

# Automation in Developing Countries

Plant-level Evidence From Indonesia

**VERY PRELIMINARY AND INCOMPLETE, PLEASE DO NOT  
QUOTE OR CIRCULATE**

Giorgio Presidente\*<sup>1</sup>

<sup>1</sup>Paris School of Economics and World Bank

December 29, 2019

This paper provides novel evidence on the impact of industrial automation in a developing country. Matching industry-level data on shipments of industrial robots to a large sample of manufacturing plants in Indonesia, we find that robotisation, both at the plant-level and through interindustry linkages, boosts plant performance. We interpret the evidence as suggesting that adoption of robots allows producers to upgrade their position in the global value chain.

---

\*The views expressed in this paper are of the author and do not necessarily reflect those of the affiliating institutions.

# 1. Introduction

Do automation technologies present an opportunity or a threat to developing countries? Based on the argument that automation in high-income countries might undermine the competitive advantage of labor-abundant economies, the literature has so far focused on the indirect effect of adoption in advanced economies. Instead, this paper investigates the direct impact of robotisation in a large sample of manufacturing plants in Indonesia. Our results suggest that automation in a developing country can be beneficial to plants' productivity, trade and even employment performance. We interpret the findings as to suggest that robotisation helps firms in developing countries to produce higher value-added products and upgrade their position on the global value chain (GVC).

Such hypothesis has been formulated elsewhere, but it has never been empirically tested. For instance, Rodrik (2018) argues that as technology evolves, products become more complex and consumers more demanding, GVC integration requires strict production requirements, which are hard to satisfy using manual work. Maloney and Molina (2016) discuss how machines can improve the quality of the intermediate products demanded by producers in advanced economies. The World Development Report 2020 makes a similar point and suggests that when different parts of the same product are produced by different firms in different countries, adopting technologies that are also in use in advanced countries contributes to improve compatibility.

Our main methodological challenge is matching data on robots - available at the two digit industry-level, to the plants populating each industry. To do so, we exploit information on the educational attainments of the workforce in each plant. We follow Autor et al. (2003) in assuming that industrial robots are best suited to perform manual routine tasks. In Indonesia, production tasks are usually carried out by workers with compulsory education.<sup>1</sup> Thus, we assign robots in each two-digit industry and year proportionally to plants' shares of midskill labor in the base year.<sup>2</sup> We show that the industry data on robots are positively and robustly correlated to investment in machinery at the plant-level (but not to investment in land and vehicles). Using such a methodology, we construct plants' exposure to automation (ETA).

Even after including the appropriate control variables and fixed effects, trends in robots adoption for a given industry might still be endogenous to plant-level outcomes. For that reason, we construct an instrumental variable using robotisation in the rest of the world (excluding Indonesia). A similar strategy is employed by Faber (2018), and Acemoglu and Restrepo (2017). Our data allows us to construct interindustry linkages by exploiting detailed information on intermediate inputs' use. For each plant, we construct an index of robotisation in upstream and downstream industries. This is important because global shipments of robots might affect plant performance through their impact on supplying and purchasing industries. Therefore, failing to include interindustry indicators would undermine the exclusion restrictions. This is the first

---

<sup>1</sup>Individuals with no education tend to perform unskilled services occupations, while those with non-compulsory education tend to have non-production jobs.

<sup>2</sup>Our procedure inevitably generates measurement errors. However, the strategy allows us to exploit the only available source of information on industrial robots and match it for the first time to plant-level data.

paper to bring the issue to light. Our results can be summarised as follows. Higher ETA reduces product-level marginal costs but boosts markups. This is consistent with improved technical efficiency and higher-quality products. ETA also increases the number of products produced by the plant and the probability of exporting. Moreover, ETA is associated to higher employment of production and non-production workers. While some displacement due to robotisation might be warranted, adopting advanced technologies allows plants to produce higher value-added products, which can translate in a higher demand for labor. Automation in upstream industries boosts intangible investment and lowers the inventories-to-sales ratio, which is consistent with robots allowing for leaner production. Stronger automation in inputs-supplying industries lowers the probability of importing them, which is consistent with the idea that better domestic intermediates' quality reduces dependency from advanced economies. Finally, robotisation in downstream industries is found to have a performance-enhancing impact on product-level outcomes, as well as on the probability of switching the main product produced by a plant.

The results in this paper suggest that the reasons leading firms to invest in robots might be different in advanced and developing countries. In high-income countries, firms invest in automation to lower production costs and increase productivity. Robots can operate 24/7 and can be faster than humans in carrying out routine production tasks. However, another advantage of using industrial robots, is improving the quality and variety of products (De Backer et al. (2018)). In a developing country, achieving product customisation and higher quality standards might have large returns, because it allows firms to upgrade along a value chain. Using robots might also be cheaper than training in overcoming skills' shortage. For instance, Maloney and Molina (2016) report that in an effort to increase the quality of its export, China increased investment in industrial robots, rather than trying to upgrade the workforce. Therefore, it might make sense to invest in automation even if cheap labor is available.

Indeed, while the vast majority of industrial robots are shipped to high-income countries, the IFR (2016) reports that investment in industrial robots grew by double digit figures not only in advanced economies and China, but also emerging economies in South East Asia. Indonesia is no exception to this trend. The number of industrial robots in use in the manufacturing sector grew from none to over 2700 units between 1993 and 2015. Most robots are used in the plastic and rubber industry, where the number of robots per thousand workers was around 6 in 2015. Many advanced economies have similar robot- density in high-automation industries.

