ECB3AMT 2021/22 GROUP ASSIGNMENT

ECB3AMT_paper2122_group3

Stiebale, J, Südekum, J and Woessner, N. 2020. 'Robots and the rise of European superstar firms'. London, Centre for Economic Policy Research. https://cepr.org/active/publications/discussion_papers/dp.php?dpno=15080.

Members
Lan Nguyen | 6687547 | m.nguyenhoangmylan@students.uu.nl
Lisa Verlare | 6929699 | l.verlare@students.uu.nl
Nguyen Thuy Duong | 6827209 | d.nguyenthuyduong@students.uu.nl

Word count*: 2355*

Statement of Originality

We, the aforementioned students, herewith declare to have written this document and that we am responsible for the content of it. We declare that the text and the work presented in this document is original and that no sources other than those mentioned in the text and its references have been used in creating it.
Utrecht University School of Economics is responsible solely for the supervision of completion of the work, not for the content.

Section 1. Research Question

The paper "Robots and the Rise of European Superstar Firms" by Südekum, Stiebale and Woessner (2020) highlights the effects of new technology – industrial robots – on the dispersion, within industries, of firm-level productivity and markups. Although the phenomenal adoption of robotics within industry has led to investigations in the firm's keys to success, the paper stands out by focusing on firm-level outcomes, instead of generalized findings at the industry-level, hence providing new and more specific insights for both businesses and economists.

Scientifically, the research therefore helps to solve the issues of broadly generalized outcomes documented by current studies, such as the surveys by Syverson (2011, as cited in Südekum, Stiebale and Woessner, 2020), Bloom and Van Reenen (2010, as cited in Südekum, Stiebale and Woessner, 2020). It does not substantially focus on industry-level outcomes, but instead observes variables at firm-level. Since the analysis is partly performed at firm-level, heterogeneity between firms can be taken into account. Purposefully, it reveals the underlying elements driving the emergence of superstar firms, in particular robotic technologies.

The paper is also supplementary and in line with the new findings by Kehrig and Vincent (2018, as cited in Südekum, Stiebale and Woessner, 2020) regarding the declining aggregate labor share witnessed in the past decades. Instead of aggregating the equilibrium impact of technology on productivity and markup, it provides profound evidence on why productivity tends to benefit the superstars rather than those with lower productivity and how increased robotization in those highly productive firms will decrease the labor share even further.

In reality the findings are beneficial as the gap between theoretical findings and application is narrowed down. Since their explanations are based on business' viewpoints, the results are more relevant for firms to apply in specific business contexts. Companies operating in automated manufacturing industries such as Samsung or Boeing, can have stronger evidence to expand their robotic adoption in the production process. However, the winners-take-most phenomenon raises concerns for market competition, authorities and workers' bargaining power, probably leading to market inefficiency. The authors indeed validate this statement by stressing that "productive firms typically pay higher wages in absolute terms, it may further push up the wages of top earners in these firms, leading to a widening dispersion in household incomes" (Südekum, Stiebale and

Woessner, 2020, p. 34). Fortunately, these concerns can be diminished as more leading firms are incorporating social responsibilities into their long-term strategies, meaning that companies with increased market shares can contribute a larger portion of their revenues to stakeholders.

Section 2. Identification Problem

Simply regressing either productivity or markup on implementation of industrial robots might not give a good indication of what influences those outcome variables, as it causes systematically incorrect causation. This could lead to bad business decisions and, perhaps, negative consequences for workers since firms make decisions based on this incomplete information.

The authors highlight two noticeable identification problems, namely omitted variable bias and reverse causality.

Omitted Variable Bias

Researchers are likely to run into OVB by regressing the outcome on the (prespecified) treatment variable(s) when the effects of some important variables are not accounted for in the model, which could bias the results.

It depends on the missing variable whether the simple OLS regression will over- or underestimate the effect and how this term ends up:

$$\text{plim}\,\hat{\beta} = \beta + \delta \frac{Cov(D_i, M_i)}{Var(D_i)}$$

A potential omitted variable, according to the authors, could be international trade, for which they control by calculating import and export variables for industries. International trade can be correlated with both firm performance and robots.

It could be that there is a large amount of international trade. In automated industries this would require more robots for higher productivity and thus the correlation between the treatment variable (*i.e. change in robotization)* and the omitted variable (*i.e. high international trade)* is positive.

$$\frac{Cov(D_i, M_i)}{Var(D_i)} > 0$$

In addition, large international trade could positively affect firm performance because they are likely to gain more profits from higher sales. Thus:

$$\delta > 0$$

To sum up, omitting large international trade leads to an upward bias, meaning that OLS overestimates the effect.

