A Meta-Analysis of the Effect of Robotics upon Labour Employment

by

Mitchell Stephen Tier

A research dissertation forming a part of the requirements for the degree of Master of Commerce

University of Canterbury

February 2022

Abstract

Using 2438 estimates collected over 32 studies, I conduct a meta analysis to investigate the relationship between robotics and employment. Using both fixed and random effects weighted least squares regressions, I find little evidence of a clear relationship between changes in robot use and employment, nor do I find evidence of publication bias within this literature. Using Bayesian model averaging and backwards stepwise regressions I explore factors that may influence the size of the robot-employment relationship. I find evidence that variables for the level at which analysis is conducted, as well as source of data used are very likely to feature in the true robot-employment specification, but none of these meet the minimum threshold for being categorized even as a small effect size. I also find evidence that controls for population size, gender shares and ethnicity shares are likely to be present in the true robot-employment specification, and are each found to have a small effect size. Specifications controlling for population size or ethnicity shares tend to have a smaller estimated effect of the impact of robots on employment, while specifications controlling for age shares tend to give larger estimates of the robot-employment effect. Finally, I find that there exists a pronounced interest within the literature for analysis of already developed nations, but only few papers on developing countries, which limits the generalisability of our findings. One further needs to recognise that there exists some significant flaws in the IFR dataset, which is the data source for the majority of analyses estimating the robot-employment effect size. Such flaws threaten the validity of the results of both empirical studies using IFR data and dependent studies, such as this meta analysis.

Table of Contents

I

Introduction

The potential for disruption in global labour markets due to the growing level of automation into the workplace has been an area of major concern for some time, simultaneously at the individual, firm, industry, and national level. As early as 1930, John Maynard Keynes (1930) references the growth of *"Technological unemployment"* rising from humanity's *"discovery of means of economising the use of labour outrunning the pace at which we can find new uses for labour".* Similarly, Baldwin and Shultz (1955) express concerns regarding the uncertain effect upon employment associated with growing industrialisation and automation. More recently, this topic has gained much media attention. In an article from the BBC, Cellan-Jones (2019) suggests "20 million manufacturing jobs around the world could be replaced by robots by 2030", while in an article written for Forbes, Kelly (2020) argues that robots will cause an additional unemployment shock, on top of the shock induced by the COVID-19 pandemic.

Over time, many economists have attempted to estimate the effect of automation on employment. Terzidis et al. (2019) reviews this literature, analysing 77 studies with empirical estimates of the relationship between automation and employment, concluding that automation is generally beneficial at the firm and industry level in terms of employment. However, the study also finds that automation increases labour displacement at the occupation level, and is more likely to displace low-skilled labour.

In the related literature, ‘automation’ has generally been the focus of analysis for those analysing the influence of technological development upon employment. However, more recently, there is a growing interest in understanding the effect of robotics. In a seminal paper originally written in 2017, Acemoglu and Restrepo (2020) focus on the effect of industrial robots upon local US labour markets. They find robots have a generally negative effect on employment, suggesting that 'this time' it might be different: while earlier automation was not bad for overall employment, the latest wave of automation through the use of industrial robots, could be. The academic response to this paper has been substantial with a multitude of related papers being released in subsequence, each estimating this niche relationship, often citing the Acemoglu and Restrepo (2020) estimation method as the foundation for their respective statistical inference.

In this thesis, I analyse this recent strand of the literature using a meta analysis. Although having a wide range of empirical research is beneficial for understanding an effect of interest in a given research area, each paper individually examines only a specific form of a relationship of interest. Understanding both the ‘average’ strength of a relationship as well what determines the strength of this relationship, is the purpose of meta analysis. Through the process of meta analysis, I seek to collate all quantitative information regarding a relationship of interest, throughout the entire body of research. Following this, I systematically analyse the collected information in order to uncover an underlying broad effect of interest, to find the variables which influence this relationship, and to assess the degree of publication bias within the literature.

This paper conducts a meta analysis of the recent literature attempting to understand the effect of **robotics** (or 'robots') upon **employment**. Overall, I find no evidence that there is a meaningful empirical relationship between the use of robots and employment. Similarly, I find no evidence of publication bias within the literature. Using Bayesian Model Averaging and backwards stepwise regressions, I find that estimates controlling for population size or ethnicity shares tend to find more negative estimates of the size of the impact of robots on employment; while estimates controlling for age shares tend to find more a positive robot-employment relationship. But even those effect sizes are small, according to the partial correlation coefficient effect size guidelines outlined in Doucouliagos (2011). I further find that a control variable for the data source on which analysis is based (International Federation of Robotics (IFR) data/ non-IFR) is likely to be present in the true robot-employment specification, but its effect does not meet the minimum threshold for a small effect, according to Doucouliagos (2011). Finally, I also find that controls for the level of analysis (for example, regional versus industry level) undertaken are likely to feature in the true robot-employment specification. But while there appears some variation in the estimated effect sizes of different levels of analysis, none of these items meet the Doucouliagos (2011) threshold for a small effect.

Finally, I also recognise two potentially disruptive issues which may threaten the validity of our results. First, there exists much criticism of the accuracy of the IFR dataset, the data source of 84.5% of the estimates collected for this study. There exists a series of issues related to this dataset which may damage the results of studies in many different areas and contexts. Secondly, I find a strong tendency of literature within this field to focus their analysis on labour markets within developed nations, making global generalisations of our findings potentially dubious.

The remainder of the paper is structured as follows. The methodology through which relevant papers are collated, alongside a literature review is conducted in section II. In section III, I describe the techniques to be used in our analysis. Section IV is used to describe the coded data and conduct analyses. Conclusions of analysis are given in section V.

II

Literature Review

This literature review begins in section II.1 by describing the method by which I collected the relevant papers to include in our study. PRISMA (2021) provides guidelines which I use to narrow down broad literature searches into a collection of exclusively directly relevant studies to the meta analysis at hand. I then explore some significant findings in the more broad literature which analyses the influence of automation upon employment (section II.2). Following this, in section II.3, I recognise the literature analysing some of the alternative effects beyond employment that both automation and robots have been tested to have. In section II.4, I focus our attention on the literature surrounding the effect of robots on employment. I split this section into 3 parts, first giving an overview of the findings and connections between recent publications within this field. I then comment on the observed sensitivity of estimated effect sizes between related literature, and finally, I recognise and bring to light some of the issues noticed throughout the literature related to a major data source used by the majority of our collected papers. Lastly, in section II.5 I first recognise some of the foundational literature in the field of meta analysis, alongside some meta analysis literature related to this study, before describing the basic technique used to make effect size estimates between studies comparable.

II.1- Literature collection

Information for this meta analysis was collated through a series of labour intensive steps, directed by the PRISMA methodology for data collection. The first step being the identification of useful databases through which to collect literature. I selected four separate search engines to ensure searches resulted in a broad selection of potentially relevant papers. These search engines are Google Scholar, IDEAs, EBSCO, and SCOPUS, which are all widely used in economic research. Next, it is necessary to select our keywords for searching on these databases. I found the most effective search terms in retrieving relevant-appearing results to be “employment” “jobs” and “robots”. All our searches use “robots” and one of “jobs” and “employment” as search terms. From these searches, I collected 521 potentially relevant papers.

To create a bibliographic database with information about relevant papers, I use web-scraping programs, alongside the integrated exporting functions of some of these literature databases. I exported our search results into separate excel files for convenience and ensuring consistency over time. However, prior to eliminating papers based on irrelevant content (i.e. papers which do not, or appear extremely unlikely to, provide some empirical estimate of the effect of robots upon employment), I ensured that each of our files contained only unique papers (to avoid the issue of including the same paper's regressions multiple times in our meta analysis). This required searching within and across excel files for keywords (e.g. author name, title, abstract segment) unique to each individual paper. In the case where a paper does appear multiple times, excess copies (i.e. not the first appearance of a paper) were removed from our file and are not examined further. This process reduced the bank of search results from 521 papers to 386.

Following this, each of the abstracts of these 386 unique papers were read separately by 2 reviewers. If the abstract suggested that the paper would present content on the relationship between the use of robotics and its effect on labour employment, I would approve and download the paper for further examination. In the case where the paper is entirely irrelevant to the topic at hand or the abstract suggests it is very unlikely to examine the particular relationship between robots and employment, I reject the paper for further analysis. This process reduced the result bank from 386 papers to 112. I then accessed the full copies of each of the remaining studies to determine whether the paper reports any estimates of the effect of robots upon employment. In the case where a paper does include such a regression, I include the paper in our final list to be coded for our meta analysis (further detail of coding process given in section IV.1). Papers that do not include such a regression are rejected and are not used in any quantitative analysis. This process reduced the bank of papers from 112 to 32. Each of the remaining 32 papers will be included in all further quantitative analysis made on this thesis. I have compiled these studies into a separate file which includes PDFs of each of these papers.

The PRISMA procedure is summarised in Figure 1 below

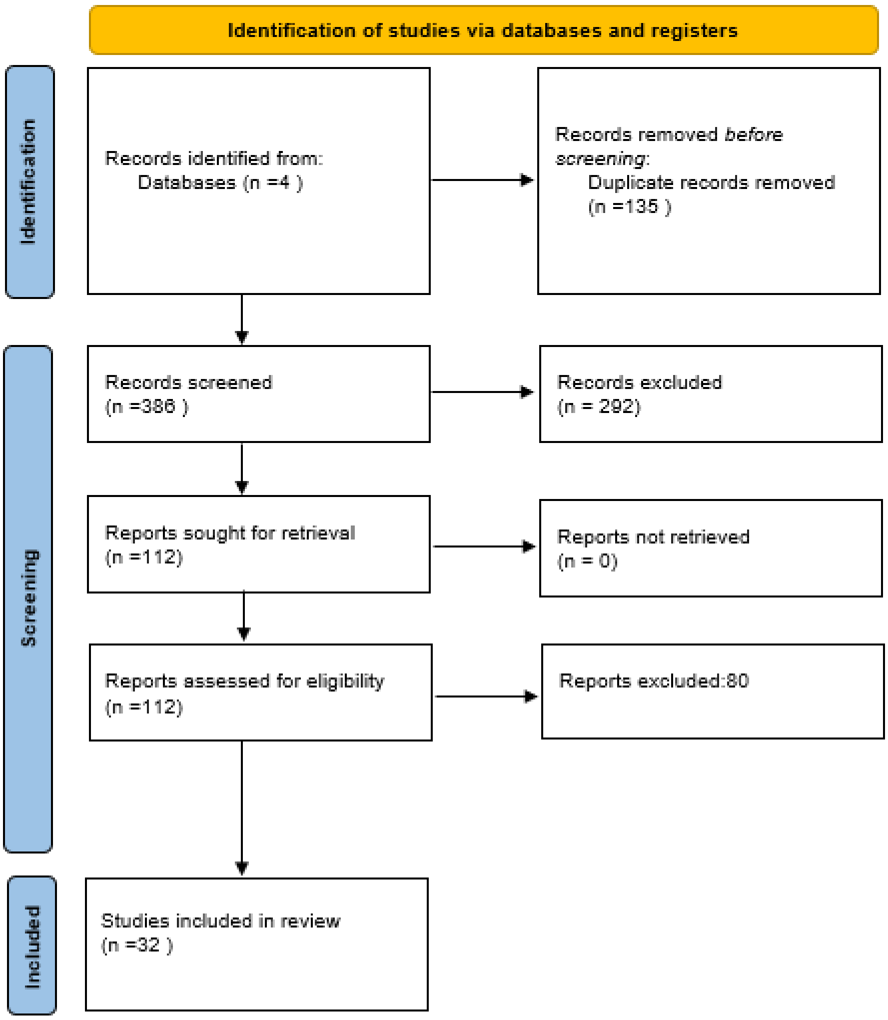

Figure 1

II.2

Broad Automation and Employment.

The area of interest of this study is part of a broader field of research, that of the effect of automation upon the labour market.

Two key forces dictate the size of the effect of automation upon employment; namely the substitution effect, and the income effect. The purpose of introducing robots into the workplace is to improve upon the efficiency of human labour, substituting generally expensive human labour for comparatively cheap mechanical labour, thereby reducing the demand for human labour. Increases in production efficiency typically translate into a reduction in production costs, which may incentivise a producer or service provider to subsequently lower prices in order to increase competitiveness. Where goods and services follow normal demand behaviour, price decreases will stimulate additional demand, which a profit maximising good/service provider meets most likely through the acquisition of additional human and mechanical labour. In this general case, automation thus generates two opposing effects: while the substitution effect will have a negative influence on employment, the income effect will have a positive effect on employment. I cannot determine the dominant effect without an empirical investigation.

The following section seeks to identify a selection of relevant contributions within this broader field, to provide a wider context to the meta analysis conducted later in this paper.

A number of studies have made attempts to establish the risk to various occupations of becoming automated, without estimating an actual automation-employment effect. Frey and Osborne (2017) conduct a study examining the susceptibility of 702 occupations to computerisation. Of these occupations, the paper finds that approximately 47% of the US labour force is at risk of computerisation (expected within 10-20 years). More specifically, the study finds that the majority of employment in transportation, logistics, office administration support, and service occupations are highly susceptible to computerisation. Arntz et al. (2016) and Pouliakas (2018) both perform similar studies, analysing the risk of automation associated with jobs in countries in the OECD and EU respectively. In Roux (2018), a model is formed for predicting the impact adopting "increasingly advanced computing technologies" may have on the labour market in South Africa. The paper analyses the correlation between a 'risk of computerisation' index and the rates of employment change across various industries, ethnicities, and labour types (skill levels), to find the proportion of South African workers whose jobs are in imminent threat of computerisation. The paper finds that after analysing 285 occupations, 35% of South African labour as of 2014 has the potential to be computerised in the near future. Additionally, about 27% of the working South African population was deemed to be in occupations at very high risk of computerisation within the near future.

