Oxford Martin Working Paper Series on Economic and Technological Change

Automation and Manufacturing Performance in a Developing Country

Massimiliano Calì and Giorgio Presidente

Working Paper No. 2021-1

Automation and Manufacturing Performance in a Developing Country

Massimiliano Calì*[1] and Giorgio Presidente†[2]

[1] *World Bank*
[2] *Oxford Martin School, Oxford University*

Abstract

This paper provides novel evidence on the economic impact of industrial automation in a large developing economy. It combines labor force survey and manufacturing plant level data from Indonesia over the period 2008-2015, when the country experienced a rapid increase in robots import. The findings show a positive impact of robots on various measures of plants' performance and integration in Global Value Chains. In contrast to existing evidence on advanced and emerging economies, these plant level impacts result in an increase in manufacturing and services employment at the local level. Such employment effects are consistent with evidence of positive employment spillovers from downstream robots-adopting plants, which help extend the benefits of automation to non adopting plants. The spillover effects may provide a rationale to incentivize manufacturing firms to adopt industrial robots. The results also suggest that the gains from automation are not equally shared: robots' adoption is associated with a reduction in the labor share in value added and an increase in skill wage premia.

Keywords: Robots, Automation, Development, GVC, Employment, Productivity, Indonesia

JEL classification: O14,J23,J24,L11,F63

*The authors thank Erhan Artuc, Paulo Bastos, Aufa Doarest, John Morrow, Bob Rijkers and Pengfei Zhang for helpful comments, Nabil Ryandiansyah for outstanding research assistance and Mauro Testaverde for sharing automatability data in south-East Asian countries. Giorgio Presidente gratefully acknowledges funding from Citi. The views expressed in this paper are of the authors and do not necessarily reflect those of the affiliating institutions.

†Corresponding author: giorgio.presidente@oxfordmartin.ox.ac.uk

1 Introduction

Are automation technologies an opportunity or a threat to developing countries? The literature has examined this question by studying automation in advanced economies - which have been early adopters - focusing on the impact on imports from developing countries (Kugler et al. (2020); Artuc et al. (2019); Faber (2018); Artuc et al. (2018)).[1] Much less is known about the impact of automation in developing countries on their own economies. This is an important gap for at least two reasons. First, given the differences in the structure of production and the type of skills in the labor force, it is not clear that the evidence on automation in high-income countries may provide useful guidance for developing countries. Second, firms in low and middle income countries have begun to invest in automation technologies, whose penetration is expected to grow over the next decades (Hallward-Driemeier and Nayyar (2017)).

To help fill the gap, this paper focuses on industrial robots, an important class of automation technology. It examines empirically the impact of robots on firms and local labor markets in Indonesia, which is a suitable context for this analysis. The number of robots in the country was very limited before the beginning of our sample in 2008 and accelerated stiffly thereafter. By the end of the sample in 2015, the penetration of robots in the most automated industries was similar to advanced economies. Therefore, the experience of Indonesia - an early adopter among developing countries - should be informative about other large developing economies, in which adoption rates are still limited today and are expected to grow. Indonesia provides also a rich set of high quality data including a large panel of manufacturing plants and labor force surveys, which our analysis can leverage along with data on robots' import.

Our key contribution is to document a positive impact of automation on employment in Indonesia (at least until 2015). Consistently with the predictions of a task-based model (Acemoglu and Restrepo, 2018b), the plant level results suggest that the positive impact on employment is driven by the productivity gains of automating plants. Our result differs

[1] A commonly tested hypothesis, for instance, is that adoption of robots in advanced economies induced re-shoring.

from the existing literature based on advanced (Graetz and Michaels (2018); Acemoglu et al. (2020); Acemoglu and Restrepo (2019); Koch et al. (2019); Dauth et al. (2017)) and emerging economies (Giuntella and Wang (2019); Artuc et al. (2019)), which document a negative impact of robots on employment. Unlike this literature, our results suggest that for labor the productivity-enhancing effect of robot adoption prevails over the replacement effect. As a further departure from the evidence in advanced economies, we also find that automation generates positive employment spillovers from down-stream robots-adopting plants, which help extend the benefits of automation to non-adopting plants as well. We discuss potential underlying reasons for the difference in findings between Indonesia and other more advanced economies in the concluding section.

By increasing revenues and the demand for domestic inputs, robots generate positive employment spillovers from downstream automating industries. Such spillovers are not limited to manufacturing, but extend to services and construction, thus boosting the demand for labor across sectors. This finding is in line with the increasingly important role of services inputs in manufacturing production, including in Indonesia (Hallward-Driemeier and Nayyar (2017); Duggan et al. (2013)).

Our second contribution is to examine in depth the plant level impacts of robot adoption. We provide evidence that plants more exposed to robots lowered real marginal costs, but they also increased markups, export and import. This suggests that robots penetration in Indonesia allowed firms to produce higher quality products and upgrade in global value chains (GVCs).[2]

We focus the analysis on the impact of robots at two levels of aggregation: plant and local labor market. The local labor market analysis exploits labor force survey data to calculate the shares of workers at high risk of automation in each industry of a local market in the base year. The shares are then used to aggregate robots import by industry and year within the local labor market.[3] We study long-run trends and account for the

[2]Rodrik (2018) argues that GVC integration requires strict production requirements, which are hard to satisfy using manual work. Similarly, World Bank (2020b) and De Backer et al. (2018) suggest that by adopting technologies used in advanced economies, developing countries could boost trade relationships. Maloney and Molina (2019) discuss how automation in developing countries can improve the quality of inputs demanded by advanced economies.

[3]The methodology in this part of the analysis is similar to Acemoglu and Restrepo (2017) and Dauth

potential correlation of residuals across local markets with similar industry composition by calculating shift-share errors (Adao et al. (2019)).

For the plant level analysis, the main challenge is that we do not observe robots' adoption at the plant level. Instead, we follow Autor et al. (2003) and assume that industrial robots are best suited to perform manual routine tasks. Since we observe the educational attainments of employment at the plant level, we calculate the typical level of education of routine task-intensive occupations based on the Indonesian labor force survey. This turns out to be either Junior or Senior High School Diploma ("secondary education" hereafter).[4] We measure plant level exposure to robots by interacting yearly imports of robots by 2-digit industry to plants share of secondary education workers in the base year. Under the assumption that within robot-adopting industries, plants with larger shares of secondary education have greater opportunities for automation, our exposure measure can be used to estimate the impact of robots by comparing the outcomes of firms with large and small shares of employment with secondary education.

We perform a battery of tests to assess the validity of our exposure measure and we show that it is strongly correlated with plant level investments in machinery and equipment. Our empirical methodology addresses concerns related to unobserved firm heterogeneity and time-varying industry characteristics that may be related with robots import and firm outcomes. We also provide robustness tests instrumenting Indonesia's robots import with OECD imports, an approach similar to Acemoglu and Restrepo (2017), Dauth et al. (2017), Faber (2018) and Giuntella and Wang (2019).

Our paper is related to the growing literature on the impact of industrial automation on advanced economies at the industry and local labor market-level (Graetz and Michaels (2018); Acemoglu and Restrepo (2017); Dauth et al. (2017); Giuntella and Wang (2019); Artuc et al. (2019); Kugler et al. (2020)), and at the firm-level (Acemoglu et al. (2020); Koch et al. (2019); Stapleton and Webb (2020)). These contributions suggest that differently to Indonesia, robots have an overall negative impact on labor demand in high-income

et al. (2017), among others.

[4] Our definitions of routine task-intensive occupations follow Graetz and Michaels (2018) and Frey and Osborne (2017). The details are presented in Section 2.1.

as well as large emerging economies.

A related body of literature focuses the determinants and impact of reshoring, providing mixed results. Artuc et al. (2019), Kugler et al. (2020) and Faber (2018) look at local labor markets and find that automation in the United States reduces exports from Mexico and Colombia to the United States. 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 an advanced economy leads to higher imports and exports from and to developing economies.

The paper is also related to the literature on firm upgrading in developing countries (see Verhoogen (2020) for a review). It contributes to this literature by providing indirect evidence on the factors driving adoption of an increasingly important class of technology. It also complements existing studies documenting the (large) returns to technology adoption not only at the firm level (Bloom et al. (2013); Cai and Szeidl (2018)) but also across firms via backward linkages. In so doing it provides evidence that robots' adoption could be an effective way to reduce the gap in the growth-age gradient between manufacturing firms in developing countries vis-à-vis the United States (Hsieh and Klenow, 2014). This in turn could reduce the large gap in aggregated productivity across many developing and high-income countries.

The rest of the paper is organised as follows: Section 2 presents plant level methodology and data; Section 3 presents the plant level results; Section 4 presents the aggregate analysis at the level of both industry (4.2) and local labor markets (4.3); Section 5 explores the distributional impact of automation, and Section 6 concludes.

2 Plant Level Analysis

The empirical methodology is based on increasing levels of aggregation. We start at the most refined level by examining the impact of automation on plant level outcomes.

2.1 Measuring Plant Level Automation

Unlike some other studies in high-income countries (e.g. Acemoglu et al. (2020) and Stapleton and Webb (2020)), we do not observe directly robots' use by plants. Instead, we match data on Indonesian imports of industrial robots by industry with plants' observable characteristics to build various plausible measures of plant level exposure to automation.

The key assumption underlying our proxies of plant exposure is that routine task-intensive occupations are the most likely to be automated (Autor et al., 2003). As we do not observe the tasks performed by workers in each plant, our baseline measure of exposure proxies routine-intensive tasks with the educational level of each plant's workforce, which we observe in 2006 when our manufacturing plants' data becomes a census.

Specifically, we identify the production occupations most exposed to being automated using two alternative definitions of occupations' "replaceability". The first is the definition used in Graetz and Michaels (2018) (GM). GM exploit information on the applications for which robots are used. They look at 3-digit occupational classifications in the United States and assign a replaceability value of 1 to an occupation if its title corresponds to at least one of the IFR application categories and 0 otherwise. For instance, if robots perform the task "handling materials", they deem an occupation to be replaceable if the occupational title includes the word "handling". The second is the probability of computerization calculated by Frey and Osborne (2017) (FO henceforth), which update the probability of computerization by occupation originally computed by Autor et al. (2003) and David and Dorn (2013). Compared to these studies, FO has the advantage of accounting for the more recent technological progress.

Using Indonesian labor force survey data (described in Section 4.3), we compute the distribution of workers in occupations at high risk of automation by their educational attainments in 2007, the first year such information is available.[5] Using the GM definition, secondary education is the typical educational attainment of workers in occupations at risk of automation in Indonesia (column 1 of Table 1). The results in column (2) of

[5]As in FO, we deem an occupation as highly likely to be automated if the probability of computerisation is larger than 0.7.

Table 1 confirm that using the FO definition, it is still the case that the typical level of education of workers in occupations at risk of automation is secondary education.

Table 1: Share of employed workers in occupation at high risk of automation, by educational attainment.

	(1)	(2)
Educational attainments	**GM**	**FO**
Primary	.39	.42
Secondary	.59	.56
Tertiary	.02	.02

The table reports the share of employment in production occupations at high risk of automation, by the educational attainments of Indonesian workers in 2007. Primary education includes up to completed primary school. Secondary education includes junior and senior high-school. Tertiary education includes education levels from diplomas to PhD. FO indicates that the list of occupations at risk of automation is based on the methodology in Frey and Osborne (2017). GM indicates that the list of occupations at risk of automation is based on the methodology in Graetz and Michaels (2018). Sources: Sakernas (LFS); Frey and Osborne (2017); Graetz and Michaels (2018).

The finding that occupations at risk of automation are dominated by secondary educated workers is consistent with two pieces of evidence. First, the literature on employment polarization suggests that automation technologies tend to replace occupations with an intermediate level of skills.[6] While we do not observe the skill level of plants' workforce directly, secondary education is both the median and the mean level of education of workers in manufacturing plants in 2006.

Second, an analysis of the characteristics of production occupations compiled by the World Bank for Indonesia, Thailand and Malaysia, suggests that secondary education is the typical educational attainment for all of the occupations at greater risk of automation across countries with similar economic characteristics.[7] The information provided in the World Bank occupation profiles is similar to the Occupational Information Network database (O*NET) for the United States, but it is based on analyses of labor force survey data for comparable Asian economies. An example of Occupation Profile for "Welders and Flame Cutters" is provided in Figure A3. According to the World Bank, this is an occupation at high risk of automation. The key piece of information provided by the Occupation Profiles is the typical educational attainments of the occupation: Junior High

[6]E.g. Goos et al. (2014); David and Dorn (2013); Goos et al. (2009); Goos and Manning (2007)).

[7]We thank Mauro Testaverde for sharing these unpublished profiles with us. More details on the Occupation Profiles is presented in Section B.

School and Vocational High School, which fall into our category of secondary education (see Section 2.3).

Formally, we define exposure to robots for plant f, $ETR_{f,t}$, as:

$$ETR_{f,t} \equiv ETR_{i,t} \times secondary_{f,t_0} \tag{1}$$

where:

$$ETR_{i,t} \equiv \frac{R_{i,t}}{L_{i,t_0}} \tag{2}$$

In (2), $\frac{R_{i,t}}{L_{i,t_0}}$ is the number of industrial robots shipped to industry i (measured at 2-digit ISIC code) in year t, divided by the number of workers in industry i (in thousands). To minimize potential endogeneity concerns, as in Acemoglu and Restrepo (2017) we fix the denominator of (2) to the 2006 number of industry workers.[8]

2.2 Robots Data

Data on imports of industrial robots are obtained from the 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-digit industry and year.

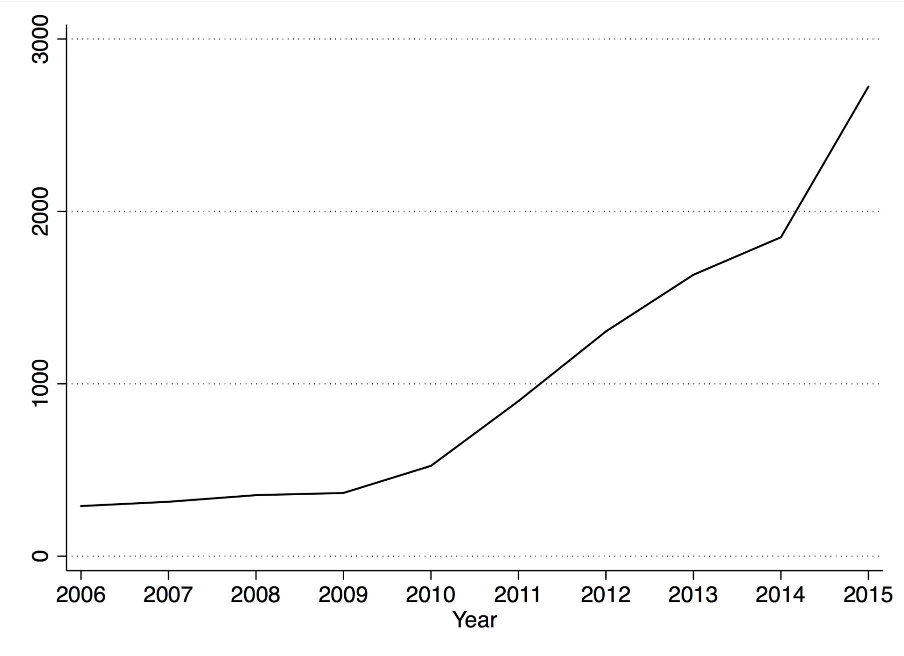
Figure 1 shows the total number of robots shipped to the Indonesian manufacturing sector, from 2006 to 2015. Total robots' counts from IFR are consistent with Comtrade. However, unlike IFR data, Comtrade does not provide an industry breakdown. Imports in Indonesia were roughly constant up to 2009, with approximately 350 units used every

[8]We use the 2006 as the base year to be consistent with the share of secondary education workers, which is available for the 2006 census year.

year. From 2009, the total number of industrial robots imported began to increase to reach almost 3000 units in 2015. Our analysis exploits this large jump in robots' adoption.

