THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?*

Carl Benedikt Frey[†] and Michael A. Osborne[‡]

September 17, 2013

.

Abstract

We examine how susceptible jobs are to computerisation. To assess this, we begin by implementing a novel methodology to estimate the probability of computerisation for 702 detailed occupations, using a Gaussian process classifier. Based on these estimates, we examine expected impacts of future computerisation on US labour market outcomes, with the primary objective of analysing the number of jobs at risk and the relationship between an occupation's probability of computerisation, wages and educational attainment. According to our estimates, about 47 percent of total US employment is at risk. We further provide evidence that wages and educational attainment exhibit a strong negative relationship with an occupation's probability of computerisation.

I. INTRODUCTION

In this paper, we address the question: how susceptible are jobs to computerisation? Doing so, we build on the existing literature in two ways. First, drawing upon recent advances in Machine Learning (ML) and Mobile Robotics (MR), we develop a novel methodology to categorise occupations according to their susceptibility to computerisation.[1] Second, we implement this methodology to estimate the probability of computerisation for 702 detailed occupations, and examine expected impacts of future computerisation on US labour market outcomes.

Our paper is motivated by John Maynard Keynes's frequently cited prediction of widespread technological unemployment "due to our discovery of means of economising the use of labour outrunning the pace at which we can find new uses for labour" (Keynes, 1933, p. 3). Indeed, over the past decades, computers have substituted for a number of jobs, including the functions of bookkeepers, cashiers and telephone operators (Bresnahan, 1999; MGI, 2013). More recently, the poor performance of labour markets across advanced economies has intensified the debate about technological unemployment among economists. While there is ongoing disagreement about the driving forces behind the persistently high unemployment rates, a number of scholars have pointed at computer-controlled equipment as a possible explanation for recent jobless growth (see, for example, Brynjolfsson and McAfee, 2011).[2]

The impact of computerisation on labour market outcomes is well-established in the literature, documenting the decline of employment in routine intensive occupations – *i.e.* occupations mainly consisting of tasks following well-defined procedures that can easily be performed by sophisticated algorithms. For example, studies by Charles, *et al.* (2013) and Jaimovich and Siu (2012) emphasise that the ongoing decline in manufacturing employment and the disappearance of other routine jobs is causing the current low rates of employment.[3] In addition to the computerisation of routine manufacturing tasks, Autor and Dorn (2013) document a structural shift in the labour market, with workers reallocating their labour supply from middle-income manufacturing to low-income service occupations. Arguably, this is because the manual tasks of service occupations are less susceptible to computerisation, as they require a higher degree of flexibility and physical adaptability (Autor, *et al.*, 2003; Goos and Manning, 2007; Autor and Dorn, 2013).

At the same time, with falling prices of computing, problem-solving skills are becoming relatively productive, explaining the substantial employment growth in occupations involving cognitive tasks where skilled labour has a comparative advantage, as well as the persistent increase in returns to education (Katz and Murphy, 1992; Acemoglu, 2002; Autor and Dorn, 2013). The title "Lousy and Lovely Jobs", of recent work by Goos and Manning (2007), thus captures the essence of the current trend towards labour market polarization, with growing employment in high-income cognitive jobs and low-income manual occupations, accompanied by a hollowing-out of middle-income routine jobs.

According to Brynjolfsson and McAfee (2011), the pace of technological innovation is still increasing, with more sophisticated software technologies disrupting labour markets by making workers redundant. What is striking about the examples in their book is that computerisation is no longer confined to routine manufacturing tasks. The autonomous driverless cars, developed by Google, provide one example of how manual tasks in transport and logistics may soon be automated. In the section "In Domain After Domain, Computers Race Ahead", they emphasise how fast moving these developments have been. Less than ten years ago, in the chapter "Why People Still Matter", Levy and Murnane (2004) pointed at the difficulties of replicating human perception, asserting that driving in traffic is insusceptible to automation: "But executing a left turn against oncoming traffic involves so many factors that it is hard to imagine discovering the set of rules that can replicate a driver's behaviour [...]". Six years later, in October 2010, Google announced that it had modified several Toyota Priuses to be fully autonomous (Brynjolfsson and McAfee, 2011).

To our knowledge, no study has yet quantified what recent technological progress is likely to mean for the future of employment. The present study intends to bridge this gap in the literature. Although there are indeed existing useful frameworks for examining the impact of computers on the occupational employment composition, they seem inadequate in explaining the impact of technological trends going beyond the computerisation of routine tasks. Seminal work by Autor, *et al.* (2003), for example, distinguishes between cognitive and manual tasks on the one hand, and routine and non-routine tasks on the other. While the computer substitution for both cognitive and manual routine tasks is evident, non-routine tasks involve everything from legal writing, truck driving and medical diagnoses, to persuading and selling. In the present study, we will argue that legal writing and truck driving will soon be automated, while persuading, for instance, will not. Drawing upon recent developments in Engineering Sciences, and in particular advances in the fields of ML, including Data Mining, Machine Vision, Computational Statistics and other sub-fields of Artificial Intelligence, as well as MR, we derive additional dimensions required to understand the susceptibility of jobs to computerisation. Needless to say, a number of factors are driving decisions to automate and we cannot capture these in full. Rather we aim, from a technological capabilities point of view, to determine which problems engineers need to solve for specific occupations to be automated. By highlighting these problems, their difficulty and to which occupations they relate, we categorise jobs according to their susceptibility to computerisation. The characteristics of these problems were matched to different occupational characteristics, using O∗NET data, allowing us to examine the future direction of technological change in terms of its impact on the occupational composition of the labour market, but also the number of jobs at risk should these technologies materialise.

The present study relates to two literatures. First, our analysis builds on the labour economics literature on the task content of employment (Autor, *et al.*, 2003; Goos and Manning, 2007; Autor and Dorn, 2013). Based on defined premises about what computers do, this literature examines the historical impact of computerisation on the occupational composition of the labour market. However, the scope of what computers do has recently expanded, and will inevitably continue to do so (Brynjolfsson and McAfee, 2011; MGI, 2013). Drawing upon recent progress in ML, we expand the premises about the tasks computers are and will be suited to accomplish. Doing so, we build on the task content literature in a forward-looking manner. Furthermore, whereas this literature has largely focused on task measures from the Dictionary of Occupational

Titles (DOT), last revised in 1991, we rely on the 2010 version of the DOT successor O∗NET – an online service developed for the US Department of Labor.[4] Accordingly, O∗NET has the advantage of providing more recent information on occupational work activities.

Second, our study relates to the literature examining the offshoring of information-based tasks to foreign worksites (Jensen and Kletzer, 2005; Blinder, 2009; Jensen and Kletzer, 2010; Oldenski, 2012; Blinder and Krueger, 2013). This literature consists of different methodologies to rank and categorise occupations according to their susceptibility to offshoring. For example, using O∗NET data on the nature of work done in different occupations, Blinder (2009) estimates that 22 to 29 percent of US jobs are or will be offshorable in the next decade or two. These estimates are based on two defining characteristics of jobs that cannot be offshored: (a) the job must be performed at a specific work location; and (b) the job requires face-to-face personal communication. Naturally, the characteristics of occupations that can be offshored are different from the characteristics of occupations that can be automated. For example, the work of cashiers, which has largely been substituted by self- service technology, must be performed at specific work location and requires face-to-face contact. The extent of computerisation is therefore likely to go beyond that of offshoring. Hence, while the implementation of our methodology is similar to that of Blinder (2009), we rely on different occupational characteristics.

The remainder of this paper is structured as follows. In Section II, we review the literature on the historical relationship between technological progress and employment. Section III describes recent and expected future technological developments. In Section IV, we describe our methodology, and in Section V, we examine the expected impact of these technological developments on labour market outcomes. Finally, in Section VI, we derive some conclusions.

II. A HISTORY OF TECHNOLOGICAL REVOLUTIONS AND EMPLOYMENT

The concern over technological unemployment is hardly a recent phenomenon. Throughout history, the process of creative destruction, following technological inventions, has created enormous wealth, but also undesired disruptions. As stressed by Schumpeter (1962), it was not the lack of inventive ideas that set the boundaries for economic development, but rather powerful social and economic interests promoting the technological status quo. This is nicely illustrated by the example of William Lee, inventing the stocking frame knitting machine in 1589, hoping that it would relieve workers of hand-knitting. Seeking patent protection for his invention, he travelled to London where he had rented a building for his machine to be viewed by Queen Elizabeth I. To his disappointment, the Queen was more concerned with the employment impact of his invention and refused to grant him a patent, claiming that: "Thou aimest high, Master Lee. Consider thou what the invention could do to my poor subjects. It would assuredly bring to them ruin by depriving them of employment, thus making them beggars" (cited in Acemoglu and Robinson, 2012, p. 182f). Most likely the Queen's concern was a manifestation of the hosiers' guilds fear that the invention would make the skills of its artisan members obsolete.[5] The guilds' opposition was indeed so intense that William Lee had to leave Britain.

That guilds systematically tried to weaken market forces as aggregators to maintain the technological status quo is persuasively argued by Kellenbenz (1974, p. 243), stating that "guilds defended the interests of their members against outsiders, and these included the inventors who, with their new equipment and techniques, threatened to disturb their members' economic status."[6] As pointed out by Mokyr (1998, p. 11): "Unless all individuals accept the "verdict" of the market outcome, the decision whether to adopt an innovation is likely to be resisted by losers through non-market mechanism and political activism." Workers can thus be expected to resist new technologies, insofar that they make their skills obsolete and irreversibly reduce their expected earnings. The balance between job conservation and technological progress therefore, to a large extent, reflects the balance of power in society, and how gains from technological progress are being distributed.

The British Industrial Revolution illustrates this point vividly. While still widely present on the Continent, the craft guild in Britain had, by the time of the Glorious Revolution of 1688, declined and lost most of its political clout (Nef, 1957, pp. 26 and 32). With Parliamentary supremacy established over the Crown, legislation was passed in 1769 making the destruction of machinery punishable by death (Mokyr, 1990, p. 257). To be sure, there was still resistance to mechanisation. The "Luddite" riots between 1811 and 1816 were partly a manifestation of the fear of technological change among workers as Parliament revoked a 1551 law prohibiting the use of gig mills in the wool-finishing trade. The British government however took an increasingly stern view on groups attempting to halt technological progress and deployed 12,000 men against the rioters (Mantoux, 2006, p. 403-8). The sentiment of the government towards the destruction of machinery was explained by a resolution passed after the Lancashire riots of 1779, stating that: "The sole cause of great riots was the new machines employed in cotton manufacture; the country notwithstanding has greatly benefited from their erection [and] destroying them in this country would only be the means of transferring them to another [...] to the detriment of the trade of Britain" (cited in Mantoux, 2006, p. 403).

