



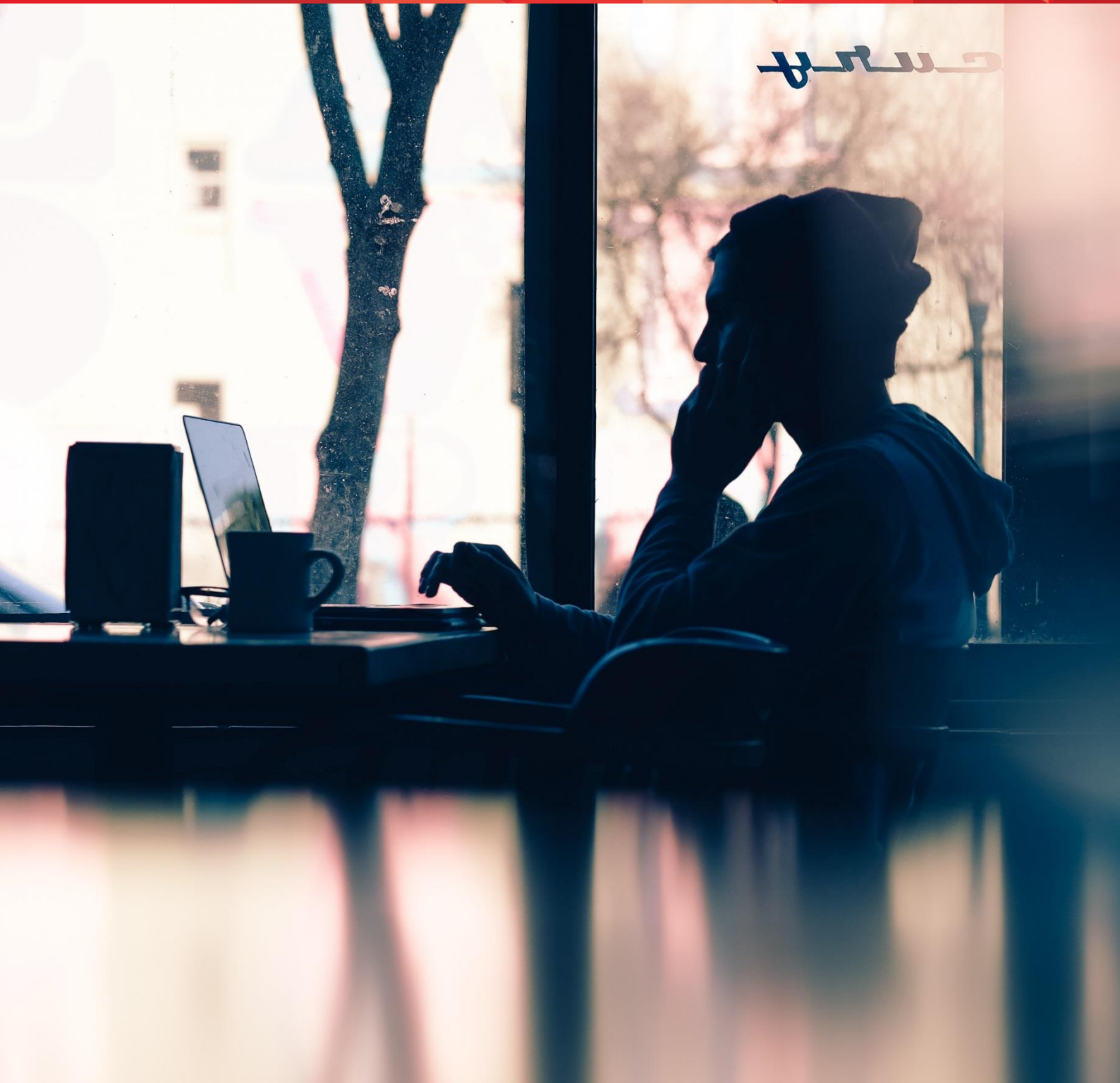
THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

The Evolving Role of ICT in the Economy

A Report by LSE Consulting for Huawei

Mirko Draca, Ralf Martin & Rosa Sanchis-Guarner

June 2018



This research was undertaken by the Centre for Economic Performance at the London School of Economics and Political Science. The study was commissioned via LSE Consulting which was set up to enable and facilitate the application of its academic expertise and intellectual resources.

LSE Consulting

LSE Enterprise Limited
London School of Economics and Political Science

Houghton Street
London
WC2A 2AE

(T) +44 (0)20 7955 7128
(E) lseenterprise.consulting@lse.ac.uk
(W) lse.ac.uk/consultancy



Contents

1. Chapter 1: The Evolving Role of ICT in the Economy	4
Executive Summary	4
<i>The Evolving Role of ICT</i>	4
<i>ICT as a Generator of Ideas</i>	4
<i>The Great Diffusion: Studying Broadband Roll-out in the UK</i>	5
<i>The Next Wave? Automation and the Future of Work</i>	6
1.1. Introduction	8
1.2. ‘Stop-Start’ Productivity and the ICT Revolution(s)	8
1.2.1. <i>Tracking the Current Productivity Slowdown</i>	9
1.2.2. <i>Mismeasurement and Consumer Surplus</i>	11
1.3. Policy Implications and the Economics of ICT	13
1.3.1. <i>Business, Innovation and Infrastructure</i>	13
1.3.2. <i>The Labour Market</i>	13
2. Chapter 2: ICT and the Production of Ideas: Tracking Knowledge Spillovers Across Technology Areas	23
2.1. Introduction	23
2.2. Key Findings	24
2.2.1. <i>Results for Patent Rank Spillovers</i>	24
2.2.2. <i>Results for Country-level Innovation per Capita</i>	24
2.3. Measurement of Spillovers	25
2.4. Innovation Patterns Across Countries	26
2.5. Spillovers	27
2.5.1. <i>Comparing Technologies</i>	27
2.5.2. <i>Spillover Gaps Over Time</i>	28
2.6. Conclusion	29
2.7. Tables and Figures	33
3. Chapter 3: The Great Diffusion: What Factors Drive the Spread of Broadband Technologies?	41
3.1. Introduction	41
3.2. The International Picture on Broadband	42
3.3. The Economic Effects of Broadband ICT	44
3.4. Studying the UK’s Great Diffusion	45
3.4.1. <i>The Roll-Out of ADSL</i>	45
3.4.2. <i>The Arrival of the Superfast Broadband</i>	49
3.5. Conclusions	50
3.6. Tables and Figures	52
4. Chapter 4: The Next Wave? Automation and the Future of Work	70
4.1. Introduction	70
4.2. Key Findings	71
4.3. Understanding Automation	72
4.3.1. <i>The Task Approach</i>	72
4.3.2. <i>The Varied Impacts of Automation</i>	72
4.4. Empirical Evidence from the ‘First Wave’	73
4.4.1. <i>Routine-Biased Technical Change?</i>	73
4.4.2. <i>Economic Challenges of Automation</i>	75
4.5. US Occupational Employment Trends in the 2000s and 2010	77
4.5.1. <i>Measuring Task Content</i>	77
4.5.2. <i>Employment, Wages and Task Content in the 2000s and 2010s</i>	77
4.6. Conclusion	78
4.7. Tables and Figures	81
5. Bibliography	94

1. Chapter 1: The Evolving Role of ICT in the Economy

Executive Summary

The Evolving Role of ICT

- Since the ICT-related surge in the 1990s and early 2000s there has been a major, persistent slowdown in productivity growth across a range of economies. The most likely explanation for this is that the slowdown is part of a cyclical process underpinned by technology, in particular ‘waves’ that relate to the development of General Purpose Technologies (GPTs).
- The other major explanation of the slowdown relates to mismeasurement, specifically the idea that a large fraction of the consumer surplus generated by internet-related goods and services is not being captured effectively in the national accounts. However, rigorous estimates of internet-related consumer surplus struggle to account for the gap implied by the productivity slowdown.
- In this report we put forward some original research on: (1) the role of ICT-related research in generating ideas via knowledge spillovers, (2) a study of the history and economic consequences of UK broadband since 2000, and (3) an analysis of recent US employment data in relation to automation.
- As we outline below, the two main policy implications of our research are that (1) the high knowledge spillovers that occur with respect to ICT-related innovations suggests that there is a window for the government to produce social returns through support for R&D in this area, and (2) the automation challenge will need to be met by a comprehensive skills policy and that the tax treatment of skills investment (for example, via a Skills and Training Tax Credit) could be a central part of this.

ICT as a Generator of Ideas

- We outline a framework for studying knowledge spillovers between patents. These spillovers represent the cumulative process of inventions influencing the development of other inventions.
- Specifically, this framework is based on a new measure that we call ‘Patent Rank’ which applies the principles of Google’s Page Rank algorithm for web search to the network of global patent citations. This lets us evaluate not just the number of cites that accrue to individual patents but also their influence in terms of the extended chain of citations.
- ICT technologies generate on average substantially larger knowledge spillovers than those generated by other technology areas and this holds even when ICT is compared to other frontier fields such as biotech and clean energy.
- For instance, as part of the results below we find that wireless technologies generate on average more than 50% higher knowledge spillovers than other comparable fields. This calculation is

based on a monetary estimate of spillovers that extends the basic framework of the Page Rank-based analysis.

- The growth in average global knowledge spillovers flattens out for ICT from the late 1990s onwards. This involves the tapering of a sharp spike in knowledge spillovers for ICT earlier in the 1990s. However, ICT spillovers are still around 25-50% higher than the pooled set of all other technology areas at the end of the period considered.
- The high knowledge spillovers that accrue in ICT suggest a role for enhanced, government-supported R&D investment in this area. This is because the knowledge spillovers are ‘social’ and hence private actors are not incentivised to invest in line with these social returns. The Patent Rank indicator provides a metric for evaluating comparative advantage and the current state of knowledge that could inform government R&D policy.
- Other research has also documented a surge in AI-related research since 2009, particularly in the field of learning. The fact that this surge is so recent means that information about the pattern of spillovers is yet accumulate. However, the scale of this surge means that average knowledge spillovers are likely to rise in the near future, but it is hard to predict the size of this effect.

The Great Diffusion: Studying Broadband Roll-out in the UK

- The diffusion of ADSL access was slower than has been previously understood. While 40% of the population (mainly concentrated in highly urbanised areas) gained access to ADSL by early 2001 it took another 6 years for the roll-out to be completed. A significant phase of the roll-out did not get underway until 2003-2004 when nearly 50% of telephone exchanges were ‘switched on’ for ADSL. The population coverage gap between the London/South-East area and regions such as the North and the Midlands was as high as 20% until late 2003.
- The roll-out of ADSL mainly tracks population density on a within-region basis with socio-economic factors correlated with the pattern of roll-out most strongly at the regional level. There is a very steep gradient in terms of population density in the pattern of roll-out that occurs even at the middle range. For example, on average an area in the 6th decile of population density received ADSL 400 days earlier than an area in the 3rd decile.
- This staggered roll-out makes it feasible to consider some economic effects of ADSL since some areas received access to the technology much earlier. We consider long run effects occurring over the period of 1991-2011, where detailed Census data is available for 1991, 2001 and 2011. The share of ‘nonroutine’ professional workers increased more in the post- ADSL period for areas that had early access to ADSL relative to trends in the pre-ADSL decade.
- This effect is equivalent to a 15% larger professional workforce by 2011 in the early receiving ADSL areas. In line with this finding for nonroutine workers there is also a decline in the share of workers in ‘routine manual’ occupations (machine operators and manual labourers) in the early receiving ADSL areas. In fact, there are sharp switches in the direction of economic trends relative to the pre-ADSL decade.

- In addition to these long-run effects, the modelling of short-run wage and income trends indicates that the diffusion of ADSL was correlated with higher incomes in areas where ADSL penetration was higher. Even after controlling for a range of area-level factors the estimates indicate that a typical 30% annual increase in an area's ADSL penetration rate is associated with a 0.6% higher wages.
- The arrival of superfast broadband in the 2010s assisted the regional convergence of broadband quality. For example, while the North and South were comparable in average speeds in the baseline year of 2012 there was a distinct 'digital divide' in maximum speeds. By 2017 differences were limited with some regions increasing maximum speed by 8-10 MB/s compared to 3MB/s for London. The correlation of area income with average broadband speed also fell over this period.
- The arrival of superfast broadband saw a massive jump in available speeds in some areas as the distance-to-exchange limitations of ADSL were overcome. In likelihood, this has led to a qualitative shift in the nature of internet usage in areas with slower speed during the mid- to late 2000s. Households will have access to new consumer services and businesses in these areas now have the ability to exploit the internet at a much higher level. A plausible focus for future research is therefore isolating these upgraded areas and studying their differential economic outcomes into the 2010s.

The Next Wave? Automation and the Future of Work

- Existing work has identified what can be characterised as a 'first wave' of automation based around the displacement of low skill 'routine' jobs and growth of high-skill 'non-routine' jobs. The 1980s and 1990s saw the disappearance of middle skill routine jobs. Importantly, the relative growth of high skill non-routine jobs slowed down markedly in the 2000s while low skill 'non-routine manual' jobs increased.
- Our main analysis relates to a study of US Occupational Employment Statistics (OES) data for the 1999-2016 period. This data allows for an analysis of 600-700 occupations on an annual basis. Further to this, these occupations have been matched to measures of task content data from the occupational skills database O*Net. This allows us to define a full set of occupational task groups across routine, non-routine, manual, analytic and inter-personal task types.
- In our main exercise, we test whether employment and wage patterns moved according to the level of different task inputs. For example, did occupations that are highly intensive in non-routine analytic tasks grow faster than occupations that were still high skill but slightly less intensive in these tasks? Since we are able to control for broad occupational groups our analysis is able to distinguish between relatively fine differences in task structure.
- Over the period from 2000-2010 there is pervasive evidence that employment patterns moved in line with initial levels of task content. For example, a shift in task content from the 35th percentile of the non-routine analytic task distribution to the 70th percentile is associated with a 5.7% difference in US occupational employment growth over 10 years. A similar shift in routine manual task content was associated with a 5.2% decline in employment. Amongst low

skill occupations growth was strongest for occupations with high levels of manual interpersonal tasks.

- The growth of wages was flat with respect to most measures of task content over the 2000-2010 period. This is consistent with other recent studies that have reported sluggish wage growth across the labour market in this period.
- Considering the decade of the 2010s up to the year 2015, there is no evidence that employment or wages are growing in line with task content measures. However, this does not imply that the influence of technological change on the labour market has ‘slowed down’. An analysis of year-by-year changes shows that developments in the 2010s were driven by reallocation during the recession years of 2009 and 2010. The fact that a major recession has not occurred in the 2010s (yet) is likely to account for the secular pattern of job growth in this period.
- Based on recent employment trends, there is no evidence that a turning point has been reached in the pattern of jobs that could be affected by advances in automation. However, the past evidence indicates that structural transformations in employment take place during recessions so the most likely starting point for a ‘second wave’ of automation affecting a new class of jobs is the next downturn.
- In terms of platform-based ‘gig’ format work, the current best estimate is that 0.5% of the US workforce is engaged in this type of work. This represents a pattern of high growth from a low base and the estimates suggest that gig economy growth can account for a large fraction of US employment growth since the last recession. The most likely scenario for the future is continued strong growth and the question is whether it will level off, perhaps as a result of regulatory action (for example, limitations on rideshare services). A point of interest at the moment is whether current strong employment conditions will translate into a rebound in traditional jobs or as an expansion of contingent work.

1.1. Introduction

In this report we review and contribute to the evidence on the economics of information and communication technologies (ICT). By common consensus, the influence of ICT on the economy appears to be pervasive and decisive. Arguably, a driving factor behind this consensus is the ubiquity of ICT. The launch of ‘desktop’ personal computing in 1982 introduced a large section of the public to regular ICT usage in both their working and personal lives. This was followed up by at least two further major ICT diffusion events - the emergence of the commercial internet in the mid-1990s and the rise of mobile computing devices in the late 2000s.

Despite this ubiquity a number of major puzzles surround the role of ICT in the economy. The influence of ICT on productivity and economic growth was very slow to appear and then seemed to dissipate after the early 2000s. This has led to ‘ICT realists’ such as Robert Gordon to renew their questioning of the structural importance of ICT relative to earlier ‘General Purpose Technologies’ (GPTs) such as the steam engine, electrification and the automobile (Gordon, 2012). In addition, the penetration of computers and related technologies into the labour market via the perceived displacement of jobs through automation has contributed to a wave of economic pessimism around ICT.

In this report we review this debate and provide some new contributions. These cover the technological influence of ICT as a generator of ‘knowledge spillovers’ (Chapter 2), an analysis of the now long history of broadband diffusion in a data-rich setting of the UK (Chapter 3), and a discussion of recent employment trends and the challenge of automation.

1.2. ‘Stop-Start’ Productivity and the ICT Revolution(s)

The central theme of economic discussions of ICT is the puzzle of its impact on the statistics we use to measure the economy. This was first raised during the productivity slowdown of the 1980s as part of the so-called ‘Solow Paradox’. At this time it was observed that the presence of computers seemed to be ‘everywhere but in the productivity statistics’.

The Solow Paradox was temporarily resolved in the mid-to-late 1990s by the acceleration of productivity growth, particularly in the United States. As we show below, productivity growth slowed down again from the mid-2000s and this seems to be associated with a number of important structural changes in productivity patterns across firms.

In the discussion below we follow Brynjolfsson et al. (2017) in converging on the conclusion that the most likely explanation of the current slowdown lies in the notion of an implementation lag. Just as the 1990s productivity surge was preceded by multiple decades of firm investments in both ICT capital stock and complementary human and organisational resources, our current era is plausibly a time when firms are working out strategies for utilising the full potential of artificial intelligence (AI) and robotics.

The main alternative explanations of the current slowdown are also credible and policy relevant, but current economic evidence does not suggest that they provide a complete explanation. The first alternative explanation lies in mismeasurement, in particular the notion that the full value of recent innovations in goods and services is not being counted. As a result, this leads to the underestimation of GDP and productivity. The core issue here is the measurement of consumer surplus, that is, the

difference between the price actually paid for a good or service and the price that a consumer is willing to pay. As we argue in this report, a focus on issues relating to consumer surplus should be a key new focus of policy-making but current estimates of consumer surplus across a range of major innovations literally do not add up as an explanation of the post-2004 productivity slowdown.

The second potential alternative explanation is fundamentally based on distributional questions. In short, the benefits of ICT-induced gains in productivity are indeed present in the economy but they are highly concentrated. In addition to this, the benefits of ICT-induced innovations are plausibly subject to dissipation as firms compete to capture the benefits that are exclusive or ‘rivalrous’. The evidence on these ‘winner takes all’ dynamics is less mature compared to the work on consumer surplus but again current calculations do not indicate that such changing dynamics can account for the slowdown.

Given this evidence, the main scenario we put forward in this report is the prospect of a productivity resurgence in 10-15 years lead by the implementation of innovations relating to AI and robotics. One of the main questions for policy is then how the economy is likely to absorb a set of potential ‘super-innovations’ such as autonomous vehicles and AI-staffed call centres or professional services. Will these super-innovations hit the economy as a cluster of disruptions that wipe out whole job classes and industries or will natural forces of economic adjustment work smoothly to absorb them? A second question is also: if the ‘great leap forward’ in productivity is not likely to occur for 10-15 years then what do we do until then? We tackle this question in detail later in the report in the context of the labour market and the matter of shifting impacts of automation.

1.2.1. Tracking the Current Productivity Slowdown

The ‘Big Picture’ of Post-WWII Productivity

In Figure 1.1 we present some stylised facts for labour productivity in leading economies over the post-WWII period. These measures are taken from the Bergeaud et al. (2016) ‘Long-Term Productivity Database’ and the specific measure used here is GDP per hour worked measured in US dollars normalised to a base set in 2010.

The pattern seen in the data here is typically summarised as an initial period of surging productivity after WWII followed by a slowdown in the mid-1970s that coincided with the well-known oil shocks of that decade. A less discussed fact is the clear and steady catch-up of European economies during the post-WWII period. This continued from the end of the war until the early 1990s when convergence was nearly complete. However, US labour productivity accelerated more quickly after 1995 such that a significant gap was apparent again by the mid-2000s. The period since the mid-2000s has been characterised by flat labour productivity growth across the US, Euro Area and Japanese economies that we plot here.

A similar pattern is evident for total factor productivity, calculated in the Bergeaud et al. (2016) database by accounting for changes in labour composition and capital deepening. European catch-up is nearly complete in the early 1990s but then the US accelerates and re-opens the gap. The flattening of TFP growth is especially marked for the Japanese economy.

A closer look at the major European economies is available in Figures 1.2 and 1.3, where the series for Great Britain, France, Germany, Italy and Spain are plotted for the period after 1970. This shows Britain

consistently lagging France and Germany in labour productivity but roughly comparable on TFP. Slow productivity growth across both measures is evident for Italy and Spain from the mid-1990s.

Dissecting the Recent Slowdown

What have been the main determinants or factors behind this ‘stop-start’ pattern to productivity growth? Table 1.1 shows a detailed breakdown of US productivity growth over different periods of time. In the US case, IT investment takes a major role in shaping the path of productivity through two main channels. Firstly, IT investment helped drive capital deepening during the 1995-2004 productivity surge, with the growth contribution of IT increasing from 0.41% points to 0.78%. Since the overall capital deepening effect in 1995-2004 was 1.22% this means that IT capital accounted for 64% of the total effect. In terms of components, IT capital deepening was driven by an approximate doubling in computer hardware (from 0.18% to 0.38%) and software (from 0.16% to 0.27%). The biggest IT-related clue to the post-2004 slowdown lies in the fall in computer hardware contributions from 0.38% to 0.12%). It appears that the 1990s and the early 2000s were a uniquely strong period in the build-up of IT-related capital stock.

The second channel of effect for IT investment lies in its dedicated production sectors. The TFP contributions of the IT Producing Sectors are reported in panel (C) of Table 1.1. To be clear, this represents the direct contributions of the hardware and software producing industries to propping up economy-wide TFP growth. The strongest shift here lies in semi-conductors where the TFP contribution jumped from 0.36% to 0.72% between the 1974-1995 and 1995-2004 intervals, but fell away sharply from the mid-to-late 2000s.

Measuring TFP changes in the IT producing industries is a key methodological problem here and is sure to affect the picture we have of trends over time. For example, a number of recent growth accounting efforts (Byrne et al., 2013; Byrne and Corrado, 2017) trace problems in the pricing of activity in the semi-conductor and software sectors that in turn affects productivity estimates.

However, as we discuss in the next sub-section, it is hard to fully reconcile a mismeasurement story with the sharp turn in trends we see occurring in the mid-2000s.

Table 1.2 provides a breakdown of European labour productivity changes derived from Van Ark and Jäger (2017). This shows the heavy impact of the financial crisis in 2008-2010 which saw labour productivity growth decline from 3.7% to 0.7% in manufacturing and 1.4% to -0.5% in market services. If we compare figures for the pre-crisis (2002-2007) and post-crisis (2011-2015) periods it is apparent that the sources of the slowdown are evenly spread but it is notable that the contribution of IT (both in terms of capital deepening and TFP contributions) halved between the two intervals.

The Long View on TFP

Current debates about the evolution of productivity are well-served by a long-term analysis. In Figure 1.4 we plot the path of TFP for the US, Europe and Japan for the full period since 1900. This reveals some evidence of potential ‘waves’ in the evolution of aggregate TFP. The literature has associated these waves with the emergence of major technological changes that induced economy wide changes in activity. In particular, Gordon (2015) identifies the interwar period in the US as a ‘big wave’

characterised by the arrival and consolidation of a set of important inventions such as electricity, internal combustion, industrial chemistry and mass electronic communication (radio, TV).

Statistical analysis of these TFP patterns does support the idea of systematic cyclical shifts. Bergeaud et al. (2016) characterise this cycle as potentially based on a ‘big wave’ in the US after the end of WWI in 1918 when US leadership was first established. A similar wave then appears in Europe after WWII as the continent plausibly catches up with and absorbs the technological revolution that the US previously benefited from.

An important point with this historical analysis is that technologies such as electricity and the internal combustion engine were well established within the body of technological knowledge before 1918 but not necessarily contributing to economy-wide productivity improvements. Hence this feeds into an analysis based on ‘General Purpose Technologies’ that are defined by long implementation and gestation periods (Gordon, 2012; Gordon, 2015).