The aggregate figures hide a large amount of industry heterogeneity in robots' use. Figure 2 plots (2) by group of industries and shows that the number of robots per thousand workers used in manufacturing was substantially higher in Motor vehicles and Rubber and plastics industries. While Motor vehicles is by far the most automated industry worldwide, the high concentration of robots in Rubber and plastics is more peculiar to Indonesia. This is one of the largest manufacturing industries in Indonesia particularly for auto-parts, such as tires, which often employ state-of-the-art technologies of production.

Figure 1: Total number of robots used in the Indonesian manufacturing sector.

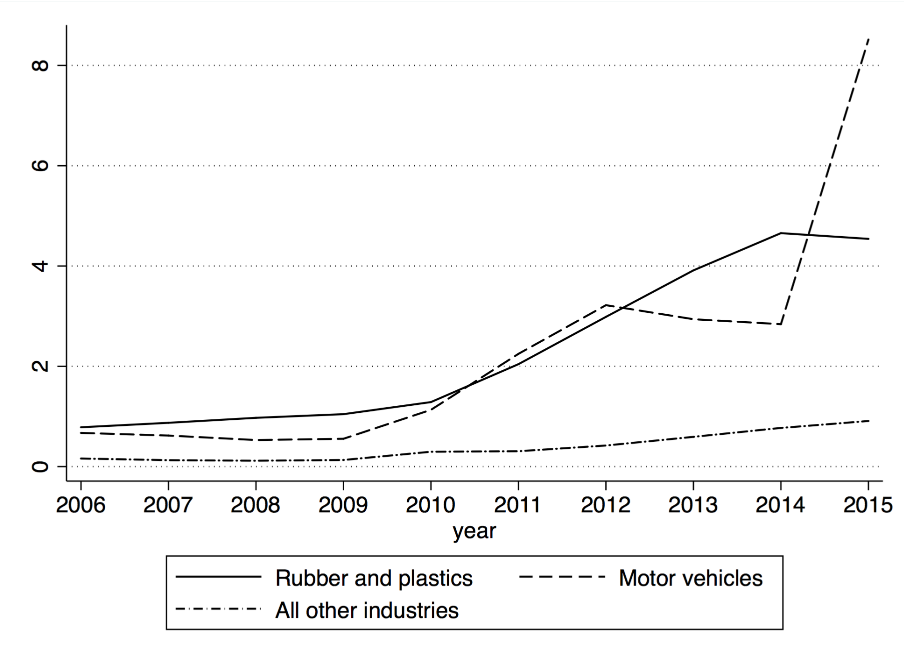


The figure shows the total number of industrial robots used in the manufacturing sector over the years of the sample. Source: IFR.

2.3 Plant Level Data

Plant level data are taken from the Indonesian survey of manufacturing plants with at least 20 employees (Statistik Industri, SI). The survey is administered by the Indonesian statistical office (BPS) and its coverage is extensive. In fact it becomes an actual census in 2006 and it is very close to a census in the remaining years, hence ensuring high representativeness even at very low levels of aggregation. Importantly, the 2006 Indonesian

Figure 2: Robots per thousand workers used in the Indonesian manufacturing sector, by industry.



The figure shows the number of industrial robots per thousand employees used in selected industries over the years of the sample. Source: IFR, SI.

census includes also plant level information on employment by educational attainments, which allows to construct the plant-based measures of exposure to robots as described above. In particular, the educational categories in our plant level data 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 define as "secondary" educational attainments falling into the categories iii) and iv).

In addition to plant level output, capital, labor and export, SI data provide information also on the quantity and value of 9-digit products produced and input used (domestically produced and imported) by each plant.[9]

We match SI and IFR data by constructing an industry crosswalk. Industry classifications in both SI and IFR data are roughly equivalent to ISIC Rev. 4. However, in some cases SI industries are more granular than IFR. Thus, we group together some SI industries to ensure maximum compatibility across the two datasets. Around 8.5 percent of the plants in our sample switch industry over the period of analysis. To avoid poten-

[9]In our sample, each plant produces on average 2 products and 25% of the plants produce more than one product. On average, each plant uses four different varieties of raw inputs.

tial endogeneity of industry choice to robots' adoption, we drop the observations after a plant switches industry. However, we obtain very similar results if we do not drop such observations, or adopt alternative strategies.[10]

We keep only plant-year cells with non-missing observations for all the dependent variables involved in the analysis. Our final plant level (unbalanced) panel includes 22,288 plants operating in thirteen 2-digit manufacturing industries between 2008 and 2015, for a total of 55,417 observations.[11]

Table 2: Summary statistics of the variables involved in the plant level analysis.

	(1) N	(2) mean	(3) sd	(4) min	(5) max
Industry number of robots (1000s of workers)	55,417	0.181	0.676	0	4.802
OECD-average industry number of robots (1000s of workers)	46,467	3.257	5.474	0.171	54.95
Innovation-intensity index	55,417	0.210	0.268	0	1
Share of secondary education workers	55,417	0.611	0.361	0	1
Share of primary education workers	55,417	0.375	0.368	0	1
Share of tertiary education workers	55,417	0.0135	0.0486	0	1
Principal component of: secondary, size, foreign capital, openness	55,413	-1.30e-09	1.298	-1.418	9.513
Downstream exposure to robots (1000s of workers)	42,773	0.108	0.363	0	4.234
Real investment in machinery and equipment (log)	55,417	1.659	3.467	0	19.03
TFPQ (log)	55,417	-7.015	15.94	-51.91	25.99
Real marginal cost (log)	55,417	2.245	1.732	-2.644	6.831
Real markup (log)	55,417	2.625	1.534	-0.792	7.168
Real value added (log)	55,417	12.58	1.828	5.925	22.85
Real revenue (log)	55,417	13.64	1.890	6.961	23.42
Exports (share of revenue)	55,417	0.118	0.295	0	1
Imports (share of revenue)	55,417	0.0383	0.143	0	1.009
Labor share in value added	55,417	0.615	2.616	0.000136	481.2
Gross surplus (share of renveue)	55,417	0.385	2.616	-480.2	1.000
Employment (log)	55,417	3.933	0.934	2.197	9.458
Real average wage (log)	55,417	7.588	1.107	0.131	16.43
Production employmeent (log)	55,417	3.735	0.952	0	8.821
Real average production wage (log)	55,417	7.533	1.139	0	16.34
Non-production employment (log)	55,417	1.609	1.428	0	8.827
Real average non-production wage (log)	55,417	6.253	3.402	-2.913	16.59
Real expenditure on domestic inputs (log)	55,417	12.68	2.114	2.772	22.59
Real expenditure on services (log)	55,417	1.563	3.588	0	19.70

[10] The first alternative strategy is assigning robots to plants based on the industry in which it operates the year of the first available information. The second strategy is dropping altogether all switching plants.

[11] In principle, we could run the plant level analysis starting in 2007. However, representative data by regency and industry, which are discussed in Section 4.3, are only available from 2008. Therefore, to be consistent across levels of analysis, we start from 2008.[12] Tables 2 presents summary statistics for the variables used in the analysis.

2.4 Empirical Specification

We use equation (1) for the plant level analysis using the following model:

$$Y_{f,i,t} = \gamma_0 + \gamma_1 ETR_{f,t} + \eta_f + u_{i,t} + I_f \times \theta_t + \epsilon_{f,t} \tag{3}$$

where $Y_{f,i,t}$ is an outcome of plant f in industry i at time t; η_f are plant fixed effects; $u_{i,t}$ are industry-year effects defined using the same classification as IFR robots' import data; I_f is a plant-specific innovation intensity index (measured in the 2006 census year); and θ_t are time effects.

The plant fixed effects should absorb the impact of unobserved time-invariant characteristics correlated with the available opportunities for automation i.e. plants' 2006 share of secondary education and the outcome variables. That may be the case for instance as foreign owned plants are usually more productive and have a larger share of routine labor.

The industry-year effects are based on the IFR industrial classification, thus capturing any industry-specific, time varying shocks, which may be related to both the outcome variables and the adoption of robots, such as changes in international trade patterns, demand or supply shocks. This mitigates the concern that exposure to robots in an industry may be endogenous to the outcomes of plants operating in that industry. At the same time it also absorbs the variation of $ETR_{i,t}$, which therefore cannot be estimated.

Estimates of γ_1 in (3) quantify the differential impact of robots' adoption within an industry in a given year across plants with different degrees of exposure to robots.

The inclusion of a rich set of time invariant and time-varying fixed effects in (3) should minimise the potential bias afflicting estimates of ETR. However, it might still be the case that the *interaction* of industry characteristics and their share of routine labor might be correlated to unobserved variables having an impact on plants' outcomes. In particular, one concern is that the share of routine labor might be systematically correlated with innovation and adoption of other non-automation technologies in industries with high exposure to robots. In that case, our estimates would be biased.[13] To mitigate such concerns, we exploit 2006 census year information on R&D units, product and process innovation, use of computers and the Internet to construct an index of "innovation intensity". Then we interact the index with year effects. To the extent that innovation activity and IT use are correlated to plants' propensity to adopt non-automation technologies, this approach should allow us to purge the estimates from their impact.

2.5 Is the Plant Level Approach Solid?

This section presents several pieces of evidence on the robustness of the key assumption underlying equation (3), i.e. that secondary education is a good proxy for routine task-intensive employment. If that is the case, plants with a larger share of workers with secondary education should be more likely to adopt robots.

First, we check whether the industries with an initial large share of secondary education workers have adopted relatively more robots in subsequent years. Figure 3 shows that this is indeed the case, thus providing some initial support to our approach. In particular, the two industries with the highest penetration of robots, Motor vehicles and Rubber and plastics, are among the industries with the largest share of secondary education workers.

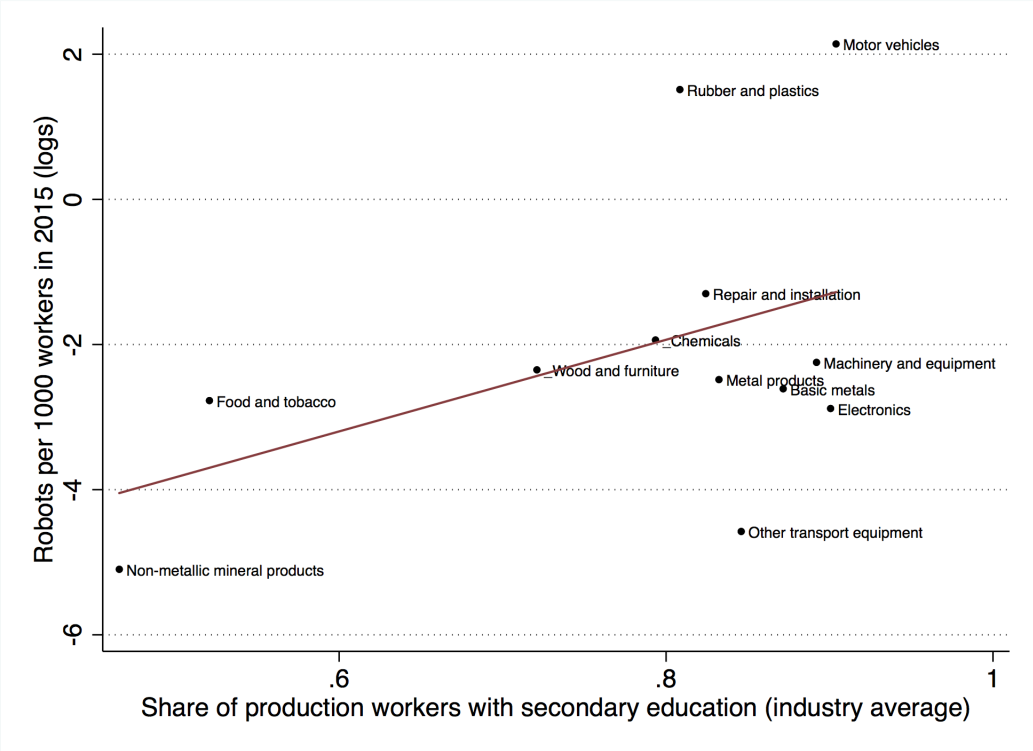
Next, we exploit the fact that in our manufacturing data plant level investment in machinery and equipment should also include investment in robots. Consistently with this assumption, Figure 4 shows that imports of industrial robots shipped to each manufacturing industry in a year are positively related to total investment in machinery and equipment in the same industry-year pair.[14]

Having established a positive link between imports of robots and aggregate investment in the data, an important check of the validity of our plant level proxy for exposure to automation is whether $ETR_{f,t}$ in (3) is positively correlated with machinery and equipment investments at the plant level. This is done in column (1) of Table 3, which reports the results of regressing log-real investment in machinery and equipment on $ETR_{f,t}$ con-

[13] For instance, logistics management software might be particularly valuable in the motor vehicle industry, which is also a strong adopter of robots. If large plants are more likely to have a large share of secondary education labor, the estimates of γ_1 in (3) might capture the impact of software.

[14] The figure excludes industry-year cells with zero imports.

Figure 3: Correlation between the 2006 shares of workers with secondary education and robots in use in 2015, by industry.