There are at least two possible explanations for the shift in attitudes towards technological progress. First, after Parliamentary supremacy was established over the Crown, the property owning classes became politically dominant in Britain (North and Weingast, 1989). Because the diffusion of various manufacturing technologies did not impose a risk to the value of their assets, and some property owners stood to benefit from the export of manufactured goods, the artisans simply did not have the political power to repress them. Second, inventors, consumers and unskilled factory workers largely benefited from mechanisation (Mokyr, 1990, p. 256 and 258). It has even been argued that, despite the employment concerns over mechanisation, unskilled workers have been the greatest beneficiaries of the Industrial Revolution (Clark, 2008).[7] While there is contradictory evidence suggesting that capital owners initially accumulated a growing share of national income (Allen, 2009a), there is equally evidence of growing real wages (Lindert and Williamson, 1983; Feinstein, 1998). This implies that although manufacturing technologies made the skills of artisans obsolete, gains from technological progress were distributed in a manner that gradually benefited a growing share of the labour force.[8]

An important feature of nineteenth century manufacturing technologies is that they were largely "deskilling" – *i.e.* they substituted for skills through the simplification of tasks (Braverman, 1974; Hounshell, 1985; James and Skinner, 1985; Goldin and Katz, 1998). The deskilling process occurred as the factory system began to displace the artisan shop, and it picked up pace as production increasingly mechanized with the adoption of steam power (Goldin and Sokoloff, 1982; Atack, *et al.*, 2008*a*). Work that had previously been performed by artisans was now decomposed into smaller, highly specialised, sequences, requiring less skill, but more workers, to perform.[9] Some innovations were even designed to be deskilling. For example, Eli Whitney, a pioneer of interchangeable parts, described the objective of this technology as "to substitute correct and effective operations of machinery for the skill of the artist which is acquired only by long practice and experience; a species of skill which is not possessed in this country to any considerable extent" (Habakkuk, 1962, p. 22).

Together with developments in continuous-flow production, enabling workers to be stationary while different tasks were moved to them, it was identical interchangeable parts that allowed complex products to be assembled from mass produced individual components by using highly specialised machine tools to a sequence of operations.[10] Yet while the first assembly-line was documented in 1804, it was not until the late nineteenth century that continuous-flow processes started to be adopted on a larger scale, which enabled corporations such as the Ford Motor Company to manufacture the T-Ford at a sufficiently low price for it to become the people's vehicle (Mokyr, 1990, p. 137). Crucially, the new assembly line introduced by Ford in 1913 was specifically designed for machinery to be operated by unskilled workers (Hounshell, 1985, p. 239). Furthermore, what had previously been a one-man job was turned into a 29-man worker operation, reducing the overall work time by 34 percent (Bright, 1958). The example of the Ford Motor Company thus underlines the general pattern observed in the nineteenth century, with physical capital providing a relative complement to unskilled labour, while substituting for relatively skilled artisans (James and Skinner, 1985; Louis and Paterson, 1986; Brown and Philips, 1986; Atack, *et al.*, 2004).[11] Hence, as pointed out by Acemoglu (2002, p. 7): "the idea that technological advances favor more skilled workers is a twentieth century phenomenon." The conventional wisdom among economic historians, in other words, suggests a discontinuity between the nineteenth and twentieth century in the impact of capital deepening on the relative demand for skilled labour.

The modern pattern of capital-skill complementarity gradually emerged in the late nineteenth century, as manufacturing production shifted to increasingly mechanised assembly lines. This shift can be traced to the switch to electricity from steam and water-power which, in combination with continuous-process and batch production methods, reduced the demand for unskilled manual workers in many hauling, conveying, and assembly tasks, but increased the demand for skills (Goldin and Katz, 1998). In short, while factory assembly lines, with their extreme division of labour, had required vast quantities of human operatives, electrification allowed many stages of the production process to be automated, which in turn increased the demand for relatively skilled blue-collar production workers to operate the machinery. In addition, electrification contributed to a growing share of white-collar nonproduction workers (Goldin and Katz, 1998). Over the course of the nineteenth century, establishments became larger in size as steam and water power technologies improved, allowing them to adopt powered machinery to realize productivity gains through the combination of enhanced division of labour and higher capital intensity (Atack, *et al.*, 2008*a*). Furthermore, the transport revolution lowered costs of shipping goods domestically and internationally as infrastructure spread and improved (Atack, *et al.*, 2008*b*). The market for artisan goods early on had largely been confined to the immediate surrounding area because transport costs were high relative to the value of the goods produced. With the transport revolution, however, market size expanded, thereby eroding local monopoly power, which in turn increased competition and compelled firms to raise productivity through mechanisation. As establishments became larger and served geographically expended markets, managerial tasks increased in number and complexity, requiring more managerial and clerking employees (Chandler, 1977). This pattern was, by the turn of the twentieth century, reinforced by electrification, which not only contributed to a growing share of relatively skilled blue-collar labour, but also increased the demand for white-collar workers (Goldin and Katz, 1998), who tended to have higher educational attainment (Allen, 2001).[12]

Since electrification, the story of the twentieth century has been the race between education and technology (Goldin and Katz, 2009). The US high school movement coincided with the first industrial revolution of the office (Goldin and Katz, 1995). While the typewriter was invented in the 1860s, it was not introduced in the office until the early twentieth century, when it entered a wave of mechanisation, with dictaphones, calculators, mimeo machines, address machines, and the predecessor of the computer – the keypunch (Beniger, 1986; Cortada, 2000). Importantly, these office machines reduced the cost of information processing tasks and increased the demand for the complementary factor – *i.e.* educated office workers. Yet the increased supply of educated office workers, following the high school movement, was associated with a sharp decline in the wage premium of clerking occupations relative to production workers (Goldin and Katz, 1995). This was, however, not the result of deskilling technological change. Clerking workers were indeed relatively educated. Rather, it was the result of the supply of educated workers outpacing the demand for their skills, leading educational wage differentials to compress.

While educational wage differentials in the US narrowed from 1915 to 1980 (Goldin and Katz, 2009), both educational wage differentials and overall wage inequality have increased sharply since the 1980s in a number of countries (Krueger, 1993; Murphy, *et al.*, 1998; Atkinson, 2008; Goldin and Katz, 2009). Although there are clearly several variables at work, consensus is broad that this can be ascribed to an acceleration in capital-skill complementarity, driven by the adoption of computers and information technology (Krueger, 1993; Autor, *et al.*, 1998; Bresnahan, *et al.*, 2002). What is commonly referred to as the Computer Revolution began with the first commercial uses of computers around 1960 and continued through the development of the Internet and e-commerce in the 1990s. As the cost per computation declined at an annual average of 37 percent between 1945 and 1980 (Nordhaus, 2007), telephone operators were made redundant, the first industrial robot was introduced by General Motors in the 1960s, and in the 1970s airline reservations systems led the way in self-service technology (Gordon, 2012). During the 1980s and 1990s, computing costs declined even more rapidly, on average by 64 percent per year, accompanied by a surge in computational power (Nordhaus, 2007).[13] At the same time, bar-code scanners and cash machines were spreading across the retail and financial industries, and the first personal computers were introduced in the early 1980s, with their word processing and spreadsheet functions eliminating copy typist occupations and allowing repetitive calculations to be automated (Gordon, 2012). This substitution for labour marks a further important reversal.

The early twentieth century office machines increased the demand for clerking workers (Chandler, 1977; Goldin and Katz, 1995). In a similar manner, computerisation augments demand for such tasks, but it also permits them to be automated (Autor, *et al.*, 2003).

The Computer Revolution can go some way in explaining the growing wage inequality of the past decades. For example, Krueger (1993) finds that workers using a computer earn roughly earn 10 to 15 percent more than others, but also that computer use accounts for a substantial share of the increase in the rate of return to education. In addition, more recent studies find that computers have caused a shift in the occupational structure of the labour market. Autor and Dorn (2013), for example, show that as computerisation erodes wages for labour performing routine tasks, workers will reallocate their labour supply to relatively low-skill service occupations. More specifically, between 1980 and 2005, the share of US labour hours in service occupations grew by 30 percent after having been flat or declining in the three prior decades. Furthermore, net changes in US employment were U-shaped in skill level, meaning that the lowest and highest job-skill quartile expanded sharply with relative employment declines in the middle of the distribution.

The expansion in high-skill employment can be explained by the falling price of carrying out routine tasks by means of computers, which complements more abstract and creative services. Seen from a production function perspective, an outward shift in the supply of routine informational inputs increases the marginal productivity of workers they are demanded by. For example, text and data mining has improved the quality of legal research as constant access to market information has improved the efficiency of managerial decision-making – *i.e.* tasks performed by skilled workers at the higher end of the income distribution. The result has been an increasingly polarised labour market, with growing employment in high-income cognitive jobs and low-income manual occupations, accompanied by a hollowing-out of middle-income routine jobs. This is a pattern that is not unique to the US and equally applies to a number of developed economies (Goos, *et al.*, 2009).[14]

How technological progress in the twenty-first century will impact on labour market outcomes remains to be seen. Throughout history, technological progress has vastly shifted the composition of employment, from agriculture and the artisan shop, to manufacturing and clerking, to service and management occupations. Yet the concern over technological unemployment has proven to be exaggerated. The obvious reason why this concern has not materialised relates to Ricardo's famous chapter on machinery, which suggests that labour-saving technology reduces the demand for undifferentiated labour, thus leading to technological unemployment (Ricardo, 1819). As economists have long understood, however, an invention that replaces workers by machines will have effects on all product and factor markets. An increase in the efficiency of production which reduces the price of one good, will increase real income and thus increase demand for other goods. Hence, in short, technological progress has two competing effects on employment (Aghion and Howitt, 1994). First, as technology substitutes for labour, there is a destruction effect, requiring workers to reallocate their labour supply; and second, there is the capitalisation effect, as more companies enter industries where productivity is relatively high, leading employment in those industries to expand.