Reverse causality

The paper emphasizes the causal effect of increasing robotization on firm performance, especially in automated industries. This inference is subject to strict exogeneity conditions by which the treatment variable (*i.e. robotization)* is completely independent of the outcome variable (*i.e. firm performance).* The authors cast doubt on the reverse causality issue caused by general technological trends. Intuitively, an increased number of large firms with high technological adoption rate for the upscaling of production is likely to positively influence robot stock, implying robot densification at industry-level.

Under OLS regression, this reverse causation can be summarized as follows:

$$\begin{cases} Performance = \beta_0 + \beta_1 Robot + \varepsilon_1 \\ Robot = \beta_0 + \beta_1 Performance + \varepsilon_2 \end{cases}$$

This equation system can be simplified as

$$Performance = \beta_0 + \beta_1 f(Performance) + \varepsilon_1$$

where $Cov(f(Performance), \varepsilon_1) > 0$ as robotization is expected to positively influence firm performance. This leads to β_1 being biased upward, hence, the simple OLS model would overestimate the effect on average. Consequently, failing to account for different causal pathways may cause systematic errors not only in the magnitude of the causal effect but also in determining the direction of causal link.

Section 3. Identification Strategy

The main identification strategy that Südekom, Stiebale and Woessner (2020) used to address concerns of omitted variable biases in the OLS is the Cobb-Douglas functional form. To avoid this, they add in interaction terms and several controls and test their model on those variables, which include industry-level changes, different effects that robots might have, countries and other technologies.

An advantage of the Cobb-Douglas is that it keeps the parameters limited to labor and capital, which makes it simple to interpret the resulting estimates. Nevertheless, several assumptions should hold.

(1) The market is perfectly competitive and output elasticities are identical for firms, which will not change over time.
(2) In the short run capital and labor are fixed for the most part, attempting to change them will have high costs.
(3) The number of materials, i.e. robots, can be changed at will.
(4) When determining the number of materials, their decision is based on observations of the other inputs in a previous period.
(5) Material demand, i.e. demand for robots, depends on more than one factor. Not only does it depend on the changes in log wages, but also on for which country the analysis is done, in which year and what the log change in robots is.

Assumptions 2-5 are very plausible. Assumption 2 is based on labour economic theory that in the short run firms will not change their capital stock due to firm capabilities, i.e., available resources. Additionally, assumption 3 is plausible, assuming that robots are always readily available. Also, assumption 4 makes sense, because firms will base their material needs on how they did in previous periods, which is an indication of their firm performance. Finally for assumption 5 it is plausible to assume that the material demand is heterogeneous among countries, industries, and years.

For assumption 1 the authors have done robustness tests. Assuming a perfectly competitive market is unrealistic, as there are imperfections caused by differences in firms’ market power, regulations. The same holds for identical output elasticities, which has now been relaxed. In reality even firms producing homogenous goods with varied resources and capacity have differing demand for inputs. Relaxing these assumptions could lead to biased estimates, as the authors warn, because then they would no longer control for these variables. However, the authors show that when these assumptions are relaxed and they do not control for a perfectly competitive market and identical output elasticities, their main conclusion still holds. If so, these assumptions are trivial to the model.

Reverse Causality

Despite the efficiency of using Cobb-Douglas function in controlling for omitted variables, increasing superstar firms resulting from more robot adoption could be improbable due to reverse causality. By using an IV approach, exogenous variation on robot stock caused by general trends are captured, hence isolating the heterogeneous effects of industrial robots on firm performance.

Particularly, the authors exploit two sets of instruments:

- Instrument set A = industry-level robot installations in the US and the UK
- Instrument set B = industry-level robot installations in the US, the UK, Norway, Belgium, Austria, and Portugal

A valid instrument has to satisfy three identifying assumptions

(1) Relevance

The instrument sets have to matter for the industry-level robot stock. Therefore, changes in robotization in the instrumental countries must influence the robot adoption rate in sample countries. Generally, we find this plausible since the causation can be attributed to advancements in the global robotics technological frontier due to domestic shocks (Acemoglu and Restrepo, 2017).

(2) Exogeneity

Robot installations in instrumental countries can only affect firm performance through robot stock in sample countries. This might be plausible as the decision of robot adoption at firm-level is ideally independent of decisions of other competitors. Firms decide to install more robots only when it does help them gain a higher level of efficiency and productivity, which translates to internally improving performances.