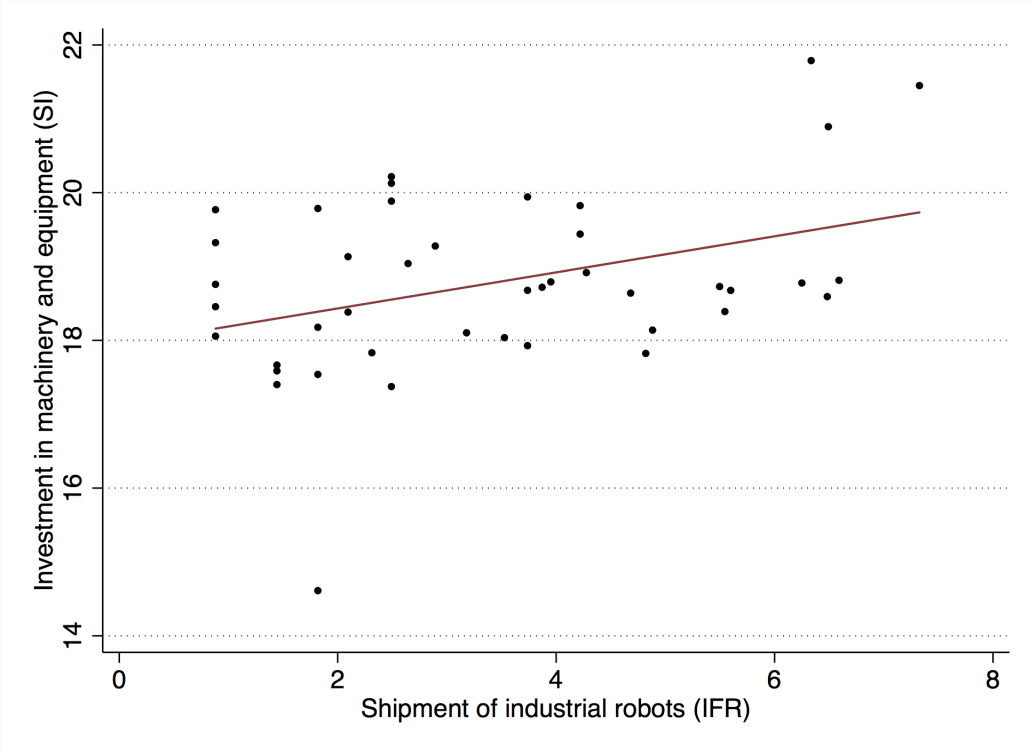
On the horizontal axis there is the industry-average share of workers with secondary education, computed from plant level data in 2006. On the vertical axis there is, for each industry, the log number of robots per thousand employees in 2015. Sources: IFR, SI.

trolling for plant and industry-year fixed effects.[15] Reassuringly, the coefficient is positive and statistically significant. It suggests that one additional robot per thousand workers in a given industry is associated with a nine percent increase in machinery and equipment investment by plants one standard deviation above the average share of secondary education workers in that industry.

As a placebo test, Table A1 presents results of interacting industry exposure with the share of plants' employment with at most primary education (column 1). The coefficient in column (1) is negative and significant, which suggests that robots adoption is discouraged by a large share of workers with low educational attainments. This might reflect the fact that relatively uneducated workers typically perform non-routine manual tasks, which are hard to automate. In column (2), we interact industry exposure to robots with plants' share of workers with tertiary education (Diploma or above, see Section B for de-

[15]Given the lumpiness of investment at the plant level, for the dependent variable we employ a log-transformation (see Section C). Qualitatively identical results are obtained with a dummy variable taking value 1 if, in a given year, a plant has positive investment.

Figure 4: Correlation between IFR data on robots and SI data on investment, by industry and year.



The figure shows the correlation between industry-level IFR data on robots and plant level investment in machinery and equipment. On the horizontal axis there is the number of industrial robots shipped to each manufacturing industry and year in Indonesia (dropping observations with zero value). On the vertical axis there is the industry-year-average of plant level investment in machinery and equipment. Sources: IFR, SI.

tails on education data). The coefficient is not statistically significant, consistently with the observation that workers with tertiary education usually perform cognitive and non-routine problem-solving and complex communications tasks (Autor et al. (2003)). These placebo tests provide further support to the assumption that secondary education is a reasonable proxy for routine task-intensity in the contest of Indonesian manufacturing.

2.6 Alternative Measures of Exposure to Automation

While using workers' education to identify exposure to automation is consistent with the consensus in the literature, there are also other potential ways to define plant level exposure. For instance, Acemoglu et al. (2020) and Koch et al. (2019) provide evidence from advanced economies that firms that are large, have foreign capital, and export and import are more likely to adopt robots. We use this evidence to construct an alternative measure of exposure to robots. We employ a principal component analysis to identify the relative importance of these factors to predict automation in Indonesia. Specifically, we focus on the 2006 values of the following variables: i) firm (employment) size; ii) a dummy for having more than fifty percent foreign capital; iii) the sum of revenue shares of import and export, and iv) the share of secondary education workers.[16] Column (2) of Table 3 presents the results of replacing the share of secondary education workers with this variable.[17] The positive correlation of this alternative exposure measure with machinery investments is consistent with the existing evidence for advanced economies and adds to the thin evidence on the drivers of technology adoption in developing countries (Verhoogen, 2020). In light of this result, we can use this measure in the next section to check the robustness of the plant level results obtained with our preferred exposure measure.[18]

Table 3: Correlations between plant level exposure to robots and plant level investment.

	(1) Investment	(2) Investment
ETR × secondary	0.089***	
	(0.011)	
ETR × PC (secondary, size, foreign capital, openness)		0.074**
		(0.021)
Observations	53,447	53,447
R-squared	0.658	0.658
Plant FE	yes	yes
Year FE	yes	yes
Industry-year FE	yes	yes
Other technologies	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots and investment. The dependent variable is the log of plant level investment in machinery and equipment. Secondary is the 2006 share of plants' employment with secondary education, normalised to have zero mean and unitary standard deviation. The PC term is the first principal component of the 2006 values of the following variables: i) firm (employment) size; ii) a dummy for having more than fifty percent foreign capital; iii) the sum of revenue shares of import and export, and iv) the share of secondary education workers. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

[16] The principal component analysis reveals that only the first component has an eigenvalue larger than 1, equal to 1.69. The first component explains 42 percent of the variation in the four variables.

[17] As for the share of secondary education workers, we normalize this variable such that it has zero mean and unitary standard deviation in the sample.

[18] Including the share of secondary education workers in the principal component analysis increases substantially the predictive power of the component, which in turn provides additional support for our preferred specification.

2.7 Estimation of Plant Level Productivity, Marginal Costs and Markup

To measure plants' productivity, marginal cost and productivity, we estimate production functions using the panel data of manufacturing plants (Statistik Industri - SI) from 2008 to 2015. The details of the estimation procedure can be found in Section F of the Appendix. Unlike most empirical work using productivity estimates, we observe plant level physical output and input quantities and values. To deflate the total revenue of multi-product plants, capital and intermediate inputs expenditure (including energy consumption), we construct plant-specific output and input price indexes, as in Eslava et al. (2004). This is an advantage over contributions employing industry-level price deflators, which essentially assume that all plants within an industry face similar inputs' cost and charge the same price.[19] Our approach allows us to obtain a measure of quantity-total factor productivity (TFPQ). Exploiting quantity information, the latter reflects changes in pure technical efficiency, rather than in revenues.

A key challenge in the measurement and identification of productivity relates to the endogeneity of the firm's optimal choice of inputs. We follow Ackerberg et al. (2015) and adopt a control function approach with the objective of proxying productivity known by plants' managers, but unobserved by us, with energy expenditures.[20] Following De Loecker and Warzynski (2012), we adapt the methodology of Ackerberg et al. (2015) and allow plants' exposure to robots to affect the expected value of future productivity, which in turn might determine differences across exposed and non-exposed plants and affect the estimation results.[21] After estimating the the parameters of the production functions industry by industry, we follow De Loecker and Warzynski (2012) and obtain a measure of plants' marginal cost and markup, based on the plants' first order conditions

[19] Foster et al. (2008) discuss the bias arising when using plant revenue deflated by industry deflators. De Loecker et al. (2016) extend the analysis in the contest of unobserved variation in input prices.

[20] We use energy rather than total intermediate inputs to edge against the potential input adjustment costs, which would be inconsistent with the assumptions on which the estimator is built. A detailed discussion is presented in Section F.

[21] Similarly, De Loecker et al. (2016) study the impact of trade reforms and include export dummies and import tariffs; De Loecker (2007) includes export quotas; Doraszelski and Jaumandreu (2013) include R&D expenditure, and Konings and Vanormelingen (2015) include measures of workforce training.

and flexible inputs' choice.

3 Plant Level Impact of Automation

All results presented in this Section are obtained by estimating equation (3) with OLS. Since $ETR_{i,t}$ in (1) varies at the highest level of aggregation, i.e. 2-digit industry and year, unless differently stated we cluster standard errors accordingly.[22] As argued above, we interpret the relationship between automation and plant level outcome variables as causal, conditional on the full set of fixed effects employed.

3.1 Productivity

We start by examining the impact of automation on plant productivity, markup, value added and revenue. Table 4 presents the results.[23] The coefficients in columns (1) and (2) suggest that robots' adoption is associated to higher TFPQ and lower marginal costs. This reflects higher technical efficiency as a result of the adoption of robots. This is in line with the firm-level evidence in Acemoglu et al. (2020), who find a positive productivity impact of robots. It is also consistent with the productivity effect predicted by task-based models (Acemoglu and Restrepo (2018b))

Column (3) shows that robots have a positive and significant impact on markup as well. The result is consistent with automation of production enabling higher product quality, for example by increasing the precision of the assembly of parts and components. Consistently with these impacts, robots' adoption also increases the plant's real value added (column 4) and revenues (Column 5) as well. In particular, the coefficient suggests that 1 additional robot per thousand workers in a plant with a mean share of secondary educated workers raises the plant's revenues by 2.5 percent.

[22] We also experiment with clustering at the plant level, and we show that the results are almost identical.

[23] In Table 4 and all other tables of this section, the number of observations is slightly lower than the total available of 55,417. This is due to the presence of singletons, i.e. sample units observed only once. Due to the structure of the panel, such observations are dropped during the estimation.

Table 4: Plant level productivity, markup, value added and revenue.

	(1) TFPQ	(2) Marginal cost	(3) Markup	(4) Value added	(5) Revenue
ETR × secondary	0.029***	-0.086***	0.051***	0.028***	0.041***
	(0.005)	(0.020)	(0.010)	(0.006)	(0.004)
Observations	53,517	53,517	53,517	53,517	53,517
R-squared	0.999	0.747	0.744	0.885	0.914
Plant FE	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes
Other technologies	yes	yes	yes	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots, productivity, markup, value added and revenue. The dependent variables are TFPQ (1) the log of plant level real marginal cost (2), markup (3) real value added (4) and real revenue (5). Secondary is the 2006 share of plants' employment with secondary education. Other technologies are capture by an index of innovation activities in 2006, interacted with year fixed effects. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

3.2 Employment and Wages

Having provided evidence of a significant productivity effect of robots, a key question in evaluating the economic impact of automation is whether the improved performance translates into increased labor demand at the plant level. In the terminology of task-based models, the question is whether the productivity effect is large enough to offset the displacement effect of automation (Acemoglu and Restrepo (2018b); Acemoglu and Restrepo (2019)). Table 5 presents the results of estimating equation (3) using the log of total employment and the log of real average wages as dependent variables. The impact of robots on total plant level employment is negative and small in absolute magnitude (column 1). An additional robot per thousand workers in a plant with a mean share of secondary educated workers reduces the plant's employment by 0.4 percent. While significant this effect is particularly small considering that the sample mean of robots is 0.18 per thousand. On the other hand automation does not have any significant effect on wages (column 2).

As robots are mainly used in production, we also test separately for the impact on employment of workers engaged in production versus other workers. Columns 1 and 2 in Table 6 present the results for *production* workers, and columns 3 and 4 for *non-production* workers. Exposure to robots is associated with lower production employment, but higher non-production employment, which includes a higher share of white-collar occupations such as accountants, sales representatives and analysts.[24] The asymmetric impact of automation is in line with the hypothesis that industrial robots substitute for workers performing production occupations, but complement some non-production occupations. For example, an industrial robot performing production tasks may send more precise, real-time information on the production process that would improve the performance of non-production workers, such as process analysts.

Table 5: plant level impact of robots on employment and wages.

	(1) Employment	(2) Wages
ETR × secondary	-0.006**	-0.018
	(0.002)	(0.012)
Observations	53,517	53,517
R-squared	0.939	0.653
Plant FE	yes	yes
Industry-year FE	yes	yes
Other technologies	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots and employment. The dependent variables are the log of plant level employment and average wages. Secondary is the 2006 share of plants' employment with secondary education. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table 6: Plant level impact of robots on production and non-production employment and wages

	(1) Production employment	(2) Production wages	(3) Non-production employment	(4) Non-production wages
ETR × secondary	-0.023**	0.008	0.050***	0.060
	(0.009)	(0.012)	(0.009)	(0.051)
Observations	53,517	53,517	53,517	53,517
R-squared	0.904	0.596	0.829	0.657
Plant FE	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes
Other technologies	yes	yes	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots and employment by type (columns 1-4). The dependent variable is the log of plant level (total, production and non-production) employment. Secondary is the 2006 share of plants' employment with secondary education. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

[24] Indeed non production workers have higher education than production workers, with 12 years of education on average versus 10 years for production workers.

3.2.1 Backward Linkages

The negative plant level impact of robots on employment contrasts with the results of other studies in high-income countries, which find a positive employment impact at the firm-level (Acemoglu et al. (2020); Koch et al. (2019)). Our result could be reconciled with a positive employment effect if non adopting plants benefited from significant positive spillover from adopting plants.[25]

As we do not observe robots' utilization at the plant level, we cannot test directly for the effects of automating on non-automating plants. Instead, we try to test for spillover effects of automation across plants by measuring inter-industry backward linkages. The idea is to examine whether improved performance of automating industries may translate into increased demand for upstream plants.

We follow two approaches. First, we test for the impact of robots' adoption on plant level demand for domestic inputs. To do so we employ equation (3) using the value of domestic inputs purchased by each plant as dependent variable. The result - reported in column (1) of Table 7 - suggests a positive and significant impact: an increase in 1 robot per thousand worker in a plant with the mean share of secondary educated workers raises the plant's purchase value of domestic inputs by 4 percent.

Second, we construct inter-industry linkages by exploiting information on inputs purchased by each plant at the 9-digit level (see Appendix C.4). We first assign the 9-digit codes to the corresponding 2-digit industry. Then, for each 5-digit buying industry, we compute the base year share of expenditure on two-digit selling industries.[26] The result is a detailed five-by-two-digit input-output table based on disaggregated information. We use the table to compute downstream exposure to robots in each five-digit industry.

Let $\sigma_{l,i}$ be the share of inputs bought by five-digit industry l from industry i. Then,

[25]In fact Acemoglu et al. (2020) and Koch et al. (2019) find evidence of negative spillovers from automation on non adopters through competition effects. This result suggests that the positive employment coefficient at the firm level captures a relative rather than an absolute effect and it helps reconcile the positive employment effect at the firm-level with the negative effect in the local labor markets in the literature.

[26]This procedure leads to the loss of observations as reliable input data is not available for all plants in all years.

we calculate downstream exposure to robots as:

$$ETR^d_{l,t} = \sum_i \sigma_{l,i} ETR_{i,t} \tag{4}$$

where $ETR_{i,t}$ is defined as in (1).

We use $ETR^d_{l,t}$ as an additional regressor in (3) to measure the impact of automation on plants via backward linkages. As we are interested in the spillovers on non-adopting plants, we also add an interaction of $ETR^d_{l,t}$ with the plant level exposure measure. The results in column 2 of Table 7 suggest that automation has a positive, albeit not significant, effect on employment in upstream non automating plants, while the effect on automating plants is precisely zero. The larger employment effect of backward linkages on non adopters is consistent with the idea that automating plants upstream would absorb the increased demand for inputs at least in part through robots as opposed to workers. As a result of the inclusion of these backward linkage variables, the coefficient of exposure to automation turns positive and significant. This result provides suggestive evidence that positive spillovers from automation on upstream non automating plants may drive the negative impact of automation on employment in 5. This is a plausible channel also considering that the bulk of manufacturing inputs are used by plants in the same 2-digit industries as the input-producing plants. Hence most of these beneficiary upstream plants are part of the relevant control group in (3).