Although the capitalisation effect has been predominant historically, our discovery of means of economising the use of labour can outrun the pace at which we can find new uses for labour, as Keynes (1933) pointed out. The reason why human labour has prevailed relates to its ability to adopt and acquire new skills by means of education (Goldin and Katz, 2009). Yet as computerisation enters more cognitive domains this will become increasingly challenging (Brynjolfsson and McAfee, 2011). Recent empirical findings are therefore particularly concerning. For example, Beaudry, *et al.* (2013) document a decline in the demand for skill over the past decade, even as the supply of workers with higher education has continued to grow. They show that high-skilled workers have moved down the occupational ladder, taking on jobs traditionally performed by low-skilled workers, pushing low-skilled workers even further down the occupational ladder and, to some extent, even out of the labour force. This raises questions about: (a) the ability of human labour to win the race against technology by means of education; and (b) the potential extent of technological unemployment, as an increasing pace of technological progress will cause higher job turnover, resulting in a higher natural rate of unemployment (Lucas and Prescott, 1974; Davis and Haltiwanger, 1992; Pissarides, 2000). While the present study is limited to examining the destruction effect of technology, it nevertheless provides a useful indication of the job growth required to counterbalance the jobs at risk over the next decades.

III. THE TECHNOLOGICAL REVOLUTIONS OF THE TWENTY-FIRST CENTURY

The secular price decline in the real cost of computing has created vast economic incentives for employers to substitute labour for computer capital.[15] Yet the tasks computers are able to perform ultimately depend upon the ability of a programmer to write a set of procedures or rules that appropriately direct the technology in each possible contingency. Computers will therefore be relatively productive to human labour when a problem can be specified – in the sense that the criteria for success are quantifiable and can readily be evaluated (Acemoglu and Autor, 2011). The extent of job computerisation will thus be determined by technological advances that allow engineering problems to be sufficiently specified, which sets the boundaries for the scope of computerisation. In this section, we examine the extent of tasks computer-controlled equipment can be expected to perform over the next decades. Doing so, we focus on advances in fields related to Machine Learning (ML), including Data Mining, Machine Vision, Computational Statistics and other sub-fields of Artificial Intelligence (AI), in which efforts are explicitly dedicated to the development of algorithms that allow cognitive tasks to be automated. In addition, we examine the application of ML technologies in Mobile Robotics (MR), and thus the extent of computerisation in manual tasks.

Our analysis builds on the task categorisation of Autor, *et al.* (2003), which distinguishes between workplace tasks using a two-by-two matrix, with routine versus non-routine tasks on one axis, and manual versus cognitive tasks on the other. In short, routine tasks are defined as tasks that follow explicit rules that can be accomplished by machines, while non-routine tasks are not sufficiently well understood to be specified in computer code. Each of these task categories can, in turn, be of either manual or cognitive nature – *i.e.* they relate to physical labour or knowledge work. Historically, computerisation has largely been confined to manual and cognitive routine tasks involving explicit rule-based activities (Autor and Dorn, 2013; Goos, *et al.*, 2009). Following recent technological advances, however, computerisation is now spreading to domains commonly defined as non-routine. The rapid pace at which tasks that were defined as non-routine only a decade ago have now become computerisable is illustrated by Autor, *et al.* (2003), asserting that: "Navigating a car through city traffic or deciphering the scrawled handwriting on a personal check – minor undertakings for most adults – are not routine tasks by our definition." Today, the problems of navigating a car and deciphering handwriting are sufficiently well understood that many related tasks can be specified in computer code and automated (Veres, *et al.*, 2011; Plötz and Fink, 2009).

Recent technological breakthroughs are, in large part, due to efforts to turn non-routine tasks into well-defined problems. Defining such problems is helped by the provision of relevant data: this is highlighted in the case of handwriting recognition by Plötz and Fink (2009). The success of an algorithm for handwriting recognition is difficult to quantify without data to test on – in particular, determining whether an algorithm performs well for different styles of writing requires data containing a variety of such styles. That is, data is required to specify the many contingencies a technology must manage in order to form an adequate substitute for human labour. With data, objective and quantifiable measures of the success of an algorithm can be produced, which aid the continual improvement of its performance relative to humans.

As such, technological progress has been aided by the recent production of increasingly large and complex datasets, known as big data.[16] For instance, with a growing corpus of human-translated digitalised text, the success of a machine translator can now be judged by its accuracy in reproducing observed translations. Data from United Nations documents, which are translated by human experts into six languages, allow Google Translate to monitor and improve the performance of different machine translation algorithms (Tanner, 2007).

Further, ML algorithms can discover unexpected similarities between old and new data, aiding the computerisation of tasks for which big data has newly become available. As a result, computerisation is no longer confined to routine tasks that can be written as rule-based software queries, but is spreading to every non-routine task where big data becomes available (Brynjolfsson and McAfee, 2011). In this section, we examine the extent of future computerisation beyond routine tasks.

III.A. Computerisation in non-routine cognitive tasks

With the availability of big data, a wide range of non-routine cognitive tasks are becoming computerisable. That is, further to the general improvement in technological progress due to big data, algorithms for big data are rapidly entering domains reliant upon storing or accessing information. The use of big data is afforded by one of the chief comparative advantages of computers relative to human labor: scalability. Little evidence is required to demonstrate that, in performing the task of laborious computation, networks of machines scale better than human labour (Campbell-Kelly, 2009). As such, computers can better manage the large calculations required in using large datasets. ML algorithms running on computers are now, in many cases, better able to detect patterns in big data than humans.

Computerisation of cognitive tasks is also aided by another core comparative advantage of algorithms: their absence of some human biases. An algorithm can be designed to ruthlessly satisfy the small range of tasks it is given. Humans, in contrast, must fulfill a range of tasks unrelated to their occupation, such as sleeping, necessitating occasional sacrifices in their occupational performance (Kahneman, *et al.*, 1982). The additional constraints under which humans must operate manifest themselves as biases. Consider an example of human bias: Danziger, *et al.* (2011) demonstrate that experienced Israeli judges are substantially more generous in their rulings following a lunch break. It can thus be argued that many roles involving decision-making will benefit from impartial algorithmic solutions.

Fraud detection is a task that requires both impartial decision making and the ability to detect trends in big data. As such, this task is now almost completely automated (Phua, *et al.*, 2010). In a similar manner, the comparative advantages of computers are likely to change the nature of work across a wide range of industries and occupations.

In health care, diagnostics tasks are already being computerised. Oncologists at Memorial Sloan-Kettering Cancer Center are, for example, using IBM's Watson computer to provide chronic care and cancer treatment diagnostics. Knowledge from 600,000 medical evidence reports, 1.5 million patient records and clinical trials, and two million pages of text from medical journals, are used for benchmarking and pattern recognition purposes. This allows the computer to compare each patient's individual symptoms, genetics, family and medication history, etc., to diagnose and develop a treatment plan with the highest probability of success (Cohn, 2013).

In addition, computerisation is entering the domains of legal and financial services. Sophisticated algorithms are gradually taking on a number of tasks performed by paralegals, contract and patent lawyers (Markoff, 2011). More specifically, law firms now rely on computers that can scan thousands of legal briefs and precedents to assist in pre-trial research. A frequently cited example is Symantec's Clearwell system, which uses language analysis to identify general concepts in documents, can present the results graphically, and proved capable of analysing and sorting more than 570,000 documents in two days (Markoff, 2011).

Furthermore, the improvement of sensing technology has made sensor data one of the most prominent sources of big data (Ackerman and Guizzo, 2011). Sensor data is often coupled with new ML fault- and anomaly-detection algorithms to render many tasks computerisable. A broad class of examples can be found in condition monitoring and novelty detection, with technology substituting for closed-circuit TV (CCTV) operators, workers examining equipment defects, and clinical staff responsible for monitoring the state of patients in intensive care. Here, the fact that computers lack human biases is of great value: algorithms are free of irrational bias, and their vigilance need not be interrupted by rest breaks or lapses of concentration. Following the declining costs of digital sensing and actuation, ML approaches have successfully addressed condition monitoring applications ranging from batteries (Saha, *et al.*, 2007), to aircraft engines (King, *et al.*, 2009), water quality (Osborne, *et al.*, 2012) and intensive care units (ICUs) (Clifford and Clifton, 2012; Clifton, *et al.*, 2012). Sensors can equally be placed on trucks and pallets to improve companies' supply chain management, and used to measure the moisture in a field of crops to track the flow of water through utility pipes. This allows for automatic meter reading, eliminating the need for personnel to gather such information. For example, the cities of Doha, São Paulo, and Beijing use sensors on pipes, pumps, and other water infrastructure to monitor conditions and manage water loss, reducing leaks by 40 to 50 percent. In the near future, it will be possible to place inexpensive sensors on light poles, sidewalks, and other public property to capture sound and images, likely reducing the number of workers in law enforcement (MGI, 2013).

Advances in user interfaces also enable computers to respond directly to a wider range of human requests, thus augmenting the work of highly skilled labour, while allowing some types of jobs to become fully automated. For example, Apple's Siri and Google Now rely on natural user interfaces to recognise spoken words, interpret their meanings, and act on them accordingly. Moreover, a company called SmartAction now provides call computerisation solutions that use ML technology and advanced speech recognition to improve upon conventional interactive voice response systems, realising cost savings of 60 to 80 percent over an outsourced call center consisting of human labour (CAA, 2012). Even education, one of the most labour intensive sectors, will most likely be significantly impacted by improved user interfaces and algorithms building upon big data. The recent growth in MOOCs (Massive Open Online Courses) has begun to generate large datasets detailing how students interact on forums, their diligence in completing assignments and viewing lectures, and their ultimate grades (Simonite, 2013; Breslow, *et al.*, 2013). Such information, together with improved user interfaces, will allow for ML algorithms that serve as interactive tutors, with teaching and assessment strategies statistically calibrated to match individual student needs (Woolf, 2010). Big data analysis will also allow for more effective predictions of student performance, and for their suitability for post-graduation occupations. These technologies can equally be implemented in recruitment, most likely resulting in the streamlining of human resource (HR) departments.

Occupations that require subtle judgement are also increasingly susceptible to computerisation. To many such tasks, the unbiased decision making of an algorithm represents a comparative advantage over human operators. In the most challenging or critical applications, as in ICUs, algorithmic recommendations may serve as inputs to human operators; in other circumstances, algorithms will themselves be responsible for appropriate decision-making. In the financial sector, such automated decision-making has played a role for quite some time. AI algorithms are able to process a greater number of financial announcements, press releases, and other information than any human trader, and then act faster upon them (Mims, 2010). Services like Future Advisor similarly use AI to offer personalised financial advice at larger scale and lower cost. Even the work of software engineers may soon largely be computerisable. For example, advances in ML allow a programmer to leave complex parameter and design choices to be appropriately optimised by an algorithm (Hoos, 2012). Algorithms can further automatically detect bugs in software (Hangal and Lam, 2002; Livshits and Zimmermann, 2005; Kim, *et al.*, 2008), with a reliability that humans are unlikely to match. Big databases of code also offer the eventual prospect of algorithms that learn how to write programs to satisfy specifications provided by a human. Such an approach is likely to eventually improve upon human programmers, in the same way that human-written compilers eventually proved inferior to automatically optimised compilers. An algorithm can better keep the whole of a program in working memory, and is not constrained to human-intelligible code, allowing for holistic solutions that might never occur to a human. Such algorithmic improvements over human judgement are likely to become increasingly common.