Unfortunately, findings from these papers are unusable for the purpose of our meta analysis for two key reasons. The first being that the papers fail to run a regression specifying the predicted effect upon employment (these papers only give the risk of computerisation), and most prominently, the independent variable that is used to predict employment risk is not of robotics, but rather of computerisation.

Another strand of the literature focuses specifically on the effect of automation on employment. Micco (2019), for example, conducts a study incorporating automation risk levels to provide causal evidence on the effect of automation on labour markets. It provides these estimates using an automation risk index value that is associated with each occupation held by an individual worker at the time of the study. The paper finds that employment within jobs deemed to be at risk of automation has declined by 2.0%-2.5% greater than employment in jobs deemed not at risk. Additionally, the paper finds that industries which contained a high share of occupations at risk of automatisation experienced relatively low employment growth during the period from 2002-2016. Additional examples exploring the effect of automation at the occupation level include Leotief and Duchin (1984); Fuei (2017); and Bessen (2019).

The literature on the employment effect of automation is a broad and varied range of niche studies, making it difficult for interested readers to interpret any general trends. Recognising this, Terzidis et al. (2019) conducts a meta-analysis of their own, aggregating and subsequently analysing the estimates generated for the effect of automation upon employment (and wages), across the entire branch of literature. The paper collects 77 studies, and 1158 estimates. The paper makes commentary of the automation effect on a multitude of employment forms and geographical areas. The paper finds a positive automation-employment relationship at the firm and industry levels. However, the study additionally finds a positive automation-labour displacement relationship at the occupation level. Terzidis et al.

(2019) also highlights the degree of heterogeneity in the automation effect that exists between different areas (those papers conducted using European data were more likely to find a positive employment effect relative to their American counterparts); skill levels (finding the effect of technology upon employment for 'low-skilled' labour to be predominantly negative, while this effect is reversed for high skilled labour); and with the use of different proxies for automation (finding that 'Research and Development investment' and 'Factor based Technical Change' tended to have significantly positive estimated effects upon employment).

II.3

Automation and Alternative Effects

When considering the effect of automation on labour markets, researchers are often interested in more than just the effect on employment. Issues such as the effect upon wages are also of interest. Micco (2019) provides an example of such a paper. The paper finds that at the occupation level, the risk of automation of a given job is correlated with a reduction in wages. Further, Acemoglu and Restrepo (2020) finds a reduction of wages of 0.77% for each additional robot per 1000 workers at the commuting zone level in the USA. Similarly, Stemmler (2019) makes predictions of the effect of robots/automation upon both wages and national exports in Brazil. The paper finds that wage inequalities which exist between high skill and low skill labourers in Brazil are exacerbated as a cause of increases in domestic automation use, suggesting that automation may lessen the bargaining power of low skilled labour, while increasing the demand for highly skilled individuals who can operate automation machinery.

Some papers focus on the indirect effects of automation. Anelli et al. (2019) explores the effect of automation on voting behaviour in 14 western European countries; the motivation behind this being that the introduction of automation in a given region or occupation is likely to have some effect on its labour market, thereby inducing some psychological effect on affected individual which may alter voting behaviour. The paper finds that greater exposure to automation results in “poorer perceived economic conditions and well-being, lower satisfaction with the government and democracy, and a reduction in perceived political self-efficacy”, which thereby leads to a greater support of ‘radical-right’ and nationalist parties. Im et al. (2019) also makes commentary on the voting effects of automation, focussing on the effect within western European nations. As opposed to automation, the paper uses ‘robot adoption’ as its primary explanatory variable. Similar to the findings of Anelli et al. (2019), Im et al. (2019) finds that greater regional exposure to robotics causes an increase in support for nationalist and radical-right parties.

It is evident from the literature that in many cases, the scope of interest extends past that which this paper seeks to analyse. For the purpose of conciseness and to provide a clear direction of our research, this paper does not make further commentary on the matter of wages or these alternative automation effects.

II.4

Robots and employment

II.4.A

Robot-Employment literature overview

I now turn to the most directly relevant literature to the study at hand – that which studies the robot-employment effect. Using the PRISMA method, I have collected 32 usable studies which make at least one estimate of the robot-employment effect. The depth and general interest of these papers vary widely within this set, yet the following section seeks to give an account of this set of studies, giving review of the key ideas and findings related to the robot-employment effect throughout the literature[1].

Most studies focus on single countries like the United States, France, Spain, Germany or Japan, following the example of Acemoglu and Restrepo (2020), originally written in 2017, which analyses the effect of industrial robots upon United States employment between 1993 and 2007. Conducting its analysis at the commuting zone level (an approximation for local labour markets), the paper concludes that one additional robot per thousand workers reduces the aggregate employment-to-population ratio by 0.2 percentage points (or rather, 3.33 workers per robot given current robot stock) and a decline of the local labour market employment-to-population ratio of 0.39 percentage points. The paper has proved particularly

[1] Although many of these papers' analyses are not limited to the robot-employment effect, this discussion relates specifically to contributions in this form.

influential in this field of literature, having been cited in all other items to be used in this analysis, although itself takes inspiration, and builds upon the model developed in Zeira (1998), which analyses the relationship between technological innovations and various economic growth variables. In addition, beyond simply sparking interest in quantitatively analysing the robot-employment effect, the model developed in Acemoglu and Restrepo (2020) appears to at least some degree in 24 of the 32 collected papers. Using the same model developed in Acemoglu and Restrepo (2020), Sequeira, Garrido and Santos (2020) provide an opposing view, finding that over the period 1990 to 2007 robots can potentially have a positive influence on United States employment beyond a particular robot penetration threshold within a given industry. In other words, the paper suggests that there exists some non-linear (U-shaped) relationship previously undiscovered, as opposed to the linearly negative relationship proposed by Acemoglu and Restrepo (2020). Borjas and Freeman (2019), although using a different methodology, and examining the robot-employment effect over a different time period (2004-2016), finds a very similar general effect size estimate to that of Acemoglu and Restrepo (2020). Anelli, Giuntella and Stella (2019) and Micco (2019) provide further estimates of the robot-employment effect in the United States, each finding generally negative effects.

Focusing on France, Aghion, Antonin, & Bunel (2019) study the influence of robots on French employment over the 1994-2014 period, and use the model developed in Acemoglu and Restrepo (2020) and Zeira (1998). Although the paper primarily seeks to uncover the effect of artificial intelligence, it forms employment effect estimates using a robotisation independent variable. The paper finds that increases in robotisation result in a decrease to employment at the commuting zone level, and finds further an accentuated negative effect on non-educated employment. Acemoglu, Lelarge, & Restrepo (2020) focus on French manufacturing firms at both the market level and the firm level. The key findings of the paper are similar to Acemoglu and Restrepo (2020) and Aghion, Antonin, & Bunel (2019), finding generally that firms which adopt robots experience declines in the share of production employment. In contrast Kariel (2021), which also follows the Acemoglu and Restrepo (2020) model, estimates the effect of industrial robot adoption upon employment in the UK, and finds a generally positive robot-employment relationship. In certain cases however, Kariel (2021) finds the effect can be negative, particularly in high-tech manufacturing employment.

Camiña, Díaz-Chao and Torrent-Sellens (2020) form estimates of the robot-employment effect in Spanish manufacturing firms over the period 1991-2016. The paper finds that although robots have the effect of replacing human labour, this effect is overshadowed by the complimentary (income effect) factor of an increase in robot use, which results in an overall long term increase in employment. Koch, Manuylov and Smolka (2019) conducts a very similar study, examining the robot-employment effect in Spanish manufacturing firms over the period 1990-2016. As might be expected, it finds similar results to that of Camiña, Díaz-Chao and Torrent-Sellens (2020), estimating a net positive effect of robots upon job creation.

This paper also codes for two separate editions of Dauth, Findeisen, Suedekum and Woessner (2017/2018). The paper analyses the effect of industrial robots upon German labour markets from 1994 to 2014 using the model developed by Acemoglu and Restrepo (2020), finding no general robot-employment effect. However, when examining individual labour types, the study finds that there exists a negative employment effect on manufacturing employment, which is neutralised by employment gains in the service sector.

Of particular interest to this field of study, is the influence of robots upon employment in Japan, given the country's extremely high robot density (robots per worker), approximately 10 times greater than in the United States according to Dekle (2020). Analysis within such a high robot usage nation may give a more reliable indication of the expected global long term robot-employment effect. This is the focus of Dekle (2020), who uses industry level data from 1979-2012 to analyse the robot-employment effect in Japan. The paper also derives its model from that used in Acemoglu and Restrepo (2020). The robot-employment effect is analysed in three separate components, namely: the negative displacement effect, the positive productivity effect, and the positive general equilibrium effect. The general finding of Dekle (2020) is that, at the industry level, the positive employment effects brought upon by robots significantly outweigh the negative displacement effect in Japan. Like Dekle (2020), Adachi et al. (2020) also estimates the size of the robot-employment effect in Japanese labour markets at the industry level, making use of a data set covering the time period 1978-2017. Again, the paper bases its methodology on that designed in Acemoglu and Restrepo (2020). As is the case of Dekle (2020), the paper finds that there exists a positive robot-employment relationship.

Ni & Obashi (2021) also focuses its analysis on Japan, particularly on the effect of industrial robots upon Japanese manufacturing employment. As opposed to simply estimating the relationship between robots and employment, Ni & Obashi (2021) breaks down the equation further, estimating the effect of robots upon both job creation and job destruction. The paper finds that robots positively affect both job creation and destruction. The effect on job destruction is found to dominate that of job creation, thus leading to an overall negative employment effect of robots upon Japanese manufacturing employment. Finally, Eggleston et al. (2021) provides an analysis of the robot-employment effect in Japan using establishment-level data. The paper has a specific focus on the robot-employment effect in Japanese nursing homes. Unlike most other papers included within this meta-study, this paper does not use a panel data set, and instead forms its estimates using a 2017 survey on long-term care work in Japan. As is the case of both Dekle (2020) and Adachi et al. (2020), Eggleston et al. (2021) finds a generally positive robot-employment relation.

Each of the countries analysed so far in this review was analysed by more than one study. For some countries, however, only one paper is available. Giuntella and Wang (2019) investigates the effect of robots upon Chinese employment, also using the methodology developed in Acemoglu and Restrepo (2020) over the period 2000-2016. Using city-individual level data, the study finds a strong negative robot-employment effect, while the effect is exacerbated for employment of low-skilled, male, prime-age and older workers. Dotorri (2020) examines the effect in Italian labour markets from 1991 to 2016, finding a small positive robot-employment effect. Dixon, Hong and Wu (2020) study the effect at the firm level in Canada, but do not find a uniform general relationship. The paper instead finds that robots have a positive effect upon employment of non-managerial employees (enhanced effect for those employees with robot complimentary skills), and a negative relationship on that of managerial employees.

An early and highly influential contribution in this field, Graetz & Michaels (2018) originally written in 2015, assesses the effect of industrial robots upon employment between 1993 and 2007. Rather than focusing on a single country however, it analyses data from 17 developed countries. On aggregate, the paper finds no significant relationship between introduction of robots and employment, although their analysis does find that robot adoption corresponds with slight decreases in the employment share of individuals with low incomes, and slight increases in the employment of individuals with middle incomes.

Blanas, Gancia, and Lee (2019) provides another multi-country investigation of the effect of industrial robots upon employment, this time in 10 high-income countries, over the period 1982-2005, and at the country-industry level. As opposed to many studies in this field, the authors base their empirical model on that of Acemoglu and Autor (2011), a paper which develops a model for the effect of technology upon various labour market outcomes. Similar to the results of Graetz and Michaels (2018), the paper finds that industrial robots are associated with a decline in employment of low-skill workers. The paper also finds industrial robots are associated with declines in the employment of medium-skill workers, young people (age 15-29), female workers, and manufacturing sector workers. However, the study also finds industrial robots have a positive relationship with high-skill, male, and service industry employment.

Anton et al. (2020), Klenert, Fernandez-Macias and Anton (2020) and Chiacchio, Petropoulos and Pichler (2018) each examine the influence of changes in robot exposure upon European employment. Each paper however comes to distinctly different general conclusions. Anton et al. (2020) analyses the impact of European robot adoption between 1995-2015, and finds that the influence of robotics has changed over time. A small negative general effect is uncovered for the period 1995-2005, while a positive general effect is found for the 2005-2015 period. Over the entire period of analysis, the paper finds only a small and ambiguous effect. Klenert, Fernandez-Macias and Anton (2020) also studies the 1995-2015 period, assessing the effect of robot adoption on employment in Europe at the industry level. The paper however finds a positive general robot-employment effect over the entire period of analysis. The paper further dismisses that there exists a relatively poor effect of robots upon low-skilled employment as compared to the effect on general employment. Chiacchio, Petropoulos and Pichler (2018) again assesses robot-employment effect using data from 6 European Union nations which together account for 85.5% of all industrial robots in the EU. The paper finds a strong negative general robot-employment effect. The negative effect was found to be exacerbated for both young workers, and workers with 'middle' education.