This paper is related as well to the literature on the impact of industrial automation on advanced economies (Graetz and Michaels (2018); Acemoglu and Restrepo (2017); Dauth et al. (2017)), and on reshoring. Evidence of reshoring is mixed. Artuc et al. (2019), and Faber (2018) look at local labor markets and find that robotisation in the US reduces exports from Mexico to the US. De Backer et al. (2018) exploit a sample of developing and advanced economies, while Oldenski et al. (2015) a sample of US firms: they both find only limited evidence of reshoring. On the contrary, Artuc et al. (2018) find that greater robot intensity in a advanced economies leads to higher imports and exports from and to less developed ones.

The rest of the paper is organized as follows: Section 2 presents the data; Section 3 discusses the empirical methodology; Section 4 presents the results, and Section 5 concludes.

## 2. Data

This section describes data sources and variables' construction.

### 2.1. Manufacturing Data

Plant level data are taken from the Indonesian survey of manufacturing plants with at least 20 employees (Statistik Industri) administered by the Indonesian statistical office (BPS). The coverage of the survey is extensive; in fact it becomes an actual census in 1996 and 2006 and it is very close to a census in the remaining years, hence ensuring high representativeness even at the provincial level. Plants are grouped into 5 digits sectors following the definition Klasifikasi Baku Lapangan Usaha Indonesia (KBLI). The KBLI classification has been adjusted to be consistent over the whole sample, ranging from 1990 to 2015. One issue is that in converting codes from KBLI rev.3 to KBLI rev.4, some industries are split in more than one industry, or viceversa. For such reason, we only keep those KBLI codes that have an unambiguous one to one mapping across the two revisions.<sup>3</sup>

The plant level data provide information on several variables such as output, capital stock, employment, and expenditure on industrial services, R&D and HR resources. However, one challenge of the Statistik Industri data is the lack of complete series of capital stock. Earlier studies tried to re-construct capital stock series applying the perpetual inventory method (PIM) to the first year of capital stock data reported by the plant (Amiti and Konings (2007); Javorcik and Poelhekke (2017)). However this imputation method crucially relies on the capital value self-reported by the plant the first year this data is available, which is not necessarily accurate.<sup>4</sup> One potential advantage of using PIM is that purchase and sales data might be more accurate relative to self-reported value of the stock, requiring an appropriate calculation of market values and depreciations. However, PIM needs to rely on measures of capital depreciation, which are difficult to accurately estimate. To mitigate such tradeoff, we have adopted a hybrid strategy. We first clean the self-reported adopting an algorithm which keeps only observations that fulfil a battery of tests, which are described in Section A of the appendix. Then, we apply the PIM only to fill the gaps between the missing observations and reapply the same battery of tests to ensure consistency of the series.

Output price deflators are constructed by matching wholesale BPS price indexes available at the 5 digits level IHPB classification (Indeks Harga Perdagangan Besar) with KBLI. Moreover, we are able to obtain different capital deflators depending on the type of asset. We distinguish general price deflators from machinery and equipment, land, vehicles, and buildings. For all deflators, 2010 is used as a base year.

---

<sup>3</sup>We experiment with a looser conversion including more industries, but that does not change our main results.

<sup>4</sup>In particular, there is no a priori reason to believe that the quality of the self-reported capital stock the first year is necessarily better than the value in other years.

## 2.2. IFR Data and Matching to SI Data

Data on shipments of industrial robots are obtained from the International Federation of Robotics (IFR). Industrial robots are defined by ISO 8373:2012 as *an automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes, which can be either fixed in place or mobile for use in industrial automation applications*. The IFR collects data from each national robotics association. Since almost all robots suppliers are members of national associations, the dataset includes virtually all robots used worldwide. An advantage of the data is that the IFR has a common protocol to count robots, so that it ensures consistency across countries and years. Information is available for each country, 2 digits industry and year.

There are two main issues with IFR data. The first is that for several countries, particularly in the early years of the sample, a breakdown of shipments by sector is not available and they are grouped under the label “unspecified”. For these countries, shares by sectors are estimated using information for the years in which the breakdown is available.<sup>5</sup> The resulting shares are used to construct the deliveries by sector. As in Graetz and Michaels (2018), the construction of the stock of operational robots is obtained by assuming a yearly depreciation rate of 10% and applying the perpetual inventory method, using 1993 estimates of the existing stock by the IFR as initial values.<sup>6</sup> The second issue is that shipments are counted in “units”. Therefore, in the paper robots are assumed to have a similar impact irrespective of their size or complexity.

The original IFR industry classification has been converted to obtain eighteen industries, roughly corresponding to 2 digit-level ISIC rev.4. These are: Agriculture, Food and tobacco, Textiles, Paper, Wood and furniture, Chemicals, Rubber and plastics, Non-metallic mineral products, Basic metals, Metal products, Electronics, Machinery and equipment, Motor vehicles, Other transport equipment, Repair and installation of machinery, Construction, and Education and R&D, and Utilities.

Figure 1 shows the total number of robots shipped to Indonesia. Shipments to Indonesia take off after 2002.<sup>7</sup> Figure 2 shows that the lion share of robots was in plastic and rubbers, followed by the motor vehicle industry, which is generally highly automated world-wide.

### 2.2.1. Matching SI and IFR Data

The KBLI classification of SI data is very similar to the ISIC Rev. 4 coding of the IFR data. However, in some cases the SI data are more detailed than the IFR ones. Thus, we group together some KBLI industries to ensure maximum compatibility across the two datasets. The correspondence is shown in Table A1 in Appendix B.