Although the extent of these developments remains to be seen, estimates by MGI (2013) suggests that sophisticated algorithms could substitute for approximately 140 million full-time knowledge workers worldwide. Hence, while technological progress throughout economic history has largely been confined to the mechanisation of manual tasks, requiring physical labour, technological progress in the twenty-first century can be expected to contribute to a wide range of cognitive tasks, which, until now, have largely remained a human domain. Of course, many occupations being affected by these developments are still far from fully computerisable, meaning that the computerisation of some tasks will simply free-up time for human labour to perform other tasks. Nonetheless, the trend is clear: computers increasingly challenge human labour in a wide range of cognitive tasks (Brynjolfsson and McAfee, 2011).

III.B. Computerisation in non-routine manual tasks

Mobile robotics provides a means of directly leveraging ML technologies to aid the computerisation of a growing scope of manual tasks. The continued technological development of robotic hardware is having notable impact upon employment: over the past decades, industrial robots have taken on the routine tasks of most operatives in manufacturing. Now, however, more advanced robots are gaining enhanced sensors and manipulators, allowing them to perform non-routine manual tasks. For example, General Electric has recently developed robots to climb and maintain wind turbines, and more flexible surgical robots with a greater range of motion will soon perform more types of operations (Robotics-VO, 2013). In a similar manner, the computerisation of logistics is being aided by the increasing cost-effectiveness of highly instrumented and computerised cars. Mass-production vehicles, such as the Nissan LEAF, contain on-board computers and advanced telecommunication equipment that render the car a potentially fly-by-wire robot.[17] Advances in sensor technology mean that vehicles are likely to soon be augmented with even more advanced suites of sensors. These will permit an algorithmic vehicle controller to monitor its environment to a degree that exceeds the capabilities of any human driver: they have the ability to simultaneously look both forwards and backwards, can natively integrate camera, GPS and LIDAR data, and are not subject to distraction. Algorithms are thus potentially safer and more effective drivers than humans.

The big data provided by these improved sensors are offering solutions to many of the engineering problems that had hindered robotic development in the past. In particular, the creation of detailed three dimensional maps of road networks has enabled autonomous vehicle navigation; most notably illustrated by Google's use of large, specialised datasets collected by its driverless cars (Guizzo, 2011). It is now completely feasible to store representations of the entire road network on-board a car, dramatically simplifying the navigation problem. Algorithms that could perform navigation throughout the changing seasons, particularly after snowfall, have been viewed as a substantial challenge. However, the big data approach can answer this by storing records from the last time snow fell, against which the vehicle's current environment can be compared (Churchill and Newman, 2012). ML approaches have also been developed to identify unprecedented changes to a particular piece of the road network, such as roadworks (Mathibela, *et al.*, 2012). This emerging technology will affect a variety of logistics jobs. Agricultural vehicles, forklifts and cargo-handling vehicles are imminently automatable, and hospitals are already employing autonomous robots to transport food, prescriptions and samples (Bloss, 2011). The computerisation of mining vehicles is further being pursued by companies such as Rio Tinto, seeking to replace labour in Australian mine-sites.[18]

With improved sensors, robots are capable of producing goods with higher quality and reliability than human labour. For example, El Dulze, a Spanish food processor, now uses robotics to pick up heads of lettuce from a conveyor belt, rejecting heads that do not comply with company standards. This is achieved by measuring their density and replacing them on the belt (IFR, 2012a). Advanced sensors further allow robots to recognise patterns. Baxter, a 22,000 USD general-purpose robot, provides a well-known example. The robot features an LCD display screen displaying a pair of eyes that take on different expressions depending on the situation. When the robot is first installed or needs to learn a new pattern, no programming is required. A human worker simply guides the robot arms through the motions that will be needed for the task. Baxter then memorises these patterns and can communicate that it has understood its new instructions. While the physical flexibility of Baxter is limited to performing simple operations such as picking up objects and moving them, different standard attachments can be installed on its arms, allowing Baxter to perform a relatively broad scope of manual tasks at low cost (MGI, 2013).

Technological advances are contributing to declining costs in robotics. Over the past decades, robot prices have fallen about 10 percent annually and are expected to decline at an even faster pace in the near future (MGI, 2013). Industrial robots, with features enabled by machine vision and high-precision dexterity, which typically cost 100,000 to 150,000 USD, will be available for 50,000 to 75,000 USD in the next decade, with higher levels of intelligence and additional capabilities (IFR, 2012b). Declining robot prices will inevitably place them within reach of more users. For example, in China, employers are increasingly incentivised to substitute robots for labour, as wages and living standards are rising – Foxconn, a Chinese contract manufacturer that employs 1.2 million workers, is now investing in robots to assemble products such as the Apple iPhone (Markoff, 2012). According to the International Federation of Robotics, robot sales in China grew by more than 50 percent in 2011 and are expected to increase further. Globally, industrial robot sales reached a record 166,000 units in 2011, a 40 percent year-on-year increase (IFR, 2012b). Most likely, there will be even faster growth ahead as low-priced general-purpose models, such as Baxter, are adopted in simple manufacturing and service work.

Expanding technological capabilities and declining costs will make entirely new uses for robots possible. Robots will likely continue to take on an increasing set of manual tasks in manufacturing, packing, construction, maintenance, and agriculture. In addition, robots are already performing many simple service tasks such as vacuuming, mopping, lawn mowing, and gutter cleaning – the market for personal and household service robots is growing by about 20 percent annually (MGI, 2013). Meanwhile, commercial service robots are now able to perform more complex tasks in food preparation, health care, commercial cleaning, and elderly care (Robotics-VO, 2013). As robot costs decline and technological capabilities expand, robots can thus be expected to gradually substitute for labour in a wide range of low-wage service occupations, where most US job growth has occurred over the past decades (Autor and Dorn, 2013). This means that many low-wage manual jobs that have been previously protected from computerisation could diminish over time.

III.C. The task model revisited

The task model of Autor, *et al.* (2003) has delivered intuitive and accurate predictions in that: (a) computers are more substitutable for human labour in routine relative to non-routine tasks; and (b) a greater intensity of routine inputs increases the marginal productivity of non-routine inputs. Accordingly, computers have served as a substitute for labour for many routine tasks, while exhibiting strong complementarities with labour performing cognitive non-routine tasks.[19] Yet the premises about what computers do have recently expanded. Computer capital can now equally substitute for a wide range of tasks commonly defined as non-routine (Brynjolfsson and McAfee, 2011), meaning that the task model will not hold in predicting the impact of computerisation on the task content of employment in the twenty-first century. While focusing on the substitution effects of recent technological progress, we build on the task model by deriving several factors that we expect will determine the extent of computerisation in non-routine tasks.

The task model assumes for tractability an aggregate, constant-returns-to-scale, Cobb-Douglas production function of the form

$$Q = (L_{\text{S}} + C)^{1-\beta} L_{\text{NS}}^{\beta}, \quad \beta \in [0, 1], \tag{1}$$

where L_{S} and L_{NS} are susceptible and non-susceptible labor inputs and C is computer capital. Computer capital is supplied perfectly elastically at market price per efficiency unit, where the market price is falling exogenously with time due to technological progress. It further assumes income-maximizing workers, with heterogeneous productivity endowments in both susceptible and non-susceptible tasks. Their task supply will respond elastically to relative wage levels, meaning that workers will reallocate their labour supply according to their comparative advantage as in Roy (1951). With expanding computational capabilities, resulting from technological advances, and a falling market price of computing, workers in susceptible tasks will thus reallocate to non-susceptible tasks.

The above described simple model differs from the task model of Autor, *et al.* (2003), in that L_{NS} is not confined to routine labour inputs. This is because recent developments in ML and MR, building upon big data, allow for pattern recognition, and thus enable computer capital to rapidly substitute for labour across a wide range of non-routine tasks. Yet some inhibiting engineering bottlenecks to computerisation persist. Beyond these bottlenecks, however, we argue that it is largely already technologically possible to automate almost any task, provided that sufficient amounts of data are gathered for pattern recognition. Our model thus predicts that the pace at which these bottlenecks can be overcome will determine the extent of computerisation in the twenty-first century.

Hence, in short, while the task model predicts that computers for labour substitution will be confined to routine tasks, our model predicts that computerisation can be extended to any non-routine task that is not subject to any engineering bottlenecks to computerisation. These bottlenecks thus set the boundaries for the computerisation of non-routine tasks. Drawing upon the ML and MR literature, and a workshop held at the Oxford University Engineering Sciences Department, we identify several engineering bottlenecks, corresponding to three task categories. According to these findings, non-susceptible labor inputs can be described as,

$$L_{\mathrm{NS}} = \sum_{i=1}^{n} \left(L_{\mathrm{PM},i} + L_{\mathrm{C},i} + L_{\mathrm{SI},i} \right) \tag{2}$$

where L_{PM}, L_{C} and L_{SI} are labour inputs into perception and manipulation tasks, creative intelligence tasks, and and social intelligence tasks.

We note that some related engineering bottlenecks can be partially alleviated by the simplification of tasks. One generic way of achieving this is to reduce the variation between task iterations. As a prototypical example, consider the factory assembly line, turning the non-routine tasks of the artisan shop into repetitive routine tasks performed by unskilled factory workers. A more recent example is the computerisation of non-routine manual tasks in construction. On-site construction tasks typically demand a high degree of adaptability, so as to accommodate work environments that are typically irregularly laid out, and vary according to weather. Prefabrication, in which the construction object is partially assembled in a factory before being transported to the construction site, provides a way of largely removing the requirement for adaptability. It allows many construction tasks to be performed by robots under controlled conditions that eliminate task variability – a method that is becoming increasingly widespread, particularly in Japan (Barlow and Ozaki, 2005; Linner and Bock, 2012). The extent of computerisation in the twenty-first century will thus partly depend on innovative approaches to task restructuring. In the remainder of this section we examine the engineering bottlenecks related to the above mentioned task categories, each in turn.