Both Compagnucci et al. (2019) and Jung and Lim (2020) also conduct multi-national analyses, with a scope limited primarily to developed nations. Compagnucci et al. (2019) analyses a set of 16 OECD nations over the period 2011-2016 at the industry level, finding an increase in robot use to reduce the growth rate of employment. Jung and Lim (2020) uses data from 42 nations over the period 2001-2017, and confirms the analysis of Compagnucci et al. (2019), finding high robot use to be associated with a reduction in employment growth rate. The analysis further finds high robot use to reduce the proportion of low-skilled labour employment.

Carbonero, Ernst, and Weber (2020) and de Vries et al. (2020) both analyse the robot-employment effect across both low and high income nations, to give a more globally generalisable estimate over the periods 2005-2015 and 2005-2014 respectively. Carbonero, Ernst, and Weber (2020) utilises data from 43 countries, and generalises its conclusions to a world-wide scale, finding a negative global robot-employment effect. The paper finds heterogeneity in the effect between nations, estimating a small negative employment effect in developed nations, in contrast to the relatively large negative effect discovered in developing nations. In contrast, de Vries et al. (2020) utilises data from 19 industries and 37 countries, but does not find any significant general relationship between robot adoption and employment, yet it does still find some interesting niche relationships. Analytic jobs employment sees a positive influence from robot adoption, while routine manual jobs are found to experience declines in employment due to robotics. Such insights are found to be robust to a series of potentially disruptive control variables (i.e. the found relationships do not appear to suffer from omitted variable bias).

Fu, Bao, Xie, & Fu (2021) extends the scope of analysis further, covering 74 nations between 2004-2016. The scope of the paper gives useful insight into the difference in the employment effect of industrial robots between developed and developing nations. The paper opposes the findings of studies such as Acemoglu and Restrepo (2020) and Aghion, Antonin, & Bunel

(2019), finding that industrial robots are generally associated with significant gains in employment in developed nations. The paper does not find any significant general robot-employment relation related to developing nations, although individuals with at least a ‘middle’ level of education within such nations do benefit from increases in industrial robot use in terms of employment.

Finally, while most papers connect domestic robot usage to domestic employment, some papers look at the impact of robots usage abroad on domestic employment. Faber (2020) extends the model presented by Acemoglu and Restrepo (2020), to estimate the influence of robots upon Mexican labour markets. The paper however further recognises the potential influence of robots employed in the United States upon local labour markets, given the powerful influence the demands of the United States have over Mexican (alongside many other nations) labour activity. Although there does not appear to be a relationship between changes in domestic robot use and Mexican employment, there does in fact appear to be a strong negative relationship between US robots and Mexican employment, presumably as the United States automates tasks that were previously exported from Mexico. The effect was found to be highly robust, confirmed by testing for pre-trends and producing estimates controlling for many potentially disruptive covariates. Stemmler (2019) examines the effect of both foreign and domestic robots upon employment in Brazilian labour markets. The paper finds a generally negative effect of domestic robots upon employment, meanwhile foreign robots are found to have some distinct effects in certain industries, particularly in the manufacturing sector (negative relationship) and the mining sector (positive relationship). Kugler, Kugler, Ripani and Rodrigo (2020) assesses specifically the effect of US robotics upon

Columbian labour markets. US robots are found to have a negative influence on employment in Colombian industries which have high robot use in their corresponding US industries. The negative effect is exacerbated for women, older workers and workers employed in small and medium sized enterprises. Further, the employment effect of US robots is found to be most pronounced within Columbian labour markets that export the most to the US, suggesting a replacement of Columbian export industry employment for US robots.

Summarising this literature review, there is a wide range of studies that attempt to estimate the impact of robots on human employment using data from different countries and data at different levels of aggregation. Moreover, one can find studies that show positive, negative and null effects. The variation in study designs and outcomes calls for a meta analysis.

II.4.B

Estimate sensitivity

How robust are the findings within our coded set of papers to changes in specification and circumstance? In the following section, I seek to recognise cases of papers coded for this analysis which show high within study sensitivity, and further examine the consistency in estimates between papers which conduct their studies in similar environments and using similar methodologies (i.e. between study sensitivity).

Faber (2020) provides an example of a paper whose' results appear particularly sensitive to changes in specification, reporting a positive effect in 62 estimates of the robots-employment effect, and a negative effect in 203. Although the abstract of the paper claims *"US robots have a sizeable negative impact on employment in Mexico".* The significant mixture of positive estimated employment effects within this study suggests that this general claim made by the paper is not particularly robust, even though the average effect found is indeed negative. Stemmler (2019), in their analysis of the automation effect in Brazilian labour markets, produces 234 negative estimates of the employment effect of some measure of robots, but also finds a positive effect in 173 estimates, however this paper does not make any clear statement of the general employment effect of robots in Brazil. Similar trends are noticeable in, but not limited to Dauth et al. (2017/2018), Giuntella and Wang (2019), and Dekle (2020).

I can further assess estimate sensitivity by making comparisons within the literature between relatively similar studies. I should expect that the results of similar studies are not significantly different from each other. Five studies used within this meta analysis produce estimates of the effect of robots upon US employment, namely Acemoglu and Restrepo (2020), Borjas and Freeman (2019), Micco (2019), Sequeira, Garrido and Santos (2020) (2020) & Anelli, Giuntella and Stella (2019). Partial correlation coefficient[2] averages of each of these papers' estimated robot-employment effects are -0.2188, -0.02026, -0.0266, -0.01762 and -0.00725

[2] Measure the strength and direction of a relationship, varying between [-1,1] – discussed further in section **II.3.b**

respectively. There are clearly some notable differences in the effect size estimates of these papers.

Interestingly, Anelli, Giuntella and Stella (2019) produce their estimates using the same regional exposure to robots as used in Acemoglu and Restrepo (2020). A key difference in the nature of the two papers however, is the time period over which their respective analyses are conducted, with the latter examining the period 1993-2007, while the former focuses on the more modern and perhaps more relevant time period of 2005-2016. Further, Anelli, Giuntella and Stella (2019) provides very general analysis of the explicit relationship of interest, providing only 2 usable regressions, relative to the 349 different specifications tested in Acemoglu and Restrepo (2020). It appears as though the general relationship uncovered in Acemoglu and Restrepo (2020) is sensitive to these contextual and specification changes, given the sizable difference in average PCC estimates between the two papers (-0.2188 for Acemoglu and Restrepo (2020), and -0.00725 for Anelli, Giuntella and Stella (2019))

Sequeira, Garrido and Santos (2020) frames itself as a *"replication exercise"* of Acemoglu and Restrepo (2020), using its own model to determine whether the results found in the latter are heavily dependent on its own specifications. For these reasons, comparisons of these two papers are particularly useful in assessing the fragility of the results found in Acemoglu and Restrepo (2020). A key conclusion of this paper is that the estimates of Acemoglu and Restrepo (2020) are not robust to the introduction of a squared robot term as an explanatory variable. Sequeira, Garrido and Santos (2020) claims there is "compelling evidence" of a 'U' shaped robot-employment relationship (i.e. marginal increases in robotization cause a slowing of the rate of employment decreases due to robots prior to the bottom of the ‘U’, and increase the rate of employment increase due to robots beyond this point), suggesting the findings in Acemoglu and Restrepo (2020) suffer from some degree of functional form/ omitted variable bias. Therefore, the general conclusions of Acemoglu and Restrepo (2020) are to at least some degree contingent on the assumption of the shape of the robot-employment relationship. This claim is backed by the sizable difference in our average PCC estimates of Acemoglu and Restrepo (2020) and Sequeira, Garrido and Santos (2020), which are -0.2188 and -0.01762 respectively.

Micco (2019) utilises the same robot measure as used in Acemoglu and Restrepo (2020) to analyse US employment, using a broad proxy for sector-level robot penetration, instrumented using average robot penetration in 15 EU nations, but analyses a more recent time period (2004-2016, as compared to 1993-2007). Like Anelli, Giuntella and Stella (2019), the scope of the relevant analysis in Micco (2019) is limited, providing only three usable regression items. I find again substantial differences in the average PCC estimates of these two papers, -0.2188 for Acemoglu and Restrepo (2020) and -0.0266 for Micco (2019), once again suggesting the results of Acemoglu and Restrepo (2020) are quite sensitive to changes in specifications and contexts.

Borjas and Freeman (2019) provide one final analysis of the robot-employment effect specific to the United States, examining the robot-employment effect over a longer time period, and using a different methodology from that of Acemoglu and Restrepo (2020). Once again, I find substantial differences in average estimated PCCs (-0.2188 for Acemolgu and Restrepo (2020), and -0.02026 for Borjas and Freeman)), highlighting the potential sensitivity of empirical estimates to changes in specification and context, even when studies are attempting to uncover the same general relationship.

It is clear that estimate sensitivities exist both within and between studies, each estimate frames the problem in a different way, and hence finds different effect sizes. Through a meta analysis, I can discover, and estimate the size of the factors that might explain these differences in estimated outcomes. In the next section, I provide more detail on meta analysis itself.

II.4.C

IFR dataset

The papers used in this meta analysis make frequent use of the International federation of Robotics (IFR) dataset (25 of 32 papers). The dataset seeks to provide a global overview on both industrial[3] and service[4] robot use at the country, application, and industry level. Given the large size of this dataset, and its potential to be applied directly to empirical problems related to robot use, the IFR dataset is evidently popular among researchers in related fields. Given the heavy influence of this particular dataset on the results of both our coded papers

---

[3] *Industrial robot defined as automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes*

[4] *Service robot defined as a robot **that performs useful tasks for humans or equipment excluding industrial automation applications***

which use the IFR dataset, and results of the following meta analysis, it is worth acknowledging some criticism of the quality of data provided by this dataset.

Several papers used in this analysis take issue with various shortcomings of the IFR dataset. Acemoglu and Restrepo (2020) notes the fact that the IFR provides 19 industry classifications that a robot may fall into, yet "*About 30% of robots are unclassified".* To account for this issue, Acemoglu and Restrepo (2020) distribute unclassified robots proportional to the current industry-robot association percentages of already classified robots. Clearly, associating robots to industries in this way is not ideal, and will lead to some degree of errors-in-variables bias. Further, Acemoglu and Restrepo (2020) note that although the IFR begins reporting US robot data from 1993, the dataset does not provide industry classifications until 2004, further hindering the authors' ability to provide detailed industry-specific analysis. Further, Acemoglu and Restrepo (2020) criticise the fact that US-specific robot stock data is not provided, instead aggregating to North American robot stock, potentially making their analysis vulnerable to heterogenous robot use practices in Canada and Mexico manipulating US robot-employment effect size estimates. The authors however disregard this as a reason for genuine concern, given the extreme dominance of robot stock held in the US relative to the remainder of all other North American nations. Giuntella and Wang (2019) also make note of the coarseness of the IFR dataset. In particular they find the sector classifications (i.e. employment types) used to be excessively broad, preferring a more disaggregated dataset, given that such limitations do not allow for examinations of the effect of robots within smaller sub-sectors. Additionally, Giuntella and Wang (2019) criticise the IFR dataset for not providing within-country robot distributions. Instead, the dataset provides robot stock for each given nation (or set of nations), thereby requiring within-country analyses to be dependent on some distribution calculation, which will be inaccurate to at least some degree.

Kariel (2021) notes that their analysis is limited to only the employment effect of industrial robots in the UK given that within the IFR dataset, there is: *"…no adequate data available on services robots for the UK".* Further, Ni and Obashi (2021) state that Japanese robot data reported before the year 2001 appears to have been "substantially manipulated" in order to obtain reported figures, additionally mentioning that the IFR offers no explanation as to the reason for this apparent manipulation. Dekle (2020) addresses an additional issue of using IFR data to address Japan-specific issues; noting that the definition of a robot used by the IFR does not align with the much broader "Japanese definition" of a robot. For this reason, Japanese IFR data will generally report a relatively low robot stock, compared to what would be reported under a hypothetical dataset using a Japanese robot definition. This issue is likely to result in Japanese robot-employment effect estimates showing a more dramatic response to robots relative to if a Japanese robot definition dataset was used. Borjas and Freeman (2019) recognise that analysis of the effect of different robot types on employment is not possible given the data is aggregated into a single robot stock measure. Hence, any analysis conducted using the IFR dataset may only make estimates of the effect of robots in general, potentially concealing some interesting individual robot type effects.

Finally, Klump et al. (2021) specifically sets out to examine the applications and limitations of the IFR dataset. The paper criticises the fact that the useful life/depreciation timeline of 12 years for robot stock used by the IFR is not in line with standard economic literature, a notion which the IFR itself recognises needs *"further investigation"*. The paper suggests a mean global robot stock depreciation rate of about 7% annually. The paper also points out that the dataset provides no way of measuring the quality or usefulness of any robot stock, in other words, every robot unit has an identical weight, irrespective of quality or usefulness of the unit. This omission means that analysis controlling for the technological progress or value of robots is not possible. Further, the dataset does not account for robots that are not sold on the open market. The paper notes the case of Amazon Robotics, who do not actually sell any robots, but rather supplies warehouse robotics (200,000+ service robots) exclusively to Amazon warehouses. Such omissions may lead to significant errors in robot-employment effect estimates, exacerbating the effect size of robots which are accounted for.