To avoid potentially confounding effects, we assign to each plant to the industry to which it belonged in the first year of observation. That mitigates the concern that due to robotisation,

---

<sup>5</sup>I experiment with two alternatives, namely taking simple averages over all the available years and using the observation for the most recent available year. Results are virtually unchanged.

<sup>6</sup>The IFR does provide estimates of the stock, but it adopts a different assumption that robots fully depreciate after twelve years.

<sup>7</sup>Data from Comtrade are consistent with such aggregate figure.

a plant might switch to another industry.

In order to evaluate the accuracy of the matching, Figure 5 plots the correlation between the average investment in machinery and equipment from SI (by two digit-industry and year), with the IFR data. The figure shows that for positive values of robots' shipment, there is a positive correlation between the two variables.

A second test is performed by regressing the average investment in machinery and equipment from SI (by two digit-industry and year) on the IFR data, including two digit-industry fixed effects and year fixed effects. Table 1 shows that there is a positive correlation between shipments of robots and investment in machinery and equipment, and buildings. The latter might be explained by the fact that proper structures are needed in order to host advanced technologies. As a placebo test, columns 3 and 4 of Table 1 show that as expected, there is no correlation between shipments of robots, land or vehicles.

### 2.3. Education

In 1996 and 2006, the Statistik Industri data are census data. In these two years, the information available is richer and it includes the number of workers at each plant by sex and educational attainments. The educational categories are: i) Not finished primary school; ii) Primary school; iii) Junior high school; iv) Senior high school; v) Diploma; vi) Bachelor; vii) Master, and viii) Ph.D. We compute the shares of employment of midskill workers in 1996 and 2006 by considering midskill workers in categories ii) and iii), which correspond to compulsory education. These are the workers most likely to perform routine tasks. Figure 3 compares the midskill shares of employments in the two years of available information, 1996 and 2006. The figure shows that the shares tend to decrease, presumably due to the adoption of automation technology between 1996 and 2006. This is especially the case for the industries with strongest trends in automation. This is confirmed in Figure 4, which plots the change in the shares between 1996 and 2006, against the 2007 stock of robots per worker.

### 2.4. Inputs Data

A key advantage of our census data is the rich description of the intermediate products and materials used in production. We link inputs to plants using the Indonesian 9 digits product classification KKI (klasifikasi komoditi Indonesia).

We first compute for each plant, the share of each input in total input consumption in the base year, which is 2006.<sup>8</sup> We then use such shares to build the plant-specific weighted average of the NIL dummy variable discussed in the previous section in the industries from which the inputs are sourced, excluding the plant's own industry. A similar strategy is used to construct downstream links. In particular, we inverted the inputs file to compute the shares of domestically sold inputs to each 5 digits-level industry, excluding intra-industry trade in

---

<sup>8</sup>Our choice of for the base year is dictated by two factors. First, the quality of input data tend to deteriorate after 2007. Second, since 2006 was a census years, it ensures maximum coverage.

order to avoid potentially confounding effects. This strategy allows us to improve on the precision of the between-industry links, which in the literature have been constructed using highly aggregate IO tables that are publicly available (e.g. Genthner and Kis-Katos (2019)).

## 2.5. Product-level Markup and Marginal Cost

Our main performance variables are product-level markups and marginal costs. We estimate such measures at the product-level using a trans-log production function, as in De Loecker et al. (2016). Using this methodology allows us to account for the fact that many plants produce more than one product (9 digits KKI classification) and purge the productivity measures from two important sources of bias: input price-bias and input allocation-bias. Input price bias might arise because, due to differences in quality or market power, plants might face different prices for the same input. Input allocation bias might arise because in the data we do not observe how inputs (e.g. capital, labor and materials) are allocated to the production of each product. Failing to account for these sources of bias can result in biased estimates.<sup>9</sup>

---

<sup>9</sup>An extensive discussion on input price and input allocation bias can be found in De Loecker et al. (2016), which also provide details on the estimation methodology.

### 3. Empirical Methodology

We observe the number of industrial robots in each two digit-industry  $i$  and year  $t$ . On average, there are  $XX$  plants  $f$  operating in industry  $i$ . To construct plants' exposure to automation,  $ETA_{f,t}$ , we assume that robots are used by plants proportionally to their share of midskill labor in total employment,

$$ETA_{f,t} \equiv \frac{R_{i,t-1}}{L_{i,t_0}} \times midskill_{f,t_0} \quad (1)$$

In (1),  $\frac{R_{i,t-1}}{L_{i,t_0}}$  is the number of industrial robots in use per worker in industry  $i$ . Since it might take some time to make a robot fully operational, we use one year lags. To avoid potential endogeneity with robots' use, in the denominator of (1) we use the base year (2006) total number of industry workers. The share of midskill workers  $midskill_{f,t_0}$  refers to 2006.

To compute the interindustry linkages, we proceed as follows. For upstream linkages, let  $\sigma_{fj}$  be the plant  $f$ ' share of inputs sourced from each industry  $j$  that is different from  $i$  (the industry in which  $f$  operates). Then,  $ETA_{f,t}^{up}$  is the exposure to automation in upstream industries,

$$ETA_{f,t}^{up} = \sum_{j \neq i} \sigma_{fj} \cdot \frac{R_{j,t-1}}{L_{j,t_0}}$$

Using a similar approach, we compute the exposure to automation in downstream industries,

$$ETA_{f,t}^{down} = \sum_{j \neq i} \lambda_{fj} \cdot \frac{R_{j,t-1}}{L_{j,t_0}}$$