In turn, the analogy with the ICT revolution presents itself directly. The ‘Solow Paradox’ period can be understood as a gestation period during which both ICT capital and a range of complementary assets (e.g.: skills, new organizational structures) were being accumulated so as to fully exploit the underlying technology. Hence investment levels led productivity growth which eventually surged in 1995. Syverson (2017) directly compares the electricity and ICT revolutions and finds that the timing of gestation and delivery match up closely.

This big wave perspective forms the implicit basis of Brynjolfsson and McAfee (2014) hypothesis of a ‘second machine age’. The first machine age in their narrative actually runs from the industrial revolution up until the development of widely deployed computing. That is, the first age encompasses initial industrialization as well as the 20th century wave of technologies such as electricity and internal combustion. Computers are posited as a qualitatively new technology because of the exponential nature of their progress and their ability to capitalise on digital information resources. The specific current hypothesis in the Brynjolfsson and McAfee (2014) narrative is then that a follow-up ICT revolution to the post-1995 surge can be plausibly envisaged. This follow-up revolution is most likely to be grounded in artificial intelligence and robotics technologies as they seem to have the capacity to decisively increase the scope of what can be done with ICT.

1.2.2. Mismeasurement and Consumer Surplus

The mismeasurement argument for the productivity slowdown is derived from the idea that the full value of the goods and services associated with the internet has not been counted. If this value was counted then output would be higher and so would labour productivity. This argument tallies with the general social intuition that internet-related services have provided pervasive personal and commercial benefits that are felt in everyday activity.

The first step to evaluating this argument is to understand how much potential ‘phantom output’ needs to be explained to account for the productivity slowdown. Syverson (2017) calculates this by extrapolating the annual rate of slowdown (just under 0.4% between 2005-2015) to an estimate of counterfactual output. This results in a figure of \$3.45 trillion per year, which more tangibly amounts to approximately \$10,800 per capita and \$27,700 per household. A lower bound estimate that incorporates the farm, government and household sectors leads to a figure of \$2.92 trillion. This is a

huge figure - equivalent to around 15-20% of US GDP - but serves to emphasize how pervasive the slowdown has been across different sectors of the economy.

The next step is to then measure the consumer surplus associated with ICT. The economic window for capturing this effect lies in the consideration of broadband internet. Since broadband internet (that is, the expansion of higher speed ADSL-style services in the early 2000s) serves as a gateway to many ICT-related goods and services, valuations derived from its pricing and consumption are a good way to summarise overall consumer surplus.

The estimates of consumer surplus from broadband internet are varied. This variation is mainly due to the methodological difficulty of measuring the gap between actual price paid and consumer willingness to pay. A central issue here is modelling consumer demand responses with respect to different pricing regimes. In an adjacent 'box' we provide a summary of recent research by Cohen et al. (2016) which analyses ride-hailing data derived from Uber's app. The setting there is unique in allowing the researchers to compare the responses of consumers across a menu of prices in clearly measurable and highly comparable conditions.

The lower end of estimates for consumer surplus from broadband internet comes in at \$17-26 billion up to 2015 (Greenstein and McDevitt, 2011). Two mid-range estimates are \$96 billion (Dutz et al., 2012) and \$54-73 billion (Rosston et al., 2010). The most methodologically complex study in this area (based on household-level data taken to an empirical model of dynamic demand) posits a \$132 billion estimate (Nevo et al., 2016).

Goolsbee and Klenow (2006) offer an approach that focuses intensively on also incorporating the time spent using the internet as an indicator of consumer value. This approach is adapted by Brynjolfsson and Oh (2012) and Syverson (2017) to construct a consumer surplus estimate of approximately \$842 billion. Returning to our basic motivating problem, the issue here is that that this generously (but rigorously) calculated figure still only amounts to about one-third of the estimated lost output over the period of the productivity slowdown (\$2.92 trillion). Hence the central problem of the mismeasurement hypothesis is that the scale of the slowdown requires too great a measurement problem to be plausible following realistic economic arguments.

A case can be made that the valuations centred on broadband internet do not capture the benefits that have accrued via mobile platforms since smart phones began to diffuse in earnest around 2007. These mobile platforms can be credibly regarded as distinct in their economic role from fixed-line systems and are particularly important in terms of generating a diverse and innovative 'app' economy. However, as we explore in the adjacent box ('The Consumer Benefits of Uber') estimates of consumer surplus directly derived from rich Uber data suggest a consumer surplus figure of just under \$6.8 billion. Hence over 20 more apps with a similar radical economic impact to Uber would be needed to explain the remainder of the slowdown not accounted for by the estimates of broadband internet reviewed above.

The final consideration to make with respect to the mismeasurement argument is a historical one. While the mismeasurement of output in relation to new innovations is a plausible contemporary problem it is also an issue that has loomed over GDP measurement since national accounts began as an explicit formal system in the 1930s. The internet and smart phone were significant innovations that affected economic measurement but so were radio and television, along with the earlier GPTs that dominated

the economy prior to World War II. Any argument for mismeasurement in the modern era must rest on the assumption that this problem is significantly worse than it was for earlier GPTs.

1.3. Policy Implications and the Economics of ICT

In this section we provide a discussion of the policy implications of ICT's evolving role in the economy. We tie these policy implications specifically to the contributions developed in the later chapters and draw on the agenda developed as part of the LSE Growth Commission (Commission et al., 2013).

1.3.1. *Business, Innovation and Infrastructure*

The role of ICT in business activity can be summarised along two lines: production and consumption. Firms can produce ICT hardware and software or benefit from the consumption of ICT goods either internally or as a result of new economies created from public adoption (e.g.: the App economy, internet commerce in general). Policy should therefore be organised along these dimensions of production and consumption.

- Our Patent Rank measure of knowledge spillovers indicates that ICT-related innovations have an enduring advantage in generating spillovers. This points to benefits from directing industrial policies (e.g.: tax credits, subsidies) towards ICT production activities.
- The basis of this argument for ICT-focused industrial policies lies in the fact that businesses do not internalise these spillovers by definition and this leads to under-investment. Since ICT-related fields demonstrate the highest spillovers this suggests that the problem is potentially most acute there.
- That said, new frameworks are needed for coherent industrial policy. Such policies should be based on clear public guidelines and oversight. There is also an argument to set up independent control or oversight of industrial policy in the same manner as the Bank of England and monetary policy. Such independent control would allow long-term planning and insulate policy-making from overt rent-seeking Commission et al. (2013).
- Our modelling of the 2000s broadband roll-out uncovers some modest effects of early roll-out on occupational structure and wages. In particular, areas that received access to ADSL broadband earlier experienced higher growth of professional, non-routine employment. These effects should be considered a lower bound since the areas with a lagged broadband roll-out have since had a chance to exploit the technology and catch-up.
- Broadband policy should be considered as an innovation policy as much as an infrastructure matter. As more households and firms gain access to higher quality broadband this creates new markets and generates consumer surplus. The UK has had some success in closing 'digital divides' in broadband speed that emerged in the 2000s and this effort should continue.

1.3.2. *The Labour Market*

- Automation - specifically the displacement of jobs by machines - is the major challenge created by ICT in the context of the labour market.
- Our estimates of employment changes across occupations for the US do not indicate that a 'new wave' that strikes at higher-skill non-routine jobs has taken root. However, this evidence also

shows that technology-inspired restructuring in jobs tends to occur during recessions. This raises the prospect that the ‘new wave’ may start at the point of the next major downturn.

- Projections of the widespread susceptibility of jobs to automation should take into account the trend for ‘partial automation’. This is the perspective that it is tasks that get automated rather than entire occupations. Estimates of job vulnerability based on this partial automation approach (Arntz et al., 2016) put the proportion of jobs-at-risk at around 10-15% versus the 40-50% of an earlier range of estimates.
- Previous research (Cortes et al., 2014) has shown that major challenges have emerged in terms of the redeployment of workers after losing employment in routine jobs. This adjustment problem has deepened with every recession since the 1980s.
- The consensus policy solution since the 1990s for worker adjustment has been investment in training and human capital in general. However, the actual shape and realisation of this type of policy has been sketchy.
- As the LSE Growth Commission (Commission et al., 2013) outlines there is effectively a policy agenda ‘on the table’ for human capital. The current tax system in many countries heavily favours investment in physical capital, computers and machines, but offers few incentives for investment in human capital. A viable option here is a Skills and Training Tax Credit, modelled after the successful R&D tax credit model that operate in many countries.
- The tax system also play a role in fostering insecure ‘gig economy’ work by encouraging self-employment classification and loopholes with respect to employee benefits. The design of the tax system should aim to be neutral with regard to forms of employment (Commission et al., 2013).

Projecting the Impact of ‘Super-Innovations’

What kind of impact could a set of AI and robotics-based ‘super-innovations’ have on employment and productivity? The scenarios below are modified versions of those sketched out by Brynjolffson, Rock and Syverson (2017).

Autonomous Vehicles: Estimates by the US Bureau of Labour Statistics (BLS) put the size of the commercial driver workforce (classified under ‘motor vehicle operators’) at 3.5 million, which is approximately 2.8% of the total US nonfarm employment circa mid-2016 (122 million). Consider a hypothetical 60% reduction of the commercial driver workforce to around 1.5 million - a sweeping but plausible change. A static calculation gives us an aggregate productivity boost of 1.7% ($=122/120$) which, spread out over a 10 year transition period, would deliver a 0.17% increase in annual productivity growth. A ‘dynamic calculation’ would need to encompass the complementary effects of autonomous vehicles within the overall economy, such as a more efficient transportation system, lower private car ownership and enhanced consumer convenience.

What about the workers? Assuming a steady shedding of employment would imply that 200,000 drivers per year would be laid off. The current volume of layoff separations is approximately 19 million (BLS 2016 JOLTS Survey) indicating that this autonomous vehicle scenario would increase total layoffs by 1% per year. As a bench- mark, the peak of the late 2000s recession saw annual layoffs increase by about 10% per year between 2007-2009 (from 22 million to 26.6 million).

AI-Staffed Call Centres: Consider a scenario where IBM Watson-style AI technology was deployed in telephone call centres. A large mass of calls take the form of routine, basic queries that an AI could successfully deal with (e.g.: ‘Have you tried turning your computer off and on?’). Official BLS estimates (BLS Occupational Employment Statistics) put the size of the office and administrative section of the Telephone Call Center (NAICS Code 561 420) industry at 327,000 in mid-2016 although other estimates that posit coverage of home-based or small site activity put the number at 2.2 million (Site Selection Group 2015).

If we adopt a ‘maximalist’ scenario and assume 60% of the 2.2 million workers could be replaced this would imply a 1% increase in productivity in total over 10 years, annualized to 0.1% per year. On the labour market side there would be an extra 130,000 layoffs per year which would represent a 0.7% annual increase in lay-offs.

The Consumer Benefits of Uber - Evidence from the App

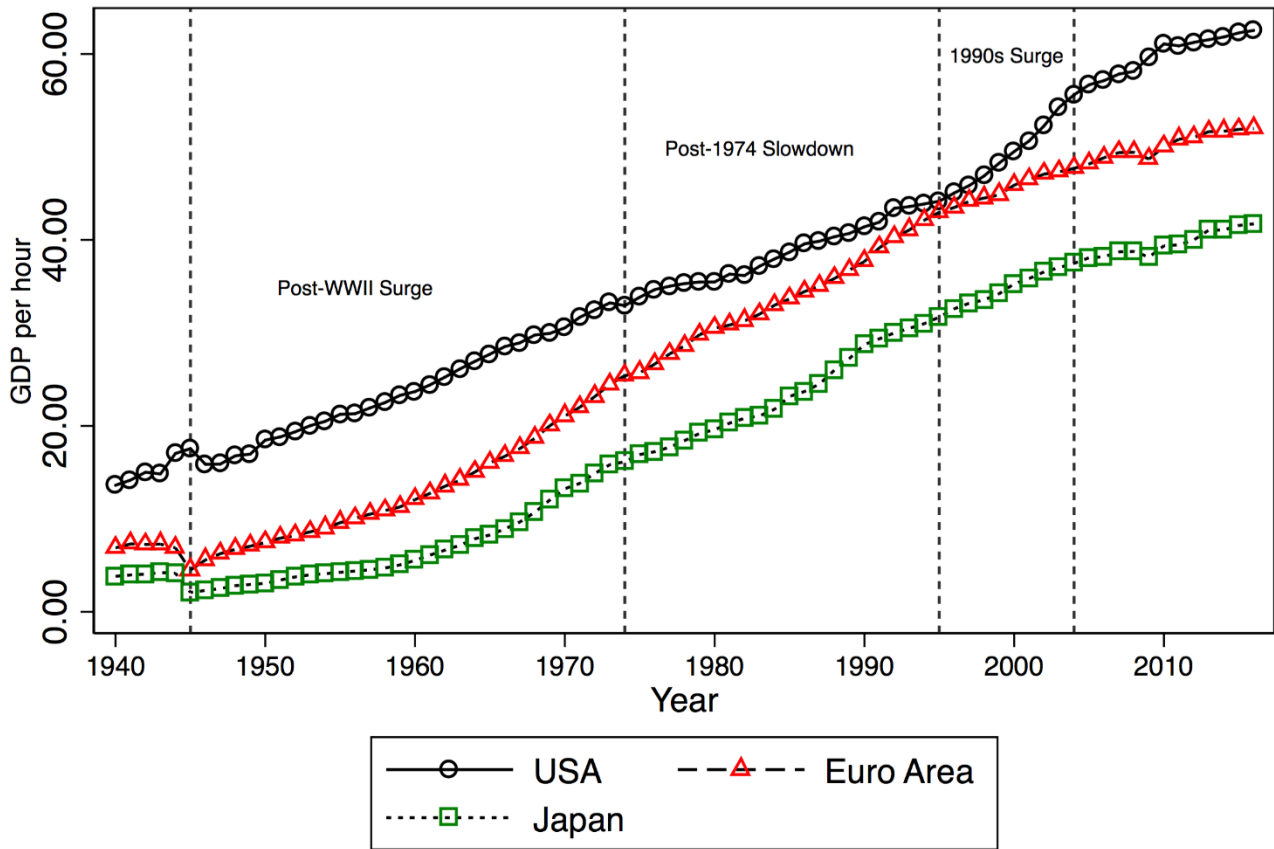
The recent study by Cohen et al (2016) uses 'in-house' data from the Uber app for the 4 largest city markets to measure the consumer surplus associated with ride-hailing technology. The innovative contribution is that this data allows the authors to understand how consumers respond to different levels of price, thereby getting at the idea of consumer surplus as the gap between actual price paid and maximum willingness to pay.

Practically, this boils down to modelling demand where random price variations due to the rounding of app price quotes creates the variation necessary to trace out consumer prices to different levels. Further to this, the study therefore uses the setting of surge pricing to get a range of price variation. Around 20% of Uber price sessions (i.e.: episodes of people opening the app to look for a ride) involve some type above-normal surge pricing. The demand elasticity is estimated as being in -0.4 to -0.6 range, which in formal terms is relatively inelastic. For example, an elasticity of -0.6 indicates that a 10% price increase is associated with a 6% decrease in demand.

Based on this, the estimate of consumer surplus for the US in total is \$6.76 billion. This figure is 1.57 times larger than consumer expenditures value at base 1.0x pricing, meaning that for every \$1 spent at the base price the consumer receives \$1.57 in surplus.

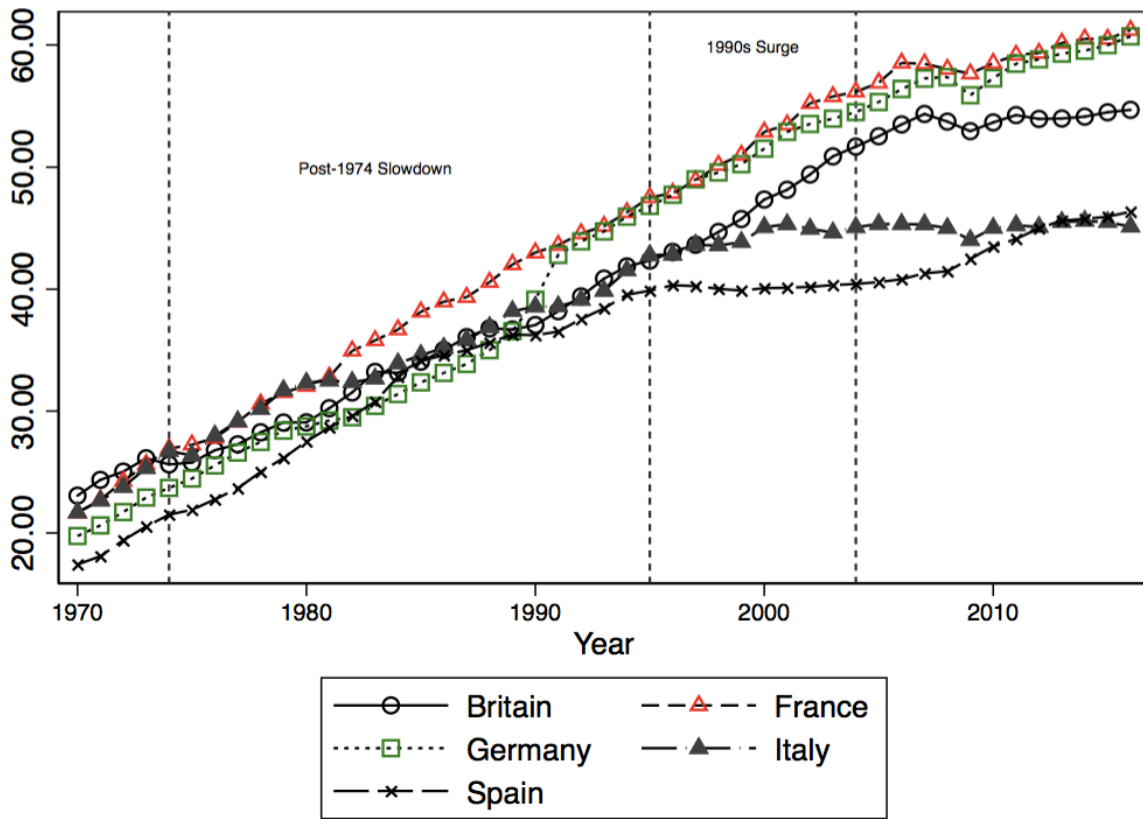
You can think of this surplus as the economic underpinning of Uber's de facto 'too big to ban' strategy (Wolhsen 2014) for roll-out. The large consumer surplus generated by ride-hailing provides a political momentum for the regulatory acceptance of the innovation. The policy choice then comes in comparing this benefit in terms of consumer surplus to the losses incurred by the existing taxi industry as a result of competition from ride-hailing services.

Figure 1.1: Labour Productivity - US, Europe and Japan, 1940-2016



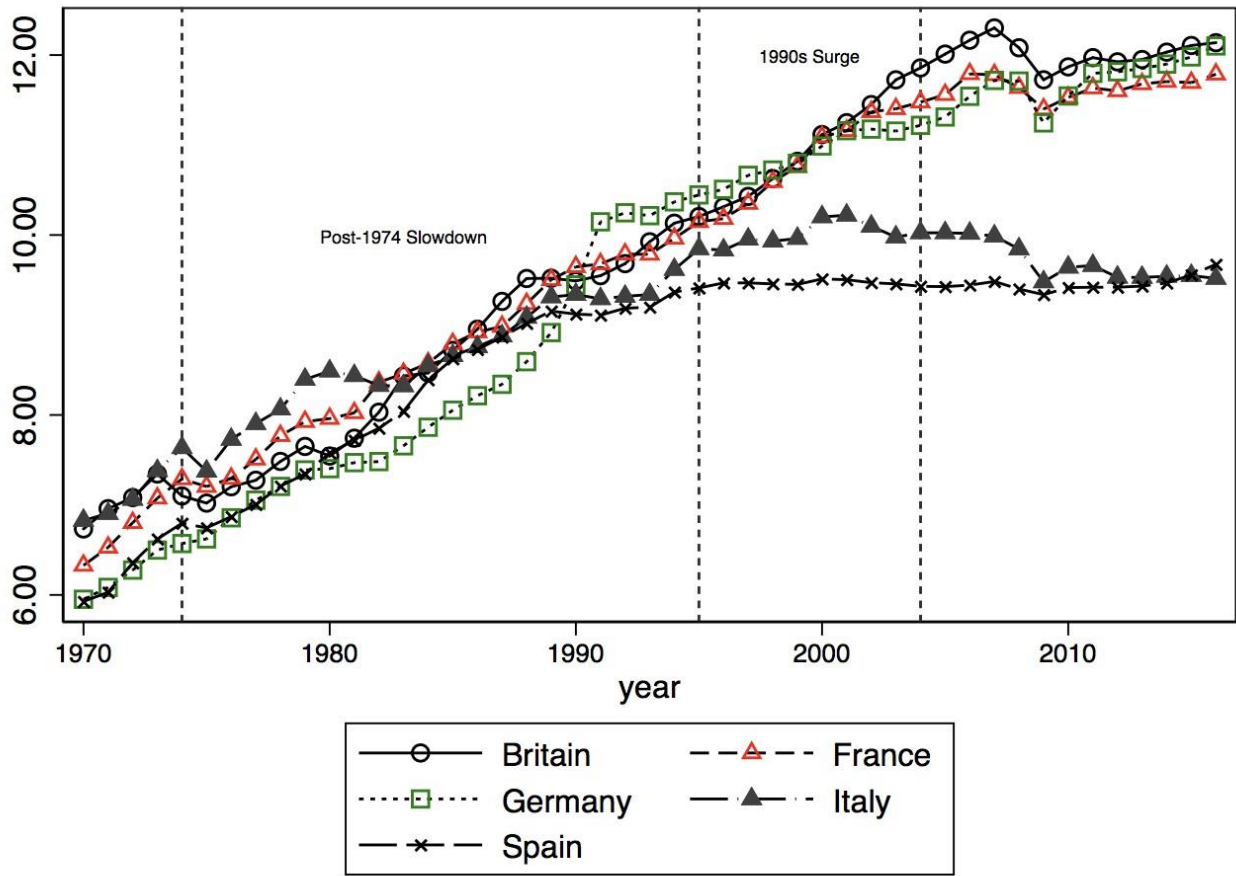
Notes: Data from Bergeaud et al. (2016) Long-Term Productivity Database.