The paper points to even more issues. First, prior to 2001, Japanese industrial robot stock was overstated, due to the inclusion of 'dedicated industrial robot' stock on top of multipurpose industrial robot stock. Secondly, there exists methodology deviations in both Japanese and Russian data from the standard IFR robot stock calculations. Klump et al. (2021) therefore notes that data from both these countries is "not consistent over time" and "difficult to use in econometric analyses" due to data inconsistencies relative to other countries. Further inconsistencies are also present in the datasets for: Austria, Taiwan, the Republic of Korea, and Australia, but are able to be corrected and included in econometric analyses.

Finally, Klump et al. (2021) makes a comparison between the IFR dataset and the Comtrade dataset, who report annual country-specific number of robots imported net of re-exports figures. One would expect these values to be similar to the corresponding annual robot installations values in the IFR dataset. However, the study finds that in general, Comtrade reports higher values than the IFR (particularly in the case of Malaysia, where Comtrade net imports are approximately 50 times greater than robot installations reported in the IFR dataset). The significant difference between the two datasets suggests that the results found in empirical studies can be heavily dependent on the dataset selected. Further, by reporting consistently lower robot values in the IFR dataset, any effect estimate found using the IFR dataset is likely to be exacerbated relative to what would have been found should the estimate have made use of the Comtrade dataset.

It is clear that the data included in the IFR dataset can have substantial degrees of measurement error. This may undermine the validity of the results of empirical studies using IFR data, and the results of this meta analysis, however, I cannot comment on the extent to which this is true.

II.5

Meta Analysis

II.5.A

Overview, use, and purpose.

Meta analysis, or the process of systematically and statistically taking review of quantitative research in a given field, is the obvious method for conducting the study at hand. Havránek et al. (2020) interprets meta analysis itself as *"a conventional tool for research synthesis*"; additionally stating that "*Research studies published in the most eminent economics journals and structural models employed by central banks now routinely rely on previously published meta-analyses…".* Meta analysis is now common practice in economic literature as a means of verifying findings of individual papers, potentially dismissing outlier relationships, and providing approximations to the true broad effect of interest that would otherwise be hidden behind a series of quantitatively disconnected literature.

Poot (2012) provides evidence of the dramatic upturn in the use of meta analysis in economic literature since 1980, citing 626 such studies that have been conducted between 1980-2012 (With an average growth rate on number of studies conducted year on year of 18% according to Stanley and Doucouliagos (2011)), the majority of these being published between 2005 and 2012. The first instance of meta analysis being used in a published journal was in Pearson (1904), in a paper titled *"Report on certain enteric fever inoculation statistics"* which quantitatively aggregated a number of prior conducted clinical reports related to typhoid inoculation. Popularity was not brought to this method of analysis in the field of Economics however, until the highly influential paper, Stanley and Jarrell (1989), provided a *"quantitative methodology for reviewing the empirical economic literature.".* The paper provides a basis through which meta analysis within economics research could be conducted, and thereby increased understandability and interpretability of subsequent meta analysis research. Stanley (2001) builds upon Stanley and Jarrell (1989), providing a "*…framework for discussing the strengths and weaknesses of meta-analysis",* which meta analysis researchers might refer to when considering the objectivity of their study, and in making refinements to their respective analyses.

Utilizing Stanley's framework, many economic researchers have sought to develop meta regression analyses in their own given field of interest. Sverke et al. (2002) provides a meta analysis of the consequences of job insecurity on the individual. The paper covers 72 studies to uncover the response of job attitudes, organizational attitudes, health, and behavioural relationships related to the organisation that a given individual is employed to in the face of job insecurity. Sverke et al. (2002) forms a quantitative overview of all relevant estimates, exposing heterogeneous effects ("…*consequences of insecurity are more detrimental among manual, as compared with nonmanual, workers*."), and in this case, allowing the authors to make general claims on the effects of interest ("…*job insecurity has detrimental consequences for employees' job attitudes, organizational attitudes, health, and, to some extent, their behavioural relationship with the organization*"). In addition to these claims, such meta research allows the authors to make recommendations for subsequent empirical analyses (i.e. Which estimation methodologies appear to be the most effective? Which specification most accurately describes the relation of interest? What control variables are most important?) Sverke et al. (2002) gives their recommendations for both selecting variables to proxy for job insecurity ("*scales capturing fear or worry of job loss best reflect the conceptual definition of job insecurity"),* and for outcome variables which are most informative to the reader (*"future studies could preferably address how job insecurity relates to job attitudes such as work intensity").*

Schaefer et al. (2016) provides an interesting parallel to the analysis at hand, in a meta analysis examining the contributing factors toward changes in human trust of automation. The paper finds human-related and automation-related factors to have moderately positive effects on human-automation trust development. Most significantly, it finds a moderate to high trust effect of both emotional and behavioural factors. Through in its overview, the paper identifies significant omissions in the literature (In particular, 'appearance-based anthropomorphism' and its relation to perceived age and gender of the robot or automation unit, and a 'three-factor model of trust with design and training'), which may be useful in more accurately determining the causes human-automation trust levels. Identification of such omissions in the literature is a key component of the value of meta-research, allowing subsequent researchers to confidently develop models which minimise the chance of omitted variable biases.

Of great significance to the analysis at hand is the aforementioned study conducted by Terzidis et al. (2019), which is perhaps the most directly relevant literature currently available to us. The study performs a meta-analysis on the effect of general automation and trade upon employment, using data from 77 papers, which provided 1158 estimates. The key finding of the study being *"Automation is beneficial at the firm level, and is more likely to displace low-skilled employment".* The paper gives commentary on the degree of heterogeneity that exists between and within multiple subsets of the general population (while an individual empirical study typically focuses its analysis on only a single subset of the population), and further recognises the non-uniformity of estimates at both the worker level, and firm level. Additionally, Terzidis et al. (2019) notes the heterogeneity that exists at the national level, acknowledging that labour market reactions that occur between nations which possess "*different labor market institutions"* are unlikely to occur in a consistent manner. Heterogeneity at the skill level is also recognised, bringing attention to the inconsistency in the reaction of labour markets with differing replaceability levels to automation. It is through meta analysis that such inconsistent effects are revealed, and hence provide the reader with a more complete view of the estimated effects of interest within a number of different contexts.

Three key benefits of meta-analysis, beyond the discovery of some broad underlying effect estimate are revealed to us through this discussion. The first, being that meta analysis can help researchers in identifying significant omissions of variables perceived to have high explanatory power. This allows researchers to minimise the degree of omitted variable bias attached to included variables within their models. Similarly, such research can be useful in effectively forming specifications which are precise in explaining the outcome effect of interest; thereby allowing researchers to consistently develop models with variables which are likely to have high explanatory power. Finally, and most significantly, meta analysis allows for highly detailed commentary on the heterogeneity imposed by selected factors on the outcome variables of interest. This is essentially the primary value of meta analysis; researchers within the field of meta research are interested in understanding how the context (i.e. geographies, industries, countries, regions, individuals etc) of data affects the reaction of the outcome variable of interest to some general explanatory variable.

II.5.B

Partial correlation coefficients

A key issue of meta-analysis is combining estimated effects across studies such that they are comparable even when different variable measurements are used within and between studies (e.g. dependent variable of the ratio of employed persons to working age population vs a simple count of employed persons). A technique to neutralise such irregularity is required. Partial correlation coefficients[5] (PCC) provides a measure of the strength and direction of the relationship between two variables and allows us to make basic comparisons on the size of the effect of interest. PCCs are given by the following formula:

$$PCC_i = \frac{t_i}{\sqrt{df_i + t_i^2}} \quad ({}^{6})$$

and can vary between [-1,1], where PCC=1 is a perfectly positive correlation, PCC=-1 is a perfectly negative relationship, and PCC=0 meaning there is no discernible linear relationship. Interpreting PCCs requires some degree of subjectivity, but I can still follow some rules such that our analysis is comparable to others. Doucouliagos (2011) used over 22,000 empirical effect sizes in economics literature and developed their own rule of thumb. Using percentile distributions at 25%, 50% and 75%, the paper defines the minimum limit for small, medium

[5] Where 'partial' refers to one continuous independent variable and one continuous dependent variable, as opposed to many potential controls that exist in actual specifications

[6] Where ti is the t-statistic of a given estimate, and dfi is the degrees of freedom related to a given estimate

and large effects at respective PCC values of 0.07, 0.17 and 0.33. The remainder of this paper uses these figures as a rule of thumb for indicating effect strength.

III

Methods of analysis

III.1

Random Effects and Fixed Effect models

A potential method of estimating the robot-employment effect is taking a simple average of every PCC value coded in our study. This is the same as using Ordinary Least Squares (OLS) to regress PCC values on a constant:

$$PCC_i^{\varepsilon} = \mu + e_i \quad i = 1,2,\dots,N$$

Where μ is the true robot-employment effect and N is the number of coded effect sizes. While the estimate will be unbiased[7] and consistent[8], so long as our coded set of estimates is representative of the true population, our estimates will however be inefficient.[9]

An alternative to this is a Fixed effect Weighted Least Squares (WLS) estimate. Such a model relies on the assumption that there exists one true effect size (in this case, the robot-employment effect) that each of the estimates collected in our dataset seeks to estimate. If

[7] Expected value of estimator equal to the true value of the population effect size of interest
[8] Convergence to true population effect size as sample size increases
[9] On average, the estimator does NOT approximate the true population with as much precision as some other estimator (i.e. does not minimize estimate variance).

there truly exists only one true effect, all deviations from the true effect are simply a result of sampling error[10]; if this is true, the more precisely[11] an estimate is given, the more likely that it will closely estimate the true effect since the more precise an estimate, the lower the potential sampling bias (i.e. the sample is more likely to accurately represent the population). This is the additional (in addition to OLS) mechanism of the WLS Fixed Effect weighting scheme, which applies weight to estimates as a function of their respective precision (FE1 in table 1). If there is one true population effect, the Weighted least squares (WLS) estimate under fixed effect weighting will produce asymptotically unbiased, consistent and efficient estimates of the true population effect (Borenstein et al. (2010)). WLS is applied as follows:

$$\frac{PCC_i^{\varepsilon}}{\omega_i} = \frac{\mu}{\omega_i} + \frac{e_i}{\omega_i} \quad i = 1,2, \dots, N$$

Where ω_i is the weight associated with each estimate, as calculated in table 1.

This study also provides random effects weightings, which follows the assumption that there exists a series of sub-population true effects as opposed to a single true population effect size. To account for this, the random effects version of WLS uses the incorporates the value t^2 into its weighting (RE1 in table 1) to represent the variance (with a standard deviation of t) of the distribution of estimated sub-population effect. This is simply a measure of heterogeneity between estimates not attributable to sampling error (Borenstien et al. (2010)). As t increases, and the field of literature further deviates from the 'single true population effect' the less weighting should be applied on the basis of estimate precision, and

[10] If a population-representative sample were collected under these circumstances, the estimate would be exactly the true population effect.
[11] Precision refers to the estimate distribution (i.e. inverse standard error size)

vice versa, given precision values indicate only the efficiency of a specification to estimate an effect related to its own population.

Further, I provide variants of the fixed and random effects our weighting schemes, which seek to recognise that estimates from particular papers may be over-represented in our meta estimates, problematic when there exists multiple true sub-population effects, or a large number of estimates from imprecise/low quality studies which may be disruptive for our overall estimates. To account for this, two augmented versions of FE1 and RE1 (FE2 and RE2, both given in table 1), reduce the weighting applied to any given estimate as the number of coded estimates from study from which it is derived ($n_{i\in s}$)[12] increases.

Finally, I also apply a funnel asymmetry test (FAT) to each of our WLS models to test for publication bias. Publication bias occurs when researchers intentionally manipulate the estimates reported in a study in order to satisfy some desired characteristic (typically statistical significance). If I do not control for publication bias, the validity of this meta-analysis is threatened. I can include these controls by observing the relationship that exists between estimate effects and their respective standard errors, calculated:

$$SE(PCC_i) = \sqrt{\frac{1-(PCC_i^2)}{df_i}}$$[13]

[12] Where *n* is a number of estimates, and *s* is a given study

[13] df_i refers to the degrees of freedom of a given estimate

Should publication bias exist, I expect there to be some relationship between effect sizes and their respective standard errors, the more imprecise (i.e. larger standard errors) the estimator, the more extreme estimates must be produced to achieve statistically significant estimates. As an estimator becomes more precise, less extreme estimates are required to achieve statistically significant results as standard errors tighten. Hence, in the case of publication bias, I expect to see a positive relationship between estimated effect size, and their related standard errors. I conduct the FAT by running each of our WLS weighting variants, but also controlling for the standard errors associated with each PCC estimate:

$$\frac{PCC_i^{\varepsilon}}{\omega_i} = \frac{\mu}{\omega_i} + \frac{e_i}{\omega_i} + \beta \frac{SE(PCC_i)}{\omega_i} \quad i = 1,2,\ldots,N$$

Table 1

Weight Calculations	
Weight	**Method**
Fixed Effect Weight 1 (FE1)	$SE(PCC_i)$
Random Effects Weight 1 (RE1)	$\sqrt{SE(PCC_i)^{2} + t^{2}}$
Fixed Effect Weight 2 (FE2)	$SE(PCC_i) \cdot \sqrt{n_{i\in s}}$
Random Effect Weight 2 (RE2)	$\sqrt{SE(PCC_i)^{2} + t^{2}} \cdot \sqrt{n_{i\in s}}$

Weights as given in Duan et al. (2020)

III.2

Bayesian Model Averaging

Our WLS regressions from section III.1 have the purpose of estimating generalised robot-employment outcome (and comment on publication bias), however it seems unlikely that this relationship is constant under all circumstances, but what are the circumstances that may cause this relationship to change? Meta analysis allows us to explore this relationship in much greater detail. By coding a number of potentially relevant characteristics (discussed in greater detail in section IV.1), I can use meta analysis techniques to attempt to uncover which of these factors may exist within the true robot-employment relationship, and the size of their influence.