As the downstream linkages are computed at the 5-digit industry-level, 2-digit industry-year fixed effects are not sufficient to address the endogeneity concerns as it was the case in the previous tables. Therefore, in column (3) of Table 7 we present 2SLS estimates instrumenting both $ETR^d_{l,t}$ and its interaction term through a variable based on robots' imports by OECD countries rather than Indonesia. As the instrument is used extensively in the aggregate analysis, we postpone its description to Section 4.1. Column (3) reports the 2SLS estimates which show a positive and significant effect of automation on employment in upstream plants with a coefficient almost double the size than that in column (2). This result is consistent with the positive impact of ETR on demand of local inputs (column 1). Again, non-automating plants experience a larger employment effect than automating plants from the increased demand for inputs, although in this case the backward linkage effect on employment is positive also for automating plants. The coefficient of exposure to automation is again positive and significant.

As the sample in columns (2) and (3) is smaller than that in table 5 (given the availability of plant level input data), to check that this is not driving the change in the ETR coefficient, in column (4) we replicate the result of table 5 (column 1) but over the smaller sample. The coefficient of exposure to automation is negative and significant and even larger in absolute term than in table 5, thus lending further support to the hypothesis that the negative ETR coefficient on employment is driven by the positive vertical linkages. The results are robust also to using directly the exogenous instruments in an OLS regression so as to avoid having to instrument two endogenous variables (column 5).

The improved performance of plants induced by automation does not only result in increased demand for goods, but also of services. This is documented in column (6) which uses services' expenditure as dependent variable. The positive and significant coefficient suggests that plants adopting robots spend additional resources on professional services, which include R&D and marketing consulting.

Table 7: Plant level impact of robots on domestic inputs' expenditures and employment: vertical linkages.

VARIABLES	(1) Domestic inputs	(2) Employment	(3) Employment	(4) Employment	(5) Employment	(6) Services
ETR × secondary	0.064***	0.028**	0.018**	-0.014**	-0.003	0.064*
	(0.010)	(0.008)	(0.006)	(0.005)	(0.004)	(0.033)
Downstream ETR		0.020	0.036**			
		(0.017)	(0.012)			
Downstream ETR × ETR × secondary		-0.018***	-0.014***			
		(0.001)	(0.003)			
Downstream ETR (oecd)					0.002**	
					(0.000)	
Downstream ETR (oecd) × ETR × secondary					-0.001***	
					(0.000)	
Observations	53,517	40,958	40,958	40,958	40,958	53,517
R-squared	0.867	0.940		0.940	0.940	0.710
Plant FE	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes
Other technologies	yes	yes	yes	yes	yes	yes
First stage F-stat			13.78			

The table presents OLS and 2SLS estimates of the relationship between plants' exposure to robots / downstream exposure, log real domestic inputs' expenditures and employment. For the 2SLS estimates, exposure to robots is instrumented with the average exposure in the OECD region. Secondary is the 2006 share of plants' employment with secondary education. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

3.3 International Trade

Next, we examine whether robots are associated to stronger engagement in GVC by looking at the shares of exported revenue and imported input expenditure. Whether robots translates into greater competitiveness in international markets is important given the relevance of export markets to fuel the sustained growth of manufacturing sectors.

The results in Table 8 suggest that the increase in turnover spurred by automation is matched by a more than proportionate increase in plants' exports so that robots' adoption raises the export share of plants' sales (column 1). This is consistent with the idea that robots' adoption improves productivity and product quality enabling plants to compete more effectively in international markets.

The increase in the export share is also accompanied by a similar increase in the import share of inputs induced by robots' adoption (column 2). This is consistent with the higher proportion of imported inputs used for exports than for domestic sale (Amiti and Konings, 2007). These results suggest that robots' adoption in a developing country increases the integration of plants into GVCs, a finding that to the best of our knowledge has not been documented before.

Table 8: Plant level engagement in international trade.

	(1) Export share	(2) Import share
ETR × secondary	0.004***	0.004***
	(0.000)	(0.001)
Observations	53,517	53,517
R-squared	0.852	0.792
Plant FE	yes	yes
Industry-year FE	yes	yes
Other technologies	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots, the share of exported output and the share of imported inputs. Secondary is the 2006 share of plants' employment with secondary education. Other technologies are capture by an index of innovation activities in 2006, interacted with year fixed effects. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

3.4 Robustness

In this section we test the robustness of the plant level results to changing the key identification assumption in (1). In particular we replace the share of secondary education workers with the principal component of key plant level observables described in Section 2.5. The results are presented in Table A2 in the Appendix and suggest that both the performance results (columns 1-2) and those on GVC participation (columns 4-5) are largely robust to the alternative specification. On the other hand, the impact on revenues is not significant. More importantly, the results on labor impact are not in line with Table 5. The employment effect is not distinguishable from zero (column 6) while the wage impact becomes negative and significant (column 7).

In a second set of robustness test, we go back to the baseline specification, but we cluster errors at the plant level instead of the 2-digit industry-year level. This might be important because our baseline specification has only 91 clusters. The results are presented in Table A3 in the Appendix. Similarly to the previous table, the results on both performance (columns 1-2) and GVC integration (columns 3-5) are robust to this test. However the effects on employment and wage are not significant (columns 6-7).

Taken together these tests provide confidence on the robustness of the positive impact of automation on plants' performance and on their integration with international markets. At the same time they cast doubts on the negative employment effects estimated in table 5, reinforcing the case made through the results in Table 7 that the true impact of automation on labor demand is in fact not negative at the plant level.

4 Aggregate Impact of Automation

Next we extend the plant level analysis to the aggregate level, both at the industry and local labor market level.

In the plant level analysis, the inclusion of industry-year fixed effects in (3) mitigates the concern that industry-level adoption might be correlated with other industry trends affecting the outcome variables. However, in both the industry-level and local labor market-level analyses presented in this section, including industry-year fixed effects is not possible.[27] Hence, to address the endogeneity concerns, we implement an instrumental

[27]In the industry-level analysis, including industry-year fixed effects would absorb all the variation of the explanatory variables. In the local labor market analysis, we aggregate industries at the local labor market level, which makes it impossible to control for industry-year fixed effects.

variable approach similar to Acemoglu and Restrepo (2017) and Dauth et al. (2017).

4.1 Instrumental Variable

This approach instruments $ETR_{i,t}$ with robots' adoption in the same industry in other countries. The idea underlying the instrument is that the use of industrial robots is induced by improvements in technology (or a reduction of their price) increasing their profitability for adopters. As such, these are trends that are largely unrelated to the specific market conditions prevailing in Indonesian industries. Similarly to Acemoglu and Restrepo (2017), we focus on countries that are ahead Indonesia in terms of robots' adoption. To that end we use data from OECD countries, and we first match IFR data with 2-digit industry employment figures from the OECD Structural Analysis Database (STAN).[28]

Then for each 2-digit industry-year pairs we compute the number of imported robots per thousand workers averaged across OECD countries. To construct the instruments at various levels of aggregation, we simply replace the density of robot imports from Indonesia with this OECD average for each industry-year pair.

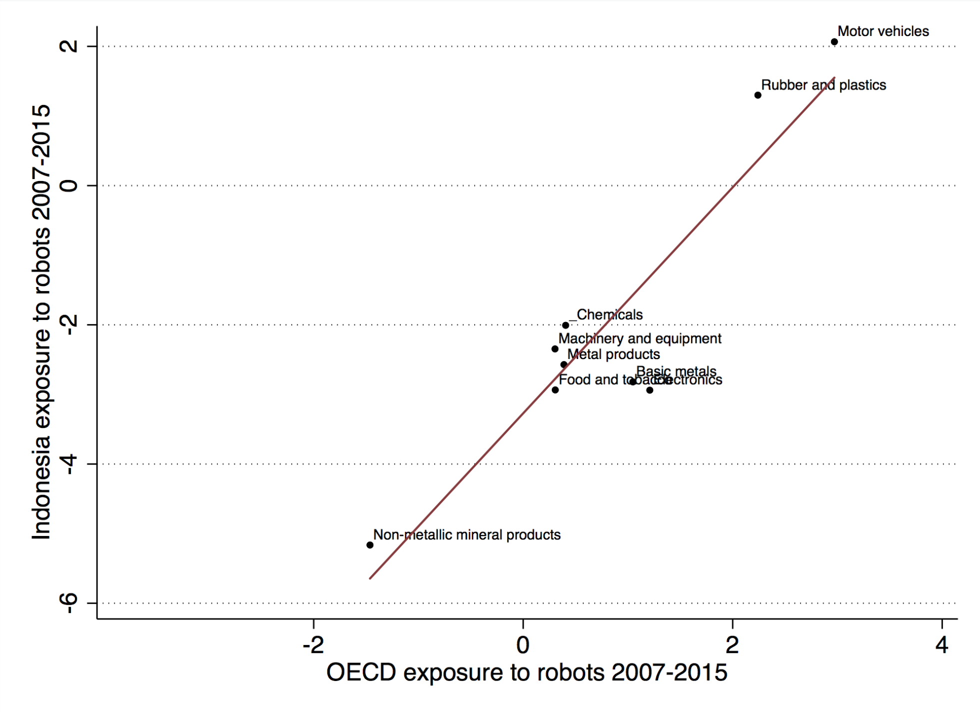
If technological trends drove robots adoption, we would expect that the industries with higher exposure of robots should be broadly the same across countries, even for countries with different levels of economic development. The strong correlation between 2007-2015 changes in $ETR_{i,t}$ between Indonesia and OECD countries is consistent with the notion that Indonesia's automation across industries is driven by technological factors (Figure 5).[29]

An endogeneity concern relates to the possible relation between robots' adoption,

[28] We are not able to construct employment for Wood and furniture, and Installation and repairs industries. Those are available in IFR data but not in STAN, which does not report a sufficiently disaggregated breakdown of employment for these 2 industries. As a result, the number of available 2-digit industries drops to twelve.

[29] As the change for Motor vehicles and Plastic and rubbers is much larger than for the other industries, the figure uses a log-scale to ease readability. The log-scale allows to include only nine of the twelve available industries as Textile and Paper did not experience any robots' adoption in Indonesia and the average change in robots per thousand workers was negative in Other transport equipment for the OECD. Figure A1 in the Appendix shows that the positive relationship holds also when using a normal scale, which allows to include Textile, Paper and Other transport equipment while it excludes Motor vehicles and Plastic and rubbers.

Figure 5: Correlation between Indonesian and OECD-average exposure to robots (log-scale).



On the horizontal axis there is the change between 2007 and 2015, of the OECD region industry-average number of robots per thousand employees. On the vertical axis, there is the change between 2007 and 2015, of the industry-level number of robots per thousand employees in Indonesia. Sources: IFR, STAN, SI.

demand shocks and labor market outcomes. For example increased demand in an industry may spur both the need to adopt more robots to improve production processes and an increase in labor demand. This type of relation would make the instrument invalid as it would not be exogenous. To mitigate this issue, we control for global demand by industry. To conform this control to our regency-based estimation, we aggregate global exports by industry (excluding Indonesian exports) to the regency-level using industries shares in each regency's total employment in the base year as weights.

4.2 Industry-level Analysis

Table 9 examines the impact of automation on an Herfindahl–Hirschman (HH) Index of concentration based on plants' sales, and the number of plants entering and exiting the markets each year, respectively. As explained we instrument Indonesia's imports of robots with OECD imports to mitigate endogeneity concerns.

The coefficient of the HH Index is positive and of similar size in the unweighted and weighted specifications of columns (1) and (4), but it is only significant in the latter. The estimates for entry and exit are always statistically significant and imply that within an industry, ten additional robot per thousand workers results in entry of 17 new plants and the exit of 41 incumbents. Therefore, the number of plants operating in an industry drops by 24 units, which is consistent with increased concentration.

In column 4-6 we weigh the estimates by the number of workers in each industry in the base year. Weighted estimates suggest that the increased concentration due to robots is stronger in the largest industries.[30] Thus, robots' adoption significantly increases business dynamism on both margins, and a higher number of exiting plants implies an increase in industry concentration, which could also be responsible for the positive impact on markup presented in Table 4.

We turn next to assess the industry-level implications for factor reallocation and aggregate productivity (Π). Based on the production function estimates discussed in 2.7. we compute quantity-total factor productivity (ω) and apply the standard decomposition of aggregate productivity of Olley and Pakes (1992):

$$\Pi_{it} = \bar{\Pi}_{it} + \sum_{f \in i} (s_{ft} - \bar{s}_{it})(\omega_{ft} - \bar{\omega}_{it}) \equiv \bar{\Pi}_{it} + COV(s_{ft}, \omega_{ft}) \tag{5}$$

where s_{ft} is the output share of plant f (in industry i) and upper bars denote industry averages. The last term in (5) is the covariance between plants' productivity and their output shares. Higher values of $COV(s_{ft}, \omega_{ft})$ implies a better allocation of resources, as the most productive plants make up a large share of the market.

Table 10 presents the results for aggregate productivity, average productivity and the covariance term. Unweighted estimates (columns 1-3) show a positive impact of robots on industries' aggregate productivity. Weighted estimates (columns 4-6) suggest a positive impact of robots on factor reallocation, as more productive plants tend to disproportionately benefit from automation in terms of output shares.

While not conclusive these results suggest a possible positive impact of automation also on the aggregate dynamism and performance of the industry.

[30]Using entry and exit rates delivers similar results, but less precisely estimated.

Table 9: Industry-level results: entry and exit.

	(1) HH index	(2) Entry	(3) Exit	(4) HH index	(5) Entry	(6) Exit
ETR	0.024	1.674*	4.089*	0.024*	4.838*	12.960*
	(0.015)	(0.848)	(2.412)	(0.014)	(2.434)	(6.532)
Observations	108	108	108	108	108	108
Weights	no	no	no	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Industry demand shifter	yes	yes	yes	yes	yes	yes
First stage F-stat	9.841	9.841	9.841	16.98	16.98	16.98

The table presents 2SLS estimates of the relationship between industry exposure to robots and business dynamism. Exposure to robots is instrumented with the average exposure in the OECD region. The dependent variables are, respectively, the number of entrants and exiting plants in an industry. Standard errors are clustered at the industry-level. Weights are constructed using 2006 (base year) industry employment. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table 10: Industry-level results: aggregate productivity and factor reallocation.