Perception and manipulation tasks. Robots are still unable to match the depth and breadth of human perception. While basic geometric identification is reasonably mature, enabled by the rapid development of sophisticated sensors and lasers, significant challenges remain for more complex perception tasks, such as identifying objects and their properties in a cluttered field of view. As such, tasks that relate to an unstructured work environment can make jobs less susceptible to computerisation. For example, most homes are unstructured, requiring the identification of a plurality of irregular objects and containing many cluttered spaces which inhibit the mobility of wheeled objects. Conversely, supermarkets, factories, warehouses, airports and hospitals have been designed for large wheeled objects, making it easier for robots to navigate in performing non-routine manual tasks. Perception problems can, however, sometimes be sidestepped by clever task design. For example, Kiva Systems, acquired by Amazon.com in 2012, solved the problem of warehouse navigation by simply placing bar-code stickers on the floor, informing robots of their precise location (Guizzo, 2008).

The difficulty of perception has ramifications for manipulation tasks, and, in particular, the handling of irregular objects, for which robots are yet to reach human levels of aptitude. This has been evidenced in the development of robots that interact with human objects and environments. While advances have been made, solutions tend to be unreliable over the myriad small variations on a single task, repeated thousands of times a day, that many applications require. A related challenge is failure recovery – *i.e.* identifying and rectifying the mistakes of the robot when it has, for example, dropped an object. Manipulation is also limited by the difficulties of planning out the sequence of actions required to move an object from one place to another. There are yet further problems in designing manipulators that, like human limbs, are soft, have compliant dynamics and provide useful tactile feedback. Most industrial manipulation makes uses of workarounds to these challenges (Brown, *et al.*, 2010), but these approaches are nonetheless limited to a narrow range of tasks. The main challenges to robotic computerisation, perception and manipulation, thus largely remain and are unlikely to be fully resolved in the next decade or two (Robotics-VO, 2013).

Creative intelligence tasks. The psychological processes underlying human creativity are difficult to specify. According to Boden (2003), creativity is the ability to come up with ideas or artifacts that are novel and valuable. Ideas, in a broader sense, include concepts, poems, musical compositions, scientific theories, cooking recipes and jokes, whereas artifacts are objects such as paintings, sculptures, machinery, and pottery. One process of creating ideas (and similarly for artifacts) involves making unfamiliar combinations of familiar ideas, requiring a rich store of knowledge. The challenge here is to find some reliable means of arriving at combinations that "make sense." For a computer to make a subtle joke, for example, would require a database with a richness of knowledge comparable to that of humans, and methods of benchmarking the algorithm's subtlety.

In principle, such creativity is possible and some approaches to creativity already exist in the literature. Duvenaud, *et al.* (2013) provide an example of automating the core creative task required in order to perform statistics, that of designing models for data. As to artistic creativity, AARON, a drawing-program, has generated thousands of stylistically-similar line-drawings, which have been exhibited in galleries worldwide. Furthermore, David Cope's EMI software composes music in many different styles, reminiscent of specific human composers.

In these and many other applications, generating novelty is not particularly difficult. Instead, the principal obstacle to computerising creativity is stating our creative values sufficiently clearly that they can be encoded in an program (Boden, 2003). Moreover, human values change over time and vary across cultures. Because creativity, by definition, involves not only novelty but value, and because values are highly variable, it follows that many arguments about creativity are rooted in disagreements about value. Thus, even if we could identify and encode our creative values, to enable the computer to inform and monitor its own activities accordingly, there would still be disagreement about whether the computer appeared to be creative. In the absence of engineering solutions to overcome this problem, it seems unlikely that occupations requiring a high degree of creative intelligence will be automated in the next decades.

Social intelligence tasks. Human social intelligence is important in a wide range of work tasks, such as those involving negotiation, persuasion and care. To aid the computerisation of such tasks, active research is being undertaken within the fields of Affective Computing (Scherer, *et al.*, 2010; Picard, 2010), and Social Robotics (Ge, 2007; Broekens, *et al.*, 2009). While algorithms and robots can now reproduce some aspects of human social interaction, the real-time recognition of natural human emotion remains a challenging problem, and the ability to respond intelligently to such inputs is even more difficult. Even simplified versions of typical social tasks prove difficult for computers, as is the case in which social interaction is reduced to pure text. The social intelligence of algorithms is partly captured by the Turing test, examining the ability of a machine to communicate indistinguishably from an actual human. Since 1990, the Loebner Prize, an annual Turing test competition, awards prizes to textual chat programmes that are considered to be the most human-like. In each competition, a human judge simultaneously holds computer-based textual interactions with both an algorithm and a human. Based on the responses, the judge is to distinguish between the two. Sophisticated algorithms have so far failed to convince judges about their human resemblance. This is largely because there is much 'common sense' information possessed by humans, which is difficult to articulate, that would need to be provided to algorithms if they are to function in human social settings.

Whole brain emulation, the scanning, mapping and digitalising of a human brain, is one possible approach to achieving this, but is currently only a theoretical technology. For brain emulation to become operational, additional functional understanding is required to recognise what data is relevant, as well as a roadmap of technologies needed to implement it. While such roadmaps exist, present implementation estimates, under certain assumptions, suggest that whole brain emulation is unlikely to become operational within the next decade or two (Sandberg and Bostrom, 2008). When or if they do, however, the employment impact is likely to be vast (Hanson, 2001).

Hence, in short, while sophisticated algorithms and developments in MR, building upon with big data, now allow many non-routine tasks to be automated, occupa tions that involve complex perception and manipulation tasks, creative intelligence tasks, and social intelligence tasks are unlikely to be substituted by computer capital over the next decade or two. The probability of an occupation being automated can thus be described as a function of these task characteristics. As suggested by Figure I, the low degree of social intelligence required by a dishwasher makes this occupation more susceptible to computerisation than a public relation specialist, for example. We proceed to examining the susceptibility of jobs to computerisation as a function of the above described

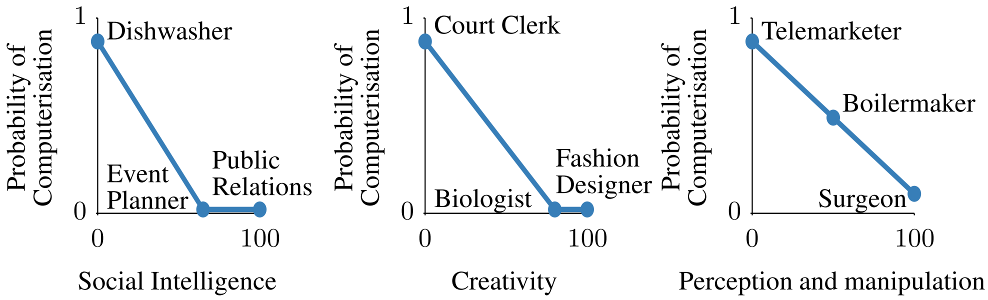

FIGURE I. A sketch of how the probability of computerisation might vary as a function of bottleneck variables.

non-susceptible task characteristics.

IV. MEASURING THE EMPLOYMENT IMPACT OF COMPUTERISATION

IV.A. Data sources and implementation strategy

To implement the above described methodology, we rely on O∗NET, an online service developed for the US Department of Labor. The 2010 version of O∗NET contains information on 903 detailed occupations, most of which correspond closely to the Labor Department's Standard Occupational Classification (SOC). The O∗NET data was initially collected from labour market analysts, and has since been regularly updated by surveys of each occupation's worker population and related experts, to provide up-to-date information on occupations as they evolve over time. For our purposes, an important feature of O∗NET is that it defines the key features of an occupation as a standardised and measurable set of variables, but also provides open-ended descriptions of specific tasks to each occupation. This allows us to: (a) objectively rank occupations according to the mix of knowledge, skills, and abilities they require; and (b) subjectively categorise them based on the variety of tasks they involve.

The close SOC correspondence of O∗NET allows us to link occupational characteristics to 2010 Bureau of Labor Statistics (BLS) employment and wage data. While the O∗NET occupational classification is somewhat more detailed, distinguishing between Auditors and Accountants, for example, we aggregate these occupations to correspond to the six-digit 2010 SOC system, for which employment and wage figures are reported. To obtain unique O∗NET variables corresponding to the six-digit SOC classification, we used the mean of the O∗NET aggregate. In addition, we exclude any six-digit SOC occupations for which O∗NET data was missing.[20] Doing so, we end up with a final dataset consisting of 702 occupations.

To assess the employment impact of the described technological developments in ML, the ideal experiment would provide two identical autarkic economies, one facing the expanding technological capabilities we observe, and a secular decline in the price of computerisation, and the other not. By comparison, it would be straightforward to examine how computerisation reshapes the occupational composition of the labour market. In the absence of this experiment, the second preferred option would be to build on the implementation strategy of Autor, *et al.* (2003), and test a simple economic model to predict how demand for workplace tasks responds to developments in ML and MR technology. However, because our paper is forward-looking, in the sense that most of the described technological developments are yet to be implemented across industries on a broader scale, this option was not available for our purposes.

Instead, our implementation strategy builds on the literature examining the offshoring of information-based tasks to foreign worksites, consisting of different methodologies to rank and categorise occupations according to their susceptibility to offshoring (Blinder, 2009; Jensen and Kletzer, 2005, 2010). The common denominator for these studies is that they rely on O∗NET data in different ways. While Blinder (2009) eyeballed the O∗NET data on each occupation, paying particular attention to the job description, tasks, and work activities, to assign an admittedly subjective two-digit index number of offshorability to each occupation, Jensen and Kletzer (2005) created a purely objective ranking based on standardised and measurable O∗NET variables. Both approaches have obvious drawbacks. Subjective judgments are often not replicable and may result in the researcher subconsciously rigging the data to conform to a certain set of beliefs. Objective rankings, on the other hand, are not subject to such drawbacks, but are constrained by the reliability of the variables that are being used. At this stage, it shall be noted that O∗NET data was not gathered to specifically measure the offshorability or automatability of jobs. Accordingly, Blinder (2009) finds that past attempts to create objective offshorability rankings using O∗NET data have yielded some questionable results, ranking lawyers and judges among the most tradable occupations, while classifying occupations such as data entry keyers, telephone operators, and billing clerks as virtually impossible to move offshore.

To work around some of these drawbacks, we combine and build upon the two described approaches. First, together with a group of ML researchers, we subjectively hand-labelled 70 occupations, assigning 1 if automatable, and 0 if not. For our subjective assessments, we draw upon a workshop held at the Oxford University Engineering Sciences Department, examining the automatability of a wide range of tasks. Our label assignments were based on eyeballing the O∗NET tasks and job description of each occupation. This information is particular to each occupation, as opposed to standardised across different jobs. The hand-labelling of the occupations was made by answering the question "Can the tasks of this job be sufficiently specified, conditional on the availability of big data, to be performed by state of the art computer-controlled equipment". Thus, we only assigned a 1 to fully automatable occupations, where we considered all tasks to be automatable. To the best of our knowledge, we considered the possibility of task simplification, possibly allowing some currently non-automatable tasks to be automated. Labels were assigned only to the occupations about which we were most confident.