Bayesian model averaging (BMA) involves estimating a randomly selected (using Monte Carlo Markov Chain (MCMC) sampling) set of all possible specifications given our coded control variables and averaging the estimated coefficients associated with each item within each specification, weighted by posterior probabilities (i.e. the probability that a given specification actually occurs).

The method (Steel (2011)) treats the model specification as a random variable, and uses the data available to conduct inferences. To describe the dataset (y), consider all the specifications that may exist given the coded data[14] $S_i, i = 1,2 \ldots I$, grouped in the space δ. In order to give a Bayesian model (a model that draws its inferences from the posterior distribution) of the problem, I specify a prior $P(S_i)$ on δ, the data will then lead to a posterior $P(S_i | y)$. I then use this posterior to determine the posterior model probabilities for each MCMC sampled specification of being true (i.e. the chance a given specification is the actual specification that describes the robot-employment relationship). As described in section III.1, I then run a series WLS regressions of each of the MCMC sampled specifications, before averaging the estimated effect of each coded variable under each specification, weighted by the posterior model probability that the specification from which a given variables' effect size estimate was derived from is accurate. The estimated effect of any given variable through BMA is therefore a probability weighted combination of WLS estimated variable effects.

III.3

Backwards Stepwise Regressions

Using the Bayesian information criterion or Schwartz information criterion (Wit et al. (2012)), I utilise a backwards stepwise regression procedure. The model works by removing variables from a specification containing all coded variables until the lowest possible BIC/SIC value is returned. In other words, until I have a model the variables most likely to exist in the true

[14] more detail on our coding process given in section IV.1

specification according to BIC/SIC. I can then run standard WLS regressions which control for these variables to determine their effect size.

BIC/SIC provides an index value for model selection, the lower the value, the better the proposed model is said to fit the true model. The equation for BIC/SIC values are given:

$$BIC = k(ln\,(n)\,) - 2ln\,(L^{\varepsilon})$$

Where k is the number of explanatory variables, n is the sample size, and L^{ε} is the maximised value of the likelihood function of the specification. Minimising the BIC corresponds to maximizing the posterior model probability, as discussed in section III.2.

I use this backwards stepwise regression, alongside a WLS specification which controls for all coded controls as a supplement for our BMA analysis. Further, I have the ability to lock in variables to our equation which I believe have particularly high explanatory value. In section IV.4, I conduct 3 separate analyses which both present a WLS specification including all variables, alongside a backward stepwise regression which locks in some variables which have an influence on the robot-employment effect that appears of particular interest given our findings in our prior analyses (i.e. BMA and WLS).

IV

Analysis

IV.1

Data characteristics

The primary interest of the present study is in discerning some overarching broad employment effect resulting from robot use. Our analysis will be conducted in the form of a quantitative review of all found literature which produces some original econometric estimation of the robot-employment relation. In obtaining the data for this study, I have attempted to collect all relevant estimates (2438 estimates) found throughout this field of study, hence, inferences made in this section are relevant to the entire branch of literature. For each of these estimates, I code a number of specification variables, which are defined in Table 2.

Table 2

Variable	**Code Description**			
	1	0	Average	No. of 1's
Exoutlier	Outliers excluded from sample	No outliers excluded	0.1493	364
Fore	Estimate controls for foreign robot exposure	Foreign robot exposure not controlled for	0.2937	716
Reglev	Analysis at regional level	Analysis not at regional level	0.5939	1448
Invlev	Analysis at individual level	Analysis not at individual level	0.1144	279

Conlev	Analysis at country level	Analysis not at country level	0.1509	368
Indlev	Analysis at industry level	Analysis not at industry level	0.4143	1010
HS	High skill specific employment	Not high skill specific employment	0.0279	68
MS	Medium skill specific employment	Not medium skill specific employment	0.0090	22
LS	Low skill specific employment	Not low skill specific employment	0.0418	102
Manufacturing	Manufacturing specific employment	Not manufacturing specific employment	0.1522	371
Automotive	Automotive specific employment	Not automotive specific employment	0.0209	51
Services	Services specific employment	Not services specific employment	0.0459	112
Allareas	Non-specific employment	Specified employment	0.5287	1289
TFF	Time fixed effects included	No time fixed effects included	0.4959	1209
IndivFF	Individual fixed effects included	No individual fixed effects included	0.0431	105
IndusFF	Industry fixed effects included	No industry fixed effects included	0.2760	673
CountryFF	Country Fixed effects included	No country fixed effects included	0.1226	299
RegFF	Region Fixed effects included	No regional fixed effects included	0.7317	1784

PopCon	Controls for population size	Does not control for population size	0.4549	1109
GenCon	Controls for gender share	Does not control for gender share	0.5915	1442
AgeCon	Controls for age share	Does not control for age share	0.4733	1154
EducCon	Controls for education shares	Does not control for education shares	0.5980	1458
Ethcon	Controls for ethnicity shares	Does not control for ethnicity shares	0.3052	744
OccCon	Controls for occupation shares	Does not control for occupation shares	0.5094	1242
SkiCon	Controls for skill level shares	Does not control for skill level shares	0.0500	122
CapCon	Controls for capital exposure	Does not control for capital exposure	0.2416	589
ImpCon	Controls for import exposure	Does not control for import exposure	0.4926	1201
WeightCon	Uses some weighting scheme	Does not use some weighting scheme	0.5406	1318
IV	Uses instrumental variable estimation method	Does not use instrumental variable estimation method	0.4516	1101
NIFR	Does not use IFR robot data	Uses IFR robot data	0.1550	378

Table 2 also reveals some interesting quantitative information about the robot-employment effect literature. I notice a strong tendency for estimates to be conducted at least partially at either the regional (59.39% of estimates) or industry level[15] (41.43% of estimates), while a relatively small proportion of estimate are conducted at the both the individual level (11.44%) and country level (15.09%), revealing a tendency of researchers to seeks seek some ‘middle ground’ between very fine, and very coarse sets of data.

There appears to be some substantial interest in analysing the specific robot-employment effect of specific skill levels, with a collective 7.87% of the estimates being specifically related to employment of a particular skill level (low, medium or high). Similarly, there is a strong interest in understanding the effect of robots upon employment specific to some given fields of labour. I coded for manufacturing, services and automotive specific employment. 21.90% of estimates are specifically related to at least one of these fields, with a particularly strong interest in manufacturing employment, which is the focus of 15.22% of estimates.

Many estimates also control for fixed effects. As might be expected for a branch of literature which almost exclusively uses panel data, many estimates control for time fixed effects (49.59%). Further, 73.17% of estimates control for regional fixed effects, to account for potential heterogenous employment effects across the regions of their respective analyses.

[15] Collective proportions of the analysis level code does not sum to 100% due to multi-level analyses

Further, many papers and estimates attempt to control for variables that may have some significant influence on the robot-employment relationship. There appears to be a strong belief within the literature that population values, gender shares, age shares, education shares, ethnicity shares, occupation shares, and import exposure (which are controlled for in 45.49% 59.15%, 47.33%, 59.80%, 30.52% 50.94% and 49.26% of estimates respectively) may have some substantial influence on the robot-employment effect, such that they are worth controlling for to avoid omitted variable biases.

Only a small proportion (15.5%) of estimates do not make use of the IFR dataset in their estimates. The literature is clearly highly dependent on this dataset, so it will be worth attempting to uncover the effect of this dataset in our quantitative analyses. Section II.4.C is dedicated to discussing some of the criticisms of this data, which may undermine the validity of both previous empirical research, and this meta study.

Finally, to make generalised claims of an effect, it is necessary that the related literature explores the effect of interest broadly. Where significant gaps in research exist, the general claims that I eventually make must be restricted. Of the 32 papers collected for this study, 22 focus their analysis specifically on the effect within a single nation, while 9 of these papers conduct their estimates for effects in either Japan or the United States exclusively (4 Japanese, 5 US). The remaining 10 studies collected do not limit themselves to only a single nation, but do typically place some restrictions on the breadth of the data they collect, often focussing their respective efforts on relatively high income nations (e.g. European/ OECD/ Developed nations). Exceptions are de Vries et al. (2020), Fu et al. (2021), and Carbonero et al. (2020) which all utilise data from a wider national income range.

There appears to be a majority interest in the literature of the effect of robotics within highly developed nations (likely due to the fact these nations are the most highly robotised, and hence have the most relevant data available), and only limited literature which seeks to analyse the effect for developing nations. Hence, I am cautious in making globally generalised claims from this study. Each of the papers coded in our analysis are listed in table 3.

Table 3

Author	Release date	Area of analysis
Acemoglu and Restrepo	2017	United States
Anelli, Giuntella and Stella	2019	United States
Borjas and Freeman	2020	United States
Micco	2019	United States
Sequeira, Garrido and Santos	2020	United States
Adachi, Kawaguchi and Saito	2020	Japan
Dekle	2020	Japan
Eggleston, Lee and Iizuka	2021	Japan
Ni and Obashi	2021	Japan
Acemoglu, Lelarge and Restrepo	2020	France
Aghion, Antonin and Bunel	2019	France
Camiña, Díaz-Chao and Torrent-Sellens	2020	Spain
Koch, Manuylov and Smolka	2019	Spain
Dauth, Findeisen, Suedekum and Woessner	2017	Germany
Dauth, Findeisen, Suedekum and Woessner	2018	Germany
Giuntella and Wang	2019	China

Stemmler	2019	Brazil
Dottori	2020	Italy
Kariel	2021	United Kingdom
Faber	2020	Mexico
Kugler, Kugler, Ripani and Rodrigo	2020	Columbia
Dixon, Hong and Wu	2020	Canada
Anton, Klenert, Fernandez-Macias, Brancati and Alaveras	2020	General (European nations)
Blanas, Gancia and Lee	2019	General (European nations)
Chiacchio, Petropoulos and Pichler	2018	General (European nations)
Klenert, Fernandez-Macias and Anton	2020	General (European nations)
Compagnucci, Gentili, Valentini and Gallegati	2019	General (OECD nations)
Carbonero, Ernst, and Weber	2020	General (Developed and Emerging economies)
de Vries, Gentile, Miroudot and Wacker	2020	General (High income and emerging economies)
Fu, Bao, Xie and Fu	2021	General (Developed and developing economies)
Graetz and Michaels	2018	General (Developed economies
Jung and Lim	2020	General (Developed economies

Next, I present some general summary statistics of the outcome variables of interest, presented in table 4. Our PCC estimates have an extremely wide distribution, with a range of [-0.927,0.744], and mean and median PCC values of -0.027 and -0.012 respectively. Thus, the distribution of estimated effects is very wide, but the average effect is negligible. Our t statistics tells a similar story, with a distribution of [-15.6, 95.9], and a mean and median of −0.6 and -0.7 respectively. Such a wide distribution of estimates creates instability when running our fixed effect and random effects models. To account for this, and to avoid significant outliers driving our results, I omit the most extreme 1% of these PCC values in all future analysis. Table 4 also presents summary statistics for our main outcome variables both before and after this restriction is applied.

Table 4

	T statistic Unrestricted	T statistic Restricted	PCC Unrestricted	PCC Restricted
Mean	-0.588	-0.845	-0.027	-0.026
Median	-0.746	-0.746	-0.012	-0.012
Minimum	-15.571	-15.571	-0.927	-0.507
Maximum	95.883	26.000	0.743	0.416
SD	5.561	3.290	0.145	0.113
1st quartile	-2.750	-2.725	-0.088	-0.083
3rd quartile	1.257	1.211	0.021	0.019
1%	-9.715	-9.823	-0.507	-0.313
99%	9.246	8.117	0.418	0.338

The following graphs (figures 2-5) are a visual representation of our estimated effect sizes both prior and post omitting these extreme values. It is clear visually that the general structure of the dataset has remained essentially identical, but has simply shortened the broad tails that exist on these distributions.