Our instrumental variable is based on global trends in robots' use. Information on industry level employment is only available for OECD countries from STAN, so we use the the OECD average number of robots per worker,  $\tilde{R}_{i,t-1}$ .<sup>10</sup>

$$XETA_{f,t} \equiv \frac{\tilde{R}_{i,t-1}}{\tilde{L}_{i,t_0}} \times midskill_{f,t_0}$$

Finally, we estimate the following system of equations,

$$ETA_{f,t} = \alpha_0 + \alpha_1 XETA_{f,t} + \alpha_2 ETA_{f,t}^{up} + \alpha_3 ETA_{f,t}^{down} + AX_{f,t} + u_f + u_{i,t} + u_{r,t} + \epsilon_{f,t} \quad (2)$$

$$Y_{f,t} = \beta_0 + \beta_1 ETA_{f,t} + \beta_2 ETA_{f,t}^{up} + \beta_3 ETA_{f,t}^{down} + BX_{f,t} + u_f + u_{i,t} + u_{r,t} + \varepsilon_{f,t} \quad (3)$$

In (2) and (3),  $X_{f,t}$  includes plant and district-level controls. These are base year values interacted with year effects of district-level: i) real gdp per capita; ii) population; iii) share of manufacturing; iv) share with population with tertiary education, and v) plant-level share of foreign equity. The variables  $u_f$ ,  $u_{i,t}$  and  $u_{r,t}$  denote, respectively, plant, industry-year and

<sup>10</sup>As in (1), the denominator of the instrument is fixed at its base year value.

region-year fixed effects. The base year (2006) is dropped from the sample prior to estimation and errors are clustered at the plant level.

## 4. Results

Table 2 presents the results. Columns 1 and 2 present the plant-product-level estimates of the impact of robots on markups and marginal costs. Plant-level robot-density (the number of robots per workers, “ETA”) increases markups and reduces marginal costs. We interpret the evidence as suggesting that robots increases the quality of the products produced and at the same time improve the technical efficiency of the plant. The same results hold for the interindustry linkages. Robotisation in upstream and downstream industries has a beneficial impact on plants’ performance. For upstream industries, that might be the result of better intermediate products and enhanced efficiency of suppliers. For downstream industries, the result might be driven by higher quality required from supplier, and possibly the impact of technology spillovers from high-tech customers. It should be noticed that the effects are generated by domestic suppliers and customers only.

ETA increases products variety (column 3). No impact from upstream and downstream industries is detected. In column 4, however, it can be seen that downstream robotisation tends to trigger switch in the main products produced by the plant, which is consistent with plants adapting to higher-technology customers.

Columns 5 and 6 look at the impact of robotisation on trade performance. ETA increases the share of production exported by plants, while it reduces the number of imported inputs. That is consistent with suppliers producing higher quality inputs and so reducing dependency from more advanced economies.

Columns 7 and 8 look at the impact on employment. The wage bill of production and non-production workers increases with ETA. Non-production workers include both white-collars and elementary occupations such as janitorial services and support activities, which include unskilled labor.

Column 9 shows that robotisation in upstream industries reduces the inventories-sales ratio, which is consistent with leaner production Holweg (2007).

Finally, column 10 shows that robotisation in upstream industries increase plant-level expenditure in R&D and other professional services, which we take as a proxy for investment in intangible capital.

## 5. Conclusion

This paper provides novel evidence on the impact of industrial automation in a developing country. Matching industry-level data on shipments of industrial robots to a large sample of manufacturing plants in Indonesia, we find that robotisation, both at the plant-level and through interindustry linkages, boosts plant performance. We interpret the evidence as suggesting that adoption of robots allows producers to upgrade their position in the global value chain. Among the possible channels driving the boost in performance, robotisation might help achieving the precision and quality needed for the compatibility of inputs in high-tech industries.

Figure 1:

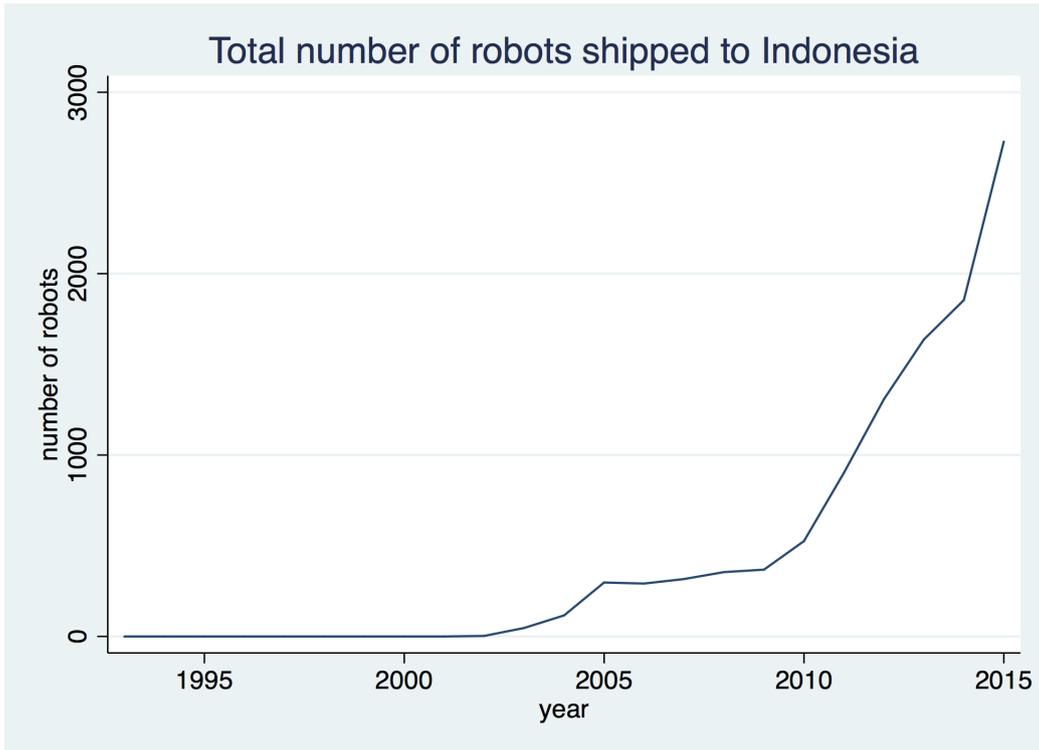


Figure 2:

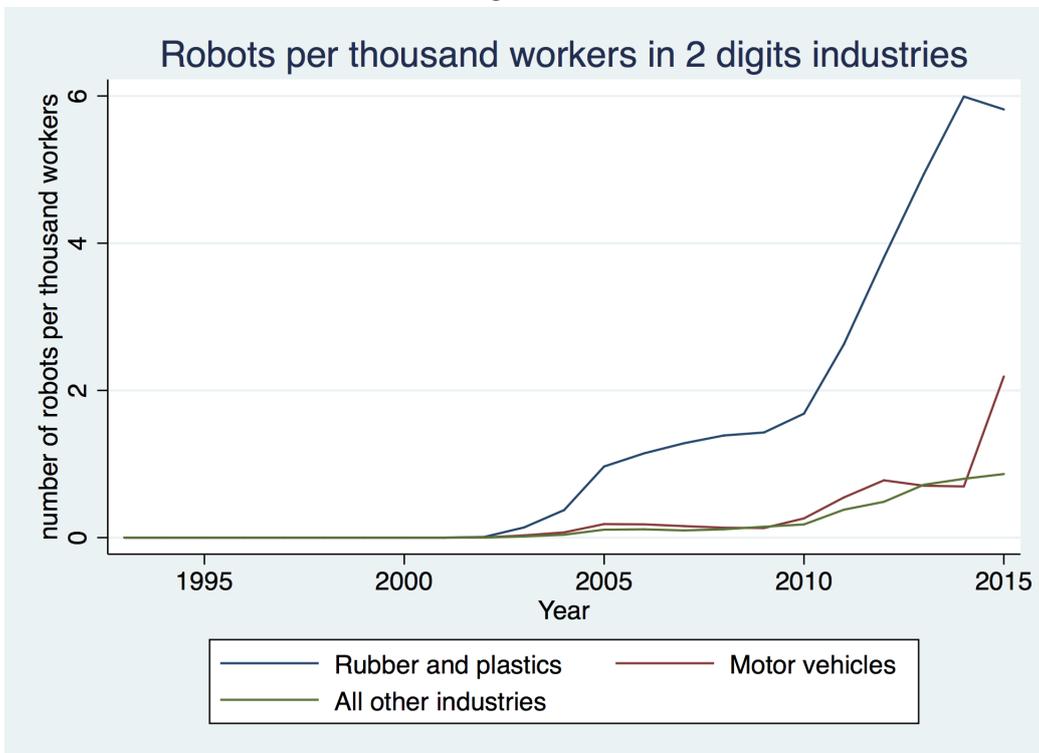


Figure 3: Comparison of midskill in 1996 and 2006.

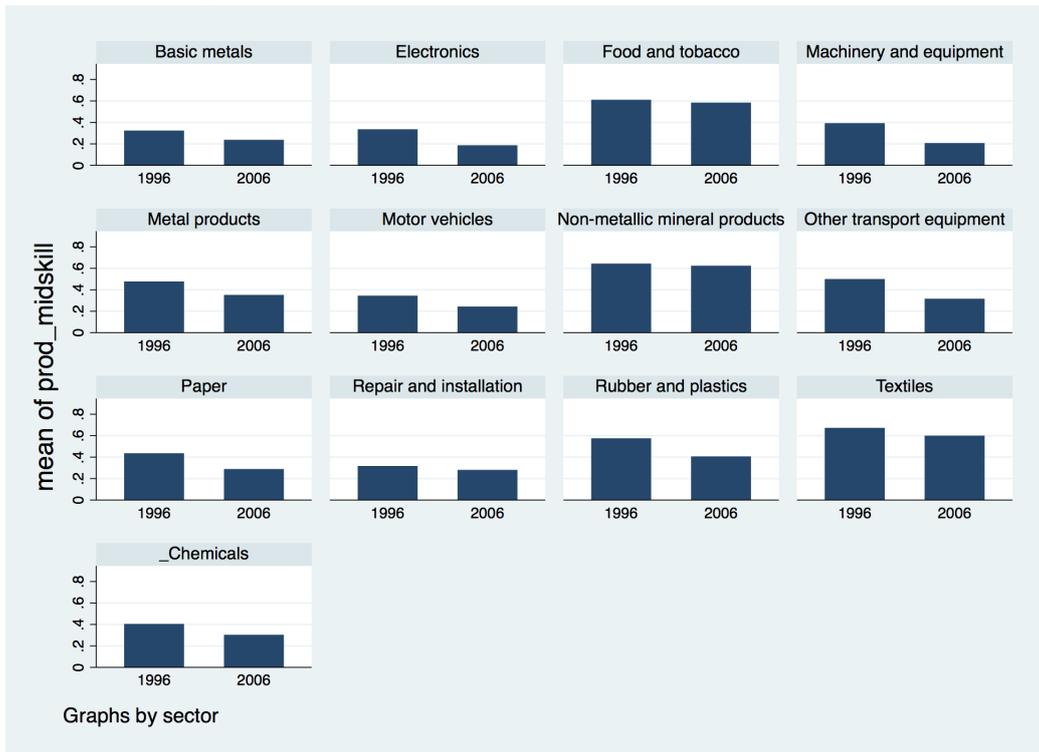


Figure 4: Change in midskill and ETA.