Figure 1.2: Labour Productivity - Major European Economies, 1970-2016



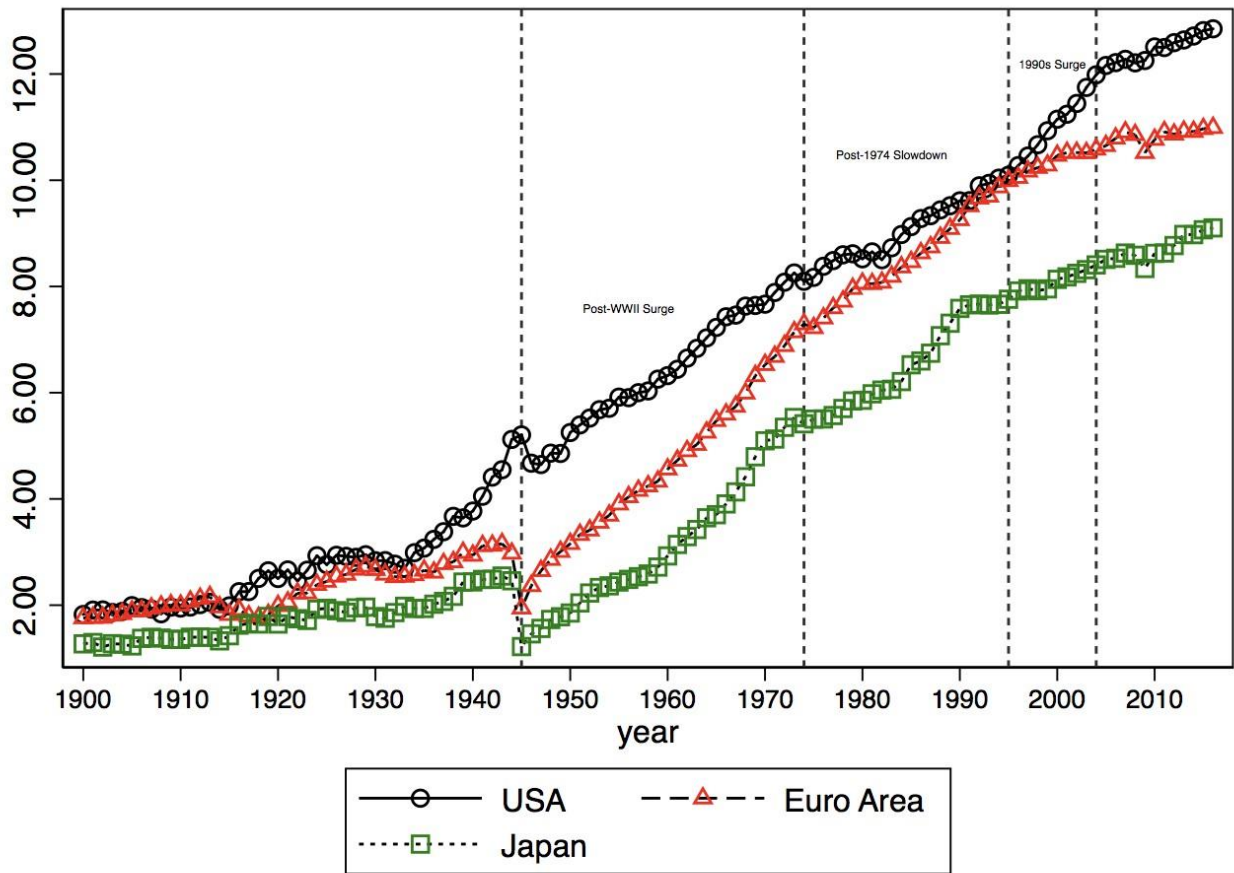
Notes: Data from Bergeaud et al. (2016) Long-Term Productivity Database.

Figure 1.3: Total Factor Productivity - Major European Economies, 1970-2016



Notes: Data from Bergeaud et al. (2016) Long-Term Productivity Database.

Figure 1.4: Total Factor Productivity - US, Europe and Japan, 1970-2016



Notes: Data from Bergeaud et al (2016) Long-Term Productivity Database.

Table 1.1: US Labour Productivity Growth, 1974-2012

Name	1974-1995	1995-2004	2004-2012
(A) Growth of Labour Productivity	1.56	3.06	1.56
Capital Deepening	0.74	1.22	0.74
Labour Composition	0.26	0.22	0.34
TFP	0.56	1.62	0.48
(B) Capital Deepening Contributions			
IT Capital (Total)	0.41	0.78	0.36
Computer hardware	0.18	0.38	0.12
Software	0.16	0.27	0.16
Communication Equipment	0.07	0.13	0.08
Non-IT Capital Deepening	0.33	0.44	0.38
(C) TFP Contributions			
IT Producing Sectors (Total)	0.36	0.72	0.28
Semi-conductors	0.09	0.37	0.14
Computer hardware	0.17	0.17	0.04
Software	0.06	0.1	0.08
Communication Equipment	0.05	0.07	0.02
Other Sectors (Total)	0.13	0.9	0.06

Notes: Based on information from Byrne et al. (2013). Rates measured as 100 times annual log difference for the indicated years. US non-farm business sector. Underlying data is via Bureau of Labor Statistics (BLS) and National Income and Product Accounts (NIPA)

Table 1.2: European Productivity Growth, 2002-2015

Name	2002-2007	2008-2010	2011-2015
(A) Labour Productivity Growth (Total)	1.1	0.2	0.7
Labour Productivity Growth - Manufacturing	3.7	0.7	1.4
Labour Productivity Growth - Market Services	1.4	-0.5	0.9
(B) Gross Value Added Growth	2	-0.6	1
Contribution of:			
Labour Input Growth	0.6	-0.2	0.4
– Total Hours Worked	0.4	-0.5	0.2
– Labour composition	0.2	0.3	0.3
Capital Input Growth	1	0.5	0.4
– ICT Capital	0.2	0.1	0.1
Non-ICT capital	0.7	0.4	0.3
TFP Growth	0.5	-1	0.2

Notes: Based on information from van Ark and Jäger (2017). Figures refer to the EU-12 member states (weighted average of: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Italy, Netherlands, Spain, and Sweden).

2. Chapter 2: ICT and the Production of Ideas: Tracking Knowledge Spillovers Across Technology Areas

2.1. Introduction

ICT is commonly referred to as a general purpose technology (GPT) (Bresnahan and Trajtenberg, 1995; David 1990). This is not surprising as the influence of ICT has diffused into nearly every area of life and economic activity. As an overall technology, it seems credible that ICT meets the main conditions of being a GPT, namely (1) a record of rapid technical improvement, (2) evidence of pervasive application across sectors, and (3) performance as a platform for a ‘swarm’ of subsidiary innovations (Helpman, 1998).

However, the research record has shown that it has been hard to quantify the evolution of GPTs across the available data. In this chapter, we put forward a metric for tracking technological influence in the context of modern patenting trends. The central concept behind our metric is the idea of knowledge spillovers. Knowledge spillovers arise if the innovation and knowledge generated by one inventor helps other inventors in advancing knowledge and coming up with further innovations. Hence, knowledge spillovers relate fundamentally to the cumulative pattern of invention.

The traditional empirical definition of spillovers has focused on identifying patterns of what could be called ‘passive intellectual influence’. For example, a common strategy in firm-level re- search (Bloom et al., 2013; Griliches, 1991) has been to construct measures of R&D capital stocks by industry or technology class and then try to detect how the innovation profiles of individual firms rise and fall in line with the knowledge that is available around them ‘in the air’.

In the following work we set out an approach for measuring these knowledge spillovers that aims to capture how this knowledge that is ‘in the air’ is evolving. This is crucial for tracking the state of scientific development across fields and quantifying the growth of knowledge at different times. Our approach is centred on a new measure of knowledge impact for patents that is based on the same principles as Google’s Page Rank algorithm. In short, our approach quantifies the network of citations that emanate from particular patents and, in aggregate, their associated technology fields.

The ‘Patent Rank’ measure that we develop allows us to characterise the comparative advantage of different countries across technology areas in terms of how rich or dense their network of invention is by scientific field. In turn the Patent Rank metric is a viable input for the design of industrial policies that are meant to assist the development of industry-relevant basic science.

In the sections below we first provide a short summary of the main findings and provide a guide to the Patent Rank methodology before reporting the detailed information on trends in knowledge spillovers.

2.2. Key Findings

2.2.1. Results for Patent Rank Spillovers

- ICT technologies generate on average substantially larger knowledge spillovers than those generated by other technology areas and this holds even when ICT is compared to other frontier fields such as biotech and clean energy.
- For instance, as part of the results below we find that wireless technologies generate on average more than 50% higher knowledge spillovers than other comparable fields. This calculation is based on a monetary estimate of spillovers that extends the basic framework of the Page Rank-based analysis.
- Across countries, the pattern of national spillovers (defined in terms of the citation network for innovations in the same originating country) basically mirrors that of the pattern of comparative advantage or specialisation according to innovation per capita. Some notable exceptions are the high average spillovers for 3D technology in the UK, robotics in Italy and wireless in China.
- The growth in average global spillovers in ICT flattens out from the early 2000s. This involves the tapering of a sharp spike in knowledge spillovers for ICT earlier in the 1990s. However, ICT spillovers are still around 25-50% higher than the pooled set of all other technology areas at the end of the period considered.
- In terms of technology areas within ICT, wireless experienced a long period of high spillovers from the early 1980s onwards. Robotics experienced some major peaks in spillovers in the early 1980s but dipped before a moderate resurgence. We note though that lags in the citation process mean that we are unable to capture the spillover contributions of the latest wave of robotics research in the late 2010s.
- The evidence over time indicates that knowledge spillovers evolve cyclically, most likely as a result of innovations waves that take place in the wake of scientific breakthroughs. The latest flattening in spillover growth preceded the latest productivity slowdown but a clear link cannot be made without a longer term analysis using data from before the 1970s.
- A pattern of falling research productivity is evident from other research in areas ranging from semi-conductors, drugs and general listed firm performance. Increased amounts of inputs (scientists, R&D expenditure) are required to maintain a constant level of innovation output. The cyclical nature of the spillovers process is a potential underpinning of this challenge in regard to sustained research productivity.
- Other research has also documented a surge in AI-related research since 2009, particularly in the field of learning. The fact that this surge is so recent means that information about the pattern of spillovers is yet to accumulate. The scale of the recent surge in the number of AI-related innovations also means that average knowledge spillovers are likely to rise in the near future, but it is hard to predict the size of this effect.

2.2.2. Results for Country-level Innovation per Capita

- Japan and Germany are the leading countries in terms of innovation per capita, that is, the number of inter-jurisdictional patents per head of population. China and Korea display the highest rate of growth in innovations per capita, with rates of 300-400%.
- Korea, the US and Japan are the main countries with a comparative advantage in ICT, defined as higher relative shares in the technology area with respect to the rest of the sample. The UK

shows a strong comparative advantage in Biotech. No country shows signs of an obvious comparative advantage in robotics.

2.3. Measurement of Spillovers

A common approach to studying knowledge spillovers is on the basis of citation records in patent data. Clearly, this type of analysis has its limitations both because not all innovations are necessarily patented and not all forms of knowledge spillover are recorded via a patent citation. However, there are still limited alternatives for empirical work in this area, in particular when the goal is to explore issues across a range of technologies and across countries.

The setting for our research is the PATSTAT database, which has been developed by the OECD and the European Patent Office (EPO) to consolidate global data on patents. PATSTAT combines information from nearly all patent offices worldwide into a single relational database going back to the 1950s. We can trace not only citations but can also identify patent families; i.e. the same underlying innovation being patented in different jurisdictions. Hence the following analysis is always at the level of common, cross-country ‘innovations’ rather than patents.

As mentioned, a common way to identify spillovers derives from citations. We build on this approach by introducing a new spillover measure which we have dubbed Patent Rank¹. It is an adaption of the Page Rank measure that has been deployed by Google to rank the importance of web pages where citations take the place of the hyperlinks that are the central focus of the search engine problem.

Page Rank boils down to a simple formula. The page rank of a given web page i is the weighted average of the page rank of all web pages j that cite web page i . As the weights we use the inverse of the number of (backward) citations web page j makes. Of course to calculate the page rank of i we need to know the page rank of all pages j . In technical terms, writing down a page rank formula for all web pages essentially yields a linear equation system that can be solved. The only challenge is that it is a rather large equation system with billions of equations in the case of web pages and millions in the case of innovations. However, there are simple recursive algorithms that can handle this in a reasonable amount of time.

Why is Page Rank a sensible approach analyse knowledge flows in patent data? There are a number of advantages to the approach we develop:

- Unlike simple citations counts it accounts for indirect connections: consider two innovations A and B with each possessing just one other innovation referring to them (As in Figure 2.1a). A citation count will rank them equally. However, innovation A is cited by an innovation that is quite heavily cited (3 times in the figure) whereas innovation B is cited by an innovation that is not cited at all. Naturally we would consider innovation A more influential.
- Page Rank adjusts for the extent of backward citations: consider Figure 2.1b. Again we have two innovations (D and E) with exactly one citation. Innovation C is only referring to innovation D. Innovation G on the other hand refers to both F and E. Hence, it is natural to assume that D had a bigger influence D compared to the influence of E on G which is what Page Rank does.

¹ See Dechezleprêtre et al. (2014) for a full exposition of the Patent Rank measure in the context of clean and dirty environmental technologies.

- It is internally consistent. The premise of looking at citations is that we presume that spillovers differ between different innovations. However, by simply counting citations we impose that the value of every citing innovation is uniform. By using an approach based on Page Rank we correctly allow for heterogeneity on both sides of the citation flow relationship.

What is the difference between Page Rank and Patent Rank? Page Rank can be interpreted as a probability. Specifically, it is the probability of a random surfer arriving at a particular web page. A random surfer is a web user that starts at a particular web page and then randomly clicks on one of the hyperlinks (citations) on that web page which brings her to another web page where she does the same. However, in developing Patent Rank, instead of accumulating probabilities over hyperlinks we load the algorithm with estimates of the private value of innovations across different steps in the citation network.

Consequently, the Patent Rank can be interpreted as the social value of a patent consisting of its private value plus a fraction of the combined private value of all innovations that refer back to it either directly or indirectly. We derive the private values from a simple regression of stock market values on innovation activity by firms; i.e. we assume that the discounted private added value of an innovation corresponds to the average marginal effect of an additional innovation on a firm's overall valuation. At present we use a uniform average value of \$250 thousand but it is straightforward to incorporate heterogeneity in private values by the type of patent.

We compute two basic types of Patent Ranks: Global and National. The global measures are exactly as described above, that is, the private value of an innovation plus a fraction of the value of all direct and indirect references to the innovation. National Patent Ranks are different for every country. When computing the National Patent Rank for country A we only take into account private values that arise to innovators in country A; i.e. for every innovation we compute the social value of the innovation to country A.

We argue that this can provide a potential indicator for national governments in identifying which technology areas to focus on when designing industrial strategies. For instance, an industrial strategy based on national Patent Rank measures would suggest a focus on those areas that are likely to generate the highest spillovers within national jurisdictions. The National Patent Rank measure therefore allows analysts to trace such national spillovers whether they arise directly or indirectly; e.g. an innovator in country A might inspire innovation in country B which then inspires innovation in country C again.

2.4. Innovation Patterns Across Countries

Before looking into spillovers this section provides some background information on innovative output across countries and the importance of ICT as a technology field. We focus on the G7 as well as China and Korea as the main countries of interest.

As Figure 2.2 shows there is considerable dispersion in innovative output across countries. Japan is the clear leader with more than 1 innovation per 1000 residents over the period from 2000 onwards. Germany, the second most innovative country, has just over one half of an innovation per 1000 residents. Panel B of Figure 2.2 then shows how per capita innovation output has grown across countries. This strongly suggests a pattern of convergence with lagging countries such as China displaying remarkable growth rates of more than 300%. However, it is Korea that displays the fastest

growth with a rate of more than 400%. In addition, this trend is notable because Korea started from a fairly high baseline level of innovative output at the beginning of the period being considered here.

Figure 2.3 explores how relevant ICT technologies are in the innovation profiles of different countries. The most ICT intensive country is Korea followed by the US and China. However, most countries are fast increasing their ICT shares, except for Korea where we see a slight decline in the ICT share when comparing the 2000s to the 1990s.

An important focus for policy makers is understanding the comparative advantage of a given economy, that is, identifying the sector or technology area where a country has the biggest relative strength. The most common approach for classifying comparative advantage usually derives from studying country-level export shares. Figures 2.4 and 2.5 examine this in terms of the number of innovations by general technology category. We subset out a number of ICT-related technology fields, such as Wireless, Robotics, 3D imaging and a general ‘other ICT’ category. These ICT fields are then compared to the Bio Tech and Clean Tech technology areas since these can be considered as ‘frontier fields’ that are most likely to rival ICT in terms of the generation of knowledge spillovers.

For our sample of leading economies we examine in what technology areas they have the biggest relative share. Practically, this means that for every country we report the share of a country’s innovation output relative to the average innovation share (across countries) in that category. For instance panel (d) tells us that the UK’s share of innovations in Biotech is 2 percentage points higher than the average innovation share in Biotech across leading economies. In turn, this suggests a strong comparative advantage in Biotech for the UK.

Interestingly, only four of the countries considered have a comparative advantage that lies within the ICT categories: Korea, the US, Japan and Canada. In designing industrial strategies governments often express their desire to support areas of national comparative advantage. This would mean only limited support for ICT. However, we argue that industrial policies should be based on areas that generate the highest amount of spillovers. We will see below that an evaluation using our Patent Rank measure of spillovers leads to a substantially different ranking of technologies.

2.5. Spillovers

This section discusses the results on spillovers for different technology types by country and over time. We report average spillover values - as described above - along those dimensions.

2.5.1. Comparing Technologies

Figure 2.6 reports average external values as computed by patent rank for different technology groups. Panel A reports the global average patent rank. This is the external value that counts the network of citations regardless of the location of the citing patent. Again, we only consider patents that are assigned to patent holders based in the G7, China and South Korea.

Wireless technologies generate the highest global spillovers with on average nearly \$900,000 in external value. This can be interpreted as an estimate of the value of the indirect knowledge spillover that arises from an individual patent. We use a baseline value of \$250,000 to price the value of a single patent based on average stock market responses to new patents at the firm-level. Hence, the figure of \$900,000

indicates that over three times more knowledge is created indirectly via the spillover channel relative to the initial patent.

Panel B on the other hand reports average national spillovers, that is, we only consider spillovers that emerge in the same country as the originating innovation. This does not change the overall picture drastically, although the 3D imaging category changes rank when national values are considered.

These average across our selected group of leading countries hide considerable heterogeneity at the national level. Figure 2.7 reports separate values by country. We see that wireless technology is the dominant category in 4 countries: China, Japan, Korea and US. In Germany by contrast, the external (national) value of clean technologies exceeds that of wireless multiple times over. Indeed wireless is not even the second most important category but is exceeded by 3D and robotics. For France on the other hand, biotechnology is (narrowly) the leading category while the UK is characterised by large spillovers in 3D imaging.

2.5.2. Spillover Gaps Over Time

In light of recent concerns that global growth might have slowed Syverson (2017) it is instructive to examine how knowledge spillovers have developed over time. Spillovers are an important factor driving economic growth and productivity. If the flow of spillovers changes this might explain changes in the overall rate of innovation which in turn could explain changes in growth.

There are a number of challenges that arise when looking at spillovers as computed here over time. Patent and citation data are naturally right censored. This means that a patent can only be cited after it has been created. Hence, the last patent that has been created will have no citations irrespective of its potential to inspire spillovers. This problem is exacerbated by time lags in the processing of patent data, particularly the lag between the application and formal granting of patents. Practically, this means that reliable estimates of the pattern of knowledge spillovers are lagged by 5-10 years.

As a result, this analysis of knowledge spillovers is not able to fully capture the impacts of the recent explosion in AI-related research. We report on this via the recent work by Cockburn et al. (2018) in a related summary ‘Box’ in this chapter. This research shows a significant upward shift in AI-related research since 2009, particularly in relation to learning applications. This burst of innovation in AI is likely to result in an analogous increase in measured knowledge spillovers once follow-on innovations (and their backward citations) manifest.

We report the evolution of Average Global Spillovers for the pre-AI period from 1970 to the mid-2000s in Figure 2.8. The spillovers for the full set of ICT-related technology areas is compared to the pooled set of all other innovations. This shows that ICT has had an enduring lead in the generation of knowledge spillovers. The growth peaked in the late 1990s but a large gap between ICT and all other technologies is still persistent in the 2000s.

An important pattern here is the slowdown in the growth of average spillovers from the early 2000s onwards. In practical terms, the slowdown and flattening of average spillovers arises because of two forces. Firstly, the new innovations that are getting added to the total stock of inter-jurisdictional patents are getting less novel or ‘inspiring’ over time. Hence they are generating (on average) a lower level of citations (and therefore spillovers) over successive yearly cohorts. Secondly, the new innovations that are getting added to the total stock every year are become more diffuse in their citation patterns. That

is, they are less concentrated in their citation patterns and this lowers the average. In overall terms, the flattening in average spillovers that is tracked in Figure

2.8 implies that the knowledge network has therefore become proportionally less dense as new innovations have been added since the early 2000s.

The evolution of spillovers by technology areas is then shown in Figure 2.9, where in each case we benchmark against a series constructed as all other technologies. This shows that Wireless and other ICT went through a cycle of high average spillovers in the 1990s followed by a slowdown that still left them with above average rates compared to all other technologies. Robotics experienced a spillover boom in the 1980s which had levelled off by the 1990s. However, it must be noted that the data here does not yet fully cover the latest wave of robotics innovations from the late 2000s and 2010s that is discussed in Cockburn et al. (2018).

What explains the general flattening of average spillovers that has occurred since the 1990s? As our discussion suggests, the pattern of knowledge spillovers can realistically be regarded as cyclical. Scientific breakthroughs will occur at different junctures and lead to waves of follow-on innovations that manifest as citations and hence higher average spillovers. These waves will naturally exhaust themselves, leading to less rapid phases of innovation until the arrival of the next major breakthrough.

That said, the secular or common flattening across the majority of technology areas that we observe since the late 1990s is an important trend. It is notable that this dip in knowledge generation precedes the productivity slowdown of the mid-2000s, thereby providing a potential explanation for that development. However, drawing the link between these patterns of knowledge generation and productivity requires explicit testing at (for example) the industry level. This would also allow for the analysis of how economic factors such as changing patterns of competition and concentration may be affecting knowledge generation.

Further to this, the falling ‘knowledge productivity’ implied by the ebbing of average spillovers is consistent with evidence on falling research productivity gathered from across the economy by Bloom et al. (2017). We discuss this evidence in an associated text box in this chapter. The specific insight offered by this analysis of knowledge spillovers is that it seems that ‘knowledge productivity’ naturally goes through cycles as the results of major breakthroughs play out. The latest evidence on trends in AI-related research indicates that a resurgence in average spillovers is likely but it remains to be seen how large this effect will be.

2.6. Conclusion

The evidence above provides both general support for the role of ICT as a general purpose technology as well as some specific insights on the pattern of its contribution. It is clear that ICT is the dominant, knowledge-generating technology area over the period being considered, with a notable boom in ICT-related knowledge occurring in the 1990s.

This is both good and bad news for current policy considerations in relation to ICT. It is good news because it means that ICT technologies can accelerate economic growth and well-being not only through their direct use but also by speeding up innovation. On the downside it means that there is not nearly enough investment being undertaken when it comes to R&D in ICT sectors, because firms do not take into account the value of knowledge spillovers when making their private investment decisions.

This creates a strong case for government intervention to encourage further R&D investment, for example, via grants or tax credits. Indeed, it provides an evidence-based case for greater support towards ICT technologies relative to other technology areas.

However, there are also some important distinctions that need to be made about the scope and structure of ICT spillovers. Firstly, there is heterogeneity within different ICT sub-technologies as well as across countries. For instance, while ICT creates substantially more spillovers than other technologies in the US the same is not true for Germany, where clean technologies are a more valuable source of spillovers. Another contrast is the UK, where biotechnologies dominate the spillover ranking. Secondly, we find a dramatic drop in the spillover advantage of ICT technologies towards the end of our sample. While we don't know (yet) the reasons for this it could provide an explanation for the drop in productivity growth rates that have been observed in recent years.

A Turning Point in AI Research?

The increased attention paid towards AI research is a recent development. Cockburn et al. (2018) set out to track the rise of AI research from 1990 onwards by assembling a database of scientific publications and patents. This portfolio of AI research is chosen both by technology/subject class definitions and keywords.

Three main fields are identified: learning, symbolic systems and robotics. They find evidence of a significant shift in learning research since 1990 but particularly from 2009. This increase has also been focused heavily on applications. Surprisingly, US leadership in learning is not assured, with Cockburn et al. (2018) noting a ‘faddish’ patterns of activity for US science relative to more consistent trends for non-US science. Growth in robotics research has been strong but as pronounced as that for learning while activity in symbolic systems has been relatively muted.

Cockburn et al. (2018) make the assessment that ‘learning-oriented AI seems to have some of the signature hallmarks of a general-purpose technology (GPT)’. Specifically, learning AI is meeting the three GPT criteria set out by David (1990), namely: a pattern of pervasive applications across sectors, rapid improvements in performance, and a capacity to spawn substantive new innovations. In particular, the pattern recognition methodology underlying learning AI shows indications of being a unique new method for scientific invention and exploration.

The Challenge of Generating New Ideas...

New ideas are a driver of economic growth and productivity but what explains the generation of ideas? Bloom et al. (2017) study research productivity (that is, the amount of new ideas that are generated by some given amount of inputs) across a range of domains: semi-conductors, crop yields, health outcomes (mortality, new drugs), and the set of listed US firms.

The banner example is semi-conductors. Moore's Law suggests a doubling of semi-conductor capacity every two years and an effective growth rate of 35%. However, it has become increasingly costly in terms of inputs to feed this process. Research effort in semi-conductors (as measured by the number of researchers) has increased by a factor of 18 since 1971.

