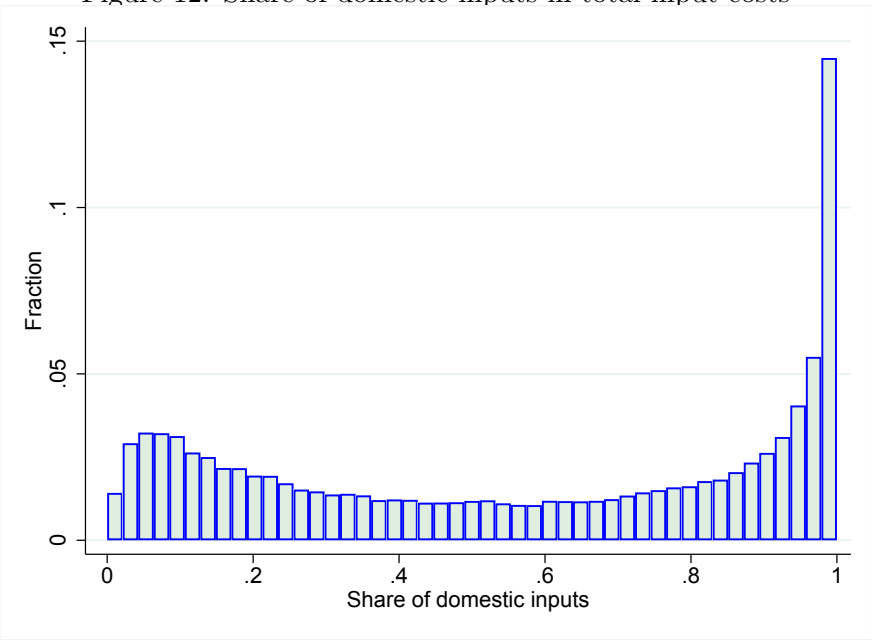
Figure 11: Share of domestic sales in total sales



Remember that the annual income of 50% of the firms that comprise the production network is US$250,000 or less (see Table 5). In other words, most exporters are small firms that export on a small scale. Note in Figure 10 that firms with income below US$250,000 export and import little to nothing.

The same pattern emerges for most importer companies. According to Figure 12, more than 50% of input purchases by a large number of firms are made within Costa Rica. Figure 12 shows the density of the domestic purchases' share for importer firms, indicating a preference for purchasing inputs in the local market. Approximately 50% of firms have domestic input shares over 70%.

Figure 12: Share of domestic inputs in total input costs



4. Conclusions

In this article, we use firm-to-firm transactions data to provide for the first time a set of stylized facts for the domestic production network of a developing country. Only few countries in the world have made this kind of dataset available for researchers and to the best of our knowledge, Costa Rica is the first developing country following this practice.

Our paper represents an initial analysis of the Costa Rican production network. Using a database from the Ministry of Finance containing transactions between 60,478 buyer and seller firms during the period of 2008 to 2015, we characterize the Costa Rican production network through a series of stylized facts. This paper relates and contributes to recent literature on production networks and their importance. The theory on the aggregate implications of the network structure of an economy has quickly grown in recent years, but empirical studies providing evidence using input-output information at the firm-level are in scarce supply (Carvalho, 2014).

Our results discover striking similarities between the network of a developing country and the networks of developed economies. These similarities – on dimensions that we deem of first order importance – suggest that our findings are not specific to Costa Rica, but rather that they extend to other domestic production networks across both the developed and the developing world. Hence, our findings suggest that there are general patterns in how production networks are organized in an economy.

This initial characterization paves the way for future research on production networks and their importance for aggregate and distributional outcomes in developing and developed countries alike.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A. and Tahbaz-Salehi, A. (2012), 'The Network Origins of Aggregate Fluctuations', *Econometrica* **80(5)**, 1977–2016.

Bernard, A. B., Dhyne, E., Magerman, G. C. G., Manova, K. and Moxnes, A. (2017), 'The Origins of Firm Heterogeneity: A Production Network Approach', *Unpublished Manuscript, University of Oslo* .

Bernard, A. B., Moxnes, A. and Saito, Y. U. (2018), 'Production Networks, Geography and Firm Performance', *Journal of Political Economy* .

Bernard, A. B., Moxnes, A. and Ulltveit-Moe, K. H. (2018), 'Two-sided Heterogeneity and Trade', *Review of Economics and Statistics* **16**(E-047), 1–55.

Carvalho, V. M. (2014), 'From Micro to Macro via Production Networks', *Journal of Economic Perspectives* **28**(4), 23–47.

Carvalho, V. M., Nirei, M. and Saito, Y. U. (2014), 'Supply Chain Disruptions: Evidence from Great East Japan Earthquake', *PRI Discussion Paper Series* **16**, 1–44.

Dhyne, E., Magerman, G. and Rubínová, S. (2015), 'The Belgian Production Network 2002-2012', *Working Paper Research No. 288* .

Lucas, R. E. (1977), 'Understanding the Business Cycles', *Carnegie–Rochester Conference Series on Public Policy* **5**(1), 7–29.

Magerman, G., De Bruyne, K., Dhyne, E. and Van Hove, J. (2016), 'Heterogeneous Firms and the Micro Origins of Aggregate Fluctuations', *Working Paper* .