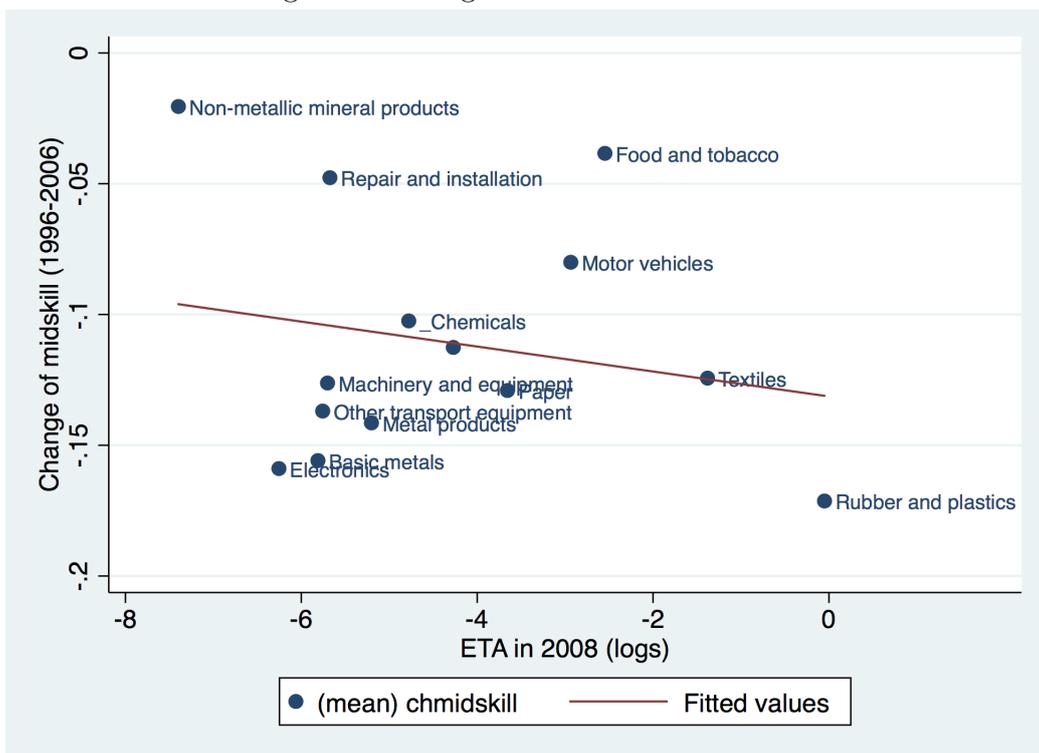


Figure 5:

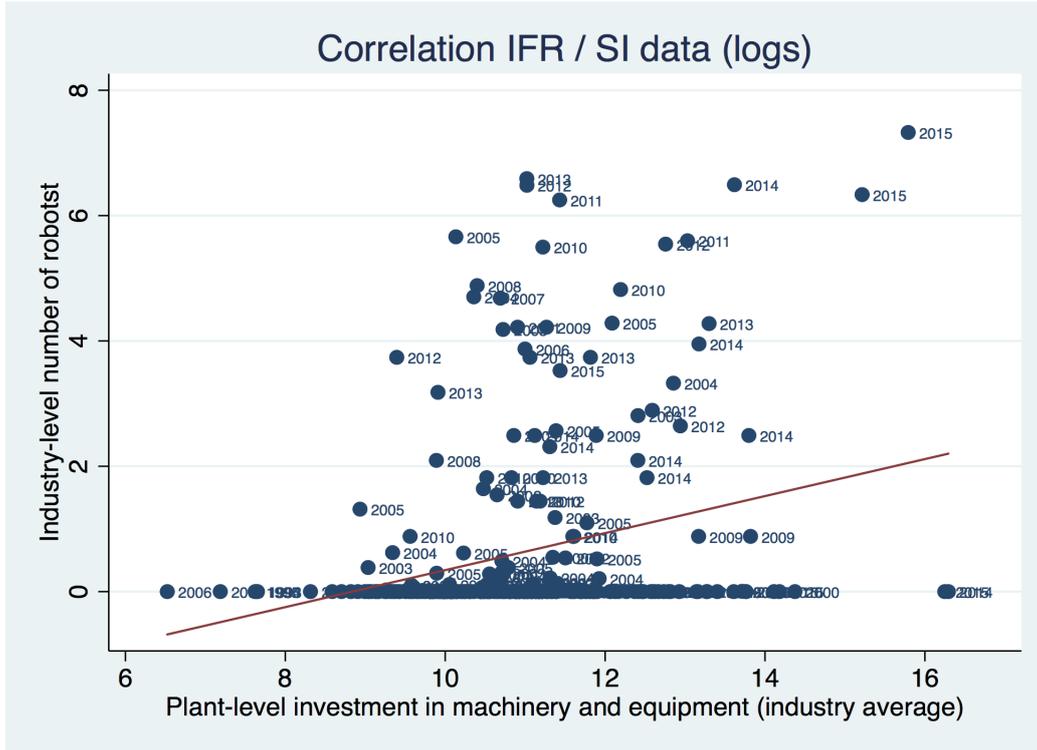


Table 1:

VARIABLES	(1) Purchase of machinery (SI)	(2) Purchase of buildings (SI)	(3) Purchase of vehicles (SI)	(4) Purchase of land (SI)
Shipments of robots (IFR)	0.132** (0.052)	0.185*** (0.061)	0.077 (0.058)	-0.046 (0.117)
Observations	308	308	308	308
R-squared	0.426	0.288	0.276	0.333
Number of industries	14	14	14	14
2 digits industry FE	yes	yes	yes	yes
Year FE	yes	yes	yes	yes

Table 2:

VARIABLES	(1) Price markup	(2) Marginal cost	(3) N.of products	(4) Product switch	(5) Exports	(6) Imports	(7) Blue-collars	(8) Other workers	(9) Inventory/sales	(10) Intangible capital
Plant-level ETA	0.235*** (0.090)	-0.179* (0.100)	0.110*** (0.037)	0.011 (0.011)	0.011* (0.006)	-0.001 (0.003)	0.152*** (0.034)	0.130*** (0.040)	0.340 (0.982)	0.114 (0.117)
Upstream ETA	0.103*** (0.038)	-0.096** (0.040)	0.004 (0.011)	-0.003 (0.003)	-0.003 (0.002)	-0.008*** (0.002)	0.004 (0.010)	-0.011 (0.015)	-0.261** (0.121)	0.078*** (0.030)
Downstream ETA	0.117*** (0.044)	-0.106** (0.048)	-0.005 (0.013)	0.013*** (0.005)	-0.001 (0.002)	-0.002 (0.001)	-0.012 (0.010)	-0.018 (0.012)	-0.843 (0.747)	-0.028 (0.035)
Observations	50,545	50,545	90,442	90,442	90,437	90,442	90,204	70,550	90,438	90,442
Number of plant-product	16,701	16,701								
Plant-9 digits-product FE	yes	yes								
2 digits industry-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Base year district-level covariates-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Base year plant-level covariates-year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
First stage F	360.8	360.8	318.5	318.5	318.5	318.5	317.7	248.8	318.3	318.5
Number of PSID			14,566	14,566	14,566	14,566	14,559	12,805	14,566	14,566
Plant FE			yes	yes	yes	yes	yes	yes	yes	yes

## References

- Acemoglu, D. and Restrepo, P. (2017). Robots and jobs: Evidence from us labor markets. *NBER working paper*, (w23285).
- Amiti, M. and Konings, J. (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review*, 97(5):1611–1638.
- Artuc, E., Bastos, P., and Rijkers, B. (2018). Robots, tasks and trade.
- Artuc, E., Christiaensen, L., and Winkler, H. J. (2019). *Does Automation in Rich Countries Hurt Developing Ones?: Evidence from the US and Mexico*. The World Bank.
- Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333.
- Dauth, W., Findeisen, S., Südekum, J., and Woessner, N. (2017). German robots-the impact of industrial robots on workers.
- De Backer, K., DeStefano, T., Menon, C., and Suh, J. R. (2018). Industrial robotics and the global organisation of production.
- De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.
- Faber, M. (2018). Robots and reshoring: Evidence from mexican local labor markets.
- Genthner, R. and Kis-Katos, K. (2019). Foreign investment regulation and firm productivity: Granular evidence from indonesia.
- Graetz, G. and Michaels, G. (2018). Robots at work. *Review of Economics and Statistics*, 100(5):753–768.
- Holweg, M. (2007). The genealogy of lean production. *Journal of operations management*, 25(2):420–437.
- Javorcik, B. and Poelhekke, S. (2017). Former foreign affiliates: Cast out and outperformed? *Journal of the European Economic Association*, 15(3):501–539.
- Maloney, W. F. and Molina, C. (2016). *Are automation and trade polarizing developing country labor markets, too?* The World Bank.
- Oldenski, L. et al. (2015). Reshoring by us firms: what do the data say? *Peterson Institute for International Economics, Policy Brief PB15-14*.
- Rodrik, D. (2018). New technologies, global value chains, and developing economies. Technical report, National Bureau of Economic Research.

# Appendix

## A. Construction of the Capital Series

In order to avoid relying on depreciation rates, we tried to preserve the self-reported original values by the plant as much as possible and applied the PIM only to fill gaps. In this paper self-reported capital series were object of an extensive cleaning algorithm aimed at mitigating measurement errors.<sup>11</sup> Our algorithm consists first in replacing zero or negative values as missing observations and then applying a two-steps procedure based on capital-labor ratios (KL). For each year, we compute the average KL in each 4 digit KBLI sector over the whole sample, but excluding the years in which the average and total values of the capital stock exhibited suspicious jumps, i.e. 1996, 2000, 2003, 2006, 2009 and 2014. An observation is dropped if the ratio of plant-KL to the sector average KL is below 0.02 or larger than 50.<sup>12</sup> Then, in a second step we compare a plant KL in a given year with the average value of the KL within the same plant but in the other years of observation. An observation is dropped if the ratio of plant-year-KL to the plant average KL is below 0.2 or larger than 5. Plants are dropped from the sample in case the cleaning procedure results in all missing values of self-reported capital.

When a plant has some but not all valid observations for self-reported capital stock, then missing values are replaced by applying a forward/backward perpetual inventory method (PIM). Being only a fraction of the total observations, we rely less on estimates of depreciation rates.<sup>13</sup> Previous studies focus on the first year of observation of a plant, without assessing the plausibility of the data point. Since PIM series are very sensitive to the choice of the initial observation, especially with relatively short time series, the resulting capital stock could be severely mis-measured. Moreover, information on purchases and sales of capital equipment, which is subject to the same measurement errors of the reported capital. For such a reason, after filling missing values with the PIM we re-apply the two stages check described above in order to minimise the possibility of mis-measurement. As a final test, we compute plant-level growth rates of KL and we check that it is reasonably distributed (Figure A1). Figure A2 compares original and clean capital stock series.