Second, we use objective O∗NET variables corresponding to the defined bottlenecks to computerisation. More specifically, we are interested in variables describing the level of perception and manipulation, creativity, and social intelligence required to perform it. As reported in Table I, we identified nine variables that describe these attributes. These variables were derived from the O∗NET survey, where the respondents are given multiple scales, with "importance" and "level" as the predominant pair. We rely on the "level" rating which corresponds to specific examples about the capabilities required of computer-controlled equipment to perform the tasks of an occupation. For instance, in relation to the attribute "Manual Dexterity", low (level) corresponds to "Screw a light bulb into a light socket"; medium (level) is exemplified by "Pack oranges in crates as quickly as possible"; high (level) is described as "Perform open-heart surgery with surgical instruments". This gives us an indication of

TABLE I. O∗NET variables that serve as indicators of bottlenecks to computerisation.

Computerisation bottleneck	O∗NET Variable	O∗NET Description
Perception and Manipulation	Finger Dexterity	The ability to make precisely coordinated movements of the fingers of one or both hands to grasp, manipulate, or assemble very small objects.
	Manual Dexterity	The ability to quickly move your hand, your hand together with your arm, or your two hands to grasp, manipulate, or assemble objects.
	Cramped Work Space, Awkward Positions	How often does this job require working in cramped work spaces that requires getting into awkward positions?
Creative Intelligence	Originality	The ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
	Fine Arts	Knowledge of theory and techniques required to compose, produce, and perform works of music, dance, visual arts, drama, and sculpture.
Social Intelligence	Social Perceptiveness	Being aware of others' reactions and understanding why they react as they do.
	Negotiation	Bringing others together and trying to reconcile differences.
	Persuasion	Persuading others to change their minds or behavior.
	Assisting and Caring for Others	Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.

the level of "Manual Dexterity" computer-controlled equipment would require to perform a specific occupation. An exception is the "Cramped work space" variable, which measures the frequency of unstructured work.

Hence, in short, by hand-labelling occupations, we work around the issue that O∗NET data was not gathered to specifically measure the automatability of jobs in a similar manner to Blinder (2009). In addition, we mitigate some of the subjective biases held by the researchers by using objective O∗NET variables to correct potential hand-labelling errors. The fact that we label only 70 of the full 702 occupations, selecting those occupations whose computerisation label we are highly confident about, further reduces the risk of subjective bias affecting our analysis. To develop an algorithm appropriate for this task, we turn to probabilistic classification.

IV.B. *Classification method*

We begin by examining the accuracy of our subjective assessments of the automatability of 702 occupations. For classification, we develop an algorithm to provide the label probability given a previously unseen vector of variables. In the terminology of classification, the O∗NET variables form a *feature vector*, denoted $\underline{x} \in \mathbb{R}^9$. O∗NET hence supplies a complete dataset of 702 such feature vectors. A computerisable label is termed a *class*, denoted $y \in \{0, 1\}$. For our problem, $y = 1$ (true) implies that we hand-labelled as computerisable the occupation described by the associated nine O∗NET variables contained in $\underline{x} \in \mathbb{R}^9$. Our *training data* is $\mathcal{D} = (X, \underline{y})$, where $X \in \mathbb{R}^{70\times 9}$ is a matrix of variables and $\underline{y} \in \{0, 1\}^{70}$ gives the associated labels. This dataset contains information about how y varies as a function of $\underline{x}$: as a hypothetical example, it may be the case that, for all occupations for which $x_1 > 50$, $y = 1$. A probabilistic classification algorithm exploits patterns existent in training data to return the probability $P(y_* = 1 \mid \underline{x}_*, X, \underline{y})$ of a new, unlabelled, *test datum* with features $\underline{x}_*$ having class label $y_* = 1$.

We achieve probabilistic classification by introducing a latent function $f : \underline{x} \mapsto \mathbb{R}$, known as a *discriminant function*. Given the value of the discriminant f_* at a test point $\underline{x}_*$, we assume that the probability for the class label is given by the logistic

$$P(y_* = 1 \mid f_*) = \frac{1}{1 + \exp(-f_*)}, \tag{3}$$

and $P(y_* = 0 \mid f_*) = 1 - P(y_* = 1 \mid f_*)$. For $f_* > 0$, $y_* = 1$ is more probable than $y_* = 0$. For our application, f can be thought of as a continuous-valued ‘automatability’ variable: the higher its value, the higher the probability of computerisation.

We test three different models for the discriminant function, f, using the best performing for our further analysis. Firstly, logistic (or logit) regression, which adopts a linear model for f, $f(\underline{x}) = \underline{w}^\intercal \underline{x}$, where the un-known weights w are often inferred by maximising their probability in light of the training data. This simple model necessarily implies a simple monotonic relationship between features and the probability of the class taking a particular value. Richer models are provided by *Gaussian process classifiers* (Rasmussen and

Williams, 2006). Such classifiers model the latent function f with a Gaussian process (GP): a non-parametric probability distribution over functions.

A GP is defined as a distribution over the functions $f\colon \mathcal{X} \to \mathbb{R}$ such that the distribution over the possible function values on any finite subset of $\mathcal{X}$ (such as X) is multivariate Gaussian. For a function $f(\underline{x})$, the prior distribution over its values $\underline{f}$ on a subset $\underline{x} \subset \mathcal{X}$ are completely specified by a covariance matrix K

$$p(\underline{f} \mid K) = \mathcal{N}(\underline{f}; \underline{0}, K) = \frac{1}{\sqrt{\det 2\pi K}} \exp\left(-\frac{1}{2} \underline{f}^{\intercal} K^{-1} \underline{f}\right). \tag{4}$$

The covariance matrix is generated by a covariance function $\kappa : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$; that is, $K = \kappa(X, X)$. The GP model is expressed by the choice of κ; we consider the *exponentiated quadratic* (squared exponential) and *rational quadratic*. Note that we have chosen a zero mean function, encoding the assumption that $P(y_* = 1) = \frac{1}{2}$ sufficiently far from training data.

Given training data $\mathcal{D}$, we use the GP to make predictions about the function values f_* at input $\underline{x}_*$. With this information, we have the predictive equations

$$p(f_* \mid \underline{x}_*, \mathcal{D}) = \mathcal{N}\big(f_*; m(f_* \mid \underline{x}_*, \mathcal{D}), V(f_* \mid \underline{x}_*, \mathcal{D})\big)\,, \tag{5}$$

where

$$m(f_* \mid \underline{x}_*, \mathcal{D}) = K(\underline{x}_*, X)K(X, X)^{-1}\underline{y} \tag{6}$$

$$V(f_* \mid \underline{x}_*, \mathcal{D}) = K(\underline{x}_*, x_*) - K(\underline{x}_*, X)K(X, X)^{-1}K(X, \underline{x}_*)\,. \tag{7}$$

Inferring the label posterior $p(y_* \mid \underline{x}_*, \mathcal{D})$ is complicated by the non-Gaussian form of the logistic (3). In order to effect inference, we use the approximate Expectation Propagation algorithm (Minka, 2001).

We tested three Gaussian process classifiers using the GPML toolbox (Rasmussen and Nickisch, 2010) on our data, built around exponentiated quadratic, rational quadratic and linear covariances. Note that the latter is equivalent to logistic regression with a Gaussian prior taken on the weights $\underline{w}$. To validate these classifiers, we randomly selected a reduced training set of half the available data $\mathcal{D}$; the remaining data formed a test set. On this test set, we evaluated how closely the algorithm's classifications matched the hand labels according to two metrics (see *e.g.* Murphy (2012)): the area under the receiver operat-

TABLE II. Performance of various classifiers; best performances in bold.

classifier model	AUC	log-likelihood
exponentiated quadratic	$\mathbf{0.894}$	$\mathbf{-163.3}$
rational quadratic	0.893	-163.7
linear (logit regression)	0.827	-205.0

ing characteristic curve (AUC), which is equal to one for a perfect classifier, and one half for a completely random classifier, and the log-likelihood, which should ideally be high. This experiment was repeated for one hundred random selections of training set, and the average results tabulated in Table II. The exponentiated quadratic model returns (narrowly) the best performance of the three (clearly outperforming the linear model corresponding to logistic regression), and was hence selected for the remainder of our testing. Note that its AUC score of nearly 0.9 represents accurate classification: our algorithm successfully managed to reproduce our hand-labels specifying whether an occupation was computerisable. This means that our algorithm verified that our subjective judgements were systematically and consistently related to the O∗NET variables.

Having validated our approach, we proceed to use classification to predict the probability of computerisation for all 702 occupations. For this purpose, we introduce a new label variable, z, denoting whether an occupation is truly computerisable or not: note that this can be judged only once an occupation is computerised, at some indeterminate point in the future. We take, again, a logistic likelihood,

$$P(z_* = 1 \mid f_*) = \frac{1}{1 + \exp(-f_*)}. \tag{8}$$

We implicitly assumed that our hand label, y, is a noise-corrupted version of the unknown true label, z. Our motivation is that our hand-labels of computerisability must necessarily be treated as such noisy measurements. We thus acknowledge that it is by no means certain that a job is computerisable given our labelling. We define $X_* \in \mathbb{R}^{702 \times 9}$ as the matrix of O∗NET variables for all 702 occupations; this matrix represents our *test features*.