The increased expense of propping up Moore's Law is well-known and it is tempting to think that this case might just be explained by a diminishing space of technological opportunities in this area. However, a consistent trend of falling research productivity is evident across all the areas that Bloom et al. (2017) examine. In the case of US listed firms - the most general example that Bloom et al. (2017) consider - research productivity drops by a factor of 3.9 - 40.3 depending on the length of time studied.

Across all the examples surveyed it is estimated that research productivity has been falling at an average rate of -5.3% which practically means that innovators need to double their research efforts every 13 years to maintain the same rate of idea generation.

2.7. Tables and Figures

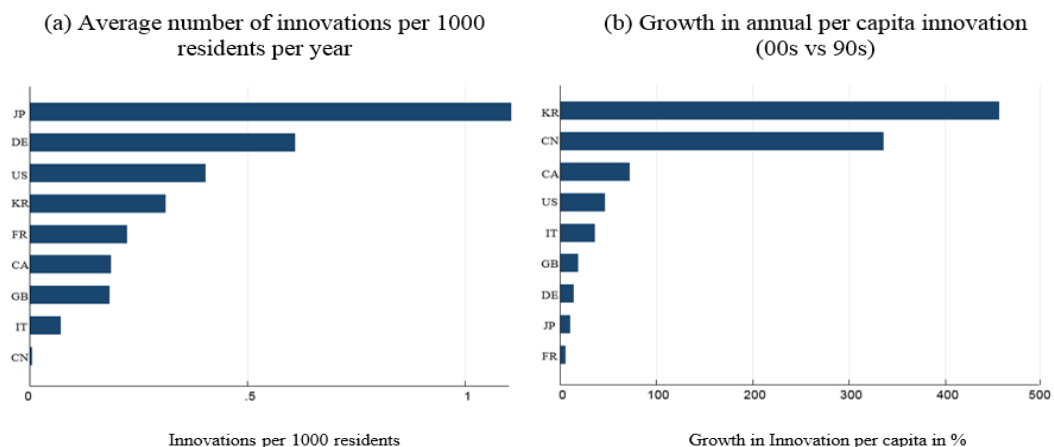
Figure 2.1: Applying Page Rank for Measuring Patent Influence



Notes: The figure illustrates how Page Rank as well as Patent Rank differ from a simple citation count.

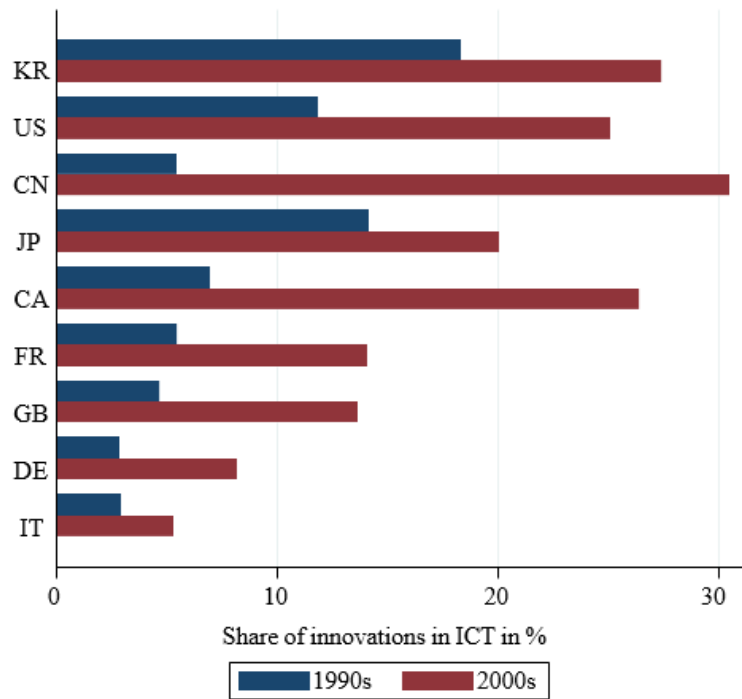
Firstly, indirect citation links are taken into account. An innovation receives a higher rank if it is cited indirectly by more innovations than another such as Innovation A in panel a. The figure shows a second order citation link. However, note that Page Rank allows for indirect links of any order. Panel b illustrates two innovations with the same direct citation count of one. However, in one case it is the citing innovations only reference in the other the citing innovation is also referencing to one other innovation. Page and Page rank would assign a lower rank to innovation E (and F) than D assuming that E had a smaller role to play in the creation of G than D in the creation of C.

Figure 2.2: Per capita Innovation over 2000 to 2010



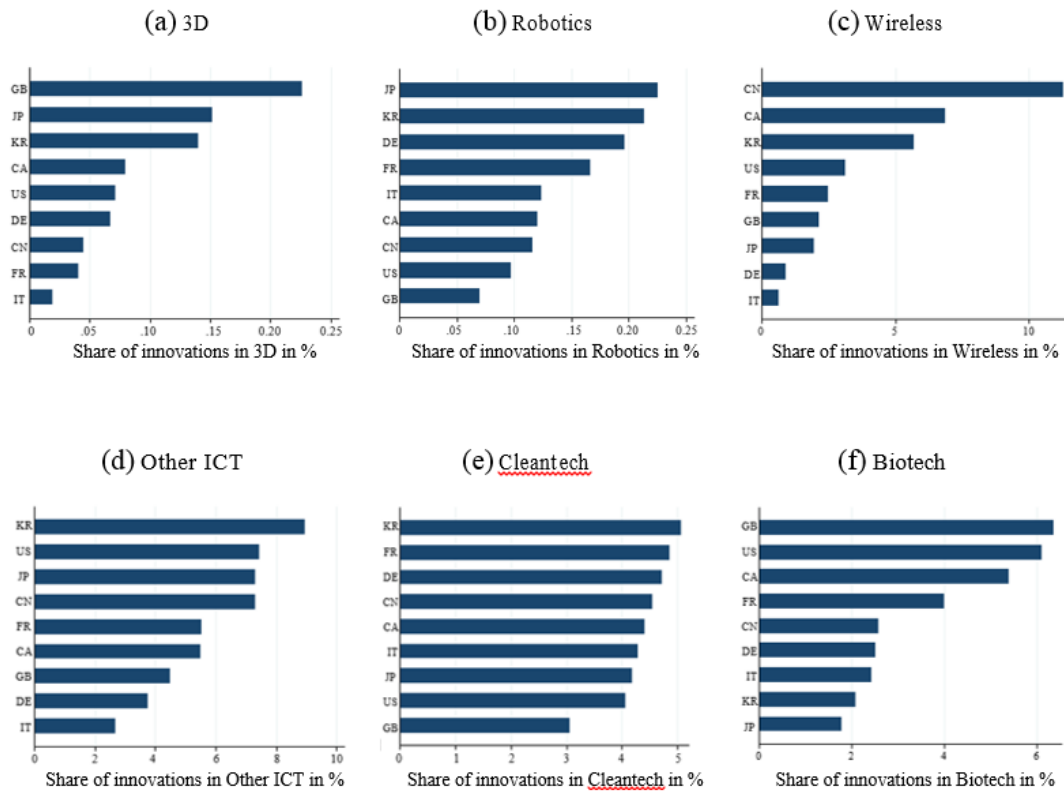
Notes: The figure (panel a) shows how many innovations were filed with patent offices worldwide by patent holders residing in a particular country per 1000 residents over the period from 2000 onwards (Per capita innovations). Note that we count innovations and not patents; i.e. multiple filings of the same underlying innovation across various countries don't lead to double counting. Panel b shows how per capita innovations have grown comparing the 1990s to the 2000s.

Figure 2.3: Share of ICT Innovations by Country



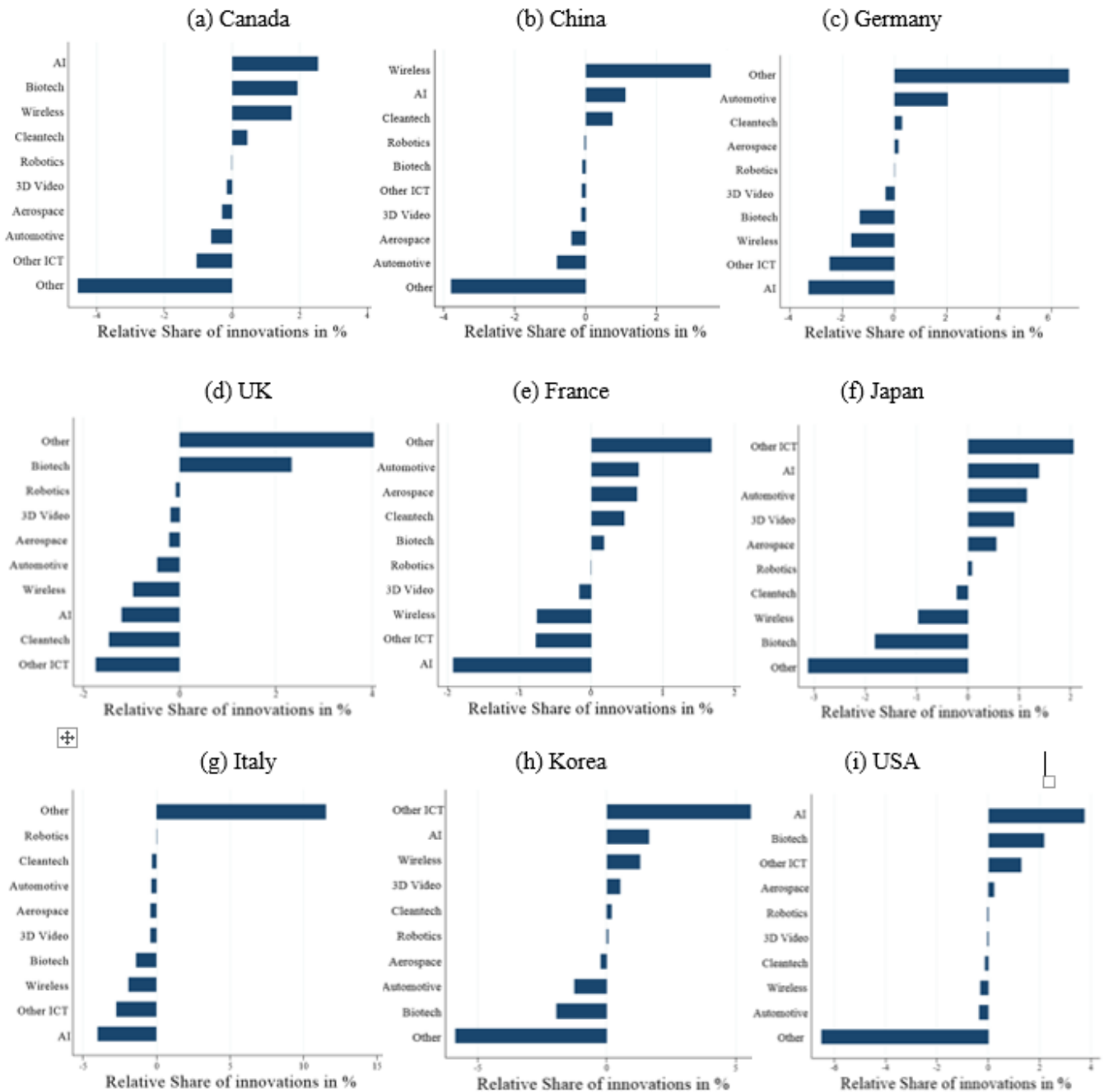
Notes: This figure shows the share of ICT innovations (defined across the sub-categories of Wireless, 3D, Robotics and All Other) in total innovation on a per country basis.

Figure 2.4: Share of Innovations Across Various Technology Categories (2000 to 2014)



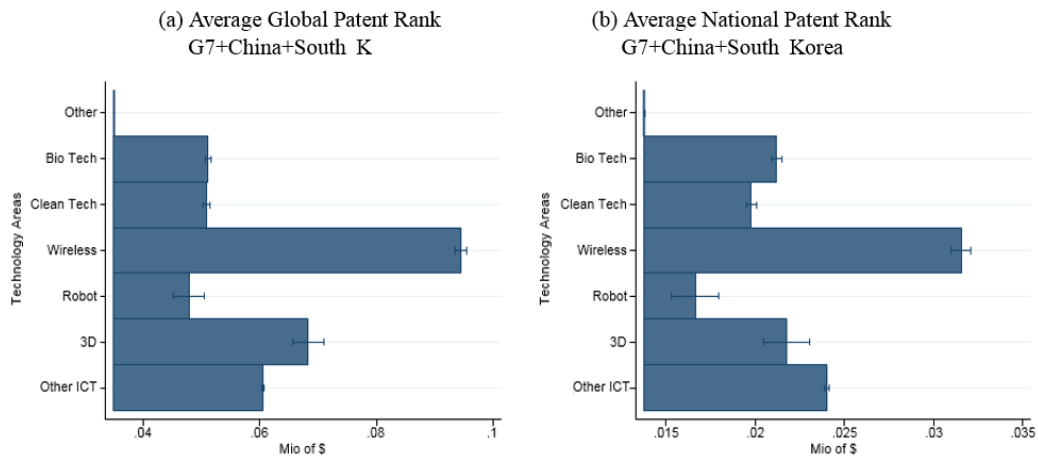
Notes: This figure shows the country-level shares in the volume of innovations by technology areas. The country codes are: DE (Germany), JP (Japan), CA (Canada), FR (France), IT (Italy), CN (China), KR (Korea), US (USA) and Great Britain (GB).

Figure 2.5: Comparative Advantage by Country (2000 to 2014)



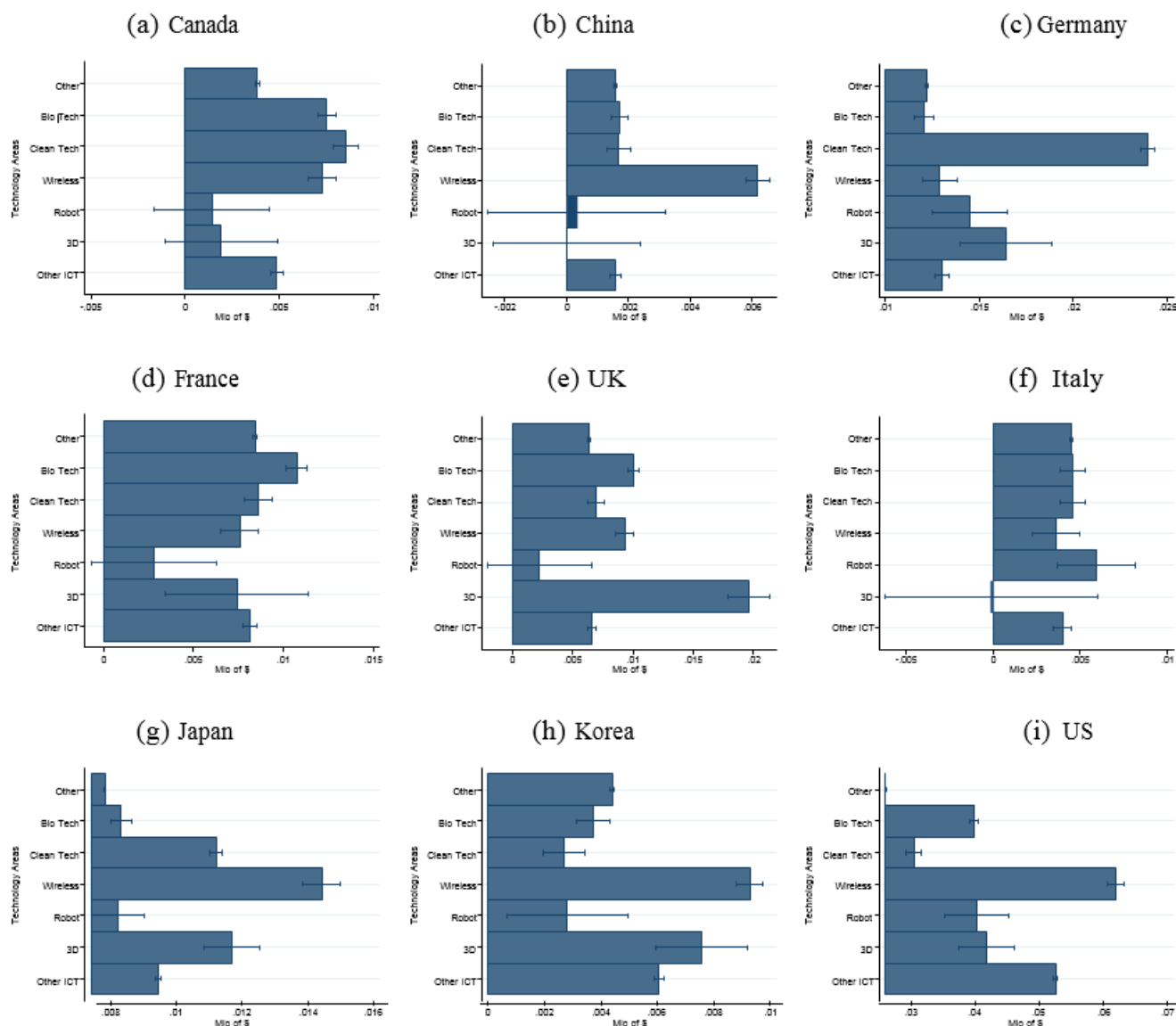
Notes: This figure shows country-level differences in relative shares of the total innovation count by technology area. That is, we calculate the difference between the country-level share in a technology area and the pooled global share of innovations in the same technology area.

Figure 2.6: Patent Rank Across Various Technology Groups (1990-2014)



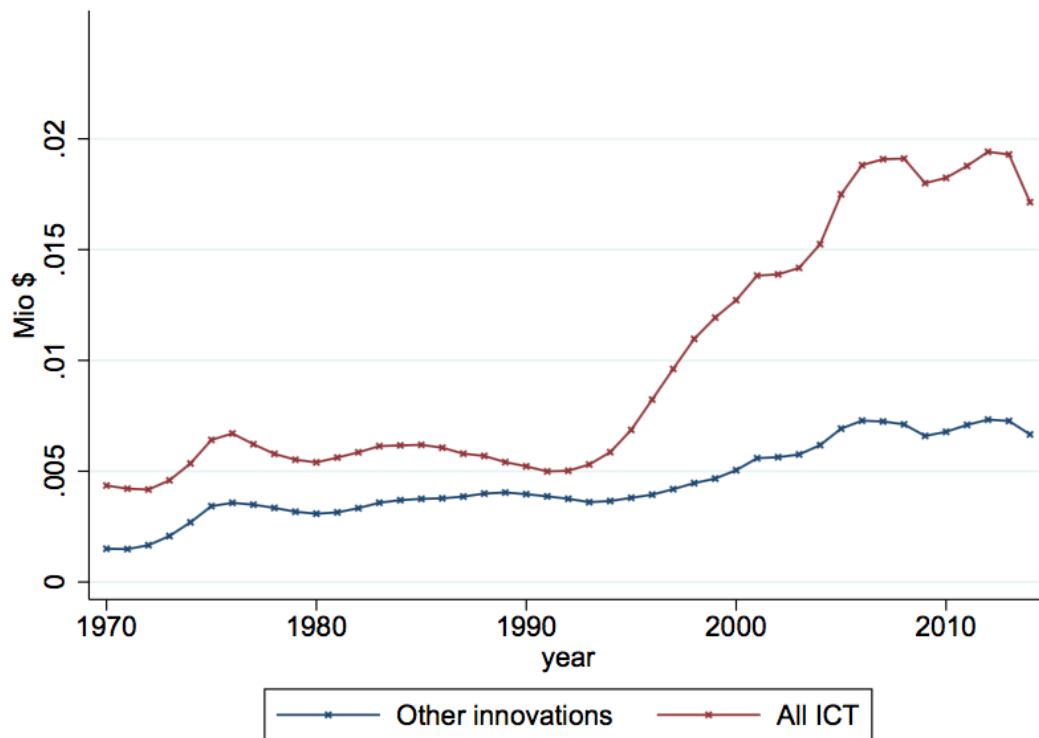
Notes: This figure reports Patent Rank measures of knowledge spillovers by technology area. In panel (a) we report Patent rank measures based on the full global profile of citations relating to a given innovation. In panel (b) we only calculate Patent Rank over citations reporting for patents in the originating country of a given innovation. The country-level means per technology measures are then averaged over all countries to obtain the ‘Average National Patent Rank’ measure reported in panel (b). Scale is in \$10 million USD.

Figure 2.7: Average National Spillovers Across Countries and Technologies (1990-2010)



Notes: This figure reports Patent Rank measures of knowledge spillovers by technology area at the country level. We report Average National Measures, that is, Patent Rank calculated on the basis of citations from patents originating in the same country as a given innovation. Scale is in \$10 million USD.

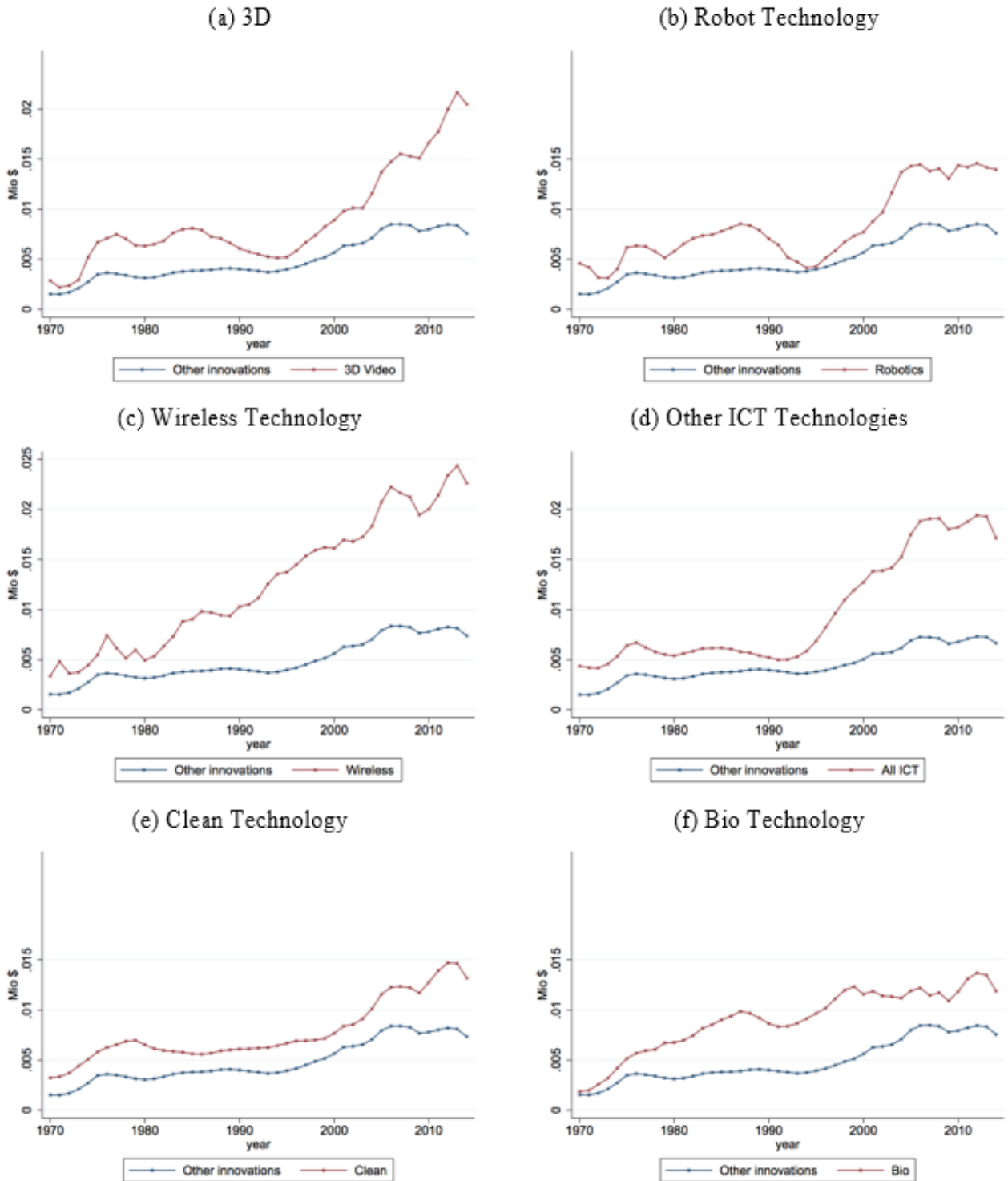
Figure 2.8: Average global spillovers within moving 5year window over time



Notes: This figure reports Patent Rank measures of knowledge spillovers across all technology areas.