(3) Independence

There should be no confounders affecting both robot stock in IV countries and firm performance in sample countries. This might be plausible because for observable omitted variables that the authors have controlled for such as international trade and other technology and innovation measures, none of them directly influence the IVs.

Section 4. Internal Validity

The identification faces three main potential threats, namely heterogeneity across firms, probability of exogeneity violation, and lack of firm-level data on robotization. The first two threats are carefully addressed in the paper, whereas the latter remains an unsolved issue.

Heterogeneity

Since variables are measured based on revenue instead of physical quantity, there is a chance that heterogeneous effects, e.g., industrial-specific shocks, across firms, are not fully covered. If these elements are correlated with the variables of interest, omitted variable bias would mislead the estimated coefficients and preclude causal inference.

Sources of heterogeneity are technology and innovation measures as well as dynamic industrial factors, such as imports and exports, the change in market share of foreign-owned firms, inconsistent capital-to-labor ratio, and wage levels (Südekum, Stiebale, and Woessner, 2020). To avoid biases, technical robustness checks allowing for heterogeneity were performed to re-evaluate the impact of robotization on firm performance. Their conclusions are then reinforced with theories supported by previous studies.

Firstly, researchers have accounted for the diversity across researched units with regards to technical efficiency (the change in the industry-level robot stock) by implementing an industry-level output deflator to control the price effects.

Secondly, they allow the log change in robots to interact with quintiles dummies of the heterogeneous effects specification, namely robot density, industry-specific trends, timing, and adequacy of the TFP.

For both steps, the checks are replicated for all specifications as shown in *Table A.1*. The work implies that the impact of robots remains similar when including heterogeneity in the equations (*Table A.3*). These results confirm their observations that the increased robot installation excessively benefits top-performing firms, which become more capable of obtaining higher firms' performance. The authors document the returns of tech adoption theory (Lileeva and Trefler, 2010; Bustos, 2011; Bertschek et al., 2015, as cited in Südekum, Stiebale and Woessner, 2020) to explain such profit-boosting behaviors of firms.

The heterogeneity across researched units is, therefore, under careful considerations and control.

Exogeneity violation

To control for a possible violation of exogeneity the authors only include the incumbents that were in both years that were sampled. Therefore, the bias that exiting and entering firms might cause is controlled for.

Additionally, in order to validate the over-identified model with log change of robot as single endogenous variable, the strong assumption of at least one instrument being exogenous has to hold. In this case, the assumption is valid for the former set but not for the latter since additional instrumental EU countries are in proximity of sample countries, suggesting the effect of endogenous factors such as the regional trend of robotization to comply with regional standards. Although having proved the predictive power of the IVs in set B via F-Tests (*Table A.2*), the authors were not able to statistically demonstrate its validity. Therefore, the possibility of endogenous IVs when adopting set B might cause misleading inference of heterogeneous effects of robots.

Lack of firm-level data on robotization

The paper exploits industry-level robot stock as the explanatory variable of interest, instead of group-level of robot stock, to conclude its causal effect on each quintile/decile. However, the use of industry-average data might give incorrect inferences for less profitable firms due to type II errors *(i.e. fail to reject the non-causation of log change on robots on improved performance when it's false).* The authors logically explained that top firms, when adopting more robots, are more likely to benefit from reduced marginal costs, therefore higher profit margin. This implies the rise in industry-level robotization is mostly due to these top-performers and using industry-level data might not be representative for the rest.

Section 5. External Validity

In their analysis the authors look at quintiles, which shows that the initial most productive top 10% firms benefit from an increase in robots by increasing their productivity even more, but no significant change occurs to the average firm when implementing more robots. Therefore, only firms in the top 10% to begin with would benefit from such a change.

In conclusion, this might not be a representative study for less- or nonautomated industries. In manufacturing there is a lot of automation, while in other industries that might not be the case. So while we find this study externally valid for the automated manufacturing industry, we do not think that is the case for other more labor-intensive industries where productivity does not depend on automation.

This study can be used as building blocks for further research with other technologies, as this paper mainly pertains to robots as a technology. The research had laid the groundwork for further research on the relationship between technological improvement and firms' economic performance in particular, and of the industry in general.

The paper implies that while productivity might have gone up due to robotization, the average wages did not as more revenue was allocated to the superstar firms with smaller labor shares. This might have further implications for economic policy to counteract the winner-take-most and income-distribution-inequality situations.

References

Acemoglu, D., & Restrepo, P. (2017, March). Robots and Jobs: Evidence from US Labor Markets (NBER Working Paper No. 23285). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w23285/w23285.pdf.