We perform a final experiment in which, given training data $\mathcal{D}$, consisting

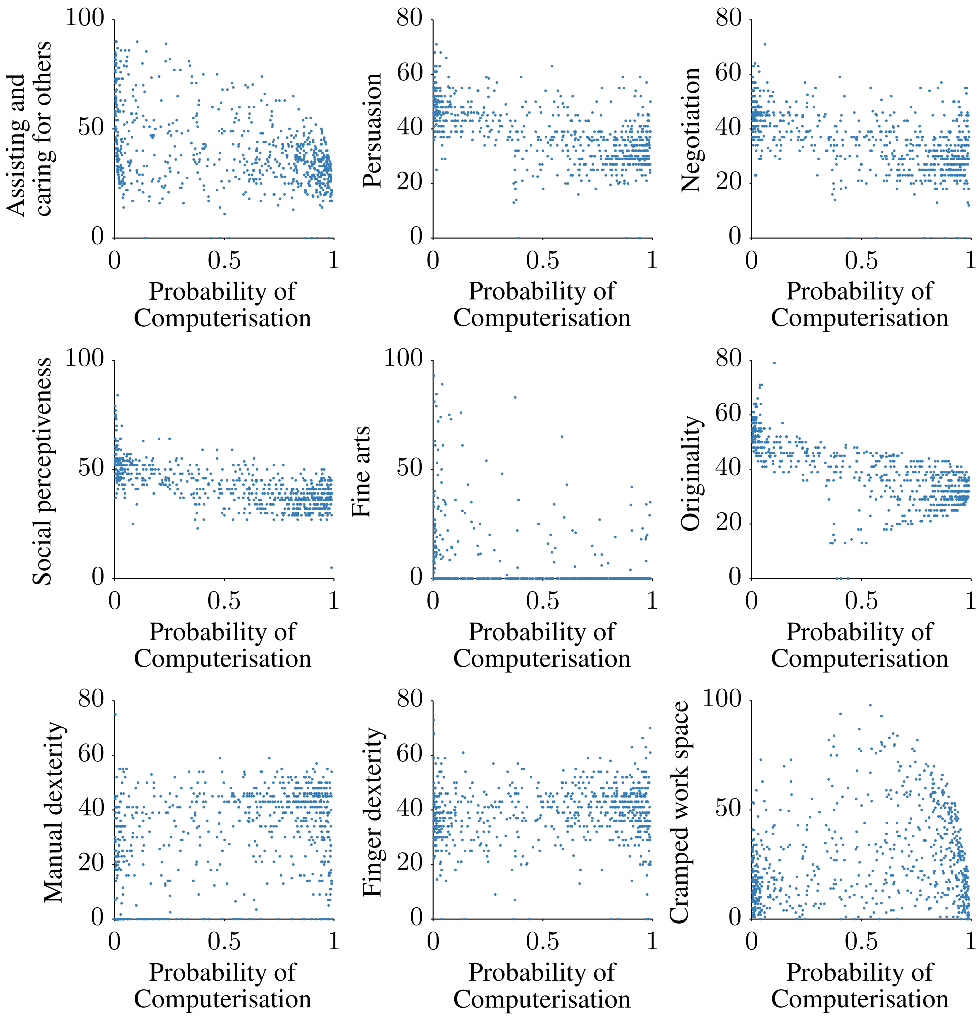

FIGURE II. The distribution of occupational variables as a function of probability of computerisation; each occupation is a unique point.

of our 70 hand-labelled occupations, we aim to predict $\underline{z}_*$ for our test features X_*. This approach firstly allows us to use the features of the 70 occupations about which we are most certain to predict for the remaining 632. Further, our algorithm uses the trends and patterns it has learned from bulk data to correct for what are likely to be mistaken labels. More precisely, the algorithm provides a smoothly varying probabilistic assessment of automatability as a function of the variables. For our Gaussian process classifier, this function is non-linear, meaning that it flexibly adapts to the patterns inherent in the training data. Our approach thus allows for more complex, non-linear, interactions between variables: for example, perhaps one variable is not of importance unless the value of another variable is sufficiently large. We report $P(\underline{z}_* \mid X_*, \mathcal{D})$ as the probability of computerisation henceforth (for a detailed probability ranking, see Appendix). Figure II illustrates that this probability is non-linearly related to the nine O∗NET variables selected.

V. EMPLOYMENT IN THE TWENTY-FIRST CENTURY

In this section, we examine the possible future extent of at-risk job computerisation, and related labour market outcomes. The task model predicts that recent developments in ML will reduce aggregate demand for labour input in tasks that can be routinised by means of pattern recognition, while increasing the demand for labour performing tasks that are not susceptible to computerisation. However, we make no attempt to forecast future changes in the occupational composition of the labour market. While the 2010-2020 BLS occupational employment projections predict US net employment growth across major occupations, based on historical staffing patterns, we speculate about technology that is in only the early stages of development. This means that historical data on the impact of the technological developments we observe is unavailable.[21] We therefore focus on the impact of computerisation on the mix of jobs that existed in 2010. Our analysis is thus limited to the substitution effect of future computerisation.

Turning first to the expected employment impact, reported in Figure III, we distinguish between high, medium and low risk occupations, depending on their

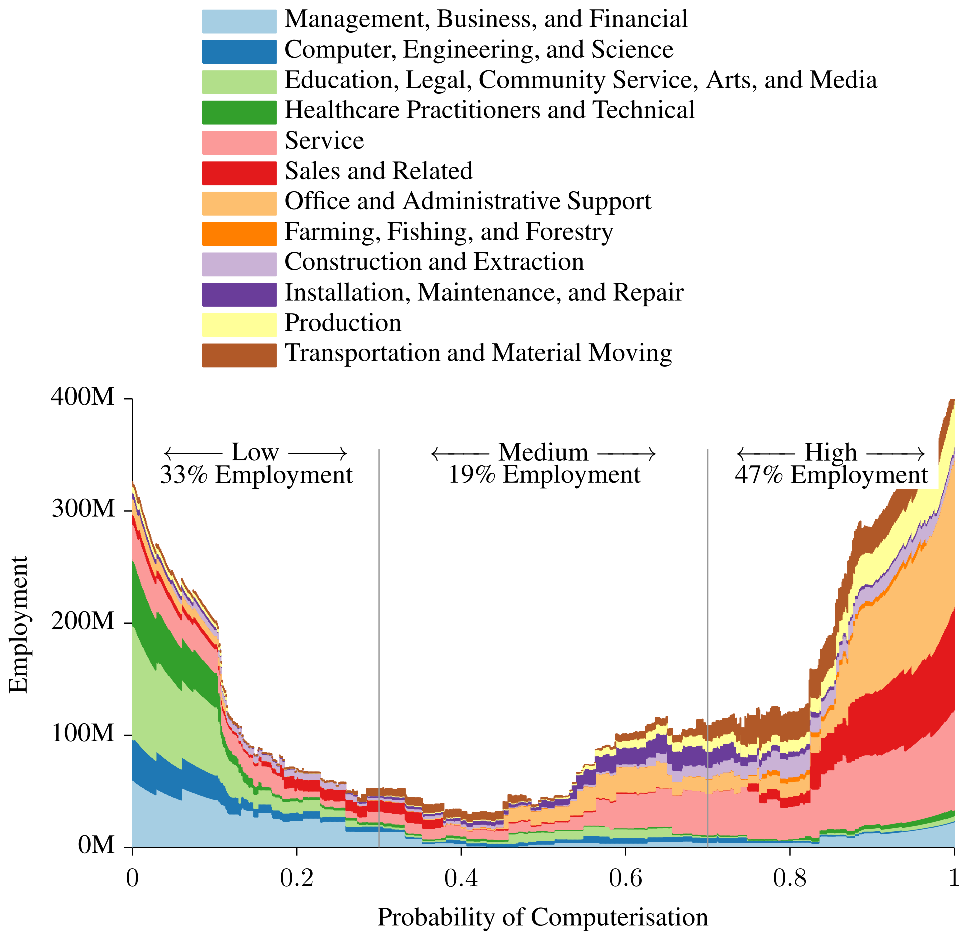

FIGURE III. The distribution of BLS 2010 occupational employment over the probability of computerisation, along with the share in low, medium and high probability categories. Note that the total area under all curves is equal to total US employment.

probability of computerisation (thresholding at probabilities of 0.7 and 0.3). According to our estimate, 47 percent of total US employment is in the high risk category, meaning that associated occupations are potentially automatable over some unspecified number of years, perhaps a decade or two. It shall be noted that the probability axis can be seen as a rough timeline, where high probability occupations are likely to be substituted by computer capital relatively soon. Over the next decades, the extent of computerisation will be determined by the pace at which the above described engineering bottlenecks to automation can be overcome. Seen from this perspective, our findings could be interpreted as two waves of computerisation, separated by a "technological plateau". In the first wave, we find that most workers in transportation and logistics occupations, together with the bulk of office and administrative support workers, and labour in production occupations, are likely to be substituted by computer capital. As computerised cars are already being developed and the declining cost of sensors makes augmenting vehicles with advanced sensors increasingly cost-effective, the automation of transportation and logistics occupations is in line with the technological developments documented in the literature. Furthermore, algorithms for big data are already rapidly entering domains reliant upon storing or accessing information, making it equally intuitive that office and administrative support occupations will be subject to computerisation. The computerisation of production occupations simply suggests a continuation of a trend that has been observed over the past decades, with industrial robots taking on the routine tasks of most operatives in manufacturing. As industrial robots are becoming more advanced, with enhanced senses and dexterity, they will be able to perform a wider scope of non-routine manual tasks. From a technological capabilities point of view, the vast remainder of employment in production occupations is thus likely to diminish over the next decades.

More surprising, at first sight, is that a substantial share of employment in services, sales and construction occupations exhibit high probabilities of computerisation. Yet these findings are largely in line with recent documented technological developments. First, the market for personal and household service robots is already growing by about 20 percent annually (MGI, 2013). As the comparative advantage of human labour in tasks involving mobility and dexterity will diminish over time, the pace of labour substitution in service occupations is likely to increase even further. Second, while it seems counterintuitive that sales occupations, which are likely to require a high degree of social intelligence, will be subject to a wave of computerisation in the near future, high risk sales occupations include, for example, cashiers, counter and rental clerks, and telemarketers. Although these occupations involve interactive tasks, they do not necessarily require a high degree of social intelligence. Our model thus seems to do well in distinguishing between individual occupations within occupational categories. Third, prefabrication will allow a growing share of construction work to be performed under controlled conditions in factories, which partly eliminates task variability. This trend is likely to drive the computerisation of construction work.

In short, our findings suggest that recent developments in ML will put a substantial share of employment, across a wide range of occupations, at risk in the near future. According to our estimates, however, this wave of automation will be followed by a subsequent slowdown in computers for labour substitution, due to persisting inhibiting engineering bottlenecks to computerisation. The relatively slow pace of computerisation across the medium risk category of employment can thus partly be interpreted as a technological plateau, with incremental technological improvements successively enabling further labour substitution. More specifically, the computerisation of occupations in the medium risk category will mainly depend on perception and manipulation challenges. This is evident from Table III, showing that the "manual dexterity", "finger dexterity" and "cramped work space" variables exhibit relatively high values in the medium risk category. Indeed, even with recent technological developments, allowing for more sophisticated pattern recognition, human labour will still have a comparative advantage in tasks requiring more complex perception and manipulation. Yet with incremental technological improvements, the comparative advantage of human labour in perception and manipulation tasks could eventually diminish. This will require innovative task restructuring, improvements in ML approaches to perception challenges, and progress in robotic dexterity to overcome manipulation problems related to variation between task iterations and the handling of irregular objects. The gradual computerisation of installation, maintenance, and repair occupations, which are largely confined to the medium risk category, and require a high degree of perception and manipulation capabilities, is a manifestation of this observation.