Average per innovation is calculated according to the full global profile of citations relating to a given innovation. We report the time series for all ICT innovations across sub-categories versus the pooled set of all other innovations. The numbers represent five-year moving averages of the annualised series. Scale is in \$10 million USD.

Figure 2.9: Average Global Spillovers within Moving 5 Year Over Time for Various Technology Types



Notes: This figure reports Patent Rank measures of knowledge spillovers by technology area over time.

Average per innovation is calculated according to the full global profile of citations relating to a given innovation. We take five-year moving averages of the annualised series. Scale is in \$10 million USD.

3. Chapter 3: The Great Diffusion: What Factors Drive the Spread of Broadband Technologies?

3.1. Introduction

Along with the PC, the internet is the core technology of the ICT era. Since its emergence as part of US scientific and defense infrastructure over a period from the 1960s to the early 1990s, the internet has gone through different phases in its development. These phases have typically been characterised by the type of technology involved, whether it be dial-up, ADSL, fibre-optic cable or mobile infrastructure.

Even just taking into account its commercial existence as a service widely available to households and firms, the internet now has an approximate 25 year history in most major economies. One major theme to the evolution of the internet has been the rapid speed at which it has developed over different stages as various technologies have diffused. This rapid development has meant that the history of diffusion has not been comprehensively tracked to the same extent as past technologies that emerged more slowly. Further to this, important data has been to some extent ‘thrown away’ or buried as one technology became superseded by another faster technology.

In the following analysis, we provide a study of the internet in the UK (specifically England, where the data is at its most comprehensive) from the advent of the ADSL era in 1999-2000 until 2017. The data that we are able to assemble is built up from the postcode level and is centred on telephone exchange catchment areas, which formed the backbone of the ADSL network. Similarly we are able to utilise Ofcom data on broadband speeds from 2012 onwards, which allows us to track the arrival of superfast technologies. This information on technological roll-out and realised download speeds is then matched to socio-economic information at the local authority area level (326 areas in England).

The work in this chapter provides some novel analysis on a series of issues related to broadband in the UK which have not been explored at this level before. While several institutions produce regular reports on the state of internet infrastructure and its uses both in the UK and internationally (such as OFCOM, the OECD, and the International Communications Union), these generally provide a static picture of the availability of different technologies or some limited analysis on the ‘potential’ economic impacts of the internet. A handful of academic papers have provided some quasi-experimental analysis on the causal impacts of broadband on a set of economic outcomes, which are reviewed below. However, this evidence base on causal effects is still developing and many policy reports put forward benefits in a certain range and present assessments via cost-benefit analysis².

Moreover, most of the policy analysis has been focused on newer technologies (mobile and superfast broadband³) while we still know very little about what determined the initial and sharper technological shock - the rollout of the ADSL network - and how this technology correlates with a series of outcomes in the UK. These two issues are important because the arrival of ADSL meant a tenfold increase in available speeds and constituted a genuine disruption in the use of ICT both at home and in the workplace. Places that got this infrastructure sooner might have been impacted in very different ways than places that were upgraded later on, and these effects might have had long-run consequences. Also, we know very little about how this technology correlates with local economic outcomes, and even less about the causal relationship between ADSL and economic performance in the UK.

² Some of these reports are OECD (2008), Qiang et al. (2009), DCMS (2013), OFCOM (2013), EIU (2013) and EIU (2014).

³ For example NESTA (2015)

A number of findings emerge from our analysis:

- The diffusion of ADSL access was slower than has been previously understood. While 40% of the population (mainly concentrated in highly urbanised areas) gained access to ADSL by early 2001 it took another 6 years for the roll-out to be completed. A significant phase of the roll-out did not get underway until 2003-2004 when nearly 50% of telephone exchanges were ‘switched on’ for ADSL. The population coverage gap between the London/South-East area and regions such as the North and the Midlands was as high as 20% until late 2003.
- The roll-out of ADSL mainly tracks population density on a within-region basis with socio- economic factors correlated with the pattern of roll-out most strongly at the regional level. There is a very steep gradient in terms of population density in the pattern of roll-out that occurs even at the middle range. For example, on average an area in the 6th decile of population density received ADSL 400 days earlier than an area in the 3rd decile.
- This staggered roll-out makes it feasible to consider some economic effects of ADSL since some areas received access to the technology much earlier. We consider long run effects occurring over the period of 1991-2011, where detailed Census data is available for 1991, 2001 and 2011.

The share of ‘nonroutine’ professional workers increased more in the post-ADSL period for areas that had early access to ADSL relative to trends in the pre-ADSL decade. This effect is equivalent to a 15% larger professional workforce by 2011 in the early receiving ADSL areas. In line with this finding for nonroutine workers there is also a decline in the share of workers in ‘routine manual’ occupations (machine operators and manual labourers) in the early receiving ADSL areas. In fact, there are sharp switches in the direction of economic trends relative to the pre-ADSL decade.

- In addition to these long-run effects, the modelling of short-run wage and income trends indicates that the diffusion of ADSL was correlated with higher incomes in areas where ADSL penetration was higher. Even after controlling for a range of area-level factors the estimates indicate that a typical 30% annual increase in an area’s ADSL penetration rate is associated with a 0.6% higher wages. This is very close to estimates using a similar modelling strategy for Norway (Akerman et al. (2015)).
- The arrival of superfast broadband in the 2010s assisted the regional convergence of broad- band quality. For example, while the North and South were comparable in average speeds in the baseline year of 2012 there was a distinct ‘digital divide’ in maximum speeds. By 2017 differences were limited with some regions increasing maximum speed by 8-10 MB/s compared to 3MB/s for London. The correlation of area income with average broadband speed also fell over this period.

3.2. The International Picture on Broadband

The internet developed from the ARPANET, which was funded by the US government to support projects within the government and at universities and research laboratories in the US. Use by a wider audience only came in 1995 when restrictions on the use of the Internet to carry commercial traffic were lifted (Segal, 2015). Initial access to the internet was achieved via Local Area Networks (LANs) or from dial-up connections using modems. Offered speeds at the end of the 1980s were 1200bits/s, and this grew to around 56Kbits/s by the end of the 1990s. In the early 2000s, broad- band technologies facilitated a significant jump in download speeds by separating the telephone and internet connections. This increased the available content and functions of the internet.

This shift from relying on established telephone networks required the roll-out of other infrastructure, most importantly submarine fibre optic cables which are able to carry very large amounts of data. The first submarine cable was installed in 1988 and installation significantly increased during the 1990s, peaking in 2002 with the end of the ‘dot com’ bubble. During this period connections were very focused on linking the US with Europe (OECD, 2014). Geographic factors and longer distances initially made

connecting Asian countries more difficult, which early on favoured Japan, though throughout the 2000s/2010s China became the focus (TC, 2012). Less economically developed countries were connected later. Much of Africa, for example, was only connected in the 2002–2009 period (OECD, 2014), though most of the Middle East and North Africa region (MENA) economies are now connected via at least two international submarine cables (World Bank, 2014). However, in the poorer areas of Africa satellite broadband services have remained important, as backbone and land network infrastructure still have some distance to catch up (CTO, 2012).

Contrary to the belief that the international digital divide is decreasing over time, internet bandwidth has historically been very unequally distributed worldwide, with increasing concentration in the digital age (Hilbert, 2016). Historically only 10 countries have hosted 70–75% of the global telecommunication capacity. In 2014, only 3 countries (China, US, Japan) host 50% of the globally installed telecommunication bandwidth potential. The US lost its global leadership in terms of installed bandwidth in 2011, being replaced by China, which hosts more than twice as much national bandwidth potential in 2014 (29% versus 13% of the global total). The bandwidth distribution among all countries is undergoing a new process of global concentration, during which North America and Europe are being replaced by Asia as the new global leader.

Figure 3.1 shows the evolution of broadband subscriptions and the percentage of population using the internet since the early 2000s in selected countries (United States, Canada, United Kingdom, France, Germany, Japan, South Korea and China). The top panel (a) shows the number of connections in millions, and we can see this increasing sharply over time, especially in larger countries such as China or the US. When we normalise by country population (panel b), France, South Korea and the United Kingdom emerge as the countries that have reached higher levels of broadband penetration (around 40–45%). Japan and the US have level of penetration of around 30% according to this measure while China has around 20%. The bottom panel (c) of Figure 3.1 shows the percentage of individuals using the internet regardless of the type of connection. This indicates that Japan, South Korea and the UK have very high rates of usage (around 90%), while the US has around 70% and China around 40%. Putting this together we see that, while the number of subscriptions and the use of internet has grown steadily across all these countries over 15 years some sharp differences persist.

In terms of how the internet is accessed, the data suggests that some structural change has taken place in recent years. Figure 3.2 shows data from the UK telecommunications regulator (OFCOM International Communications Market Report) on the average number of minutes spent online using a computer in a range of countries. It compares the number of minutes spent in 2004, 2008, 2010 and 2011. Between 2004 and 2008 there was growth in minutes spent across all countries: the country which experienced the largest growth was the UK (118%), followed by Italy (93%), France and Germany (around 60%), Spain (41%) and the US (28%). In contrast, we see that after 2008 the amount of time spent accessing the internet from a computer decreased in most of these countries (UK, France, Italy, US) and only slightly decreased in others (Germany -9%- and Spain -0.5%-).

This decrease in accessing the internet from traditional devices such as computers could be the result of the substitution of access towards other types of devices. The data presented in Figure 3.3 and Figure 3.4 explore this hypothesis. Figure 3.3 uses OFCOM data from the UK over several years (2010, 2012, 2014 and 2016) to plot the extent to which adults use different types of devices to go online. Even though PCs and laptops remained the main device for access with their usage rate increasing 20% between 2010 and 2016, access from mobile phones increased more sharply (215%) in the same period. This pattern seems replicated in the international context as figure 3.4 shows.

The top panel (a) of Figure 3.4 plots data from OFCOM on the percentage of adults in different countries that use mobile phones to access the internet. In the UK, France, Germany, Italy and the US the

proportion of individuals who use mobiles to connect to the internet more than doubled between 2008 and 2011. Among these countries, internet use on mobiles is highest in the UK, perhaps due to the wide availability and take-up of smartphones there (OFCOM, 2011). Panel (b) of Figure 3.4 shows the number of wireless mobile broadband subscriptions per 100 inhabitants in several countries using data from the OECD for the years 2009, 2011, 2013 and 2015. Between 2009 and 2015 these numbers more than doubled in the UK (138% increase), doubled in France and the US, and increased by 50% and 20% in Japan and Korea respectively. In Canada, where the series begins in 2011, the number of subscribers increased 60% between that year and 2015. These last two figures clearly indicate that internet use via mobile phone and tablet devices has increased significantly, and is likely to be substituting for time spent on the internet using a traditional computer.

3.3. The Economic Effects of Broadband ICT

The measurement of the economic impacts of the internet at the microeconomic level has faced a similar level of challenges to the work analysing productivity discussed earlier in the report. In the case of productivity the main intervening factor in understanding the effects of ICT has been implementation, that is, the idea that while a set of technologies are in place they need time to take root and induce real economic changes. Until this full implementation occurs there will appear to be null returns from sometimes large ICT investments.

The intervening factor in microeconomic studies of ICT in general, and the internet in particular, has a different source. The correlation of technology and outcomes at the individual level cannot be automatically interpreted as causal. Firms, households and individuals who are likely to gain a higher return from technology are more likely to be in possession of it. This is either because of voluntary efforts to acquire the technology or because their local environment (for example, dense urban areas with good infrastructure) is characterised by lower costs of adoption. This problem of distinguishing correlation from causation is then compounded by measurement issues, for example, those revolving around difficult-to-measure consumer surplus that is created by apps and devices that increase efficiency in diverse tasks at low cost.

The research on the economic effects of broadband internet has therefore sought to tease out causal effects by tracing out the implications of ‘accidents’ or natural experiments in the roll-out of broadband technologies. These accidents have typically taken the form of barriers to adoption or factors that limit the speed of adoption in some areas, most notably geographical factors. There are also some characteristics of broadband technologies (in particular ADSL) that produce situations where very similar households or firms benefit from different levels of broadband speed based on small differences in their location. Such phenomena greatly assist the detection of causal impacts.

A range of recent studies have addressed the challenge of determining the causal effects of broadband ICT. These effects look at both economic and social outcomes⁴. Bhuller et al. (2013) and Akerman et al. (2015) exploit the timing of the roll-out of broadband internet across Norwegian municipalities to estimate the effect on sex crime and skilled-to-unskilled wages and productivity respectively. Campante et al. (2013) interact municipality distances to the broadband backbone with the time broadband was first introduced nationally to estimate the effect on political participation in Italy. Bertsek et al. (2013) examine the impact of the ADSL expansion in Germany on firm productivity and innovation, and Bellou (2015) exploits time variation in internet penetration across US states to estimate the effect on marriage

⁴ See Bertsek et al. (2015) for a recent survey

rates. Evidence on the impact of more recent technologies, such as mobile internet, is still very scarce (an exception is Bertsek and Niebel (2016)).

In the US, Forman et al. (2012) use county-level measures of infrastructure costs, industry mix, or characteristics of nearby locations as instruments for internet investment to estimate the effect on local wages. Technological quirks are exploited by Bauernschuster et al. (2014) who use the fact that some regions in Eastern Germany use telephony networks that are not compatible with ADSL-broadband to estimate the effect on social capital. Atasoy (2013) analyses the effects of the expansion of broadband Internet access from 1999 to 2007 on labour market outcomes throughout the US and finds some positive significant correlations between access to broadband services and employment rates, especially in rural and isolated areas. The study that is closest in institutional set-up to the situation that we study in this report is Falck et al. (2014). They exploit municipality-level variation in distances to telephone exchange stations for a sample of German municipalities that do not host a telephone exchange station within their own boundary to estimate effects on political participation.

For the UK there exist a limited number of empirical papers which have looked at the impact of broadband on socio-economic outcomes. Faber et al. (2015) study the impact of changes in broadband speed on student test scores using a spatial discontinuity approach and find no impact of speed changes on educational achievement. Ahlfeldt et al. (2017) finds substantial impact of internet speed on house prices using data in the UK between 1995 and 2010. Gavazza et al. (2017) study the effects of broadband diffusion on local election outcomes and local government policies.

Stefano et al. (2014) is the only paper looking at firm outcomes, but focuses on a small area in the North of England. They find no statistical significant effect of that broadband availability affected plant employment, plant exit rates or sales.

3.4. Studying the UK's Great Diffusion

3.4.1. *The Roll-Out of ADSL*

The internet now has a long history of approximately 25 years with different eras defined by the dominant access technology of the time. Dial-up internet access was first introduced in the UK in 1992 and on average these services offered speeds of 40-50 Kbits/s. Broadband internet access based on ADSL technology was introduced commercially in 2000 across a range of exchanges owned by BT. In formal statistical terms, broadband is defined within the industry as speeds delivering greater than 500Kbits/s.

The spread of broadband ADSL in the UK was relatively sluggish in its early stages. In 2001 only 9% of UK households outside of the most dense urban areas had broadband as compared to 39% in Germany and 33% in Sweden. However, by 2004 penetration levels were similar to France and Germany. Direct population surveys of users indicate that in 2002 there were 200,000 UK internet users, but by mid-2005 already 15 million had adopted and over 16 million in 2006 (62% of whom used broadband).

The arrival of broadband internet had a large qualitative effect on personal activity. Time spent online in households with dial-up connections increased from an average of 2 hours a week in 1999 to 8 hours in mid-2004, while households with broadband connections spent an average of 18 hrs/pw. Broadband users were also more likely to use more 'advanced' features: including gaming, chatting, making purchases, watching video, downloading files and using instant messaging. They also used it more often (with more than 50% using it every day).

The Basics of Roll-Out at the Exchange Level

A salient feature of DSL-broadband technology is that the available connection speed depends on the distance of the copper wire connection between the router in the residence and the local telephone exchange station. The activation of these stations for ADSL is therefore the main indicator of broadband access.

The initial location of telephone exchange stations was determined during the roll-out of the English landline telephony network which mainly occurred before and during the Second World War. Importantly, distance to the local exchange station did not affect the quality of traditional telephone services but was highly relevant for the realised speed of ADSL internet at the household level.

We use information about the location of all English telephone exchange stations (roughly 3,900), and link the exchange station identifiers to telecom network records that provide us with information on each English postcode's assignment (in total, there are around 1.35 million English postcodes) to one of the 3,900 telephone exchange stations. The fact that we are using postcode level information on exchanges means that we are also able to match the data to detailed Census Output Areas and obtain both demographic and geographical information⁵.

Figure 3.5 shows the progress of the exchange 'switch on' dates in terms of the frequency at the level of the 3,900 exchanges. For example, the early spike in the figure indicates that about 12% of exchanges were switched on in early to mid-2000. It is notable that the mass of exchanges were not activated until a 2.5 year phase that began in mid-2003. According to this data, the last exchanges were switched on in 2007.

We provide a geographical heat map of the exchange activation in Figures 3.6(a) and 3.6(b). In line with expectations, this indicates that exchanges in urban areas upgraded the earliest and that large concentrations of business premises were covered by this early roll-out. In Figure 3.7 we report this information on a cumulative, annual basis.

The clearest picture of the 'Great Diffusion' is available in Figure 3.8, where we plot exchange roll-out weighted by the population in an exchanges' catchment area. This shows that the overall ADSL diffusion was characterised by an initial 'big bang' in 2000 when London and other major urban areas were activated. This early phase resulted in around 30% of England's population being covered by ADSL. There was then a slow climb to approximately 70% population coverage until early 2003 when a long period of non-urban roll-out began in earnest.

Arguably, the most notable feature of the ADSL diffusion was its duration. While urban areas received ADSL access relatively early in 2000 and became accustomed to higher download speeds there was long process of diffusion in the rest of the country that took 7 years to complete. In turn, this means that non-urban parts of England had a qualitatively different experience of the internet with a much more limited capacity to participate in "advanced" activities such as gaming, making purchases, watching video, downloading files, searching for jobs and using instant messaging on a regular basis.

⁵ This data is the same used in Faber et al. (2015).

Determinants of ADSL Roll-Out

What factors determined the path of the ADSL roll-out? The detailed nature of our exchange data (with its 3,900 area observations) gives us an unprecedented view into BT's implicit decision rule for rolling out ADSL.

In Table 3.1 we conduct a regression analysis where we model the correlates of 'days until roll-out' as the dependent variable. The results here indicate that being near rivers or minor roads is correlated with earlier roll-out. For example, in column (1) the coefficient on Near Rivers indicates that for every kilometre that an exchange is closer to a river then roll-out is 3.6% later relative to other areas.

However, a key finding of Table 3.1 comes in column (3), where we control for population density in the exchange area. This knocks out most of the previous correlation with rivers and minor roads since these variables were most likely proxying for sparsely populated areas. We continue this analysis in Table 3.3 where we instead focus on the correlation of roll-out with the proportion of skilled workers in an area. This indicates relatively weak effects of skilled workers on roll-out speed, especially after controlling for population density. It should be noted that while the coefficient on the Professionals variable looks large the effect in practice is small. Since the actual proportion of professionals in the sample is low, even an area at the 75th percentile of this skill measure would only get the ADSL roll-out 90 days sooner according to these estimates.

Hence, the driving factor of the roll-out at the exchange-area level is population density. We summarise this in Figure 3.9 by plotting the roll-out speed by population decile with the y-axis measuring how many days sooner that a population group received ADSL according to a basic regression model. This shows a very clear gradient with even roll-out the 6th and 3rd decile separated by around 400 days. The dominant BT strategy or decision rule for the roll-out therefore appears to population coverage, with a limited role for broader socio-economic factors once density is accounted for. However, as we will see, demographic factors play a more important role when we consider more aggregated areas. The basic story of roll-out then is that once BT entered a region or sub-region it proceeded by switching on exchanges according to population density - a 'low hanging fruit' strategy for expanding coverage quickly in population weighted terms.

Correlation with Economic Outcomes

The evidence above indicates that despite common perception (and indeed memories) the transition from dial-up to ADSL broadband was a prolonged process for a significant fraction of the country. Importantly, this means that a range of areas that received ADSL early plausibly had a major strategic advantage in exploiting internet technologies economically.

We set out some tests for such economic effects in Tables 3.3 and 3.4. In these tables we take a long-run perspective over 20 years in order to test for changes in occupational structure that could be related to the emergence of ADSL. Specifically, we use UK Census data for 1991, 2001 and 2011 and define the 10-year change in occupational shares across the different intervals. Here, it is useful to think of the 1991-2001 change as summarising economic patterns 'before' ADSL hit while the 2001-2011 change reflects what occurred 'after' in the wake of ADSL.

Our test comes in correlating the level of the in-progress ADSL roll-out across areas in 2001 against the 1991-2001 and 2001-2011 changes, controlling for area population and broad region effects. The pattern of ADSL roll-out in 2001 represents the early phase and therefore measures the groups of areas that will have been exposed to ADSL for the longest period of time when the 2011 Census data was collected. If economic trends or structure shifted across the two intervals this could be consistent with an economic influence of ADSL.

Table 3.3 first looks at occupational structure following broad definitions of routine and non-routine occupations. In panel (A) of Table 3.3 we report the evidence on the share of Professionals and Associate Professionals in an area's workforce. Across both occupational measures there is a switch in the direction of the relationship with ADSL levels in 2001. The 2001 ADSL levels are positively correlated with the changes in occupational shares for the later 2001-2011 period but negatively correlated before. This is consistent with a mechanism whereby ADSL facilitated the growth of non-routine jobs. The size of the effect is reasonable - for values at the mean ADSL levels in 2001 are associated with a 15% higher share of professionals in the local workforce in 2011.

Panel (D) of Table 3.3 then shows the shifts occurring at the other end of the occupational distribution, looking at routine manual occupations, defined as machine operators and labourers ('elementary occupations'). This again shows a flip in direction but in this case from positive to negative. That is, ADSL levels in 2001 become negatively associated with routine manual employment over the 2001-2011 period.

These results are closely related by definition. If the share for one group of occupations increases then the shares of another group must decrease. Further to this, the fact that the 2001 ADSL level is significant in some regressions for the pre-ADSL period (1991-2001) implies that the 2001 ADSL level may be proxying for some general economic characteristics and picking up trends that are not directly related to ADSL. However, even if this is the case the fact that the direction of the relationship changes is unusual and the new direction is consistent with effects of ICT on the labour market as seen in other studies (David and Dorn (2013), Akerman et al. (2015)).

Our next exercise looks at short-run correlations between the diffusion of ADSL and measures of household income, wages and hours worked. This is reported in Tables 3.4 and 3.5. In these models we test whether the regional and yearly variation in the pattern of the roll-out is correlated with household income, wages and hours worked. The regressions are run in a panel of 326 LAD areas in England between 2000-2008. We test the effect of the (lagged) level of ADSL penetration rate (i.e. the share in the LAD premises which have been already upgraded to ADSL) on the different economic outcomes. In this approach we apply a tough test of including comprehensive controls for general region-year effects as well as area-level characteristics such as the unemployment rate, economic inactivity and geographical features.

Table 3.4 shows estimates that relate the level of ADSL penetration to the levels of different outcomes. We report results for the GDHI per person (Gross Disposable Household Income), average annual pay, average weekly pay and average weekly hours worked. The results indicate that there is some positive effect of short-run ADSL diffusion evident even after including these controls. Columns (1) and (4) include a set of time-varying co-founders such as the unemployment rate, the share of population with higher education etc., and year and region dummy variables. We find significant correlations between the income variables (GDHI per person and pay) and negative correlations with hours worked (although this coefficient is very small). The estimates remain very similar when we include region-year dummies, which control for region-year specific shocks related to ADSL penetration rates and the outcomes of interest. In the most stringent models (columns (3) and (6)) we include LAD-level time-varying characteristics, and these seems to absorb most of the effect except for annual pay, where we still find a point estimate of 0.03. This implies that areas in which the penetration rate is 10% larger, annual pay would be 0.3% higher. Taking the mid-point of this range we can calculate the magnitude of this effect as implying that a potential 30% annual increase in an area's ADSL penetration rate (a realistic figure in the context of the trends discussed above) is associated with a 0.9% increase in pay or income. This is consistent with major studies of the diffusion of internet technologies and wages in the US (Forman et al., 2012) and Norway (Akerman et al., 2015). The magnitude of this effect should be seen in the context of average real wage growth during the roll-out period, which was approximately 2% in the UK.