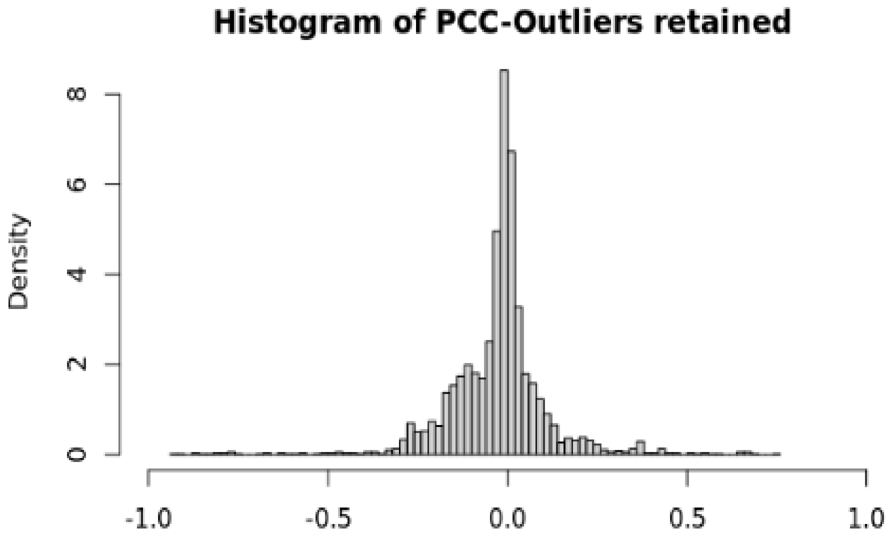

Figure 2

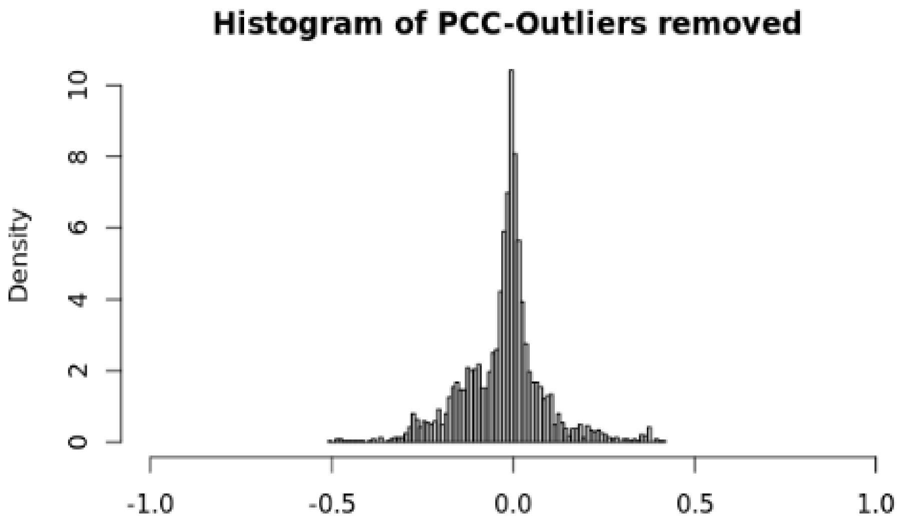

Figure 3

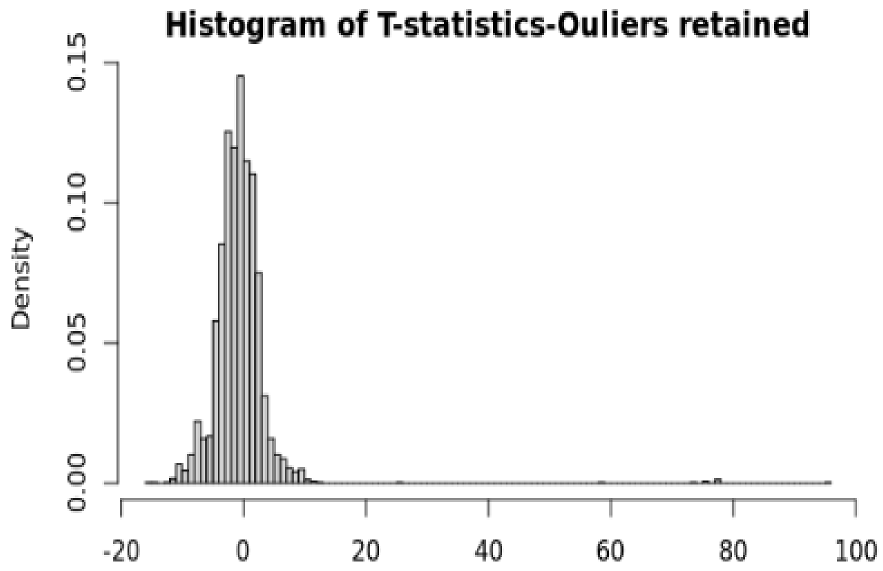

Figure 4

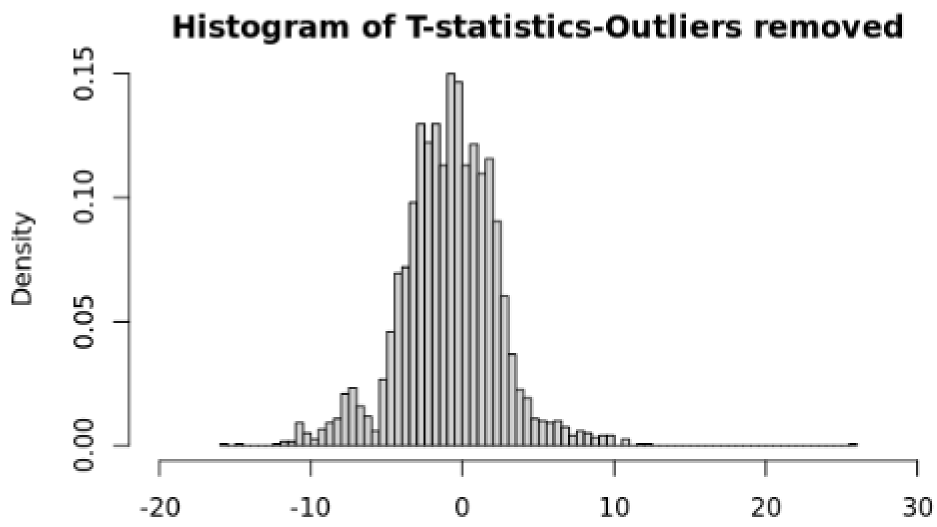

Figure 5

IV.2

Weighted Least Squares

Described in greater detail in section III.1, I present our Fixed Effect weighted WLS estimates in Table 5. Such a model relies on the assumption that there exists one true robot-employment effect size (i.e. there is no heterogenous sub-population effect sizes) that each of the estimates collected in our dataset seeks to estimate. This key assumption of Fixed Effect weighting is not realistic in this context, however. As discussed in section IV.1, the literature collected covers a diverse range of nations, regions, time periods, and employment types. It is not reasonable to assume that there exists only one true robot-employment effect that applies exactly to every form of labour. Given this, the results of a Fixed Effect estimate can be seriously misinformative due to its pure precision-based weighting. In our case for example, estimates formed in Kugler et al. (2020) make use of an extremely high number of observations, causing these estimates to be extremely precise. This essentially results in the fixed effect estimator being almost entirely dependent (99.22% weighting under FE1) on results from Kugler et al. (2020), a study which specifically estimates the short run robot-employment effect in Columbia.

Table 5

Fixed effect and Random effects estimates				
Variable	FE1	FE2	RE1	RE2
No publication bias correction (Panel A)				
Constant	-0.0000764	-0.0000688	-0.0271	-0.00909
	(0.0000119)	(0.0000173)	(0.0020)	(0.01427)
Publication bias correction (Panel B)				
Constant	0.00000536	-0.0000769	-0.0233	0.00544
	(0.0000671)	(0.0000473)	(0.0193)	(0.01538)
SE PCC	-0.848	0.0838	-0.1326	-0.46264
	(0.725)	(0.545)	(0.3755)	(0.42878)

Cluster robust standard errors given in brackets

Since the Fixed effect model is clearly not a useful estimator for this meta-analysis, I seek an alternative weighting for our WLS estimator. The random effects model, also described in greater detail in III.1, is an obvious choice, given its capacity to account for a series of sub-populations, each with their own respective true robot-employment effect.

For the sake of completeness, I present Fixed Effect estimates of the robot-employment effect in table 5, but for reasons mentioned previously, I cannot use these estimates to make any economically sound claims. The random effects estimates produced in Panel A of table 5 produce mixed results. RE1 suggests a highly statistically significant negative result, but yet the estimated effect size falls below the minimum threshold for a small effect as suggested by Doucouliagos (2011). RE2 of the random effects estimates additionally fails to achieve statistical significance, and similarly produces an estimated effect size falling below the minimum threshold for a small effect. Further, these panel A estimates do not control for potential publication bias in the literature.

In panel B of table 5, I reproduce the Fixed Effect and Random Effects estimations as in Panel A, but include standard errors as a control variable in our WLS model. This is known as a Funnel Asymmetry Test (FAT). Given the expected association between effect sizes and standard errors under publication bias (described in section III.1), this method both serves as a control for publication bias, and as a simple test for the existence of publication bias itself.

Again, Fixed Effect estimates are produced in panel B of table 5, but cannot be used for making economically sound claims. Our random effects estimates using FAT appear ineffective in uncovering some genuine robot-employment effect. Under both specifications, I find statistically insignificant estimates. Further, both estimated effect sizes fall below the minimum threshold for a small effect as suggested by Doucouliagos (2011). In other words, these tests indicate there is no general robot-employment effect The results also suggest that there does not appear to be a statistically significant presence of publication bias within the literature, further backed by the observation that our Random Effects estimates for the robot-employment effect change very little between the cases where I do, and do not control for publication bias. In other words, panel A estimates do not suffer from omitted variable bias due to not controlling for publication bias.

IV.3

Bayesian Model Averaging

So far, I have made claims of the general robot-employment effect, but as seen in our distributions t-statistic and PCC distributions, there exists much variation between individual estimates and studies. Are these differences the result of random variability or heterogeneity-causing factors between estimates? Understanding the sources of differences in estimated effects between estimates and studies is a primary value of meta-analysis. Bayesian model averaging allows us to investigate these sources of heterogeneity. The method works by estimating all possible specifications given a set of collected control variables before averaging the coefficients associated with each variable, weighted by the posterior probabilities (i.e. the probability that a given specification actually occurs) of each respective specification. Such a method attempts to recognise the uncertainty associated with every plausible scenario, and incorporate this uncertainty into our estimates, with the disadvantage that no one specific specification is examined. In this case however, due to the high number of variables, BMA uses a Monte Carlo Markov chain to sample from the set of all possible specifications. I present both FE1 and RE1 estimates using BMA in table 4.

Posterior inclusion probability (PIP) refers to the probability that a given variable is part of the actual relationship that exists. A value of 0 suggests that the variable certainly does not exist in the actual relationship, while a value of 1 suggests the opposite. Posterior mean and standard error (Post Mean and Post SD) refer to the weighted (weighted by the likelihood value of each given specification) average of a variable's estimated coefficients and standard errors values respectively. Finally, positive sign (Pos Sign) refers to the likelihood weighted probability that the relationship that a given explanatory variable holds with the outcome variable is positive. A value of 0 suggests a certainly negative relationship, while a value of 1 suggests a certainly positive relationship. Of course, the value of the suggestions made by these latter three outcomes are contingent on the value of the PIP.

Table 6

Variable	**BMA analysis**							
	FE1				RE1			
	PIP	Post mean	Post SD	Pos Sign	PIP	Post Mean	Post SD	Pos Sign
NIFR	***1.0000***	***0.0188***	***0.0038***	***1.0000***	0.9995	0.0221	0.0045	1.0000
Reglev	*1.0000*	*-0.0223*	*0.0027*	*0.0000*	*1.0000*	*-0.0473*	*0.0056*	*0.0000*
Invlev	0.9874	0.0155	0.0041	1.0000	***1.0000***	***0.0370***	***0.0069***	***1.0000***
Conlev	0.0281	0.0001	0.0013	0.9696	0.9949	0.0603	0.0136	0.0000
Indlev	0.9851	0.0138	0.0034	1.0000	***1.0000***	***0.0552***	***0.0057***	***1.0000***
Manufacturing	0.0246	0.0000	0.0000	1.0000	0.8861	0.0162	0.0075	1.0000
Automotive	0.1081	-0.0000	0.0001	0.0000	0.0605	-0.0011	0.0054	0.0000
Services	0.0271	0.0000	0.0003	1.0000	0.8947	0.0261	0.0118	1.0000

Table 6

Variable	BMA analysis							
	FE1				RE1			
	PIP	Post mean	Post SD	Pos Sign	PIP	Post Mean	Post SD	Pos Sign
NIFR	***1.0000***	***0.0188***	***0.0038***	***1.0000***	0.9995	0.0221	0.0045	1.0000
High Skill	0.0653	-0.0006	0.0025	0.0006	0.0203	0.0000	0.0016	0.3822
Medium Skill	0.1268	-0.0042	0.0126	0.0000	0.0503	-0.0012	0.0066	0.0000
Low Skill	0.8752	-0.0149	0.0073	0.0000	0.9977	-0.0039	0.0090	0.0000
All Areas	0.0224	-0.0000	0.0000	0.0000	0.1030	-0.0010	0.0034	0.0000
TFF	0.0227	-0.0000	0.0001	0.1548	***1.0000***	***0.0296***	***0.0046***	***1.0000***
InvFF	0.1224	-0.0002	0.0006	0.0056	0.0215	-0.0001	0.0016	0.0687
IndusFF	0.0588	-0.0001	0.0003	0.0091	0.0319	-0.0002	0.0012	0.0000
CountryFF	0.0376	0.0002	0.0018	1.0000	0.9943	0.0556	0.0131	1.0000
RegFF	0.8172	0.0019	0.0011	1.0000	0.0280	0.0001	0.0010	0.9923
PopCon	*1.0000*	*-0.0949*	*0.0053*	*0.0000*	*1.0000*	*-0.0839*	*0.0057*	*0.0000*
GenCon	***1.0000***	***0.0870***	***0.0071***	***1.0000***	***1.0000***	***0.0547***	***0.0087***	***1.0000***
AgeCon	0.0430	-0.0002	0.0010	0.0000	***1.0000***	***0.1020***	***0.0071***	***1.0000***
EducCon	0.9297	0.0169	0.0067	1.0000	0.2421	-0.0044	0.0087	0.0000
EthCon	*1.0000*	*-0.1080*	*0.0046*	*0.0000*	*1.0000*	*-0.1110*	*0.0067*	*0.0000*
OccCon	0.6043	-0.0058	0.0055	0.0000	0.0826	-0.0008	0.0032	0.0000
SkiCon	0.0224	-0.0002	0.0023	0.0000	0.0317	-0.0003	0.0025	0.0000
CapCon	0.0704	-0.0001	0.0004	0.0000	0.0951	-0.0009	0.0031	0.0000
ImpCon	0.0224	-0.0000	0.0000	0.0000	0.0246	-0.0001	0.0009	0.0053
WeightCon	0.9982	-0.0036	0.0010	0.0000	*1.0000*	*-0.0335*	*0.0040*	*0.0000*
IV	0.9687	-0.0141	0.0040	0.0002	0.0248	0.0000	0.0009	0.9337