---

<sup>11</sup>One important problem with the reported series is that in some years, there are plants were characterised by implausible large values of capital. Studying the behaviour of the stock within plants reveals that in some circumstances plants reported values in different units. The phenomenon is somewhat more frequent in 1996 and 2006, when the BPS conducted a wider economic census that collected information in units rather than in thousand Rupiah. For instance, in 2006 the number of surveyed firms increased by 40%. The increase in coverage required hiring unexperienced enumerators that were more likely to make mistakes, which contributed to increase measurement errors.

<sup>12</sup>We experiment with stricter thresholds which result in too many observations dropped.

<sup>13</sup>We follow Arnold and Javorcik (2009) and assume that the annual depreciation rate for buildings is 3.3 percent, for machinery 10 percent, and for vehicles and other fixed assets 20 percent. For land, we assumed no depreciation.

Figure A1: Plants' growth rate distribution of capital-labor ratio.

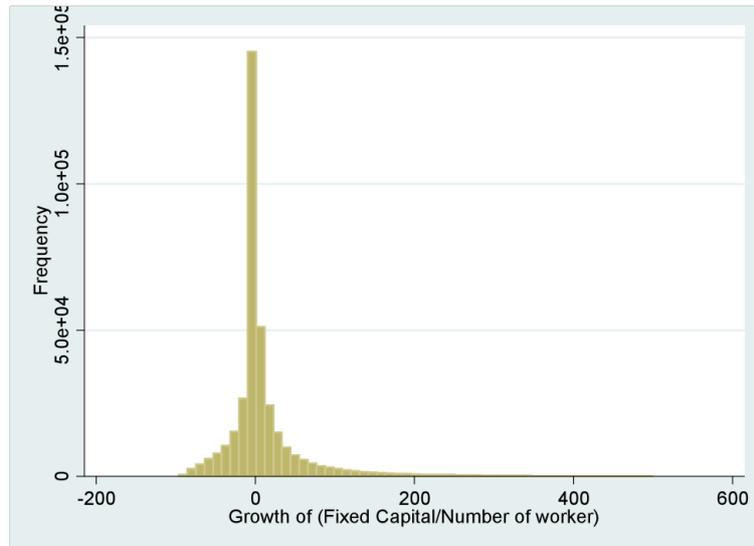
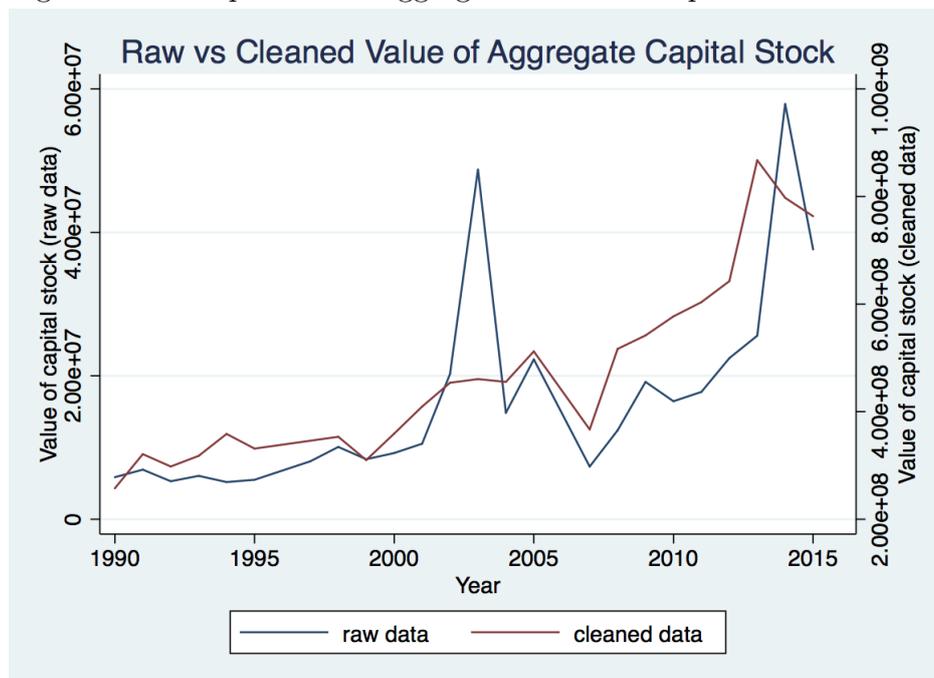


Figure A2: Comparison of Aggregate Nominal Capital Stock Series.



## B. Conversion Tables IFR / KBLI

Table A1:

<b>IFR industries</b>	<b>Description</b>	<b>KBLI industries</b>
D10T12	Food products, beverages and tobacco	10,11,12
D13T15	Textiles, wearing apparel, leather and related products	13,14,15
D16and3132	Wood,furniture,n.e.c	16,31,32
D17T18	Wood and paper products	17,18
D19T21	Chemicals	19,20,21
D22	Rubber and plastics products	22
D23	Other non-mineral products	23
D24	Basic metals	24
D25	Metal products	25
D26T27	Electronics	26,27
D28	Machinery and equipment n.e.c.	28
D29	Motor vehicles	29
D30	Other transport equipment	30
D33	Repair and Installation	33