In Table 3.5 estimates the impact of changes in the penetration rate on the growth of the out-comes variables, e.g. we estimate the models of Table 3.4 in first-differences. The first-differencing removes any LAD-specific time-invariant characteristic that might be correlated both with the outcomes of interest and the rollout pattern. In line with the results above, we find insignificant effects and very small coefficients for the estimates on GDHI per head, weekly pay and weekly hours, but we find significant and positive coefficients for the growth of annual pay. In the most demanding specification of column (6), which includes LAD dummies and all the controls we find a coefficient of 0.014, which implies that doubling the rate of penetration is associated with a higher pay growth of around 1.5%.

3.4.2. *The Arrival of the Superfast Broadband*

The ‘Big Bang’ of the 2010s

Ofcom provides postcode data level information on available broadband speed (and in some cases for number of superfast connections) from 2012. This allows us to evaluate the impact of the arrival of superfast broadband. After consolidating the information across years, the final dataset we assemble from the Ofcom sources is defined at the level of 326 geographic areas (local authority districts). In practice, the data reports speeds that reflect the mixture of ADSL and superfast technologies available to consumers at a given point in time.

We plot the distribution of average and maximum speeds across areas in Figure 3.10. This shows a remarkable transition of broadband technology over the 2010s. In 2012, during the early phase of superfast roll-out, average speeds across areas are in the 10-15MB/s range but by 2017 they sit in the 25-28 MB/s range. Further to this, while some areas had average speeds as low as 5MB/s in 2012 this moved up to around 12MB/s in 2017. Figure 3.10(b), which plots maximum available speeds, explicitly shows the role of superfast technologies in this transition. There were sharp increases in maximum speeds between 2012-2017. By 2017 most areas had speeds in the 28-30 MB/s range with only a limited mass of areas in the below 20MB/s range. In practice, the technological driver of this shift derived from the fact that the new superfast technologies were not limited by the distance-from-exchange problem that had plagued ADSL, whereby households further away from a telephone exchange experience an attenuation of the speeds. In practice, this created significant patches of areas with speeds of around 2MB/s despite the presence of ADSL.

Closing the Digital Divide?

This ‘big bang’ of increased speeds between 2012 and 2017 also appears to have had a major impact on regional differences in broadband speeds. In Figure 3.11 we plot the distribution of speeds in the North and the South of England over the two years. The notable point here is that while the two broad regions were comparable in average speeds in the baseline year of 2012 there was a distinct ‘digital divide’ in maximum speeds, with a strong Southern advantage evident in Figure 3.11(c). It is clear that, similar to the diffusion of ADSL, the South had an advantage in the roll-out of superfast technologies.

This Southern advantage was eroded by 2017. While Figure 3.11(d) still shows a big spike in maximum speeds in the South for 2017, note that the scale of this figure is different from the other panels because of the massive shift in broadband speeds. That is, Figure 3.11(d) still shows a large mass of high speeds in the 25-30MB/s range for the North. This is even clearer from the report of average and maximum speeds by region that is given in Table 3.7. Here we can note shifts such as the 8MB/s increase in maximum speeds for the North West between 2012-2017 versus the 3.5MB/s increase for London. In a similar fashion, the less population dense regions such as the South West and Yorkshire-Humber experienced increases in speeds from well below 20MB/s to around 27MB/s. The fact that superfast

technologies did not discriminate according to distance from the exchange clearly played an important role in this convergence.

More evidence on this convergence is given in Table 3.8, where we measure the demographic correlates of average broadband speed. Here we are able to ask: is area-level household income now less correlated with average broadband speed than it was in 2012? A lower correlation would imply that the socio-economic gradient in broadband (the ‘digital divide’) had flattened out. This seems to be the case as we see the coefficient relating income to speed across fall across the two time periods. A final point to note is that the models in Table 3.8 control for region effects. In turn, this means that the income sensitivity that we are measuring in these models relates to differences within a region (i.e.: amongst different areas inside the North-West) rather than the inter-regional differences that were the focus of the earlier plots of North versus South.

3.5. Conclusions

In this chapter, we have considered the determinants of the ‘great diffusion’ of internet technologies that occurred in the UK from the early 2000s onwards. This documented the a substantial regional ‘digital divide’ in ADSL access was persistent until the mid-2000s but that the later roll-out of superfast technologies managed to level out this divide in the 2010s.

The research presented here has also uncovered evidence of an economic effect of early roll-out. Changes in occupational structure that are consistent with an economic impact of broadband are evident in areas that received access to ADSL in its earliest phase. Wage effects are also evident but the pattern of effects is less decisive in statistical terms.

These estimates of economic effects due to broadband access can be considered as something of a ‘lower bound’. The advantage that the early roll-out areas received was one that was only in place for 3 years relative to other major economic regions in England. The long period of time that we consider means that the areas with lagged ADSL have had the chance to catch-up in terms of exploiting broadband.

What should be the focus of the next phase in understanding the impacts of broadband in countries such as the UK? Our analysis of the download speeds in the wake of the roll-out of superfast broadband indicates a direction here. The arrival of superfast broadband saw a massive jump in available speeds in some areas as the distance-to-exchange limitations of ADSL were overcome. In likelihood, this has led to a qualitative shift in the nature of internet usage in areas with slower speed during the mid-to-late 2000s. Households will have access to new consumer services and businesses in these areas now have the ability to exploit the internet at a much higher level. A plausible focus for future research is therefore isolating these upgraded areas and studying their differential economic outcomes into the 2010s.

International Experiences of the ‘Great Diffusion’.

What kind of economic effects have been associated with the spread of internet technology? Some indicative studies are those that have been conducted for the US and Norway.

USA, 1995-2000: ‘The Internet and Local Wages.’

Forman et al. (2012) describe the case of the internet and local wages in the US as a ‘puzzle’. They consider the earliest phase of internet diffusion over the 1995-2000 window and focus on a specific form of internet-related investment. To measure internet investment, they use data from an ICT marketing company to establish the level of local business purchases in the area of ‘advanced internet’ technologies. These technologies are mainly associated with business software (e.g.: Enterprise resource Planning (ERP), customer service, technical support systems). Crucially, they make a distinction between these advanced internet investments and other elements of ICT such as basic internet access and the number of PCs.

Their analysis is conducted at the level of approximately 2,700 US counties. The correlation between advanced internet investments and county wages is weak on average for their sample but a stronger relationship is evident for a subset of counties with high levels of income, education, population and pre-existing ICT intensity. Their estimates indicate that advanced internet investment can explain around 23% growth in this subset of counties (163 out of the full 2,700) but only explains 1% for the rest. This is not explained by outlier counties (e.g.: Santa Clara, San Francisco) driving the result and is robust to alternative instrumental variables strategies based on the patterns of earlier ARPANET infrastructure and cross-site investments amongst multi-establishment firms. As the author’s summarise: ‘In short, while Internet investment is widespread, the payoffs are not.’

Norway, 2000-2008: ‘The Skill Complementarity of Broadband Internet.’

The main focus of Akerman et al. (2015) is identifying broadband internet as a driver of skill-biased technological change. They do this by considering the roll-out of broadband across Norway in the early to mid-2000s. They provide a causal interpretation of the roll-out based on an analysis of the determinants (or lack thereof) of entry into an area.

The modelling of the roll-out in Norway then suggests that broadband enhanced the skill premium amongst workers - a 10% increase in broadband availability raised the wages of skilled workers by 0.2%. Calculated in terms of ‘counterfactual trends’ (i.e.: what would have happened to outcomes if broadband was not put in place), skilled worker wages were 1.9% higher over the period. However, the interesting contrast is with firms. There was an increase in the firm-level use of skilled labour (as measured by output elasticities in the production function) but the size of this effect is large enough to suggest that the full benefits of ICT were not passed on to workers as higher wages (a figure of only 20% is suggested). That is, firms managed to gain some economic rents from the expansion of broadband internet.

3.6. Tables and Figures

Table 3.1: Physical Barriers and ADSL Roll-Out, 1999-2007

	(1) Baseline	(2) Plus Region	(3) Plus Density
Near Rivers	0.036*** (0.007)	0.003 (0.005)	-0.006 (0.005)
Near Coast	-0.000 (0.001)	-0.000 (0.001)	0.001 (0.001)
Near Minor Road	0.417*** (0.049)	0.196*** (0.028)	0.064 (0.047)
Near Major Road	-0.022** (0.007)	-0.007** (0.003)	-0.010*** (0.002)
Near Road Junction	-0.016*** (0.002)	-0.006** (0.002)	-0.005** (0.002)
Near Railway	-0.020*** (0.003)	-0.003* (0.001)	-0.001 (0.002)
Near Airport	-0.005*** (0.001)	-0.002 (0.001)	-0.001 (0.001)
Near Town Centre	-0.056*** (0.005)	-0.014*** (0.004)	-0.015*** (0.003)
R squared	0.472	0.685	0.704
Number of Observations	3,922	3,922	3,922

Notes: Ofcom Exchange-level data matched to Census data for 2001. Dependent variable is days until the roll-out of ADSL in an exchange area. Standard error clustered by region in parentheses. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Table 3.2: Skilled Workers and ADSL Roll-Out, 1999-2007

	(1) Catchment Area Baseline	(2) Catchment Area Plus Controls	(3) LAD Baseline	(4) LAD Plus Controls
Proportion Managers	-6594.8 (4446.3)	882.9 (1418.1)		
Proportion Professionals	5920.3 (4876.2)	1856.1* (849.4)		
No Qualifications			-0.300** (0.132)	-0.189 (0.128)
Population Density			0.015*** (0.004)	0.016*** (0.004)
Rsquared	0.02	0.706	0.776	0.820
Number of Observations	3,922	3,922	2,934	2,934

Notes: Columns (1) and (2) are defined at the Ofcom exchange-level Catchment area with the dependent variable defined as days until the roll-out of ADSL in an exchange area. Columns (3) and (4) are at the (panel) LAD-level with the dependent variable defined as the ADSL penetration rate. The controls across all columns include full geographic controls plus population density. Columns (4) also includes region-year fixed effects as part of the controls. Standard error clustered by region in parentheses. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Table 3.3: Occupational Structure and Early ADSL Roll-Out

	2001	2011	2001	2011
(A) NR Cognitive ADSL 2001	Professionals -0.007* (0.004)	Professionals 0.020*** (0.005)	Assoc Prof -0.019*** (0.004)	Assoc Prof 0.004 (0.003)
(B) R Cognitive ADSL 2001	Admin -0.013** (0.006)	Admin -0.006* (0.003)	Skilled -0.014** (0.006)	Skilled -0.013*** (0.003)
(C) NR Manual ADSL 2001	Sales 0.003 (0.003)	Sales 0.006** (0.003)	Personal 0.014*** (0.005)	Personal -0.001 (0.003)
(D) R Manual ADSL 2001	Machin Op 0.012*** (0.005)	Machin Op -0.008*** (0.003)	Elementary 0.020*** (0.005)	Elementary -0.006 (0.004)
Number Obs	326	326	326	326

Notes: LAD-level data for England. Data source is UK Census for occupation and Ofcom for ADSL. Dependent variable change in the share of an occupational group between 2001-1991 and 2011-2001. Standard error clustered by region in parentheses. Level2001 represents that share of the population in an LAD with ADSL access in 2001. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Table 3.4: ADSL penetration effect on income/pay/hours – PANEL 2000 to 2008

	GDHI per head				Annual pay	
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged ADSL penetration rate	0.058** (0.026)	0.081*** (0.031)	0.009 (0.018)	0.075*** (0.019)	0.084*** (0.022)	0.030** (0.015)
Year and region dummies	Yes	No	No	Yes	No	No
Region-year dummies		Yes	Yes		Yes	Yes
Time invariant controls			Yes			Yes
Observations	2934	2934	2934	2934	2934	2934
Model F-test	145.68	27.50	25.02	222.08	37.21	43.45
R squared	0.71	0.71	0.79	0.68	0.69	0.77
	Weekly pay				Weekly hours	
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged ADSL penetration rate	0.053*** (0.015)	0.064*** (0.018)	0.013 (0.015)	-0.009** (0.004)	-0.008* (0.005)	-0.007 (0.005)
Year and region dummies	Yes	No	No	Yes	No	No
Region-year dummies		Yes	Yes		Yes	Yes
Time invariant controls			Yes			Yes
Observations	2,934	2,934	2,934	2,934	2,934	2,934
Model F-test	139.08	30.67	32.47	31.23	21.13	13.90
R squared	0.62	0.62	0.71	0.35	0.36	0.38

Notes: Data for 326 Local Authority Districts (LADs) in England 2000-2008. GDHI stands for Gross Disposable Household Income. Penetration rate and time-varying controls lagged one period with respect to the outcome variables. Outcome variables in logs. Standard error clustered by LAD and robust to heteroskedasticity in parentheses. Time varying controls included in all columns (lagged unemployment rate, lagged activity rate, lagged percentage employed with nvq4+, lagged log population per square km, lagged log jobs over population rate, lagged percentage female jobs, lagged percentage part time jobs, lagged percentage of population aged 65+). Time invariant controls included in columns (3) and (6) (log distance to central London, log distance to coastline, log number of addresses per square km, percentage of urban addresses, log mean slope of terrain, log mean elevation of terrain, log sum of night-light intensity in 1997). Time varying and time invariant controls are jointly significant. Sources: Own authors' data/OFCOM/NSPD/ASHE/LFS/APS/OS Panorama/NASA. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Table 3.5: ADSL penetration change effect on income/pay/hours growth – FD 2001 to 2008

	FD GDHI per head			FD Annual pay		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged ADSL penetration rate change	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	0.015** (0.007)	0.015** (0.007)	0.014* (0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies		Yes			Yes	
LAD dummies			Yes			Yes
Observations	2608	2608	2608	2608	2608	2608
Model F-test	3.95	54.29	61.74	0.90	1.66	2.11
R squared	0.28	0.31	0.40	0.01	0.02	0.06
	FD Weekly pay			FD Weekly hours		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged ADSL penetration rate change	0.004 (0.007)	0.004 (0.007)	0.003 (0.008)	-0.004 (0.003)	-0.004 (0.003)	-0.004 (0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies		Yes			Yes	
LAD dummies			Yes			Yes
Observations	2608	2608	2608	2608	2608	2608
Model F-test	0.58	4.31	4.90	2.52	6.26	5.67
R squared	0.03	0.03	0.07	0.06	0.06	0.09

Notes: Data for 326 Local Authority Districts (LADs) in England 2001-2008. GDHI stands for Gross Disposable Household Income. Penetration rate and time-varying controls lagged one period with respect to the outcome variables. Outcome variables and time-varying controls in first-differences. All specifications include time dummies. Standard error clustered by LAD and robust to heteroskedasticity in parentheses. Time varying controls included in all columns (lagged unemployment rate, lagged activity rate, lagged percentage employed with nvq4+, lagged log population per square km, lagged log jobs over population rate, lagged percentage female jobs, lagged percentage part time jobs, lagged percentage of population aged 65+). Time invariant controls included in all columns (log distance to central London, log distance to coastline, log number of addresses per square km, percentage of urban addresses, log mean slope of terrain, log mean elevation of terrain, log sum of night- light intensity in 1997). Time varying and time invariant controls are jointly significant. Sources: Own authors' data/OFCOM/NSPD/ASHE/LFS/APS/OS Panorama/NASA. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Table 3.6: Early ADSL Roll-out and Long-Run Wage Outcomes

	(1) 10-year	(2) 2-year	(3) 1-year	(4) 10-year	(5) 2-year	(6) 1-year
ADSL 2001	0.033 (0.027)	0.009 (0.005)	0.003 (0.003)	0.006* (0.003)	0.010* (0.005)	0.004 (0.003)
Pop density*ADSL 2001				0.0070*** (0.000)	0.003 (0.000)	0.001 (0.000)
R squared	0.063	0.063	0.057	0.095	0.066	0.058
N	326	326	326	326	326	326

Notes: Based on a cross-section of LAD-level data. ADSL 2001 is the penetration rate in 2001. Dependent variable is the change in log annual wages derived from the ASHE dataset. 10-year represents a single 10-year change between 2001-2011 while 2-year and 1-year are smoothed (average) changes calculated over the same interval. Controls include geographic controls, percentages with high and low qualifications, unemployment rate, inactivity rate, distance to London and population density. Robust standard errors in parentheses. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Table 3.7: Broadband Speeds by Region, Superfast Era, 2012-2017

	Average 2012	Average 2017	Maximum 2012	Maximum 2017
East Midlands	12.0	26.0	20.5	28.8
East of England	12.8	25.3	20.7	28
London	15.7	27.4	26.5	29.1
North East	14.9	26.8	23.8	28.9
North West	12.6	26.0	20.7	28.8
South East	13.1	26.1	21.4	28.8
South West	10.5	23.6	16.9	27.3
West Midlands	12.4	25.9	21.8	28.6
Yorkshire/Humber	11.7	23.7	18.3	26.9

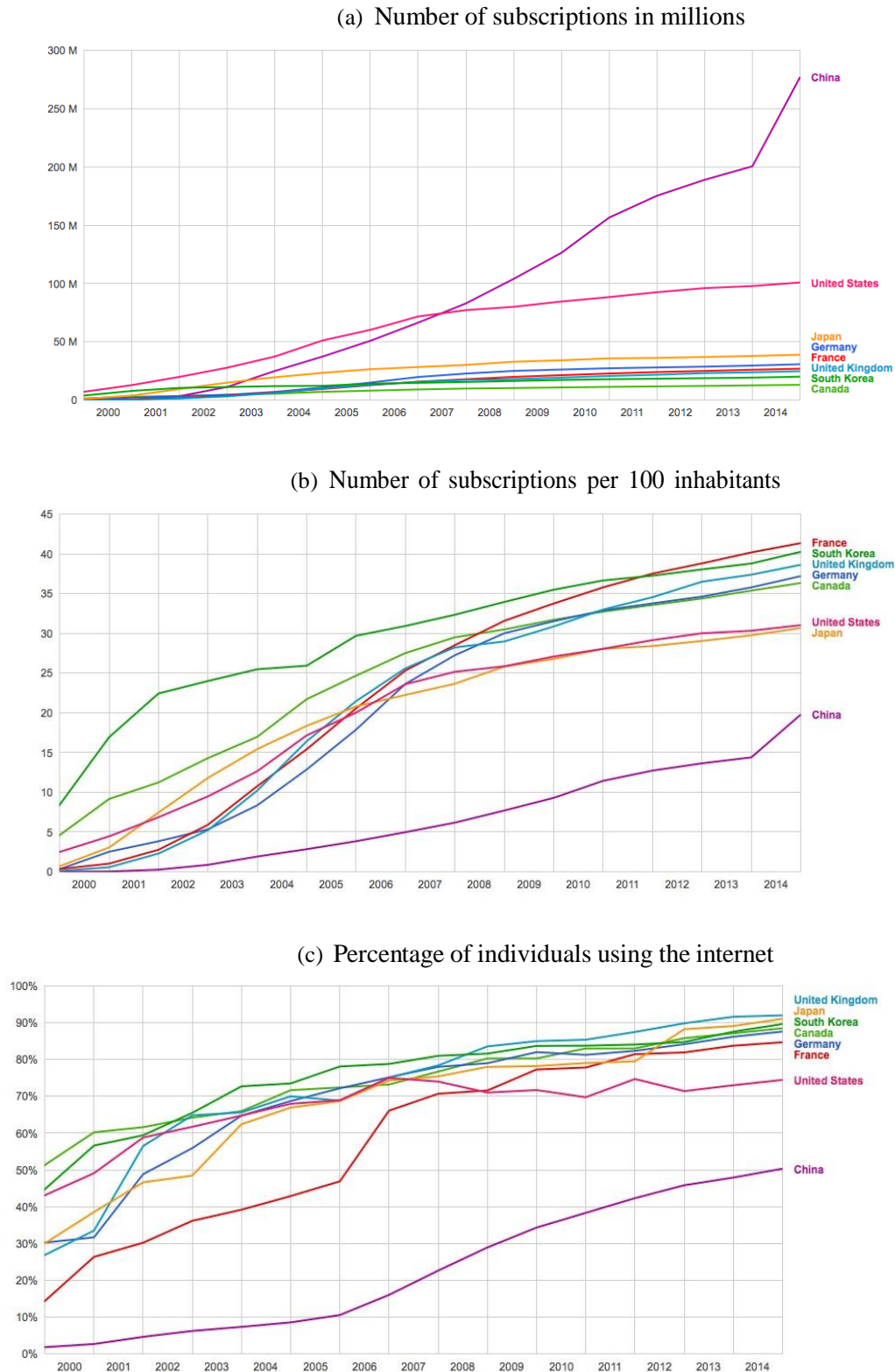
Notes: Report authors' calculations based on Ofcom data. Speeds reported in MB per second. LAD-level data. Calculations are made as unweighted averages across LADs within a major region. All speeds above 30 MB per second have been censored down to 30MB in the postcode microdata before collapsing.

Table 3.8: Correlates of Average Broadband Speed, 2012-2017

	(1) 2012	(2) 2017	(3) 2012	(4) 2017
log(Income)	2.217*** (0.458)	1.172** (0.354)	1.543** (0.549)	0.973* (0.432)
Pop Density			0.001 (0.000)	0.001 (0.000)
Activity Rate			11.324*** (3.238)	11.999*** (2.810)
Unemp Rate			22.790* (10.142)	9.016 (11.185)
High Skill Percent			1.622 (3.974)	-4.807 (4.905)
R-squared	0.24	0.17	0.36	0.33
Number Obs	326	326	326	326

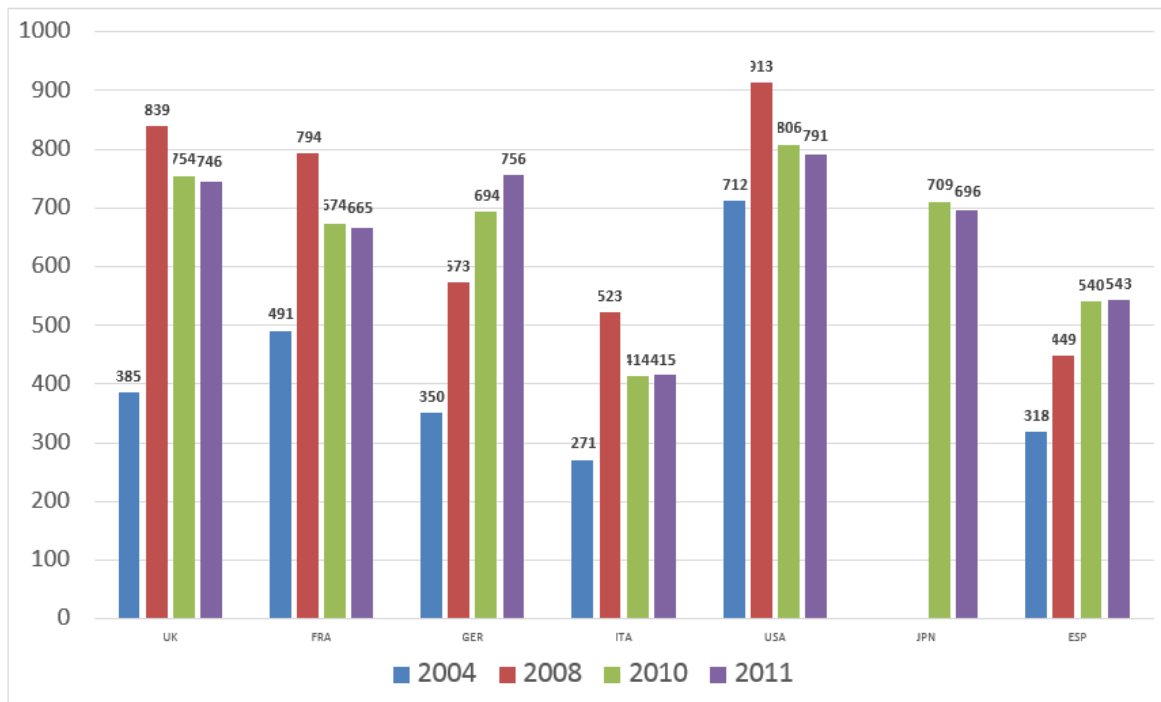
Notes: LAD-level data for England. Data source is UK Census for demographics and Ofcom for broadband speed. Dependent variable change is average broadband speed in the LAD. High Skill Percent measured as share of workers with NVQ4 or above qualifications. Standard error clustered by region in parentheses. All regressions include region fixed effects. *, **, and *** represents 10 %, 5 % and 1% significance level, respectively.