Table 6

Variable	BMA analysis							
	FE1				RE1			
	PIP	Post mean	Post SD	Pos Sign	PIP	Post Mean	Post SD	Pos Sign
NIFR	***1.0000***	***0.0188***	***0.0038***	***1.0000***	0.9995	0.0221	0.0045	1.0000
NonIV	0.9688	-0.0140	0.0039	0.0007	0.0227	-0.0001	0.0009	0.0220
ExOutlier	0.0391	-0.0000	0.0000	0.0000	0.0609	0.0005	0.0022	1.0000

Table 6 reports the results of BMA using both fixed and random effects weights. Cases which achieve a PIP value of 1 and a conditional positive sign of 1 (exists in the true specification and have a positive relationship) are highlighted in bold italics, while items that achieve a PIP of 1 and a conditional positive sign of 0 (cases which exist in the true specification and have a negative relationship) are highlighted in italics.

As suggested previously, I am unable to make use of Fixed Effect estimates to inform our results but are provided for the sake of completeness. 9 items satisfy both having a PIP value of 1 and a positive sign value of either 1 or 0 under random effects weighting, namely: Reglev, Invlev, Indlev, TFF, PopCon GenCon, AgeCon, EthCon and WeightCon. Our data level items: Reglev, Invlev, Indlev (Regional, Individual and Industry respectively), which satisfy both of the previously mentioned criteria, suggest that the level at which analysis is conducted affects the PCC of a robot-employment relationship estimate by values of -0.0473, 0.0370 and 0.0552 respectively, relative to a random effects WLS estimate conducted at some other level (i.e. a level not coded for). Although there appears to be some notable differences in the effect sizes of these data level variables, none of these estimates meet the minimum threshold for a small effect according to Doucouliagos (2011). Time fixed effects also exist in the true specification according to BMA analysis. Estimates controlling for time fixed effects have higher PCCs, by a value of 0.0296, again, failing to meet the threshold for a small effect size.

The random effects specification of our BMA analysis finds a number of controls which exist in the true specification; namely: PopCon, GenCon, AgeCon, EthCon and WeightCon (Population, Gender, Age, Ethnicity and Weight control respectively) which affect PCCs values by -0.0839, 0.0547, 0.1020, -0.1110 and -0.0335 respectively. Doucouliagos (2011) guidelines suggest PopCon (negative), AgeCon (positive) and EthCon (negative) estimated effects meet the criteria for having a small influence (absolute size of at least 0.07) on the robot-employment relationship.

Our variable for the IFR dataset comes extremely close to a PIP value of 1 (.9995) in our random effects BMA analysis, in other words, a variable for IFR/non-IFR data almost certainly exists in the true specification. Further, our analysis suggests that the effect of not estimating using the IFR data is certainly positive (i.e. Using IFR data decreases the effect size of the robot-employment relationship). The estimated effect in our analysis however does not meet the Doucouliagos (2011) criteria for a small effect, giving an effect size estimate of 0.0221.

Finally, I notice that the method by which our estimates are made does not appear to have any significant influence on effect estimates. Our random effects BMA analysis suggests that a variable for IV estimates does not exist within the true specification, and further gives an extremely small effect size estimate, well below the Doucouliagos (2011) threshold for a small effect.

IV.4

Backwards Stepwise Regressions

Using the Bayesian information criterion (BIC) or Schwartz information criterion (SIC), I can utilise a backwards stepwise regression procedure, sequentially selecting the 'best' set of variables that return the lowest BIC/SIC values, in other words, the variables most likely to exist in the true specification. I conduct 3 backwards stepwise regressions which each lock in some variables of interest, alongside a WLS specification which controls for all coded controls as a supplement for our BMA analysis. The backwards stepwise regression is designed to find the specification most likely to be the actual relationship of interest, and then simply run this WLS specification.

In table 7 I perform backwards stepwise regression, locking in each of the coded levels at which analysis is made (i.e. Reglev, Invlev, Conlev, Indlev). For this specification (and those further below) I present results based on our 4 weighting schemes (FE1, FE2, RE1 and RE2) and further lock in standard errors of effect estimates as a means of controlling for publication bias.

Table 7

Variable	FE1	FE2	RE1	RE2
		All controls		
SE	0.41930	0.29015	-0.01207	-0.47927
	(0.34508)	(0.38079)	(0.22711)	(0.30413)
Reglev	-0.02391***	-0.02106**	-0.04733***	-0.03445*
	(0.00720)	(0.00739)	(0.01303)	(0.01641)
Invlev	0.01462	0.02215.	0.03614	0.04880*
	(0.00916)	(0.01247)	(0.02263)	(0.02343)
Conlev	-0.00561	-0.04311.	-0.05700	-0.05814.
	(0.02952)	(0.02237)	(0.03952)	(0.03528)
Indlev	0.01316	0.01506	0.05343.	0.04188
	(0.01003)	(0.01073)	(0.02981)	(0.03137)
		Backwards stepwise regression		
SE	0.00845	0.31014	0.00845	0.00845
	(0.29618)	(0.32194)	(0.22222)	(0.28615)
Reglev	-0.04734***	-0.02133**	-0.04734***	-0.04734**
	(0.00576)	(0.00766)	(0.01153)	(0.01618)
Invlev	0.03747***	0.01670	0.03747*	0.03747.
	(0.00880)	(0.01058)	(0.01792)	(0.01944)
Conlev	-0.05966.	-0.04284.	-0.05966	-0.05966.
	(0.03231)	(0.02290)	(0.03878)	(0.03452)
Indlev	0.05366***	0.01266.	0.05366*	0.05366*
	(0.00833)	(0.00752)	(0.02462)	(0.02720)

Cluster robust standard errors given in brackets
(.)Statistical significance at 10% level
(*) Statistical significance at 5% level
(**)Statistical significance at 1% level
(***)Statistical significance at 0.1% level

In table 8, I lock in each of those control variables which were found to exist in the true specification, have a relationship of a definite direction, and have a strong enough relationship to qualify for at least a weak effect according to Doucouliagos (2011) (i.e. Popcon, Agecon and Ethcon).

Table 8

Variable	FE1	FE2	RE1	RE2
	All controls			
SE	0.41930	0.29015	-0.01207	-0.47927
	(0.34508)	(0.38079)	(0.22711)	(0.30413)
Popcon	-0.09334**	-0.07889**	-0.08059*	-0.06921*
	(0.02886)	(0.02507)	(0.03328)	(0.03333)
Agecon	-0.00183	0.00215	0.10143**	0.12334**
	(0.00742)	(0.00853)	(0.03562)	(0.04589)
Ethcon	-0. 10759***	-0.09423***	-0.01620**	-0.11297**
	(0.02766)	(0.02357)	(0.02516)	(0.03705)
	Backwards stepwise regression			
SE	0.00845	0.28768	0.00845	0.00845
	(0.40157)	(0.33609)	(0.22222)	(0.28615)
Popcon	-0.08359**	-0.07913***	-0.08359**	-0.08359**
	(0.02932)	(0.02316)	(0.02725)	(0.03163)
Agecon	0.09789***	0.00278	0.09789**	0.09789*
	(0.00662)	(0.00728)	(0.03260)	(0.04429)
Ethcon	-0.10759***	-0.09514***	-0.10759**	-0.10759**
	(0.02786)	(0.02403)	(0.03294)	(0.03959)

Cluster robust standard errors given in brackets
(.)Statistical significance at 10% level
(*) Statistical significance at 5% level
(**)Statistical significance at 1% level
(***)Statistical significance at 0.1% level

Finally, in table 9 I lock in the control for IFR data (i.e. NIFR). Although random effects BMA does not suggest NIFR certainty exists within the true specification (PIP= 0.9995), the IFR dataset is clearly highly influential in this literature, so is worth examining here.

Table 9

Variable	FE1	FE2	RE1	RE2
		All controls		
SE	0.41930	0.29015	-0.01207	-0.47927
	(0.34508)	(0.38079)	(0.22711)	(0.30413)
NIFR	0.01739.	0.02055	0.02234	0.03588
	(0.00963)	(0.01342)	(0.01654)	(0.02380)
		Backwards stepwise regression		
SE	0.00845	0.31014	0.00845	0.00845
	(0.39492)	(0.32194)	(0.22222)	(0.28615)
NIFR	0.02344**	0.01927.	0.02344	0.02344
	(0.00860)	(0.01089)	(0.01689)	(0.02277)

Cluster robust standard errors given in brackets
(.)Statistical significance at 10% level
(*) Statistical significance at 5% level
(**)Statistical significance at 1% level
(***)Statistical significance at 0.1% level

Are the estimate sizes of the BMA analysis, all-variable regressions, and backwards stepwise regressions in alignment when using random effects weighting? BMA analysis suggests Reglev, Invlev and Indlev should exist within the true specification, but also suggests effect size estimates which are below the Doucouliagos (2011) small effect minimum threshold. In table 7, all-variable regressions, and backwards stepwise regressions follow this claim, with none of our coded analysis level variables estimated to be beyond the minimum threshold for a small effect under both our alternative random effects weighting schemes. The results from our alternative fixed effect weighting schemes are primarily given for the sake of interest, but are also in alignment with this claim. As claimed in our BMA analysis, I find in table 8, each of the effect sizes of Popcon, Agecon and Ethcon are beyond the minimum threshold for a small effect (negatively, positively, and negatively respectively) under both random effects weighting schemes, in both our all-variable specification and our backwards stepwise regression (bar Ethcon in the all-variable specification under RE1). Further, in each of these cases, statistical significance of at least a 5% level is achieved.

Finally, in table 9 I again test the importance of using the IFR dataset on estimated robot-employment effect sizes. Our BMA analysis reported an effect of the IFR dataset not beyond the minimum threshold for a small effect according to Doucouliagos (2011). Both our all-variable specification and backward stepwise regression find similar results under both Random effects weighting schemes, and both Fixed effect weighting schemes.

In Section IV.2, WLS found no significant presence of publication bias under either of the random effects weighting schemes. Further, no model, under the all-variable specification, or under a backwards stepwise regression finds a statistically significant estimate for the SEPCC coefficient, aligning with our previous suggestion that there exists no significant evidence of publication bias within the literature.

V

Conclusion

Acemoglu and Restrepo (2020), has sparked a significant interest in analysing the effect of the use of robots on labour employment, reviving the old debate of whether automation will destroy more jobs than it creates. Their study has been followed by a series of papers, each attempting to analyse this effect for different labour markets, and using different specifications. From 32 of such studies, I have collected 2438 individual estimates of the use of robots on human employment. This study quantitatively aggregates and analyses all of this empirical literature in an attempt to uncover a generalised robot-employment effect, assess influential factors and variables, and comment on the degree of publication bias within the literature. The key finding of this paper is that the literature suggests that, in general, there exists no sizable effect of robotics use upon labour employment. Further, I find that there exists no significant presence of publication bias within the literature. In other words, there does not appear to be a tendency of publishers to manipulate their reported effect estimates in order to satisfy some condition. These conclusions are confirmed using a wide variety of analytical techniques, including weighted least squares, Bayesian model averaging and backward stepwise regressions.

Using random effects weights applied to Bayesian model averaging, a backwards stepwise regression, and a weighted least squares regression containing all coded variables, I also analyse both the likelihood of some potentially significant factors existing in the 'true' robot-employment specification, and the effect size of these factors. I find controls for population size, age shares and ethnicity shares are very likely to be in the true specification and have some significant (but also relatively small) influence (negative, positive, and negative respectively) on the size of the robot-employment effect. Hence, this suggests that it is a good idea to control for such variables in order to avoid issues of omitted variable bias.

According to Bayesian Model Averaging, I find that controls for level at which analysis is conducted (regional level, individual level, country level or industry level), alongside both the time at which analysis is conducted (time fixed effects), and the data with which the analysis is conducted with (IFR/non-IFR) are likely to be in the true specification of the robot-employment effect. However, I also find that none of these factors meet the minimum threshold, according to Doucouliagos (2011) for having a small influence on the general robot-employment estimate. Interestingly, our random effects BMA analysis also suggests that a variable for the estimate method used (IV/Non-IV technique) does not exist within the true robot-employment specification and further gives an extremely small effect size estimate.