Our model predicts that the second wave of computerisation will mainly

TABLE III. Distribution (mean and standard deviation) of values for each variable.

Variable	Probability of Computerisation		
	Low	Medium	High
Assisting and caring for others	48±20	41±17	34±10
Persuasion	48±7.1	35±9.8	32±7.8
Negotiation	44±7.6	33±9.3	30±8.9
Social perceptiveness	51±7.9	41±7.4	37±5.5
Fine arts	12±20	3.5±12	1.3±5.5
Originality	51±6.5	35±12	32±5.6
Manual dexterity	22±18	34±15	36±14
Finger dexterity	36±10	39±10	40±10
Cramped work space	19±15	37±26	31±20

depend on overcoming the engineering bottlenecks related to creative and social intelligence. As reported in Table III, the "fine arts", "originality", "negotiation", "persuasion", "social perceptiveness", and "assisting and caring for others", variables, all exhibit relatively high values in the low risk category. By contrast, we note that the "manual dexterity", "finger dexterity" and "cramped work space" variables take relatively low values. Hence, in short, generalist occupations requiring knowledge of human heuristics, and specialist occupations involving the development of novel ideas and artifacts, are the least susceptible to computerisation. As a prototypical example of generalist work requiring a high degree of social intelligence, consider the O∗NET tasks reported for chief executives, involving "conferring with board members, organization officials, or staff members to discuss issues, coordinate activities, or resolve problems", and "negotiating or approving contracts or agreements." Our predictions are thus intuitive in that most management, business, and finance occupations, which are intensive in generalist tasks requiring social intelligence, are largely confined to the low risk category. The same is true of most occupations in education, healthcare, as well as arts and media jobs. The O∗NET tasks of actors, for example, involve "performing humorous and serious interpretations of emotions, actions, and situations, using body movements, facial expressions, and gestures", and "learning about characters in scripts and their relationships to each other in order to develop role interpretations." While these tasks are very different from those of a chief executive, they equally require profound

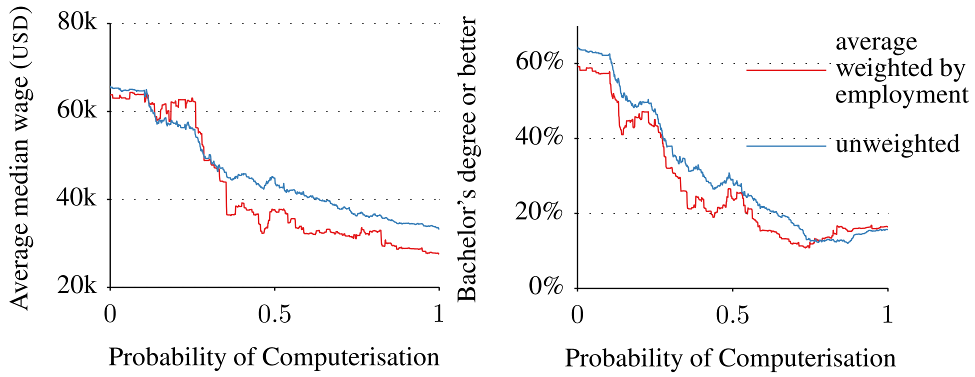

FIGURE IV. Wage and education level as a function of the probability of computerisation; note that both plots share a legend.

knowledge of human heuristics, implying that a wide range of tasks, involving social intelligence, are unlikely to become subject to computerisation in the near future.

The low susceptibility of engineering and science occupations to computerisation, on the other hand, is largely due to the high degree of creative intelligence they require. The O∗NET tasks of mathematicians, for example, involve "developing new principles and new relationships between existing mathematical principles to advance mathematical science" and "conducting research to extend mathematical knowledge in traditional areas, such as algebra, geometry, probability, and logic." Hence, while it is evident that computers are entering the domains of science and engineering, our predictions implicitly suggest strong complementarities between computers and labour in creative science and engineering occupations; although it is possible that computers will fully substitute for workers in these occupations over the long-run. We note that the predictions of our model are strikingly in line with the technological trends we observe in the automation of knowledge work, even within occupational categories. For example, we find that paralegals and legal assistants – for which computers already substitute – in the high risk category. At the same time, lawyers, which rely on labour input from legal assistants, are in the low risk category. Thus, for the work of lawyers to be fully automated, engineering bottlenecks to creative and social intelligence will need to be overcome, implying that the computerisation of legal research will complement the work of lawyers in the medium term.

To complete the picture of what recent technological progress is likely to mean for the future of employment, we plot the average median wage of occupations by their probability of computerisation. We do the same for skill level, measured by the fraction of workers having obtained a bachelor's degree, or higher educational attainment, within each occupation. Figure IV reveals that both wages and educational attainment exhibit a strong negative relationship with the probability of computerisation. We note that this prediction implies a truncation in the current trend towards labour market polarization, with growing employment in high and low-wage occupations, accompanied by a hollowing-out of middle-income jobs. Rather than reducing the demand for middle-income occupations, which has been the pattern over the past decades, our model predicts that computerisation will mainly substitute for low-skill and low-wage jobs in the near future. By contrast, high-skill and high-wage occupations are the least susceptible to computer capital.

Our findings were robust to the choice of the 70 occupations that formed our training data. This was confirmed by the experimental results tabulated in Table II: a GP classifier trained on half of the training data was demonstrably able to accurately predict the labels of the other half, over one hundred different partitions. That these predictions are accurate for many possible partitions of the training set suggests that slight modifications to this set are unlikely to lead to substantially different results on the entire dataset.

V.A. Limitations

It shall be noted that our predictions are based on expanding the premises about the tasks that computer-controlled equipment can be expected to perform. Hence, we focus on estimating the share of employment that can potentially be substituted by computer capital, from a technological capabilities point of view, over some unspecified number of years. We make no attempt to estimate how many jobs will actually be automated. The actual extent and pace of computerisation will depend on several additional factors which were left unaccounted for.

First, labour saving inventions may only be adopted if the access to cheap labour is scarce or prices of capital are relatively high (Habakkuk, 1962).[22] We do not account for future wage levels, capital prices or labour shortages. While these factors will impact on the timeline of our predictions, labour is the scarce factor, implying that in the long-run wage levels will increase relative to capital prices, making computerisation increasingly profitable (see, for example, Acemoglu, 2003).

Second, regulatory concerns and political activism may slow down the process of computerisation. The states of California and Nevada are, for example, currently in the process of making legislatory changes to allow for driverless cars. Similar steps will be needed in other states, and in relation to various technologies. The extent and pace of legislatory implementation can furthermore be related to the public acceptance of technological progress.[23] Although resistance to technological progress has become seemingly less common since the Industrial Revolution, there are recent examples of resistance to technological change.[24] We avoid making predictions about the legislatory process and the public acceptance of technological progress, and thus the pace of computerisation.

Third, making predictions about technological progress is notoriously difficult (Armstrong and Sotala, 2012).[25] For this reason, we focus on near-term technological breakthroughs in ML and MR, and avoid making any predictions about the number of years it may take to overcome various engineering bottlenecks to computerisation. Finally, we emphasise that since our probability estimates describe the likelihood of an occupation being fully automated, we do not capture any within-occupation variation resulting from the computerisation of tasks that simply free-up time for human labour to perform other tasks.

Although it is clear that the impact of productivity gains on employment will vary across occupations and industries, we make no attempt to examine such effects.

VI. CONCLUSIONS

While computerisation has been historically confined to routine tasks involving explicit rule-based activities (Autor, *et al.*, 2003; Goos, *et al.*, 2009; Autor and Dorn, 2013), algorithms for big data are now rapidly entering domains reliant upon pattern recognition and can readily substitute for labour in a wide range of non-routine cognitive tasks (Brynjolfsson and McAfee, 2011; MGI, 2013). In addition, advanced robots are gaining enhanced senses and dexterity, allowing them to perform a broader scope of manual tasks (IFR, 2012b; Robotics-VO, 2013; MGI, 2013). This is likely to change the nature of work across industries and occupations.

In this paper, we ask the question: how susceptible are current jobs to these technological developments? To assess this, we implement a novel methodology to estimate the probability of computerisation for 702 detailed occupations. Based on these estimates, we examine expected impacts of future computerisation on labour market outcomes, with the primary objective of analysing the number of jobs at risk and the relationship between an occupation's probability of computerisation, wages and educational attainment.

We distinguish between high, medium and low risk occupations, depending on their probability of computerisation. We make no attempt to estimate the number of jobs that will actually be automated, and focus on potential job automatability over some unspecified number of years. According to our estimates around 47 percent of total US employment is in the high risk category. We refer to these as jobs at risk – *i.e.* jobs we expect could be automated relatively soon, perhaps over the next decade or two.

Our model predicts that most workers in transportation and logistics occupations, together with the bulk of office and administrative support workers, and labour in production occupations, are at risk. These findings are consistent with recent technological developments documented in the literature. More surpris-ingly, we find that a substantial share of employment in service occupations, where most US job growth has occurred over the past decades (Autor and Dorn, 2013), are highly susceptible to computerisation. Additional support for this finding is provided by the recent growth in the market for service robots (MGI, 2013) and the gradually diminishment of the comparative advantage of human labour in tasks involving mobility and dexterity (Robotics-VO, 2013).

Finally, we provide evidence that wages and educational attainment exhibit a strong negative relationship with the probability of computerisation. We note that this finding implies a discontinuity between the nineteenth, twentieth and the twenty-first century, in the impact of capital deepening on the relative demand for skilled labour. While nineteenth century manufacturing technologies largely substituted for skilled labour through the simplification of tasks (Braverman, 1974; Hounshell, 1985; James and Skinner, 1985; Goldin and Katz, 1998), the Computer Revolution of the twentieth century caused a hollowing-out of middle-income jobs (Goos, *et al.*, 2009; Autor and Dorn, 2013). Our model predicts a truncation in the current trend towards labour market polarisation, with computerisation being principally confined to low-skill and low-wage occupations. Our findings thus imply that as technology races ahead, low-skill workers will reallocate to tasks that are non-susceptible to computerisation – *i.e.*, tasks requiring creative and social intelligence. For workers to win the race, however, they will have to acquire creative and social skills.