Figure 3.1: Broadband Subscriptions and Usage, selected countries, 2000–2014



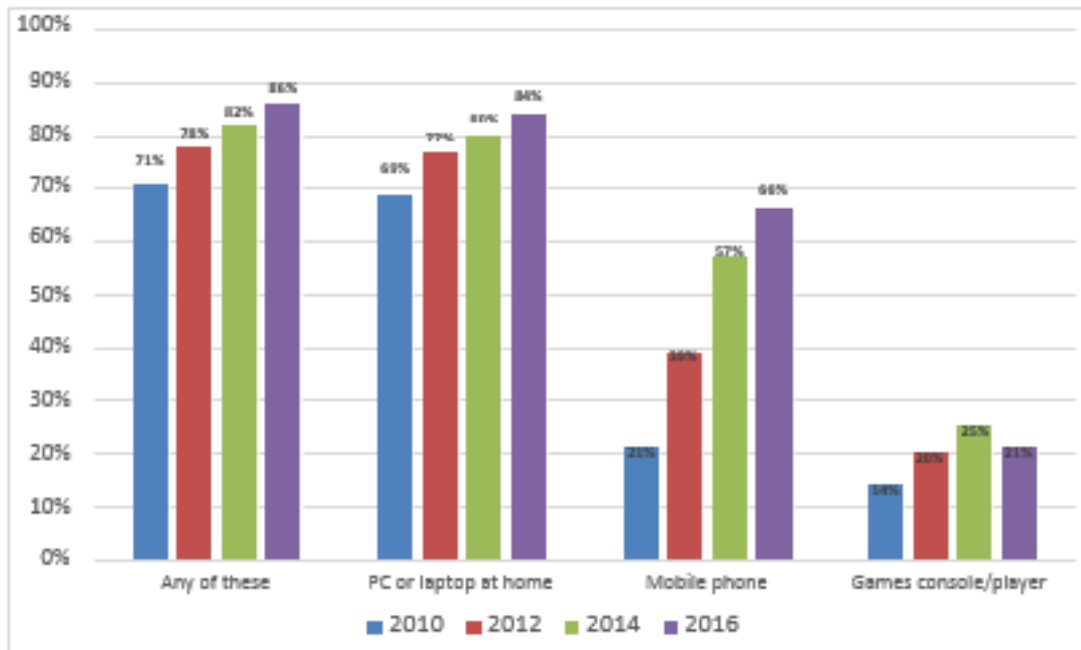
Notes: The Telecommunication Development Sector (ITU-D), <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>

Figure 3.2: Minutes per week online using a computer



Source: OFCOM International Communications Market Report from Nielsen data, several years

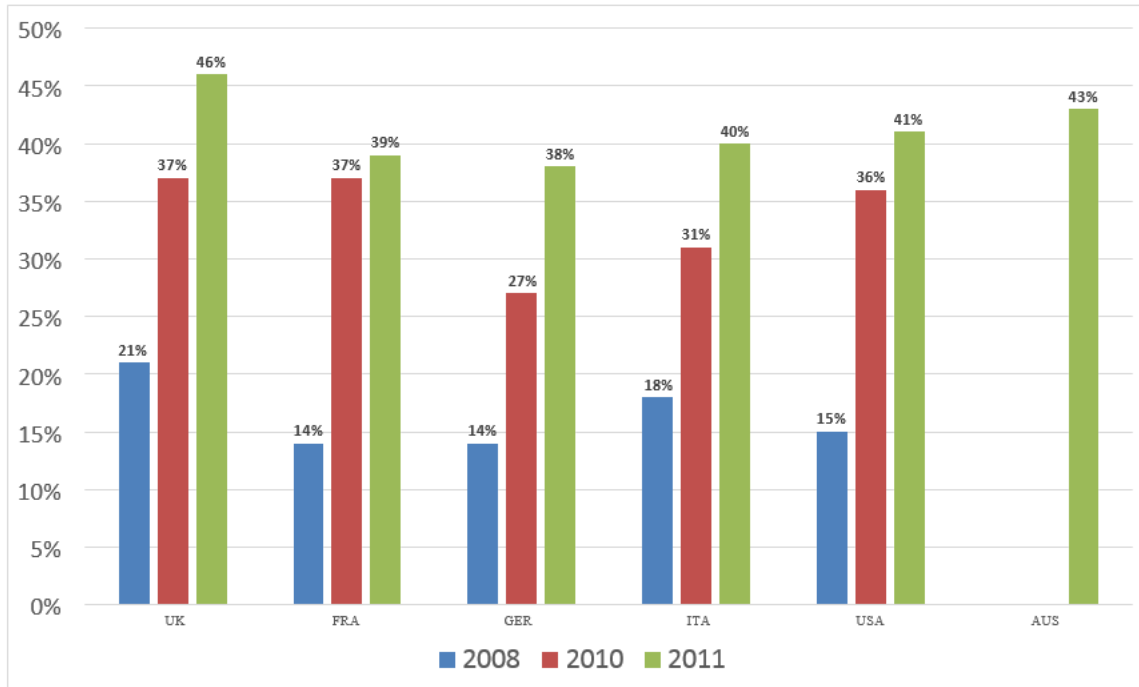
Figure 3.3: Percentage use of different devices



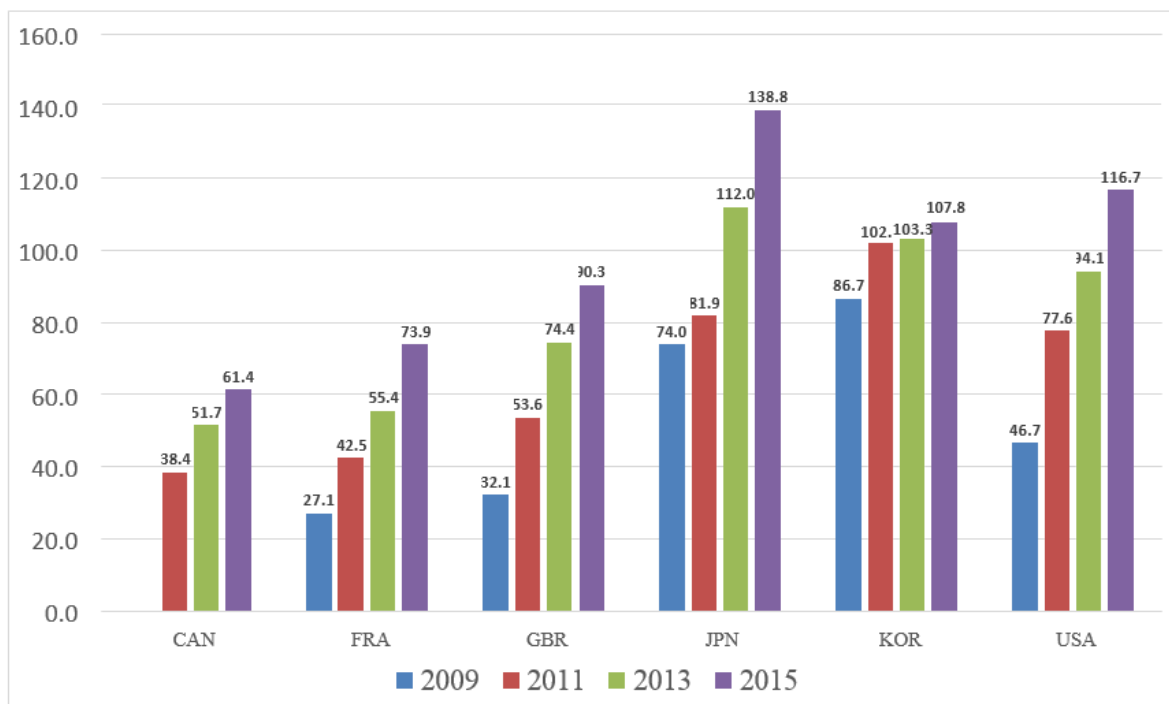
Source: OFCOM Internet use and attitudes Metrics Bulletin, several years

Figure 3.4: Use of mobile phone to access internet

(a) Percentage users

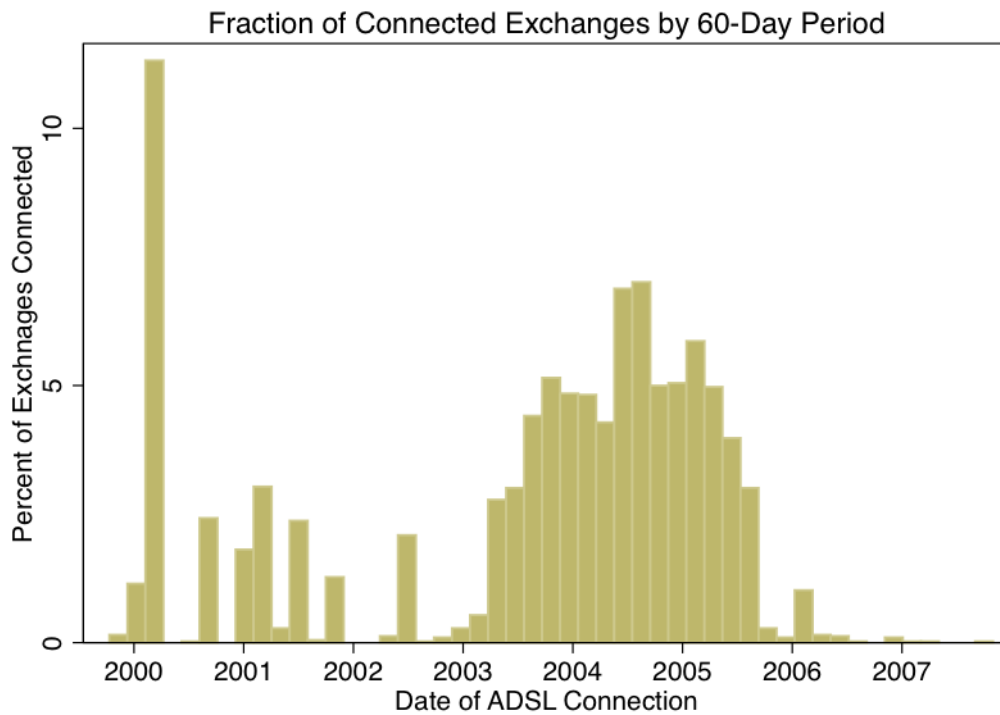


(b) Subscriptions per 100 inhabitants



Source: OFCOM International Communications Market Report and OECD

Figure 3.5: ADSL Switch-On Rates - Exchange-level

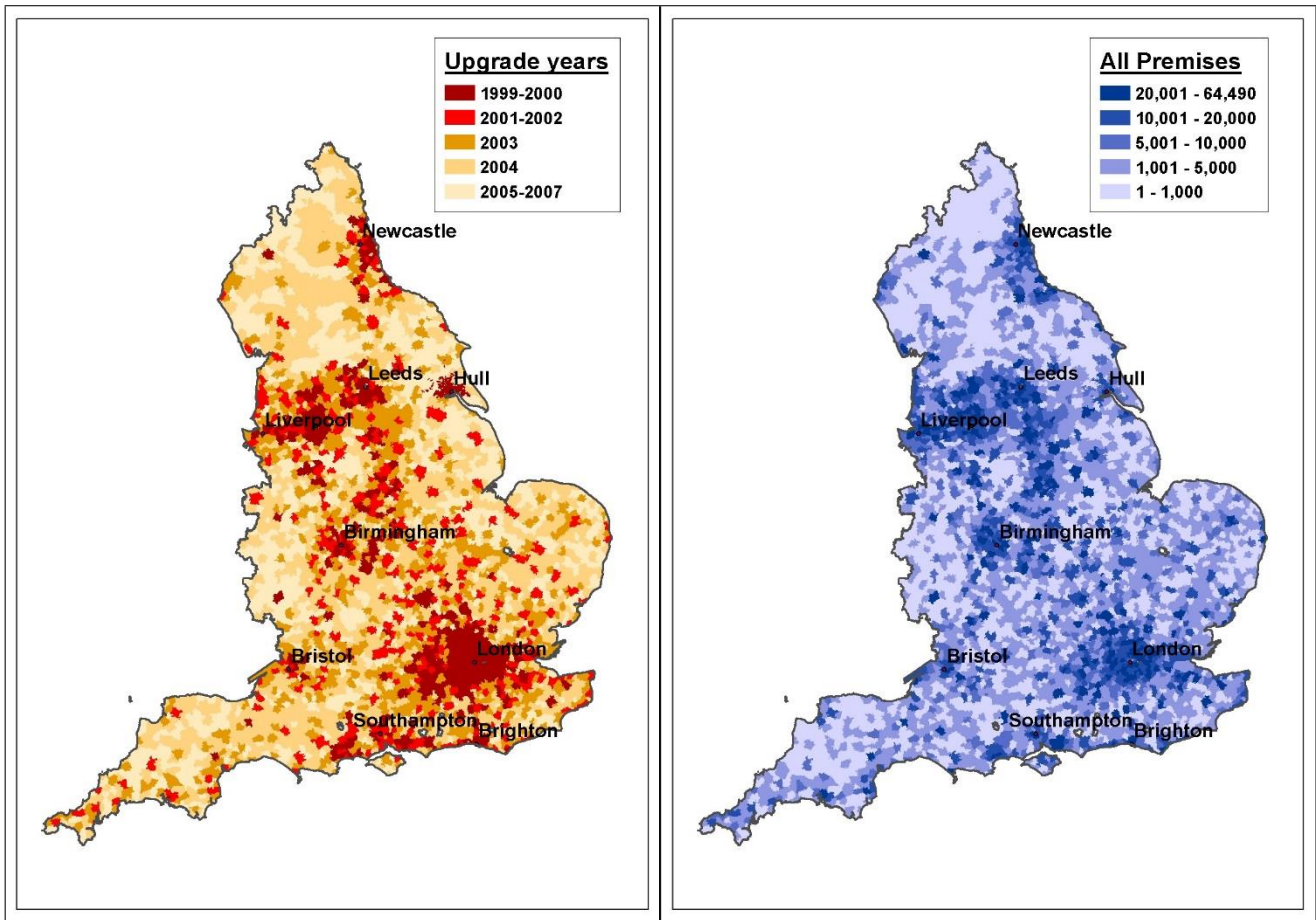


Notes: Data for 3,992 exchanges across England. Width of histogram bars is calculated as two-month intervals. Percentage connected defined solely in frequencies amongst exchanges with no population weighting in this figure.

Figure 3.6: England Upgraded Exchanges Maps

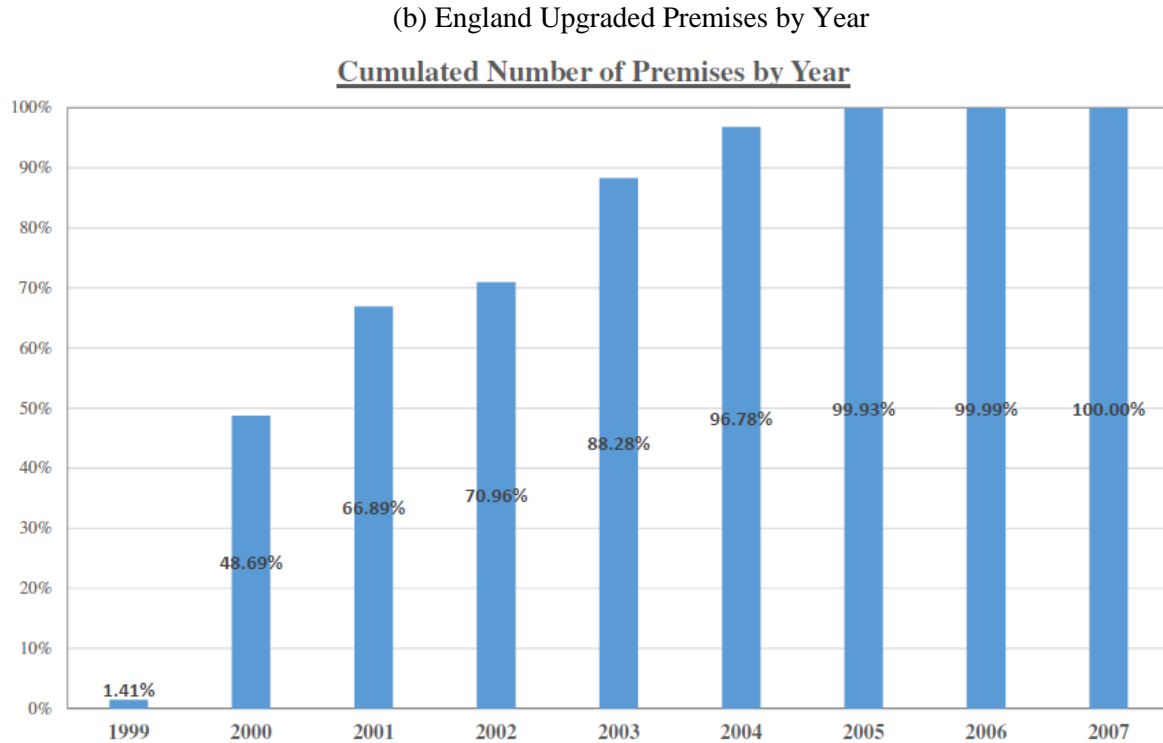
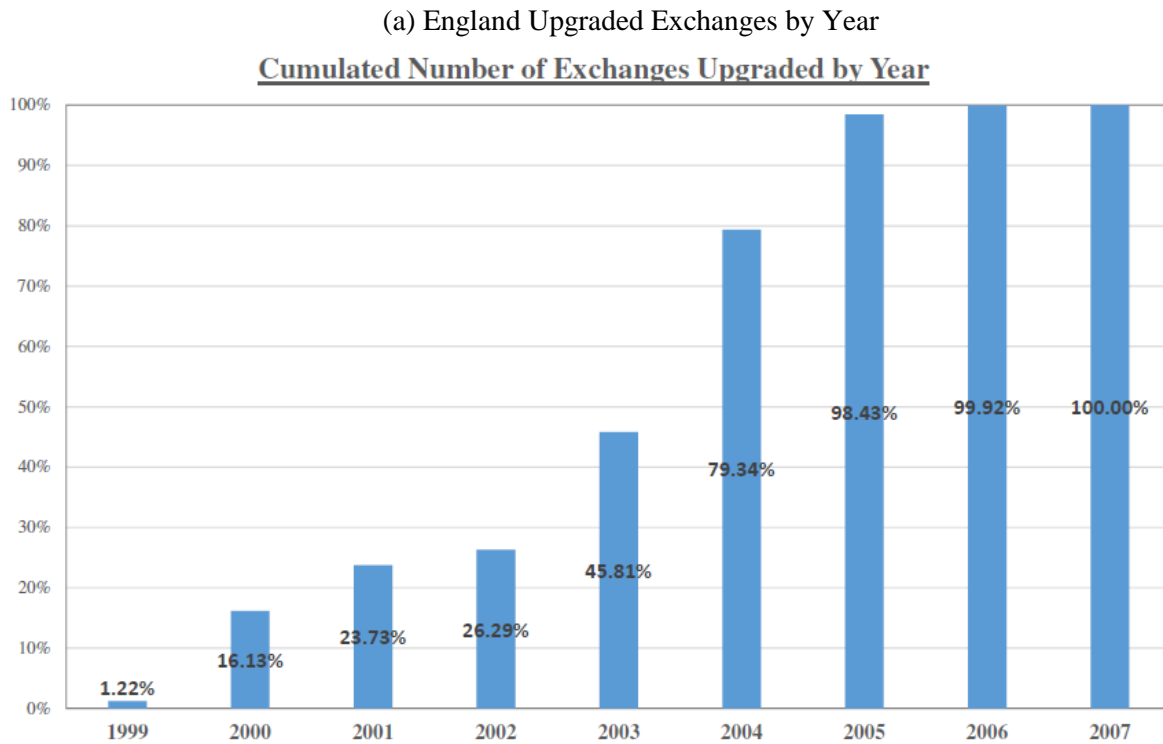
(a) England Upgraded Exchanges

(b) England Upgraded Premises



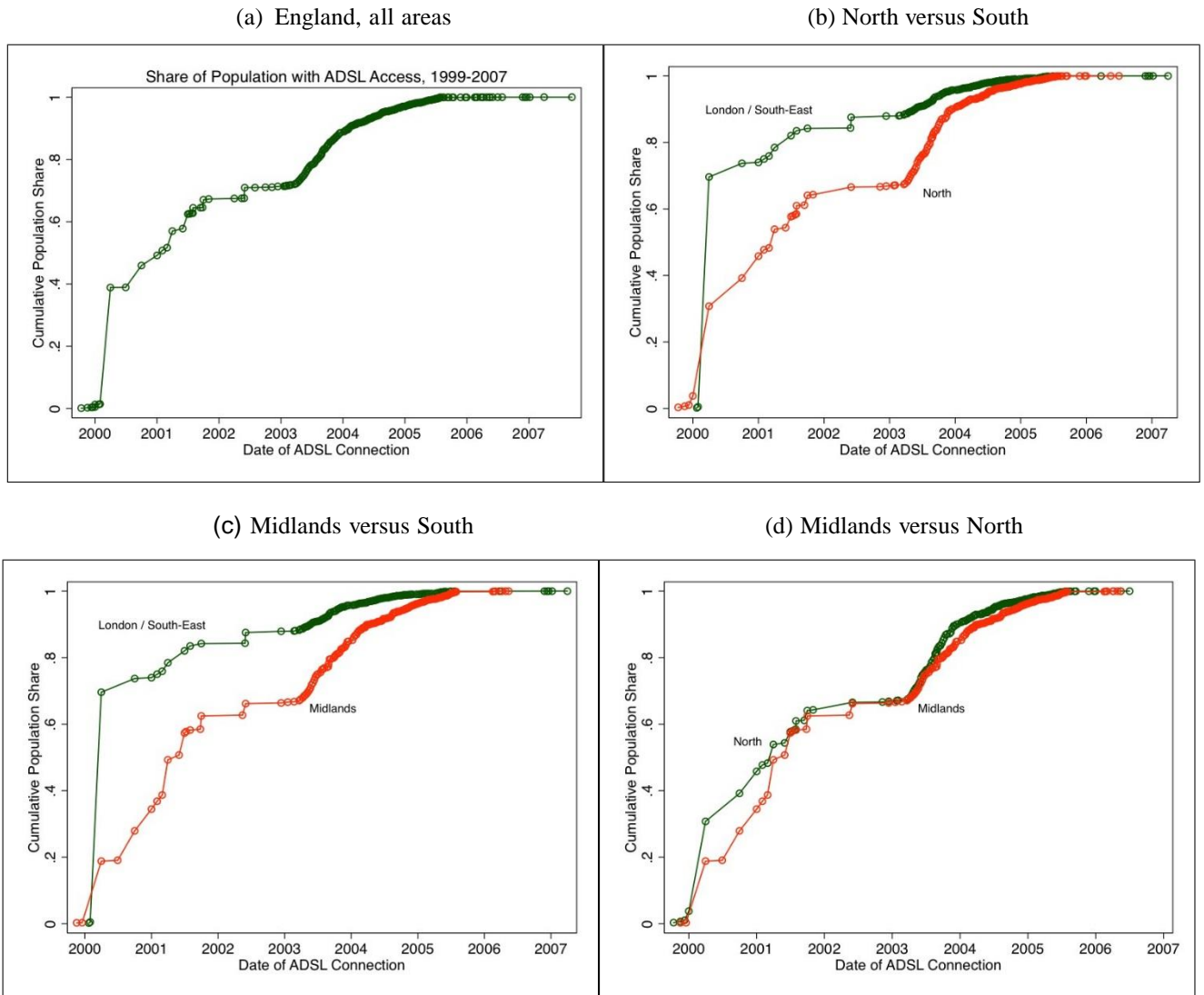
Notes: Data for 3,992 exchanges across England. Heat mapping in panel (a) indicates during which period that exchanges in an area were upgraded to ADSL. The premises information in panel (b) weights the upgrading according to the number of business premises in a 2001 UK Catchment area.