Throughout the literature, there is extensive use of the International Federation of Robotics dataset, which provides data on robotics use across countries and continents. There have been several issues uncovered related to this dataset, which may be damaging to the validity of some of our coded estimates, and hence our meta-analysis itself. This study notes these potential problems, but the nature of the study means that little can be done to mitigate these issues. I further find a tendency of authors to focus their analysis on developed nations. Hence, the claims made in this paper are mainly generalisable to developed nations.

VI

Bibliography

Acemoglu, & Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics, 4(B),* 1043-1171. doi:https://doi.org/10.1016/S0169-7218(11)02410-5

Acemoglu, & Restrepo (2020, June). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy, 128*(6). doi:https://doi-org.ezproxy.canterbury.ac.nz/10.1086/705716

Acemoglu, Lelarge, & Restrepo (2020). Competing with Robots: Firm-Level Evidence from France. *AEA Papers and Proceedings, 110*, 383-388. doi:10.1257/pandp.20201003

Adachi, Kawaguchi, & Saito (2020). Robots and Employment: Evidence from Japan, 1978-2017. *Discussion papers, 20051*. Retrieved from https://www.rieti.go.jp/jp/publications/dp/20e051.pdf

Aghion, Antonin, & Bunel (2019). Artificial intelligence, growth and employment: The role of policy. *Economie et Statistique*(510-512), 149-164. doi:10.24187/ecostat.2019.510t.1994

Anelli, Colantone, & Stanig (2019). We were the robots: Automation and voting behavior in western europe. *BAFFI CAREFIN Centre Research Paper, (2019-115).* doi:http://dx.doi.org/10.2139/ssrn.3419966

Anelli, Giuntella, & Stella (2019). Robots, Labor Markets, and Family Behavior. *IZA Discussion Papers, 12820*. Retrieved from https://docs.iza.org/dp12820.pdf

Anton, Klenert, Fernandez-Macias, Brancati, & Alaveras (2020). The labour market impact of robotisation in Europe. *JRC Working Papers on Labour, Education and Technology, 2020-06*. Retrieved from https://ec.europa.eu/jrc/sites/default/files/jrc121388.pdf

Arntz, Gregory, Zierahn, & OECD / Directorate for Employment, Labour and Social Affairs. (2016). The risk of automation for jobs in OECD countries: A comparative analysis. *OECD Publishing.* doi:https://doi.org/10.1787/5jlz9h56dvq7-en

Baldwin, & Shultz (1955). Automation: A New Dimension to Old Problems. *Monthly Labor Review*, *78*(2), 165–169. http://www.jstor.org/stable/41831145

Bessen (2019). Automation and jobs: When technology boosts employment. *Economic Policy*, *34*(100), 589-626. doi:https://doi.org/10.1093/epolic/eiaa001

Blanas, Gancia, & Lee (2019). Who is afraid of machines? *Economic Policy, 34*(100), 627-690. doi:10.1093/epolic/eiaa005

Borenstein, Hedges, Higgins, & Rothstein. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 97-111. doi:10.1002/jrsm.12

Borjas, & Freeman (2019). From immigrants to robots: The changing locus of substitutes for workers. *Russell Sage Foundation Journal of the Social Sciences, 5*(5), 22-42. doi:10.7758/rsf.2019.5.5.02.pdf

Camiña, Díaz-Chao, & Torrent-Sellens (2020). Automation technologies: Long-term effects for Spanish industrial firms. *Technological Forecasting and Social Change, 151*. doi:10.1016/j.techfore.2019.119828

Carbonero, Ernst, & Weber (2020). Robots Worldwide: The Impact of Automation on Employment and Trade. *Gender Economics, 224602*. Retrieved from https://www.econstor.eu/bitstream/10419/224602/1/vfs-2020-pid-40084.pdf

Cellan-Jones (2019). Robots 'to replace up to 20 million factory jobs' by 2030. *BBC*. Retrieved from https://www.bbc.com/news/business-48760799

Chiacchio, Petropoulos, & Pichler (2018). The impact of industrial robots on EU employment and wages: A local labour market approach. *Bruegel Working Papers, 2018/02*. Retrieved from http://hdl.handle.net/10419/207001

Compagnucci, Gentili, Valentini, & Gallegati (2019). Robotization and labour dislocation in the manufacturing sectors of OECD countries: a panel VAR approach. *Applied Economics, 51*(57), 6127-6138. doi:10.1080/00036846.2019.1659499

Dauth, Findeisen, Südekum, & Woessner (2017). German Robots - The Impact of Industrial Robots on Workers. *CEPR Discussion Papers, 12306*. Retrieved from http://www.cepr.org/active/publications/discussion_papers/dp.php?dpno=12306

Dauth, Findeisen, Suedekum, & Woessner (2018). Adjusting to Robots: Worker-Level Evidence. *Opportunity and Inclusive Growth Institute Working Papers, 13*. doi:10.21034/iwp.13

Dekle (2020). Robots and industrial labor: Evidence from Japan. *Journal of the Japanese and International Economies, 58*. doi:10.1016/j.jjie.2020.101108

Dixon, Hong, & Wu (2020). The Employment Consequences of Robots: Firm-level Evidence. *Statistics Canada*. Retrieved from https://content.tcmediasaffaires.com/LAF/lacom2019/robots.pdf

Dottori (2020). Robots and employment: evidence from Italy. *Questioni di Economia e Finanza, 572*. Retrieved from https://www.bancaditalia.it/pubblicazioni/qef/2020-0572/QEF_572_20.pdf

Doucouliagos (2011). How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics. *Deakin University, Department of Economics.*

Retrieved from https://www.deakin.edu.au/__data/assets/pdf_file/0006/897351/2011_5.pdf

Duan, Das, Meriluoto, & Reed (2020). Estimating the effect of spillovers on exports: A meta-analysis. Review of World Economics, 156(2), 219-249. doi: https://doi.org/10.1007/s10290-020-00377-z

Eggleston, Lee, & Iizuka (2021). Robots and Labor in the Service Sector: Evidence from Nursing Homes. *NBER Working Papers, 28322*. Retrieved from http://www.nber.org/papers/w28322.pdf

Faber (2020). Robots and reshoring: Evidence from Mexican labor markets. *Journal of International Economics, 127*. doi:10.1016/j.jinteco.2020.103384

Frey, & Osborne. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change, 114*, 254-280. doi:10.1016/j.techfore.2016.08.019

Fu, Bao, Xie, & Fu (2021). Diffusion of industrial robotics and inclusive growth: Labour market evidence from cross country data. *Journal of Business Research, 122*, 670-684. doi:10.1016/j.jbusres.2020.05.051

Fuei (2017). Automation, Computerization and Future Employment in Singapore. *Journal of Southeast Asian Economies*, *34(2)*, 388–399. doi:https://doi.org/10.1355/ae34-2h

Giuntella, & Wang (2019). Is an Army of Robots Marching on Chinese Jobs? *IZA Discussion Papers, 12281*. Retrieved from https://docs.iza.org/dp12281.pdf

Graetz, & Michaels (2018). Robots at Work. *The Review of Economics and Statistics, 100*(5), 753-768. doi:https://doi-org.ezproxy.canterbury.ac.nz/10.1162/rest_a_00754

Havránek, Stanley, Doucouliagos, Bom, Geyer-Klingeberg, Iwasaki, . . . Aert, v. (2020). Reporting Guidlines for Meta-Analysis in Economics. *Journal of Economic Surveys, 34*(3), 469-475. doi:https://doi.org/10.1111/joes.12363

Im, Mayer, Palier, & Rovny (2019). The “losers of automation”: A reservoir of votes for the radical right? *Research & Politics*. doi:https://doi.org/10.1177/2053168018822395

Jung, & Lim (2020). Industrial robots, employment growth, and labor cost: A simultaneous equation analysis. *Technological Forecasting and Social Change*(159). doi:10.1016/j.techfore.2020.120202

Kariel (2021). Job Creators or Job Killers? Heterogeneous Effects of Industrial Robots on UK Employment. *The Journal of Industrial Relations, 35*(1), 52-78. doi:https://doi-org.ezproxy.canterbury.ac.nz/10.1111/labr.12192

Keynes (1930). Economic possibilities for our grandchildren: published in Keynes (1972), also freely available online

Kelly (2020). U.S. Lost Over 60 Million Jobs—Now Robots, Tech And Artificial Intelligence Will Take Millions More. *Forbes*. Retrieved from https://www.forbes.com/sites/jackkelly/2020/10/27/us-lost-over-60-million-jobs-now-robots-tech-and-artificial-intelligence-will-take-millions-more/?sh=4333e4951a52

Klenert, Fernandez-Macias, & Anton (2020). Do robots really destroy jobs? Evidence from Europe. *JRC Working Papers on Labour, Education and Technology*. Retrieved from https://ec.europa.eu/jrc/sites/default/files/jrc118393.pdf

Klump, Jurkat, & Schneider (2021). Tracking the rise of robots: A survey of the IFR database and its applications. Retrieved from https://mpra.ub.uni-muenchen.de/110390/1/MPRA_paper_110390.pdf

Koch, Manuylov, & Smolka (2019). Robots and Firms. *CESifo Working Paper, 7608*. doi:http://dx.doi.org/10.2139/ssrn.3377705

Kugler, Kugler, Ripani, & Rodrigo (2020). U.S. Robots and their Impacts in the Tropics: Evidence from Colombian Labor Markets. *NBER Working Papers, 28034*. Retrieved from http://www.nber.org/papers/w28034.pdf

Leontief, & Duchin. (1984). The Impacts of Automation on Employment, 1963-2000. *Institute for Economic Analysis, New York University*. Retrieved from https://files.eric.ed.gov/fulltext/ED241743.pdf

Micco (2019). Automation, Labor Markets, and Trade. *Working Papers, wp486*. Retrieved from http://www.econ.uchile.cl/uploads/publicacion/39055f6e8a17858ef533f025562bc30d4ea20506.pdf

Ni, & Obashi (2021). Robotics technology and firm-level employment adjustment in Japan. *Japan and the World Economy, 57*. doi:https://doi.org/10.1016/j.japwor.2021.101054

Pearson (1904). Report on Certain Enteric Fever Inoculation Statistics. *British medical journal*, *2*(2288), 1243–1246. Retreived from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2355479/pdf/brmedj08194-0039.pdf

Poot (2012). Learning from the Flood of Findings: Meta-Analysis in Economics. *National Institute of Demographic and Economic Analysis*. Retrieved from https://www.nzae.org.nz/wp-content/uploads/2012/09/NZAE-2012-Poot-meta-analysis.pdf

Pouliakas (2018). Determinants of automation risk in the EU labour market: A skills-needs approach. *Institute of Labor Economics*. Available at SSRN 3253487.

PRISMA (2021). Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Retrieved from http://www.prisma-statement.org/

Roux (2018) Automation and employment: The case of South Africa. *African Journal of Science, Technology, Innovation and Development, 10(4)*, 507-517, doi: 10.1080/20421338.2018.1478482

Schaefer, Chen, Szalma, & Hancock (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. *Human Factors*, *58*(3), 377–400. https://doi.org/10.1177/0018720816634228

Sequeira, Garrido, & Santos (2021). Robots are not always bad for employment and wages. *International Economics, 167*(C), 108-119. doi:10.1016/j.inteco.2021.06.001

Stanley (2001). "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives*, *15 (3).* 131-150. doi: 10.1257/jep.15.3.131

Stanley, & Doucouliagos (2011). Meta-Regression Analysis in Economics and Business (1st ed.). *Routledge*. doi: https://doi.org/10.4324/9780203111710

Stanley, & Jarrell. (1989). Meta-Regression Analysis: A Quantitative Method of Literature Surveys. *Journal of Economic Surveys, 3*(2), 161-170. doi: https://doi.org/10.1111/j.1467-6419.1989.tb00064.x

Steel (2011). Bayesian Model Averaging and Forecasting. *Bulletin of EU and US Inflation and Macroeconomic Analysis*, 200, 30-41. Retrieved from https://warwick.ac.uk/fac/sci/statistics/staff/academic-research/steel/steel_homepage/publ/bma_forecast.pdf

Stemmler (2019). Does automation lead to de-industrialization in emerging economies? Evidence from Brazil. *Center for European, Governance and Economic Development Research Discussion Papers, 382*. Retrieved from https://www.econstor.eu/bitstream/10419/203326/1/1677065451.pdf

Sverke, Hellgren, & Näswall (2002). No security: A meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology, 7(3)*, 242-264. doi: https://doi.org/10.1037/1076-8998.7.3.242

Terzidis, Brakman, & Ortega-Argilés (2019). Labour Markets, Trade and Technological Progress: A Meta-Study. *CESifo Working Paper No. 7719*. doi: http://dx.doi.org/10.2139/ssrn.3421146

Wit, Heuvel, Jan-Willem, & Romeijn. (2012). ‘All models are wrong...’: an introduction to model uncertainty. *Statistica Neerlandica, 66*(3), 217-236. doi:10.1111/j.1467-9574.2012.00530.x

Zeira (1998). Workers, Machines, and Economic Growth. *The Quarterly Journal of Economics, 113(4)*, 1091–1117. doi:https://doi.org/10.1162/003355398555847