Stiebale, J, Südekum, J and Woessner, N. 2020. 'Robots and the rise of European superstar firms'. London, Centre for Economic Policy Research. https://cepr.org/active/publications/discussion_papers/dp.php?dpno=15080.

APPENDIX A. Appendix Tables

Table A.1. Robustness checks

Table 4: Robustness checks.

	Density	Ind. dummies	Not lagged	$\triangle_4$	$\triangle_3$	Translog	IV, set A	IV, set B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[A] $\triangle$ ln(TFP)								
$\triangle$ Robots x Quin1	-0.0009*	-0.0068	0.0009	-0.0101**	-0.0166***	0.0006	-0.0014	0.0003
	(0.000)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.008)	(0.006)
x Quin2	0.0000	-0.0002	0.0056	-0.0038	-0.0127***	0.0053	0.0036	0.0054
	(0.001)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.008)	(0.007)
x Quin3	0.0007	0.0010	0.0069	0.0007	-0.0056	0.0079	0.0040	0.0057
	(0.001)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.007)	(0.007)
x Quin4	0.0005	0.0030	0.0087*	0.0004	-0.0068	0.0085*	0.0055	0.0075
	(0.001)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.007)	(0.006)
x Quin5	0.0025**	0.0140**	0.0201**	0.0122	0.0036	0.0101*	0.0117	0.0167
	(0.001)	(0.007)	(0.008)	(0.008)	(0.007)	(0.006)	(0.014)	(0.012)
[B] $\triangle$ ln(Markup)								
$\triangle$ Robots x Quin1	-0.0032***	-0.0215**	-0.0247**	-0.0273***	-0.0257***	-0.0223***	-0.0120	-0.0179
	(0.001)	(0.009)	(0.011)	(0.009)	(0.009)	(0.008)	(0.012)	(0.011)
x Quin2	-0.0025**	-0.0260***	-0.0290***	-0.0260***	-0.0224***	-0.0154**	-0.0232*	-0.0265**
	(0.001)	(0.008)	(0.010)	(0.009)	(0.008)	(0.006)	(0.012)	(0.012)
x Quin3	-0.0025**	-0.0164**	-0.0166*	-0.0153**	-0.0124**	-0.0125**	-0.0170	-0.0189*
	(0.001)	(0.008)	(0.010)	(0.007)	(0.006)	(0.005)	(0.011)	(0.011)
x Quin4	-0.0021*	-0.0173**	-0.0175*	-0.0030	-0.0053	-0.0086	-0.0187*	-0.0177
	(0.001)	(0.008)	(0.010)	(0.006)	(0.006)	(0.006)	(0.011)	(0.011)
x Quin5	0.0022	0.0202**	0.0168**	0.0305***	0.0287***	-0.0067	0.0081	0.0142
	(0.001)	(0.008)	(0.008)	(0.008)	(0.009)	(0.006)	(0.013)	(0.011)
N	110,727	110,710	114,140	171,570	228,313	109,679	110,710	110,710

Note. Based on N firm observations. This table presents robustness checks for the heterogeneous effects of robots on TFP (Panel A) and on markups (Panel B), based on the specifications in column (5) of Table 2 respectively Table 3. Column (1) uses the change in the robot density – the change in robots per thousand workers – instead of the log change in robots as the main variable of interest. In column (2), we add industry dummies to control for industry trends. Columns (3)–(5) check the robustness with regard to timing issues, by not lagging the log change in robots (column 3), and by using four-year (column 4) and three-year (column 5) instead of five-year differences. Column (6) assumes a translog rather than a Cobb-Douglas production function in estimating TFP and markups. In columns (7) and (8), the industry-level robot stock in the sample countries is instrumented with robot installations in the US and the UK (set A), or in the US, the UK, Norway, Belgium, Portugal, and Austria (set B), and over-identified models are estimated by 2SLS. Standard errors clustered by country x industry in parentheses. Levels of significance: *** 1%, ** 5%, * 10%. Sources: Amadeus, IFR, OECD Stan, Eurostat SBS, EU KLEMS, Comtrade, own calculations.

Note. From *Robots and the rise of European superstar firms*, by J. Stiebale, J. Südekum and N. Woessner, 2020, *Centre for Economic Policy Research*, p. 25.

Table A.2. First stage results

Table A.5: First stage results.