	(1) Aggregate	(2) Average	(3) Covariance	(4) Aggregate	(5) Average	(6) Covariance
ETR	0.033**	0.018	0.015	0.032	-0.021	0.053*
	(0.015)	(0.011)	(0.010)	(0.027)	(0.026)	(0.031)
Observations	108	108	108	108	108	108
Weights	no	no	no	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Industry demand shifter	yes	yes	yes	yes	yes	yes
First stage F-stat	9.841	9.841	9.841	16.98	16.98	16.98

The table presents 2SLS estimates of the relationship between industry exposure to robots, aggregate productivity and its decomposition according to the formula $\Pi_{it} = \bar{\Pi}_{it} + \sum_{f \in i}(s_{ft} - \bar{s}_{it})(\omega_{ft} - \bar{\omega}_{it}) \equiv \bar{\Pi}_{it} + COV(s_{ft}, \omega_{ft})$. Exposure to robots is instrumented with the average exposure in the OECD region. Standard errors are clustered at the industry-level. Weights are constructed using 2006 industry employment. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

4.3 Local Labor Market Analysis

In order to examine the local labor market impact of automation we use Indonesian regencies as our key unit of observation. That represents the second level of sub-national administrative divisions (the first being the province). A number of features make it a reasonable proxy for local labor markets in Indonesia. First, the mobility of labor is limited across regencies. In 2010 for example only 5 percent of the workforce worked in a different regency than their residence.[31] Second, regencies hold significant administrative

[31] Indonesia's island geography and often underdeveloped transportation infrastructures make the hypothesis of limited mobility likely to hold.

powers following the 1999 decentralization reform in Indonesia. Those include also the minimum wage setting. We use the pre-decentralization reform division in 292 regencies to ensure the consistency of the analysis over time.

The identification strategy exploits the variation in occupational composition across regency-industry pairs.[32] Specifically, we compute regency r share of employment in industry i for occupations at high risk of automation:

$$s_{i,r} = \frac{L^{H}_{r,i,t_0}}{L_{r,i,t_0}} \tag{6}$$

For the baseline estimates of this section, we identify the occupations at high risk of automation following GM.[33] However, since our baseline plant level estimates are based on information on educational attainments, we experiment as well with shares of workers with secondary education, which deliver similar results.

To compute regency-level exposure to robots, $ETR_{r,t}$, we use the shares in (6) to aggregate industry exposure in each regency:

$$ETR_{r,t} = \sum_{i \in r} s_{i,r} ETR_{i,t} \tag{7}$$

One concern with using $ETR_{r,t}$ as a regressor is its likely endogeneity with respect to regency-level outcomes. For instance, a positive demand shock in industry i would result in higher employment in regencies where those industries concentrate, i.e. a large share computed as in (6). If investment in robots responds to the shock, for instance to keep up with higher demand, the estimated impact of $ETR_{r,t}$ will be upwardly biased. To address this concern, we use the instrumental variable presented in Section 4.1 combined with the regency-level demand shifter.

We estimate the following system of equations in first difference:

$$\begin{cases} \Delta ETR_r = b_0 + b_1 \Delta ETR^{oecd}_r + \Delta D_r + BX_r + u_p + \epsilon_r \\ \Delta Y_r = \beta_0 + \beta_1 \Delta ETR_r + \Delta D_r + BX_r + u_p + \varepsilon_r \end{cases} \tag{8}$$

[32] Unlike for the plant level analysis, in the labor force survey data we observe occupations.

[33] We chose to follow GM because their list is derived by information on what robots actually do. However, we obtain similar results if we use the list compiled by FO.

In (8), the operator Δ represents the difference of a variable between 2015 and 2008. ΔETR_r^{oecd} is the change in instrumental variable obtained by substituting OECD average robots' exposure in (7). ΔY_r is the change of an outcome variable in regency d. The change in regency-level demand shifter is denoted by ΔD_r. The vector X_r includes a series of regency-level characteristics in the 2007 base year, which may affect the outcome variables including: i) population; ii) shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the shares of manufacturing employment. Finally, equation (8) includes province fixed effects u_p to account for differences in local economic development.

This approach is in line with Acemoglu and Restrepo (2017), Dauth et al. (2017) and Giuntella and Wang (2019), which use long differences to capture aggregate long-run trends in local labor markets. That is different from the plant level regressions that instead exploit yearly variation in robots' import. We argue that the yearly analysis is appropriate for two reasons. First, running the plant level specifications in long difference would reduce the sample size by almost two-thirds relatively to the yearly regression model as it would only focus on surviving plants throughout the entire period.[34] Second, dropping plants that were in the sample in 2006 but exited before 2015, might introduce selection bias.

In the baseline specifications, standard errors are robust to arbitrary heteroskedasticity within each regency. We also experiment with two alternative specifications. First, we calculate shift-share standard errors following the methodology of Adao et al. (2019). They show that shift-share designs, such as those used in this part of the analysis, might lead to residuals that are correlated across regions with similar industry shares and so over-rejection of the null hypothesis of no impact. In the second alternative specification, we provide errors clustered at the province-level.

Table A4 presents the first stage regression of regency-level exposure on the instrument based on the OECD region. In column (1), we compute the share of employment at high risk of automation based on GM. The coefficient of OECD exposure is significant at the

[34] Labor force survey data are not subject to this problem, which allows us to estimate models in long differences.

99 percent level and the R-square of the regressions is 0.67. In column (2), we compute the share of employment at risk of automation based on educational attainments. Also in this case, the coefficient is positive and significant at the 1 percent level and the R-square is 0.65.

The main data source for the labor market analysis is the Indonesia's labor force survey (Survei Tenaga Kerja Nasional - Sakernas), which is published by Indonesia National Statistics Bureau (BPS). The Sakernas is a cross-sectional dataset with wide national representation undertaken twice a year. The August waves of the survey since 2008 are representative at the regency level, which allows us to construct regency-level labor market measures.[35] Besides information on wages, employment status, sector and work location, the survey also includes information on the occupation of the worker identified according to the Indonesian classification (Klasifikasi Baku Jenis Pekerjaan Indonesia - KBJI). This is compatible with the International Standard Classification of Occupations (ISCO), allowing us to construct exposure to robots measures at the regency level as explained above. We focus on wage employees as those are the workers more directly affected by firms' adoption of robots. This focus is also relevant from a policy perspective, as wage employment typically provides a more reliable and higher incomes than self-employment in a developing country like Indonesia.

Summary statistics for the labor market data are presented in Table 11. To facilitate the interpretation of the results, we normalise the change of regency ETR to have zero mean and unitary standard deviation in the sample.

Employment increased substantially during the period with a 12 percent mean increase across regencies. Employment growth has been particularly high in services, construction and utility sectors, which were the sectors where most jobs were created as a result of the commodity-driven growth of the 2000s in Indonesia (World Bank (2015)). At the same time, wages doubled in the average regency, partly driven by minimum wage growth, particularly since 2012.

Table 12 presents the baseline results on total and manufacturing employment and

[35] Since we do not need 2006 census year information in the regency-level analysis, in this part of the analysis we take 2007 as the base year.

Table 11: Summary statistics of the variables involved in the labor market analysis (regency-level).

	(1) N	(2) mean	(3) sd	(4) min	(5) max
Regency-level change in robots (EDU)	284	0	1.000	-0.529	10.48
Regency-level change in robots (OECD, EDU)	284	0	1.000	-3.939	9.843
Regency-level change in robots (GM)	284	0	1.000	-0.876	7.620
Regency-level change in robots (OECD, GM)	284	0	1.000	-7.682	4.789
Change in total employment (log)	284	0.122	0.128	-0.178	0.615
Change in total wages (log)	284	1.006	0.219	0	1.687
Change in manufacturing employment (log)	284	0.121	0.376	-2.839	1.325
Change in manufacturing wages (log)	283	1.026	0.349	-0.175	2.072
Change in production employment (log)	284	0.117	0.374	-2.839	1.310
Change in production wages (log)	283	1.018	0.346	-0.188	2.072
Change in non-production employment (log)	197	0.328	1.004	-2.017	3.500
Change in non-production wages (log)	142	0.972	0.716	-1.071	3.172
Change in services employment (log)	284	0.276	0.210	-0.176	1.373
Change in construction and utilities employment (log)	284	0.373	0.310	-1.049	1.564
Change in construction and utilities wages (log)	236	0.663	0.447	-0.762	1.832
Change in agriculture and mining employment (log)	284	-0.0920	0.273	-1.400	0.869
Change in agriculture and mining wages (log)	244	0.638	0.494	-1.682	2.066
Change in RoW export by regency (log)	276	0.0771	0.380	-2.913	0.829
Population in 2007	281	793,849	704,630	29,682	5.756e+06
Share of workers with tertiary education in 2007	281	0.0482	0.0324	0.00898	0.206
Share of workers with no education in 2007	281	0.113	0.0459	0.0178	0.270
Manufacturing share of output in 2007	281	0.100	0.0780	0.00110	0.447
GDP per capita in 2007 (log)	281	2.934	0.638	1.757	5.518

wages. Since we are interested in aggregate effects, we weigh the estimate by regencies' population in the base year. Due to the normalization of the independent variable, the coefficient β_1 in (8) measures the long run impact of robots in the regencies whose change in exposure over the period is one standard deviation above the mean.

The dependent variables in columns (1) and (2) are total regencies' employment and real average wages (in log). The coefficient of total regency's employment is positive, although not significant at conventional levels. The coefficient of wages is positive and statistically significant. The magnitude of the coefficients in columns (1) and (2) imply that on average, robots have increased wages by roughly 5 percentage point over the period in the most exposed regencies.

Columns (3) and (4) of Table 12 focus on manufacturing, which is the sector employing robots in production. The coefficient on employment in column (3) is positive and statistically significant, and implies a long-run average increase in employment of 6 percent. Conversely, the impact on real manufacturing wages is not statistically significant. The positive impact of automation on local employment along with the negative (or neutral) employment effect at the plant level corroborate the hypothesis that non adopting plants may benefit from adopting plants in the same industry. This is consistent with the positive inter-plant spillovers documented above. In light of these results, the spillovers are likely to be intra-industry.

Columns 5-8 of Table 12 split manufacturing employment and wages for production and non-production workers. Columns (5) and (7) show that the positive impact of robots on manufacturing employment is stronger for non-production workers. This is consistent with the plant level analysis, which reveals a positive direct impact of automation on non-production workers. For production employment, the coefficient reflects the negative direct impact vs the positive indirect impact due to vertical spillovers. The coefficient of non-production wages is also positive and significant.

Table 13 presents the results of estimating equation (8) through 2SLS for non-manufacturing sectors. The evidence supports the hypothesis that automation in manufacturing generates positive employment and wage spillovers in services. The coefficients in columns (1) and (2) imply that on average over the period, robots have increased employment by 3 percent and services wages by 5 percent in the most exposed regencies. The impact on services is consistent with the plant level results as well as with extensive evidence showing the presence of backward linkages of manufacturing and services, including in Indonesia (Hallward-Driemeier and Nayyar (2017); Duggan et al. (2013)). For Constructions and Utilities, the coefficients in columns 3 and 4 are not statistically significant.

The coefficient for employment in agriculture and mining, on the other hand, is large, negative and statistically significant. The estimated impact of robots on agriculture employment is almost 10 percent in the most exposed regencies. This is consistent with reallocation of labor towards manufacturing and services facilitated by the robots' adoption.

The results in Tables 12 and 13 are robust to different alternative specifications. First, Table A5 reports the results in which the share of employment at risk of automation is based on (secondary) educational attainments. The coefficients are in line with the baseline results, but in this specification exposure to robots has a positive and significant impact on total employment as well. Second, results are similar using the instrument directly as a regressor, in place of the endogenous ETR (see Table A6 in the Appendix).

Table A7 experiments with alternative calculations of standard errors. Square brackets include shift-share errors, computed as in Adao et al. (2019). As expected, shift-share errors are larger than the robust errors of the baseline specification, but the level of significance of the coefficients remains roughly the same.

Round brackets include province-level clustered errors, which are systematically smaller of both baseline robust and shift-share errors. This is likely to be due to the low number of clusters available of 26 provinces.[36]

Table 12: Regency-level results: employment and wages, total and manufacturing sector

Change 2008-2015 of:	(1) Total employment	(2) Total wages	(3) Manufacturing employment	(4) Manufacturing wages	(5) Production employment	(6) Production wage	(7) Non-production employment	(8) Non-production wage
Regency ETR (GM)	0.022	0.053***	0.059**	0.035	0.050*	0.021	0.169**	0.123**
(change 2008-2015)	(0.014)	(0.019)	(0.028)	(0.025)	(0.028)	(0.023)	(0.073)	(0.059)
Observations	276	276	276	275	276	275	193	139
Province FE	yes	yes	yes	yes	yes	yes	yes	yes
District demand shifter	yes	yes	yes	yes	yes	yes	yes	yes
District base year covariates	yes	yes	yes	yes	yes	yes	yes	yes
First stage F-stat	14.89	14.89	14.89	14.79	14.89	14.79	14.25	12.46

The table presents 2SLS estimates of the relationship between regency-level exposure to robots and employment. Exposure to robots is instrumented with the average exposure in the OECD region. The dependent variables are the 2008-2015 differences of log of employment (total or manufacturing) in each regency. The regency demand shifter is regency-level average global exports excluding Indonesia. Base year regency covariates include: i) population; ii) the shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the share of manufacturing employment. Standard errors are clustered at the regency-level. Weights are constructed using 2007 (base year) regency population. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table 13: Regency-level results: employment and wages in non-manufacturing sectors

Change 2008-2015 of:	(1) Services employment	(2) Services wages	(3) Cons/Util employment	(4) Cons/Util wages	(5) Agr/Mining employment	(6) Agr/Mining wages
Regency ETR (GM)	0.027*	0.052***	0.001	0.037	-0.095***	0.119
(change 2008-2015)	(0.016)	(0.017)	(0.028)	(0.044)	(0.031)	(0.108)
Observations	276	276	276	229	276	238
Province FE	yes	yes	yes	yes	yes	yes
District demand shifter	yes	yes	yes	yes	yes	yes
District base year covariates	yes	yes	yes	yes	yes	yes
First stage F-stat	14.89	14.89	14.89	12.58	14.89	12.81

The table presents 2SLS estimates of the relationship between regency-level exposure to robots and employment. Exposure to robots is instrumented with the average exposure in the OECD region. The dependent variables are the 2008-2015 differences in log of employment by sector in each regency. The regency demand shifter is regency-level average global exports excluding Indonesia. Base year regency covariates include: i) population; ii) the shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the share of manufacturing employment. Standard errors are clustered at the regency-level. Weights are constructed using 2007 (base year) regency population. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

[36] This justifies our choice of relegating such specification in the appendix.

The evidence presented in this section thus departs from previous studies, which use a similar empirical approach but document a negative local labor market impact of robots in manufacturing. Acemoglu and Restrepo (2017) and Dauth et al. (2017) find evidence of a negative impact on both employment and wages in the United States and Germany, respectively. Giuntella and Wang (2019) and Artuc et al. (2019) provide similar evidence for developing economies in China and Mexico, respectively. We discuss a possible explanation behind this difference in Section **??** below.

5 Distributional Impact of Automation

Automation seems to deliver unambiguous benefits to manufacturing plants and even to benefit workers on average. However it is important to examine how these benefits are distributed across production factors, including capital, skilled and unskilled labor. This is a relevant policy question as inequalities within countries are rising and recent research identifies automation as one of the key drivers behind this rise (Hemous and Olsen (2014); Prettner and Strulik (2020)).

We look first at the impact of automation on the labor share in value added and the profit rate (measured as gross surplus over revenue). The results - presented in Table 14 - suggest that automation reduces the share of labor in value added (column 1). Consistent with this result, column (2) shows that automation is associated with a higher profitability (column 2), although the coefficient is not statistically significant at conventional levels (p-value = 0.12). These results are consistent with the idea that most of the gains of automation are captured by the capital owners. This is also consistent with the empirical evidence based on firm-level analysis in the US (Acemoglu et al. (2020)).

Next we test for the impact of automation across workers by skills' level. To that end we estimate the following Mincer-type regression by matching IFR data and two cross sections of labor force survey data for 2008 and 2015:

$$y_{v,t} = \beta_0 + \boldsymbol{\beta X}_{v,t} + \gamma_1 Educ_{v,t} + \gamma_2 Educ_{v,t} * ETR_{r,t} + \alpha_{r,t} + \varepsilon_{v,t} \tag{9}$$

Table 14: Plant level labor share and profit rate.

	(1) Labor share	(2) Profit rate
ETR × secondary	-0.051**	0.073
	(0.017)	(0.041)
Observations	53,517	53,517
R-squared	0.500	0.244
Plant FE	yes	yes
Industry-year FE	yes	yes
Other technologies	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots, labor share in value added and gross operating surplus over revenue. Secondary is the 2006 share of plants' employment with secondary education. Other technologies are capture by an index of innovation activities in 2006, interacted with year fixed effects. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Where $y_{v,t}$ is real wages, $\boldsymbol{X}_{v,t}$ is a vector including the number of years of experience of worker v at year t (defined as years of schooling minus the worker's age), its squared term and a gender dummy, $Educ_{v,t}$ is the years of schooling, and $ETR_{r,t}$ is the regency r exposure, defined as in (7) We control for regency-time fixed effect as $\alpha_{r,t}$.

The coefficient γ_2 captures the impact of automation on schooling wage premia (a measure of wage inequality across skill levels). In particular it measures the difference in the change in return to schooling between a worker in a regency highly-exposed to automation versus a worker in a low-exposed regency. A positive coefficient indicates that automation is associated with increased wage inequality. As the main regressor varies at the district-year level, standard errors are clustered at that level too. As usual we instrument the exposure variable and its interaction terms through the same variables but replacing Indonesia's robot imports with OECD countries' imports.

Table 15 presents the results only of the coefficient of interest γ_2, which suggests a significant inequality-inducing effect of automation, both for all sectors (odd columns) and for manufacturing only (even columns). The result applies whether using monthly wages (columns 1-2) or hourly wages (columns 3-4).[37]

These results confirm the emerging evidence that automation is associated with a rise in economic inequality both through a reduction in the labor share in value added and an increase in the skill wage premia.

[37] The results are very similar if we do not instrument the exposure variable and if measure the automatability of occupations using the FO instead of GM probabilities (tables available upon request).

Table 15: Worker individual-level results: IV Estimation

	(1)	(2)	(3)	(4)
	Monthly		Hourly	
	Wage	Wage	Wage	Wage
Schooling x Regency ETR	0.442***	1.180**	0.390***	1.024***
	(0.179)	(0.523)	(0.0945)	(0.335)
Observations	458,684	52,798	450,722	51,962
R-squared	0.208	0.195	0.177	0.144
Worker Controls	Yes	Yes	Yes	Yes
District-Year FE	Yes	Yes	Yes	Yes
Sample	All	Manufacturing	All	Manufacturing

The table presents 2SLS estimates of the relationship between industry exposure to robots and worker's income. Exposure to robots is constructed as the share of workers susceptible to automation (Graetz and Michaels (2018)) at the regency-level in a year. The dependent variables are, respectively, log of real monthly wage and log of real hourly wage. Standard errors are clustered at the regency-year level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

6 Conclusions

This paper has provided novel evidence that robots increase plant level efficiency, value added and employment in a large developing country experiencing a robust increase in robots' imports. Moreover, we find evidence of positive employment spillovers from downstream automation, within manufacturing and across other sectors.

The positive employment impact of automation in Indonesia contrasts with the evidence not only from advanced economies but also from emerging ones as Mexico (Artuc et al., 2019) and China (Giuntella and Wang, 2019). In these countries the increase in labor demand driven by higher productivity is not sufficient to offset the labor displacement effect of robots. While a systematic investigation of the reasons behind this difference is beyond the scope of this paper, we discuss two candidate explanations.

The first is related to the evidence that the productivity impact of robots is subject to diminishing returns to scale (Graetz and Michaels, 2018). In a country with a low initial density of robots, the adoption of robots may result in particularly large productivity increases that could offset the labor displacement effect. This low initial density seems to apply to Indonesia, which in 2011 - the first year we have the data for the 3 countries - recorded 0.018 robots per thousand workers, versus 0.04 in Mexico and 0.1 in China.[38]

[38] These numbers are obtained by combining IFR data on robots' stock with ILO data on total employment across all sectors.

This intuition is akin to the idea of marginal decreasing returns to capital. Consistently with the low robots' density, the median Indonesian firm employs a much smaller amount of capital per worker than its counterpart in the same 2-digit ISIC sector in China and Mexico (Figure A2 in Appendix).[39]

The second (non mutually exclusive) possible explanation hinges on the lack of relevant skills. By filling skill shortages, robots may induce large productivity gains that more than compensate the labor replacement effects. In contexts where the quality of labor is relatively low, robots can provide substantial improvements to the production process as they enable to achieve higher products' standards.[40] This paper provides suggestive evidence that this type of mechanism may be at play in Indonesia where the penetration of robots significantly increases markups and international trade integration, which are enabled by higher production standards. This hypothesis is also consistent with the relatively low quality of skills in Indonesia as proxied by high-shool students' test score through the OECD Programme for International Student Assessment (PISA 2018).[41]

From a policy perspective, the evidence of positive inter-industry employment spillovers provides a possible rationale to facilitate or even incentivize the adoption of automation technologies among manufacturing firms in Indonesia. Further research will be needed to support the external validity of our findings and the related policy implications. In particular, firm-level information on the adoption of automation technologies in developing countries would be a crucial first step towards that end. In addition, why we don't see more Indonesian firms adopting robots given the high returns to adoption remains an important question for future research.[42]

[39]The computation is based on firm-level data from different waves of the World Bank Enterprise Survey. Specifically we compute the median real value of the capital stock per worker (in international dollars) for all of the manufacturing industries the data allows. The data is available for Indonesia in 2009, for Mexico in 2010 and for China in 2012.

[40]This is similar to Acemoglu and Restrepo (2018a), where robots offset the scarcity of young workers in advanced economies. Maloney and Molina (2019) also cite anecdotal evidence that Chinese firms choose to improve the quality of their export by employing robots, rather than investing in employees' training.

[41]The average reading literacy score of Indonesian students is 371, against 555 for China, 420 for Mexico and 487 for the OECD average. Similarly, science performance is 396 for Indonesia, against 590 for China, 419 for Mexico and 489 for the OECD region.

[42]Evidence in Calì et al. (2019) presents a similar puzzle also in the context of the switching from

At the same time, the analysis also documents that the gains from automation are mainly captured by the firms' owners in the form of higher profit rate, while the share of labor in value added shrinks as a result of automation. This unequal distribution of gains extends also to wage employment, with automation increasing the wage of skilled workers' relatively to that of unskilled workers. Shifting the tax burden from labor to profits could help counter the inequality-inducing impact of automation.

fuel-based to more modern electricity-base technologies in Indonesian manufacturing.

References

Acemoglu, D., C. Lelarge, and P. Restrepo (2020). Competing with robots: Firm-level evidence from france. In *AEA Papers and Proceedings*, Volume 110, pp. 383–88.

Acemoglu, D. and P. Restrepo (2017). Robots and jobs: Evidence from us labor markets. *NBER working paper* (w23285).

Acemoglu, D. and P. Restrepo (2018a). Demographics and automation. Technical report, National Bureau of Economic Research.

Acemoglu, D. and P. Restrepo (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2019). Automation and new tasks: how technology displaces and reinstates labor. *Journal of Economic Perspectives 33*(2), 3–30.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Adao, R., M. Kolesár, and E. Morales (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics 134*(4), 1949–2010.

Amiti, M. and J. Konings (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review 97*(5), 1611–1638.

Artuc, E., P. Bastos, and B. Rijkers (2018). Robots, tasks and trade.

Artuc, E., L. Christiaensen, and H. J. Winkler (2019). *Does Automation in Rich Countries Hurt Developing Ones?: Evidence from the US and Mexico*. The World Bank.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics 118*(4), 1279–1333.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). Does management matter? evidence from india. *Quarterly Journal of Economics 128*(1), 1–51.

Cai, J. and A. Szeidl (2018). Interfirm relationships and business performance. *Quarterly Journal of Economics 133*(3), 1229–1282.

Calì, M., N. Cantore, T. Hidayat, L. Iacovone, M. Pereira-López, and G. Presidente (2019). Too much energy the perverse effect of low fuel prices.

Dauth, W., S. Findeisen, J. Südekum, and N. Woessner (2017). German robots-the impact of industrial robots on workers.

David, H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–97.

De Backer, K., T. DeStefano, C. Menon, and J. R. Suh (2018). Industrial robotics and the global organisation of production.

De Loecker, J. (2007). Do exports generate higher productivity? evidence from slovenia. *Journal of international economics 73*(1), 69–98.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, markups, and trade reform. *Econometrica 84*(2), 445–510.

De Loecker, J. and P. T. Scott (2016). Estimating market power evidence from the us brewing industry. Technical report, National Bureau of Economic Research.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *American Economic Review 102*(6), 2437–71.

Doraszelski, U. and J. Jaumandreu (2013). R&d and productivity: Estimating endogenous productivity. *Review of Economic Studies 80*(4), 1338–1383.

Duggan, V., S. Rahardja, and G. Varela (2013). Service sector reform and manufacturing productivity: evidence from indonesia.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler (2004). The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from colombia. *Journal of development Economics 75*(2), 333–371.

Faber, M. (2018). Robots and reshoring: Evidence from mexican local labor markets.

World Bank (2015). *Indonesia Systematic Country Diagnostic: Connecting the Bottom 40 Percent to the Prosperity Generation*. The World Bank.

World Bank (2020a). Indonesia's critical occupations list 2018 : Occupation profiles.

World Bank (2020b). *World Development Report 2020: Trading for Development in the Age of GVCs*. The World Bank.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review 98*(1), 394–425.

Frey, C. B. and M. A. Osborne (2017). The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change 114*, 254–280.

Giuntella, O. and T. Wang (2019). Is an army of robots marching on chinese jobs?

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The review of economics and statistics 89*(1), 118–133.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in europe. *American economic review 99*(2), 58–63.

Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American economic review 104*(8), 2509–26.

Graetz, G. and G. Michaels (2018). Robots at work. *Review of Economics and Statistics 100*(5), 753–768.

Hallward-Driemeier, M. and G. Nayyar (2017). *Trouble in the Making? The Future of Manufacturing-Led Development*. The World Bank.

Hemous, D. and M. Olsen (2014). The rise of the machines: Automation, horizontal innovation and income inequality.

Hsieh, C.-T. and P. J. Klenow (2014). The life cycle of plants in india and mexico. *Quarterly Journal of Economics 129*(3), 1035–1084.

Javorcik, B. and S. Poelhekke (2017). Former foreign affiliates: Cast out and outperformed? *Journal of the European Economic Association 15*(3), 501–539.

Koch, M., I. Manuylov, and M. Smolka (2019). Robots and firms.

Konings, J. and S. Vanormelingen (2015). The impact of training on productivity and wages: firm-level evidence. *Review of Economics and Statistics 97*(2), 485–497.

Kugler, A. D., M. Kugler, L. Ripani, and R. Rodrigo (2020). Us robots and their impacts in the tropics: Evidence from colombian labor markets. Technical report, National Bureau of Economic Research.

Maloney, W. F. and C. Molina (2019). Is automation labor-displacing in the developing countries, too?: Robots, polarization, and jobs.

Mertens, M. (2019). Micro-mechanisms behind declining labour shares: Market power, production processes, and global competition. Technical report, IWH-CompNet Discussion Papers.

Oldenski, L. et al. (2015). Reshoring by us firms: what do the data say? *Peterson Institute for International Economics, Policy Brief PB15-14*.

Olley, G. S. and A. Pakes (1992). The dynamics of productivity in the telecommunications equipment industry.

Prettner, K. and H. Strulik (2020). Innovation, automation, and inequality: Policy challenges in the race against the machine. *Journal of Monetary Economics 116*, 249–265.

Rodrik, D. (2018). New technologies, global value chains, and developing economies. Technical report, National Bureau of Economic Research.

Stapleton, K. and M. Webb (2020). Automation, trade and multinational activity: Micro evidence from spain.

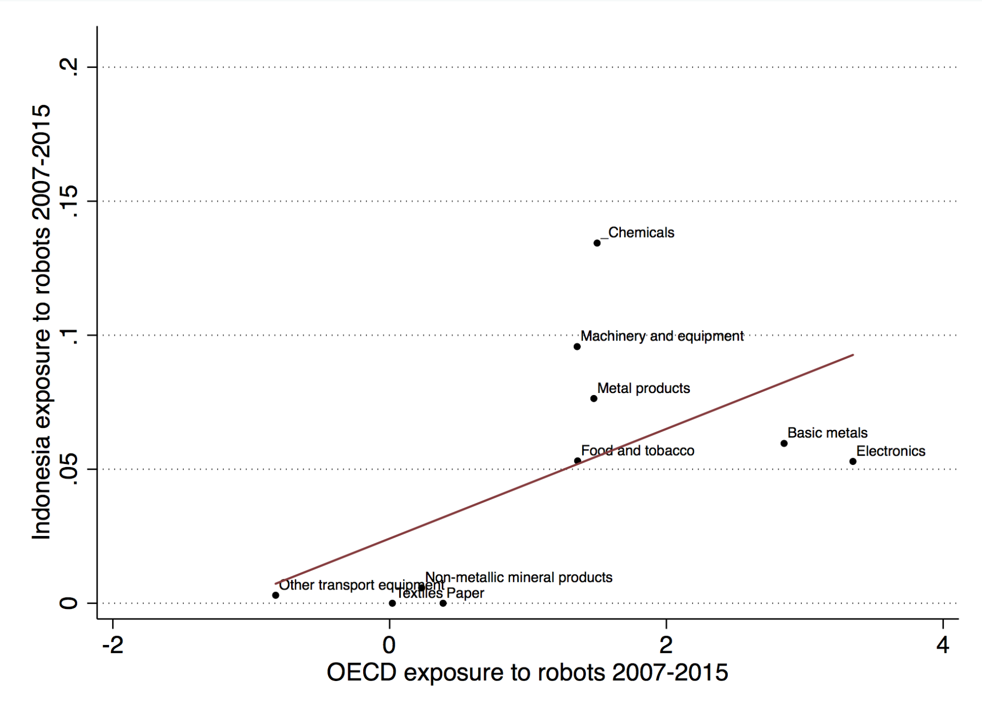
Verhoogen, E. (2020). Firm-level upgrading in developing countries. *CDEP-CGEG Working Paper 83*.

Online Appendix

(not for publication)

A Figures and Tables Appendix

Figure A1: Correlation between Indonesian and OECD-average exposure to robots excluding Motor Vehicles and Rubber and plastics.



On the horizontal axis there is the change between 2007 and 2015, of the OECD region industry-average number of robots per thousand employees. On the vertical axis, there is the change between 2007 and 2015, of the industry-level number of robots per thousand employees in Indonesia. The figure excludes two high-exposure industries in Indonesia, Motor vehicles and Rubber and Plastics. Sources: IFR, STAN, SI.