	△ ln(TFP)		△ ln(Markup)	
	IV, set A	IV, set B	IV, set A	IV, set B
	(1)	(2)	(3)	(4)
[A] F-Test on excluded instruments				
△ ln(Robots) x Quin1	8.649	17.875	9.219	22.399
x Quin2	8.652	14.428	9.103	28.069
x Quin3	8.519	16.568	7.910	19.138
x Quin4	8.198	15.038	8.496	18.284
x Quin5	8.174	14.318	9.831	15.218
[B] Kleibergen-Paap weak identification test				
	12.439	9.429	14.041	10.448
[C] Baseline specification				
C.1 F-Test on excluded instruments				
△ ln(Robots)	33.699	24.145	33.697	24.145
C.2 Kleibergen-Paap weak identification test				
	58.914	37.611	58.910	37.610

Note. Panel A shows that IVs in set B are possibly significantly strong as their F-Statistics are larger than 10.

Note. From *Robots and the rise of European superstar firms*, by J. Stiebale, J. Südekum and N. Woessner, 2020, *Centre for Economic Policy Research*, p. 46. Copyright: Jens Südekum, Joel Stiebale and Nicole Woessner

Table A.3. Heterogeneous effects of other technology and innovation measures

Table A.2: Heterogeneous effects of other technology and innovation measures.

	$\triangle_5$ ln(TFP)		$\triangle_5$ ln(Markup)	
	(1)	(2)	(3)	(4)
$\triangle_5$ ln(Robots) x Quin1	-0.0025	-0.0011	-0.0272**	-0.0280***
	(0.004)	(0.004)	(0.010)	(0.010)
x Quin2	0.0041	0.0051	-0.0317***	-0.0316***
	(0.004)	(0.004)	(0.010)	(0.010)
x Quin3	0.0054	0.0058	-0.0219**	-0.0218**
	(0.005)	(0.005)	(0.010)	(0.010)
x Quin4	0.0073	0.0069	-0.0228**	-0.0228**
	(0.005)	(0.005)	(0.009)	(0.009)
x Quin5	0.0183**	0.0157**	0.0147*	0.0154*
	(0.008)	(0.007)	(0.008)	(0.009)
$\triangle_5$ ln(ICT) x Quin1		-0.0044		0.0132**
		(0.005)		(0.006)
x Quin2		-0.0042		0.0078
		(0.004)		(0.006)
x Quin3		-0.0014		0.0059
		(0.004)		(0.007)
x Quin4		0.0026		0.0105
		(0.004)		(0.007)
x Quin5		0.0121		-0.0031
		(0.008)		(0.008)
$\triangle_5$ ln(R&D) x Quin1		-0.0007		0.0067
		(0.007)		(0.015)
x Quin2		-0.0075		0.0161
		(0.007)		(0.018)
x Quin3		-0.0138*		0.0147
		(0.007)		(0.016)
x Quin4		-0.0236***		0.0175
		(0.007)		(0.017)
x Quin5		-0.0487***		0.0168
		(0.011)		(0.019)
$\triangle_5$ ln(Software) x Quin1		-0.0032		0.0133
		(0.008)		(0.013)
x Quin2		-0.0089		0.0080
		(0.008)		(0.011)
x Quin3		-0.0098		0.0075
		(0.007)		(0.013)
x Quin4		-0.0132**		0.0155*
		(0.006)		(0.009)
x Quin5		-0.0279***		-0.0069
		(0.008)		(0.012)
Country, year dummies	✓	✓	✓	✓
$\triangle_5$ other technologies	✓		✓	
$\triangle_5$ other industry changes	✓	✓	✓	✓
Industry controls in $t-5$	✓	✓	✓	✓
Dummies for quintiles	✓	✓	✓	✓

Note. Based on 110,710 firm observations. Columns (1) and (3) replicate the results from column (5) of Table 2 and Table 3 respectively. In the columns (2) and (4), we check the robustness of the estimated heterogeneous effects of robots by additionally allowing for heterogeneous effects of the other technology and innovation variables. The log changes in ICT, R&D, and software and databases are interacted with the dummy variables for the quintiles of baseline TFP respectively markups (i.e., *Quin*1 to *Quin*5). The regression equations are estimated by OLS using overlapping five-year differences. See Table 2 for a detailed description of control variables. Standard errors clustered by country x industry in parentheses. Levels of significance: *** 1%, ** 5%, * 10%. Sources: Amadeus, IFR, OECD Stan, Eurostat SBS, EU KLEMS, Comtrade, own calculations.

Note: From *Robots and the rise of European superstar firms*, by J. Stiebale, J. Südekum and N. Woessner, 2020, *Centre for Economic Policy Research*, p. 43. Copyright: Jens Südekum, Joel Stiebale and Nicole Woessner