Table A1: Correlations between Industry- and plant level exposure to robots, and plant level investment: placebo tests

	(1) Investment	(2) Investment
ETR × none and primary	-0.091***	
	(0.018)	
ETR × tertiary		-0.006
		(0.009)
Observations	53,447	53,447
R-squared	0.658	0.658
Plant FE	yes	yes
Year FE	yes	yes
Industry-year FE	yes	yes
Other technologies	yes	yes

The table presents OLS estimates of the relationship between plants' exposure to robots and investment in machinery and equipment. The dependent variable is the log of plant level investment in machinery and equipment. Primary is the base year share of employment with at most primary education; tertiary is the base year share of employment with higher education. Both share are normalised to have zero mean and unitary standard deviation. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table A2: Plant level impact of robots: alternative specification based on principal component analysis.

	(1) Marginal cost	(2) Markup	(3) Revenue	(4) Export share	(5) Import share	(6) Employment	(7) Wages
ETR $\times$ PC (secondary, size, foreign capital, openness)	-0.030*	0.025*	-0.009	0.003	0.009***	0.001	-0.024***
	(0.014)	(0.013)	(0.010)	(0.003)	(0.001)	(0.005)	(0.005)
Observations	53,513	53,513	53,513	53,513	53,513	53,513	53,513
R-squared	0.747	0.744	0.914	0.852	0.792	0.939	0.653
Plant FE	yes	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes	yes
Other technologies	yes	yes	yes	yes	yes	yes	yes

The table presents OLS estimates of the relationship between exposure to robots and key plant level variables. The interaction term is the first principal component of the base year value of the following variables: i) firm (employment) size; ii) a dummy for having more than fifty percent foreign capital; iii) the sum of revenue shares of import and export, and iv) the share of secondary education workers. Standard errors are clustered at the 2-digit industry- and year-level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table A3: Plant level impact of robots: alternative clustering of errors at the plant level.

	(1) Marginal cost	(2) Markup	(3) Revenue	(4) Export share	(5) Import share	(6) Employment	(7) Wages
ETR × secondary	-0.086***	0.051***	0.041**	0.004*	0.004*	-0.006	-0.018
	(0.025)	(0.019)	(0.017)	(0.002)	(0.002)	(0.009)	(0.011)
Observations	53,517	53,517	53,517	53,517	53,517	53,517	53,517
R-squared	0.747	0.744	0.914	0.852	0.792	0.939	0.653
Plant FE	yes	yes	yes	yes	yes	yes	yes
Industry-year FE	yes	yes	yes	yes	yes	yes	yes
Other technologies	yes	yes	yes	yes	yes	yes	yes

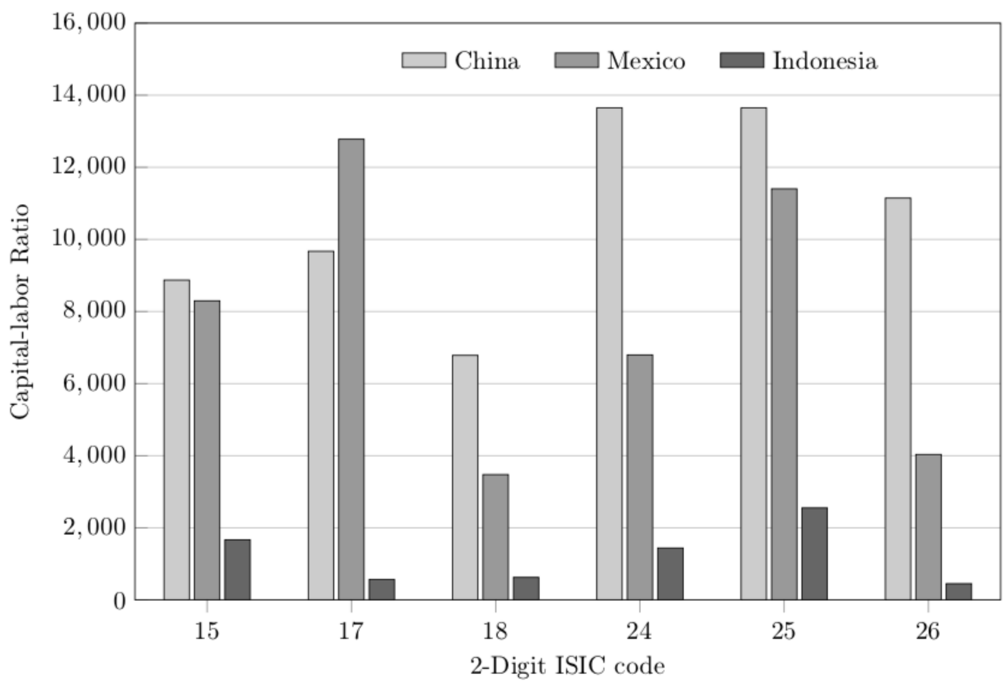
The table presents OLS estimates of the relationship between exposure to robots and key plant level variables. The interaction term is the base year share of plants' employment with secondary education. Standard errors are clustered at the plant level. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table A4: Regency-level first stage regression

Change 2008-2015 of:	(1) Regency ETR (GM)	(2) Regency ETR (EDU)
Regency ETR (OECD,GM)	0.812***	
	(0.210)	
Regency (OECD,EDU)		0.808***
		(0.310)
Observations	276	276
R-squared	0.673	0.650
Province FE	yes	yes
District demand shifter	yes	yes
District base year covariates	yes	yes

The table presents OLS estimates of the relationship between regency-level exposure to robots and regency-level exposure based on the average industries in the OECD region. The dependent and independent variables are expressed as the 2008-2015 differences in log exposure. Base year regency covariates include: i) population; ii) the shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the share of manufacturing employment. Standard errors are clustered at the regency-level. Weights are constructed using 2006 (base year) regency population. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Figure A2: Median capital-labor ratio by 2-digit ISIC industry.



The figure shows the median real value of capital-labor ratio (in international dollars) for China (2012), Mexico (2010), and Indonesia (2009). The data have been cleaned from outliers. Sources: World Bank Enterprise Survey.

Table A5: Regency-level results: shares of employment at high risk of automation based on educational attainments.

Change 2008-2015 of:	(1) Total employment	(2) Total wages	(3) Manufacturing employment	(4) Manufacturing wages	(5) Services employment	(6) Services wages	(7) Cons/Util employment	(8) Cons/Util wages	(9) Agr/Mining employment	(10) Agr/Mining wages
Regency ETR (EDU)	0.025*	0.032**	0.050**	0.030*	0.025*	0.030**	0.013	0.052*	-0.049*	0.081
(change 2008-2015)	(0.014)	(0.016)	(0.021)	(0.017)	(0.014)	(0.012)	(0.016)	(0.027)	(0.027)	(0.073)
Observations	276	276	276	275	276	276	276	229	276	238
Province FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District demand shifter	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District base year covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
First stage F-stat	6.772	6.772	6.772	6.743	6.772	6.772	6.772	6.227	6.772	6.635

The table presents 2SLS estimates of the relationship between regency-level exposure to robots and employment. Exposure to robots is instrumented with the average exposure in the OECD region. The dependent variables are the 2008-2015 differences of log of employment (total or manufacturing) in each regency. The regency demand shifter is regency-level average global exports excluding Indonesia. Base year regency covariates include: i) population; ii) the shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the share of manufacturing employment. Standard errors are clustered at the regency-level. Weights are constructed using 2007 (base year) regency population. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table A6: Reduced form regency-level results

Change 2008-2015 of:	(1) Total employment	(2) Total wages	(3) Manufacturing employment	(4) Manufacturing wages	(5) Services employment	(6) Services wages	(7) Cons/Util employment	(8) Cons/Util wages	(9) Agr/Mining employment	(10) Agr/Mining wages
Regency ETR (GM)	0.018	0.043***	0.048*	0.029	0.022*	0.042***	0.001	0.031	-0.077**	0.096
(change 2008-2015)	(0.011)	(0.014)	(0.026)	(0.021)	(0.012)	(0.015)	(0.022)	(0.039)	(0.036)	(0.087)
Observations	276	276	276	275	276	276	276	229	276	238
R-squared	0.248	0.151	0.063	0.087	0.124	0.151	0.038	0.035	0.232	0.141
Province FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District demand shifter	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District base year covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

The table presents OLS estimates of the relationship between regency-level exposure to robots and employment. The dependent variables are the 2008-2015 differences in log of employment and wages (total or by sector) in each regency. Base year regency covariates include: i) population; ii) the shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the share of manufacturing employment. Standard errors are clustered at the regency-level. Weights are constructed using 2006 (base year) regency population. The coefficients with *** are significant at the 1% level, with ** are significant at the 5% level, and with * are significant at the 10% level.

Table A7: Regency-level results: alternative standard errors

Change 2008-2015 of:	(1) Total employment	(2) Total wages	(3) Manufacturing employment	(4) Manufacturing wages	(5) Services employment	(6) Services wages	(7) Cons/Util employment	(8) Cons/Util wages	(9) Agr/Mining employment	(10) Agr/Mining wages
Regency ETR (GM) (change 2008-2015)	0.022	0.053	0.059	0.035	0.027	0.052	0.001	0.037	-0.095	0.119
	[0.019]	[0.028]	[0.044]	[0.025]	[0.011]	[0.022]	[0.046]	[0.044]	[0.062]	[0.392]
	(0.010)	(0.012)	(0.022)	(0.014)	(0.012)	(0.010)	(0.038)	(0.032)	(0.025)	(0.062)
Observations	276	276	276	275	276	276	276	229	276	238
Province FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District demand shifter	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District base year covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

The table presents 2SLS estimates of the relationship between regency-level exposure to robots and employment. Exposure to robots is instrumented with the average exposure in the OECD region. The dependent variables are the 2008-2015 differences in log of employment and wages (total or by sector) in each regency. Base year regency covariates include: i) population; ii) the shares of workers with tertiary and no education (separately); iii) real output per capita, and iv) the share of manufacturing employment. Shift-share standard errors (Adao et al. (2019)) are presented in square brackets. Province-level clustered errors are presented in round brackets.

B Occupation Profiles

The Occupation profiles are compiled by the World Bank in partnership with national governments. The Occupational profiles are used as key inputs of Critical Occupations Lists, which aim to identify shortages of certain occupations of strategic importance to the economy (e.g. World Bank (2020a)).

The methodology used by the world bank is based on the probability of computerization provided by Frey and Osborne (2017). Occupations with probability greater than 0.7 are deemed at high risk of automation.

The list of occupations at high risk of automation extracted from the World Bank Occupation profiles for Indonesia, Malaysia and Thailand are: Welders and Flame Cutter; Handicraft Workers in Wood, Textile, Leather and Related Materials; Power-Production Plant Operators; Woodworking-machine tool setters and operators; Mineral and stone processing plant operators; Cement, stone and other mineral products machine operators; Well drillers and borers and related workers; Cement, stone and other mineral products machine operators; Metal processing plant operators; Metal finishing, plating and coating machine operators; Chemical products plant and machine operators; Rubber products machine operators; Plastic products machine operators; Food and related products machine operators; Packing, bottling and labelling machine operators; Stationary plant and machine operators; Mechanical machinery assemblers; Electrical and electronic equipment assemblers.

The key piece of information provided in the Occupation Profiles is the typical educational attainments of workers in occupations at high risk of automation. An example of occupation profile for “Welders and Flame Cutters” is provided in Figure A3. As for Welders and Flame cutters, a manual inspection of all available occupations suggests that the modal educational attainment for occupations at high risk of automation is junior or senior secondary education - items iii) and iv) in the SI educational category.

Figure A3: Example of occupation profile, compiled by the World Bank.

70 | Occupations Profiles

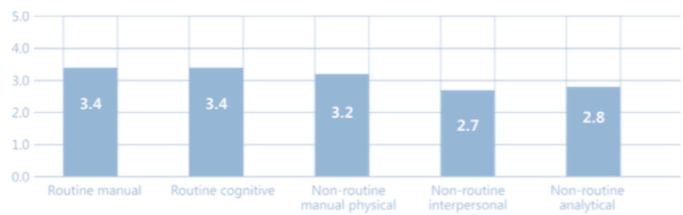
Welders and Flame Cutter

Summary	
KBJI title:	Welders and Flame Cutter
KBJI code:	7212
Job description:	Workers in this group weld and cut metal parts using gas flamesor electric flares and melt and unify the metal. Tasks include: to weld parts of metal using gas or electric flares, thermite compounds, or other methods; operate resistance welding machines; use a welding torch to make and repair important layers, pipes, floors, and other equipment; solder metal parts; cut metal plates with fire or electric incandescent; combine metal parts with hand soldering.
Position includes:	Welder (with specification on underwater welders and welders for food and beverage manufacturing)

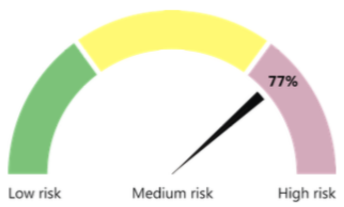
Statistics	
Gender:	1% female
Age:	33-38 years old
Monthly wage:	IDR 1,800,000 – 2,375,000
Working hours:	47-48 hours per week
Typical education:	Junior High School and Vocational High School
Field of study:	N/A
Most common sector of employment:	Manufacturing
Most common employer type:	Individual/Household business

Automatability and skills content of occupation

Task importance (1-5)



What is the probability that this occupation will be automated given current technology?



Although Welders have a very high risk of automation overall, the non-routine analytical tasks that they undertake may be more difficult to automate.

Sources: World Bank Occupation Profiles

C Data Appendix

C.1 Robots Data

One issue with IFR data is that in the early years of the sample, a breakdown of imports by sector is not available and they are grouped under the label "unspecified". In this case, shares by sectors are estimated using information for the years in which the breakdown is available. We 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 very similar. 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. The IFR does provide estimates of the stock, but it adopts a different assumption that robots fully depreciate after twelve years.

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.

C.2 Manufacturing Data

From SI data, we select the following plant level variables: Output; Fixed assets (capital); Production employment; Non-production employment; Production wage bill Non-production wage bill; Intermediate materials; Share of exported output; Share of imported inputs; Investment in machinery and equipment

From these data we derive the following variables: Total employment (production+non-production employment); Total wage bill (production+non-production wage bill); Average production wage (production wage bill / production employment); Average non-production wage (non-production wage bill / non-production employment); Share of imported materials (imported materials / intermediate materials) Labor share (wage bill / value added); Profit rate (gross profits (value added minus wage bill) / revenue).

After taking logs of employment and wage bills, we replace the observation with zero if the original variables are zero. Due to the high number of observations equal to zero, we adopt the following log-transformation for investment in machinery and equipment: $\ln[x + (x^2 + 1)^{.5}]$.

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.[43] 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 E. 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.

In order to allow the matching between SI and IFR data, we build a consistent industry classification. Plants in SI are grouped into 5-digit sectors following the definition Klasifikasi Baku Lapangan Usaha Indonesia (KBLI). A KBLI code is assigned to a plant according to the classification in which the main product produced belongs. The KBLI classification has been adjusted to be consistent over the whole sample, ranging from 2006 to 2015. One issue is that in converting codes from KBLI rev.3 (in use until 2009) to KBLI rev.4, some industries are split in more than one industry, or viceversa. For such

[43] 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.

reason, we only keep those KBLI codes that have an unambiguous one to one mapping across the two revisions. We also experimented with a looser conversion including more industries, without significant changes in our main results.

C.3 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 re 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 A8. We observe that 8.5% of plants switches to another industry during the 10 years covered by our sample. Therefore, to avoid potentially confounding effects, we assign to each plant the trends in robots' adoption of the industry to which it belonged in the first year of observation.

Table A8: Correspondence between IFR and SI industry classification.

IFR industries	**Description**	**KBLI industries**
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

C.4 Nine-digit Products and Inputs Data

Our data include information on quantities and values of the products produced and raw materials used by each plant. These are both defined at a highly granular level, namely 9-digit Klasifikasi Komoditi Indonesia (KKI). In our sample, each plant produces on average 2 products and 25% of the plants produce more than one product. We use disaggregate products information to measure the number of products produced by each plant. After computing unit prices by dividing value with quantities, we compute yearly price growth. If the price grow by more than a factor of 10 or decreases more than by a factor of 1/10, we drop the observation. Average unit prices are then used to construct plant level price deflators (see D). On average, each plant uses four different varieties of raw materials. We also have information on use of domestically produced and imported materials, which we aggregate at the plant level to measure the share of imported materials.

D Construction of plant level Price indices

The derivation of plant-specific price indices from product-level price data closely follows Eslava et al. (2004) and Mertens (2019).

These are plant level Tornqvist indices exploiting information on 9-digit products produced and inputs used by each plant.

$$\pi_{jt} = \prod_{p=1}^{n} \left(\frac{P_{pjt}}{P_{pj,t-1}} \right)^{.5(s_{pjt}+s_{pj,t-1})} \pi_{j,t-1}$$

where P_{pjt} is the price of good p and s_{pjt} is the share of this good in total product market sales of plant j in period t. Therefore, the growth of π_{jt} is the product of each plant's price growth, each weighted with the average share of sales in t and $t-1$. Wee set $\pi_{jt} = 100$ in 2006. For plants entering after 2006, we follow Eslava et al. (2004) and Mertens (2019) and use the 5-digit industry average of the plant price indices as a starting value. When price growth data are missing, we replace it with an average of product or inputs price changes within the same 5-digit industry.

E 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. One 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 inexperienced enumerators that were more likely to make mistakes, which contributed to increase measurement errors.

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 is the ratio of plant-KL to the sector average KL is below 0.02 or larger than 50. We experiment with stricter thresholds which result in too many observations dropped. 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. 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.

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 A4). Figure A5 compares original and clean capital stock series.

Figure A4: Plants' growth rate distribution of capital-labor ratio.

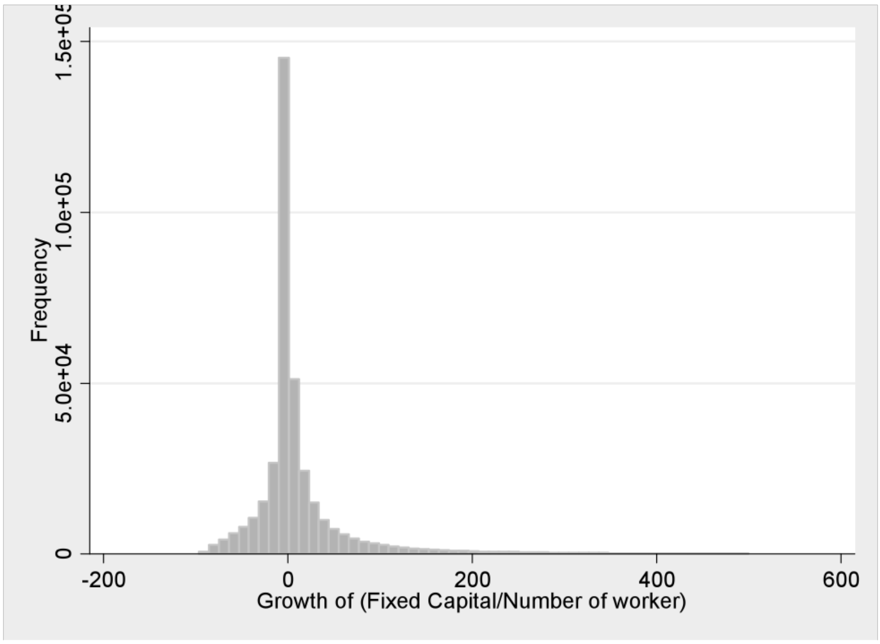
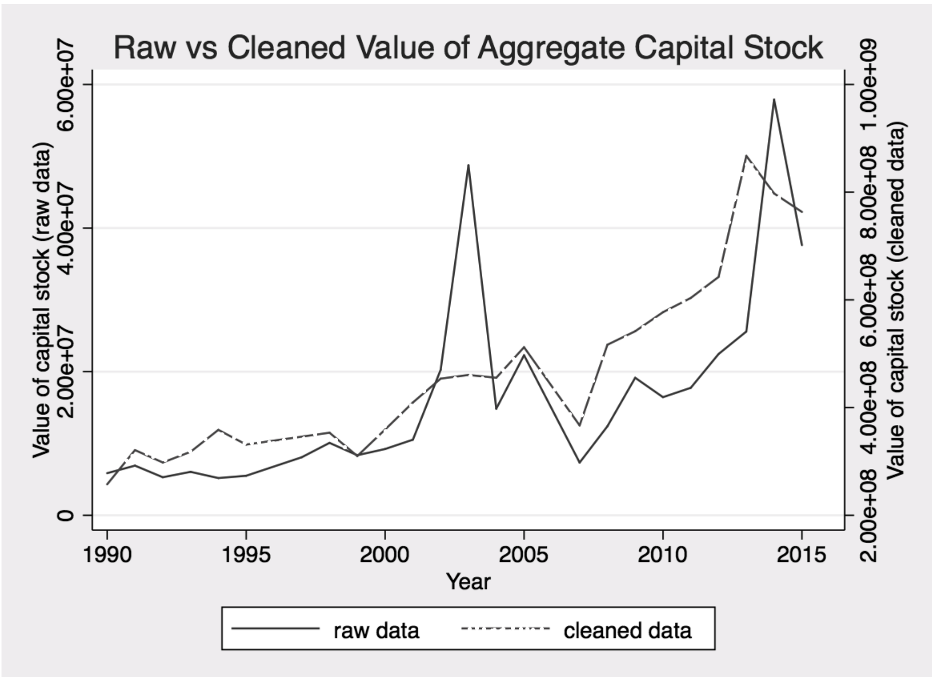


Figure A5: Comparison of Aggregate Nominal Capital Stock Series.



F Estimation of Marginal Cost and Markup

We assume that in each year t, plant f produces output Q_{ft} with the following production function:

$$Q_{ft} = \min\left\{\gamma_m M_{ft}, F(K_{ft}, L_{ft}) \cdot \Omega_{jt}\right\} \tag{10}$$

where M_{ft} are intermediate inputs, K_{ft} the capital stock and L_{ft} labor. The term Ω_{jt} represents Hicks-neutral productivity.

The production function (10) is a structural value added specification De Loecker and Scott (2016) in which capital and labor are allowed to be characterised by some degree of substitution and inputs are perfect complements to the combination of the other inputs. Given (10), a profit maximising plant sets

$$Q_{ft} = \gamma_m M_{ft} = F(K_{ft}, L_{ft}) \cdot \Omega_{ft} \tag{11}$$

Our objective is estimating plants' production function parameters, in order to obtain estimates of TFP and markups. To recover such parameters from the data, we estimate the logged version of production function (11).

$$q_{ft} = f\big(k_{ft}, l_{ft}; \boldsymbol{\beta}\big) + \omega_{ft} + \epsilon_{ft} \tag{12}$$

Recall that ω_{ft} represents the log of Hicks-neutral productivity, which is known by plants' managers but not by us. The variable ϵ_{ft} is an i.i.d. error term that captures factors such as measurement errors.

We are interested in estimating the vector of the production function parameters $\boldsymbol{\beta}$. To recover unbiased and consistent estimates of firms' production function (12), we need to address the well-known simultaneity problem deriving from the fact that ω_{ft} is correlated to labor but not to capital, which is chosen one period ahead. We build on the methodology of Ackerberg et al. (2015). In particular, we make the following timing assumptions concerning inputs' decisions: i) capital k_{ft} is chosen at $t-1$; ii) l_{ft} is chosen at $t-b$ after observing ω_{ft}, and iii) materials m_{ft} are chosen at $t-a$, with $1 < b < a$.

We then specify the materials' demand function, $m_{ft} = \tilde{h}(\omega_{ft}, k_{ft}, l_{ft}, \boldsymbol{\theta}_{ft})$. The vector $\boldsymbol{\theta}_{ft}$ includes variables that affect plant level demand for materials.

Assuming that the materials' demand function of the plant, $\tilde{h}$ is monotonically increasing and invertible in ω, we obtain a control function that proxies for unobserved productivity,

$$\omega_{ft} = h(m_{ft}, k_{ft}, l_{ft}, \boldsymbol{\theta}_{ft}) \tag{13}$$

where $h \equiv \tilde{h}^{-1}$. Adding $h(\cdot)$ to (12), we get

$$q_{ft} = f\big(k_{ft}, l_{ft}; \boldsymbol{\beta}\big) + h(m_{ft}, k_{ft}, l_{ft}, \boldsymbol{\theta}_{ft}) + \epsilon_{ft} \tag{14}$$

We follow Ackerberg et al. (2015) by approximating the right-hand-side of (14) with a third-order polynomial in all its elements, except for the elements of $\boldsymbol{\theta}$, which we enter linearly.[44] From the first stage, we obtain expected output $\hat{q}_{ft}$ and the residuals $\hat{\epsilon}_{ft}$.[45]

The next step is specifying a law of motion for productivity ω_{ft}. We assume that ω_{ft} follows a Markov process that can be shifted by plant managers' action:

$$\omega_{ft} = g\big(\omega_{f,t-1}, \boldsymbol{\Gamma}_{f,t-1}\big) + \xi_{ft} \tag{15}$$

In (15), ξ_{ft} denotes the innovation to productivity and the vector $\boldsymbol{\Gamma}$ includes variables controlled by plants' managers that influence the expected future value of productivity and state variables which determine differences in productivity dynamics across plants. In our framework, these variables capture the opportunities for automation available to each plant, which we measure as the industry-level exposure of robots, the plant level share of secondary education workers and their interactions.[46] Current expected productivity is

[44] This approach is similar to Mertens (2019).

[45] It should be noticed that in the first stage, none of the production function parameters are identified, because they enter both $f(\cdot)$ and $h(\cdot)$.

[46] In our application we impose a simple AR(1) form for (15).

then expressed as a function of the data and parameters,

$$\omega(\boldsymbol{\beta})_{ft} = \hat{q}_{ft} - f\big(k_{ft}, l_{ft}; \boldsymbol{\beta}\big) \tag{16}$$

To estimate $\boldsymbol{\beta}$, we form moments based on the innovation ξ_{ft} in the law of motion (15),

$$\xi(\boldsymbol{\beta})_{ft} = \omega(\boldsymbol{\beta})_{ft} - E\big[\omega(\boldsymbol{\beta})_{ft} | \omega(\boldsymbol{\beta})_{f,t-1}, \boldsymbol{\Gamma}_{f,t-1}\big] \tag{17}$$

The moments that identify the parameters are:

$$E\big[\xi(\boldsymbol{\beta})_{ft} \boldsymbol{M}_{ft}\big] = 0 \tag{18}$$

where the vector $\boldsymbol{M}_{ft}$ includes current capital, lagged labor, and lagged materials use.

In our empirical application, we use a flexible trans-log specification to approximate $f(\cdot)$. Moreover, our setup and timing assumptions are based on the idea that materials are the most flexible inputs in production. Therefore, to avoid problems related to the existence of raw inputs' adjustment costs, in our empirical application we use energy consumption to proxy for unobserved productivity. It should be noticed that labor and electricity consumption are both significantly correlated within a plant over time, which justify their inclusion in (18) as instruments. We deflate output and energy expenditure with the plant-specific deflators (see Appendix D). For capital, we employ asset specific price indexes, distinguishing between machinery and equipment, vehicles, buildings, and land.

We obtain the production function parameter vector $\hat{\boldsymbol{\beta}}$ by estimating 18 with GMM and bootstrapping errors over fifty repetitions.

F.1 Deriving TFPQ, Markup and Marginal Cost From Plants' Cost Minimisation

Quantity-total factor productivity (TFPQ) is obtained using (16).

We follow De Loecker and Warzynski (2012) to obtain a measure of plant level markup from the plants' first order conditions. Cost minimisation with respect to labor, which we consider a static input, implies the following first order condition:

$$\frac{\partial \mathcal{L}_{jt}}{\partial L_{ft}} = W_{ft} - \lambda_{ft} \frac{\partial F(K_{ft}, L_{ft}) \cdot \Omega_{ft}}{\partial L_{ft}} = 0$$

where $\mathcal{L}$ is plant's f Lagrangian, W_{ft} wages and λ_{ft} the Lagrangian multiplier. Rearranging terms and multiplying both sides of the previous equation by $\frac{L_{ft}}{Q_{ft}}$, we obtain

$$\frac{\partial F(K_{ft}, L_{ft}) \cdot \Omega_{ft}}{\partial L_{ft}} \frac{L_{ft}}{Q_{ft}} = \frac{1}{\lambda_{ft}} \frac{W_{ft} L_{ft}}{Q_{ft}}$$

As in De Loecker and Warzynski (2012), we define the plant's markup over the marginal cost of output λ_{ft} as

$$\mu_{ft} \equiv \frac{P_{ft}}{\lambda_{ft}}$$

where P_{ft} is the price of output produced by the plant. The previous equation yields an expression of plants' markup depending on the elasticity of output with respect to the variable input, β_l, and the inverse of the revenue share of expenditure on L_{lt}:

$$\mu_{ft} = \frac{\partial F(K_{ft}, L_{ft}) \cdot \Omega_{ft}}{\partial L_{ft}} \frac{L_{ft}}{Q_{ft}} \frac{P_{jt} Q_{jt}}{W_{jt} L_{jt}} = \beta_l \frac{P_{jt} Q_{jt}}{W_{jt} L_{jt}}$$

In our empirical application, the markup is given by

$$\mu_{ft} = \beta_l \frac{P_{ft} Q_{ft}}{W_{ft} L_{ft}} \frac{1}{\hat{\epsilon}_{ft}} \tag{19}$$

where the last term in (20) is the residual obtained from the first stage estimation of (14). As discussed in De Loecker and Warzynski (2012), including $\hat{\epsilon}_{ft}$ is important, as it allows to purge the estimated markup for variation in output not due to changes in inputs.

Finally, we recover marginal cost as

$$mc_{ft} = \frac{P_{ft}}{\mu_{ft}} \tag{20}$$

Prior to estimation, we drop the bottom and top one percent of the distribution of markup and marginal cost in order to avoid outliers driving the results. We obtain very similar coefficients if we do not trim our estimates.