Figure 3.7: England Upgraded Exchanges by Year



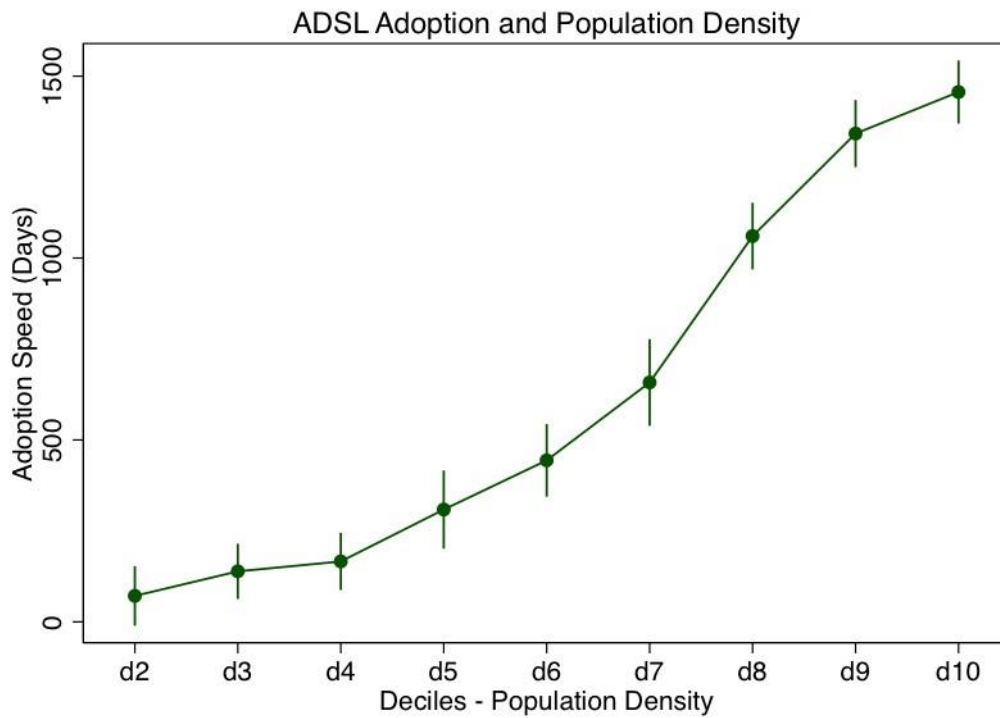
Notes: Panel (a) gives the cumulative number of exchanges connected by year (unweighted at the level of 3,992 exchanges). Panel (b) reports the cumulative number of business premises covered by the activated exchanges.

Figure 3.8: The ADSL Roll-Out of the 2000s



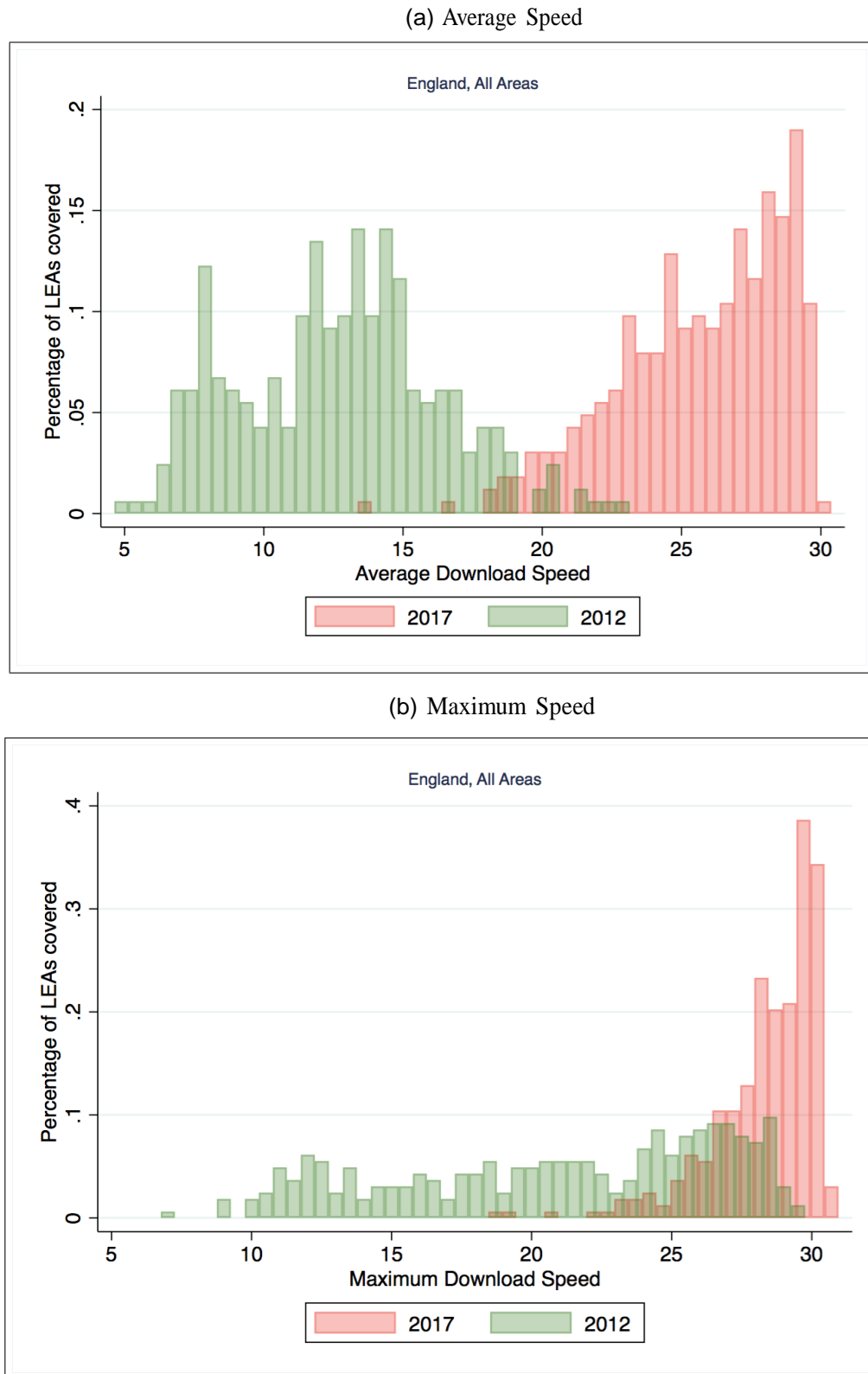
Notes: These figures show the population-weighted coverage of exchange roll-out by region. Panel (a) shows all areas, while panel (b) compares London and the South East to the combined North West and North East regions. The Midlands is defined as the combined East and West Midlands.

Figure 3.9: ADSL Roll-Out by Population Deciles



Notes: Data for 3,992 exchanges across England. The y-axis measures the coefficient for a set of dummies that define each decile of the distribution of population density. Specifically, the coefficient measures how many days sooner a decile obtained ADSL. Higher deciles indicate higher population density

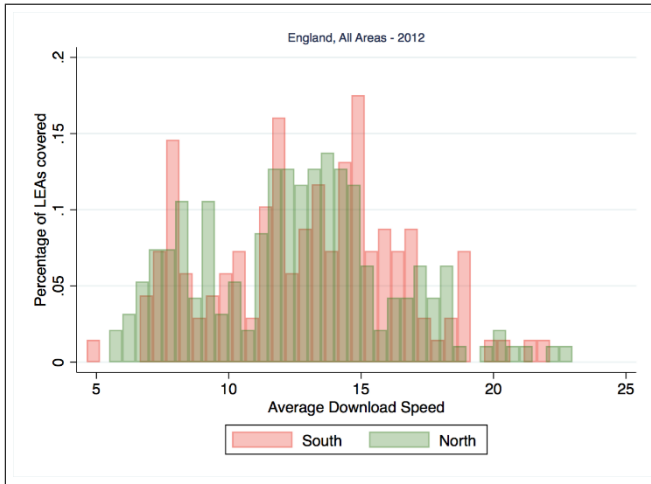
Figure 3.10: Broadband Speeds Across LADs, 2012–2017



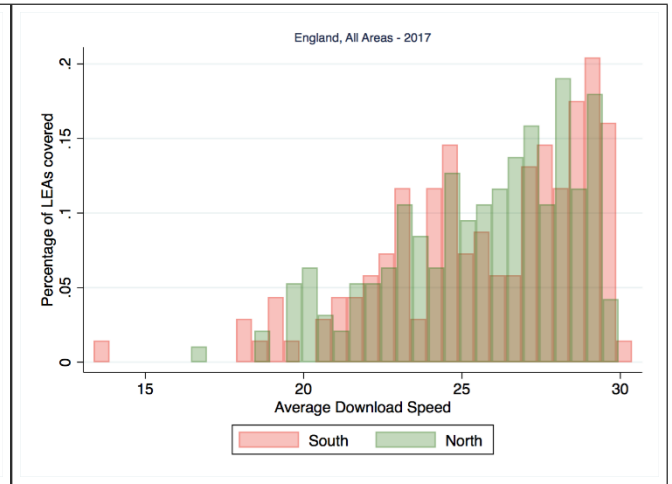
Notes: Author’s calculations based on Ofcom broadband speed data for 2012 and 2017. Data aggregated to the level of 326 Local Education Authority (LA) areas. The bars in the histogram represent bands of 0.6MB/s speed. The brown coloured sections trace out the density of the less common mass in a given section of the support.

Figure 3.11: Broadband Speeds - North versus South

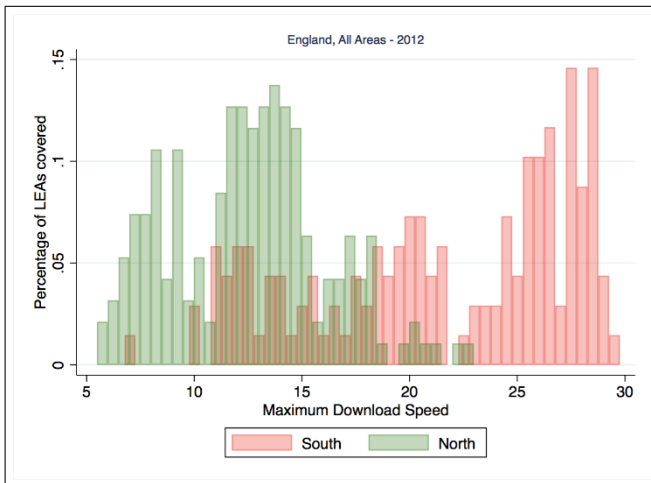
(a) Average Speed 2012



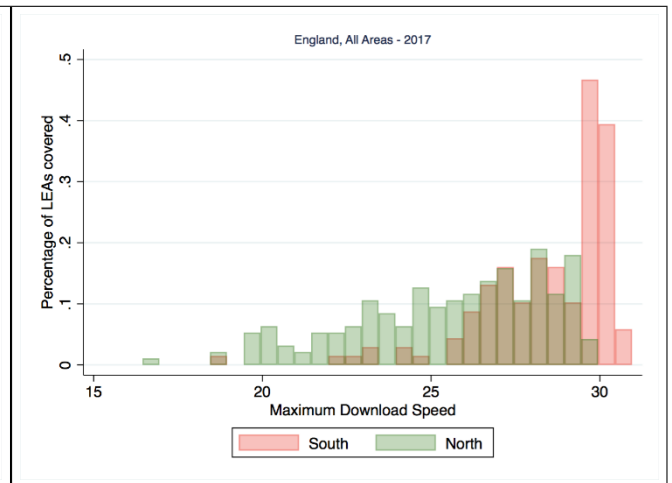
(b) Average Speed 2017



(a) Maximum Speed 2012

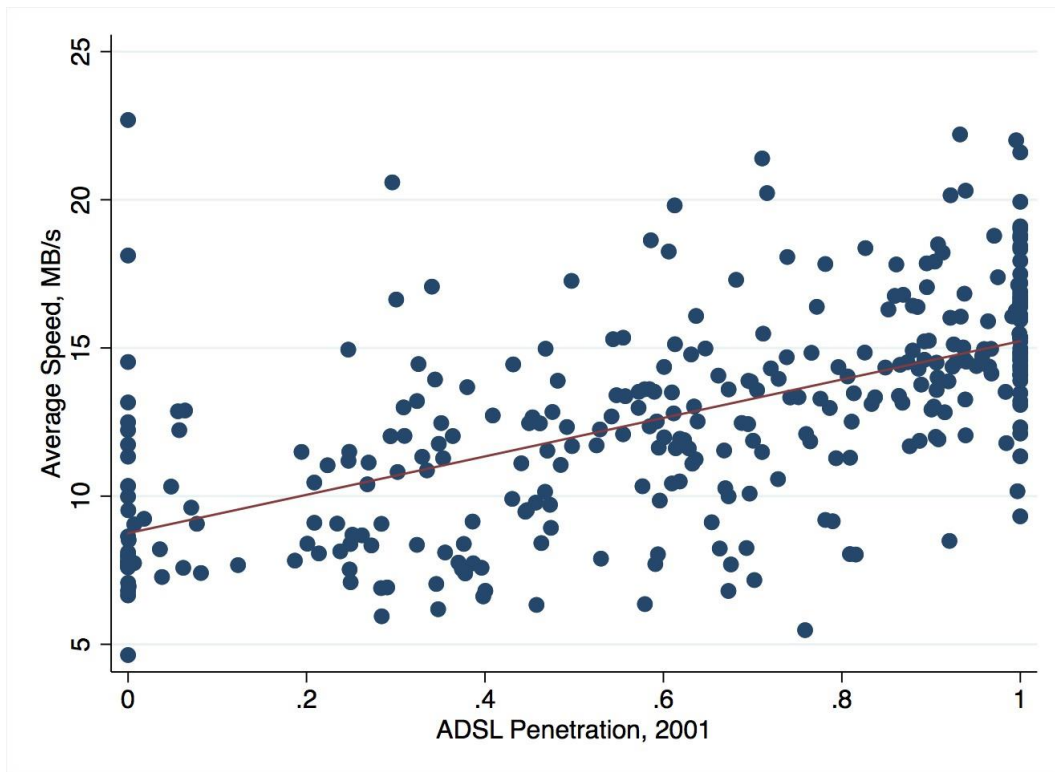


(d) Maximum Speed 2017



Notes: Author's calculations based on Ofcom broadband speed data for 2012 and 2017. Data aggregated to the level of 326 Local Authority Districts (LAD) areas. The 'South' is defined as London, the South-East and the South-West while the 'North' is defined as the North-East, the North-West and Yorkshire-Humber. The bars in the histogram represent bands of 0.6MB/s speed. The brown coloured sections trace out the density of the less common mass in a given section of the support.

Figure 3.12: Correlation Between Early ADSL Roll-Out and Average Speeds in 2012



Notes: Data for 326 LADs across England. The y-axis measures average broadband speed in 2012 while the x-axis measures the ADSL penetration rate at the LAD-level in 2001. The correlation coefficient for this relationship is 0.58.

4. Chapter 4: The Next Wave? Automation and the Future of Work

4.1. Introduction

Many societies in the late 2010s are currently gripped by what could be called ‘automation anxiety’. A recent survey from the Pew Research Center (2017) found that over 70% of American adults are ‘worried’ by the prospect of robots performing more varieties of jobs and 67% are also concerned about the use of algorithms in evaluating and hiring job candidates. Interestingly, only 54% express worries about the development of driverless cars.

However, concern about automation is not new. Consider the 1966 *Report of the National Commission on Technology, Automation and, Economic Progress*, which was commissioned directly by President Johnson to study the economic implications of automation. Concerns in this period focused on a number of radical technological changes that were then current:

‘Technological change is exemplified by the automation of a machine tool, reorganization of an assembly line, substitution of plastics for metals, introduction of a supersonic transport, discovery of a new method of heart surgery, teaching of foreign languages by electronic machines, introduction of self-service into retailing, communications by satellite, bookkeeping by electronic computer, generation of electricity from nuclear energy, introduction of frozen foods and air conditioning, and the development of space vehicles and nuclear weapons.’ Bowen (1966)

The Bowen (1966) report is rich in novel empirical evidence and stands up alongside modern research on ‘evidence-based policy’. It also takes a moderate position, arguing that technology ‘eliminates jobs not work’ and stresses that the demand for goods and services is the most important factor in determining employment. Further to this, it makes the point that technology is as related to the creation of new types of jobs as it is to the displacement of existing types.

We now have access to over fifty years of hindsight in evaluating the Bowen Report. Since 1940, agricultural employment in the US has disappeared at a rate of a 3.5% lower employment share per decade⁶. The set of occupations associated with machine operation and labouring also declined consistently with clerical/sales position turning downwards from 1980. As we discuss in this chapter, recent research has shown that these structural shifts are plausibly related to a wave of technological change associated with the displacement of ‘routine’ work. The Bowen report was correct in noting that technology creates work - the growth of professional and managerial occupations over recent decades is evidence of this - but the significant employment losses experienced in middle-occupation routine work are hard to classify as moderate.

In this chapter, we discuss the current thinking on the impact of technology on jobs and present new empirical evidence that evaluates whether a ‘next wave’ of automation is in process. In line with recent discussions (for example, Brynjolfsson and McAfee (2014), Frey and Osborne (2017)) our focus is on evaluating the potential that technologies such as AI and robotics have for affecting patterns of employment in ‘non-routine’ jobs that have previously been shielded from the effects of ICT because of the intrinsic complexity of their core tasks.

The basis of our new evidence is the little-studied US Occupational Employment Statistics (OES) database. This database has the key novel feature of covering occupational structure in much more detail than has been typical in the research - 600-700 different occupations versus the 300-400 covered in the majority of studies. The OES database also has the strength of covering these detailed occupations on

⁶ As Autor (2015) reports, agricultural employment represented 41% of the US workforce in 1900 but only 2% by 2000.

an annual basis since the late-1990s. While the occupational employment changes of the 1980s and 1990s are now well understood, the patterns that emerged in the 2000s have been subject to less analysis. Finally, as we show, the OES can be matched to explicit measure of task characteristics that allow us to test the degree to which fine measures of task intensity matter for employment and wage changes. These measures of task structure include the main measure of the probability of automation developed in the widely quoted Frey and Osborne (2017) report.

In the sections below we first give a short summary of the key findings of the new evidence and then review existing work on automation before presenting the detailed results of our analysis.

4.2. Key Findings

- Existing work has identified what can be characterised as a ‘first wave’ of automation based around the displacement of low skill ‘routine’ jobs and growth of high skill ‘non-routine’ jobs. The 1980s and 1990s saw the disappearance of middle skill routine jobs. Importantly, the relative growth of high skill non-routine jobs slowed down markedly in the 2000s while low skill ‘non-routine manual’ jobs increased.
- Our main analysis relates to a study of US Occupational Employment Statistics (OES) data for the 1999-2016 period. This data allows for an analysis of 600-700 occupations on an annual basis. Further to this, these occupations have been matched to measures of task content data from the occupational skills database O*Net. This allows us to define a full set of occupational task groups across routine, non-routine, manual, analytic and inter-personal task types.
- In our main exercise, we test whether employment and wage patterns moved according to the level of different task inputs. For example, did occupations that are highly intensive in non-routine analytic tasks grow faster than occupations that were still high skill but slightly less intensive in these tasks? Since we are able to control for broad occupational groups our analysis is able to distinguish between relatively fine differences in task structure.
- Over the period from 2000-2010 there is pervasive evidence that employment patterns moved in line with initial levels of task content. For example, a shift in task content from the 35th percentile of the non-routine analytic task distribution to the 70th percentile is associated with a 5.7% difference in US occupational employment growth over 10 years. A similar shift in routine manual task content was associated with a 5.2% decline in employment. Amongst low skill occupations growth was strongest for occupations with high levels of manual inter-personal tasks.
- The growth of wages was flat with respect to most measures of task content over the 2000-2010 period. This is consistent with other recent studies that have reported sluggish wage growth across the labour market in this period.
- Considering the decade of the 2010s up to the year 2015, there is no evidence that employment or wages are growing in line with task content measures. However, this does not imply that the influence of technological change on the labour market has ‘slowed down’. An analysis of year-by-year changes shows that developments in the 2010s were driven by reallocation during the recession years of 2009 and 2010. The fact that a major recession has not occurred in the 2010s (yet) is likely to account for the secular pattern of job growth in this period.
- Based on recent employment trends, there is no evidence that a turning point has been reached in the pattern of jobs that could be affected by advances in automation. However, the past evidence indicates that structural transformations in employment take place during recessions so the most likely starting point for a ‘second wave’ of automation affecting a new class of jobs is the next downturn.
- In terms of platform-based ‘gig’ format work, the current best estimate is that 0.5% of the US workforce is engaged in this type of work. This represents a pattern of high growth from a low

base and the estimates suggest that gig economy growth can account for a large fraction of US employment growth since the last recession.

- The most likely scenario for the future of the gig economy is continued strong growth and the question is whether it will level off, perhaps as a result of regulatory action (for example, limitations on rideshare services). A point of interest at the moment is whether current strong employment conditions will translate into a rebound in traditional jobs or as an expansion of contingent work.

4.3. Understanding Automation

4.3.1. *The Task Approach*

The key analytical concept that has driven discussion of automation is the model of routine and non-routine job tasks. This approach was introduced by Autor et al. (2003) as a framework that was compatible with a close analysis of the role of computers in the labour market and is adaptable for studying the impacts of other types of technology.

The Autor et al. (2003) task model is motivated by a ‘machine’s eye view’ of the role of computers in production. Specifically, computers occupy the technological niche of scaling the implementation of tasks that involve detailed, explicit instructions or programmable rules. By contrast, they still face operational challenges when situations require tacit knowledge that is hard to codify. This is summarised in Table 4.1, which further describes the distinction between analytic/interactive tasks and manual tasks. In turn, this results in four main classes of jobs that have been comprehensively mapped into various data sources since the work of Autor et al. (2003).

The novel aspect of the job tasks framework is that it can explain patterns of employment and wage changes that were not tractable in previous frameworks that distinguished between jobs on the basis of the education of the workers involved. In particular, the task model can explain the differential impact of ICT across the labour market. Computers are able to substitute for the tasks found in routine jobs but enhance or complement activities in non-routine analytic jobs. Finally, the (current) technical limits of the robotics branch of ICT means that a large subset of jobs that involve hard-to-codify non-routine tasks (such as cleaning or the driving of vehicles) have been insulated from the direct economic effects of computers.

4.3.2. *The Varied Impacts of Automation*

Technologies that foster the automation of work activities such as ICT have varied effects. The central effect, already canvassed above, is the displacement effect whereby technology effectively replaces workers without also increasing labour demand. The reduction in demand then feeds into lower rates of employment and wages in the technology-affected jobs. This displacement effect has been the main focus of discussions of automation. However, following the framework of Acemoglu and Restrepo (2018) automation also feeds into other economic mechanisms that typically have a countervailing influence on employment and wages.

- **Productivity Effect:** The demand for labour in non-automated tasks can increase as the cost of performing routine tasks falls due to automation. An example of this, via Bessen (2016), is the employment growth amongst bank tellers that occurred after automatic teller machines (ATMs) were introduced. ATMs reduced costs and, because demand for banking services was elastic, more branches could open and more sophisticated non-routine services could be rolled out. Other examples of occupations whose employment surged after major technological innovations include check-out cashiers and paralegals (Bessen, 2016).

- **Capital Accumulation:** This is a similar channel to the above productivity effect. In this case, automation triggers capital accumulation (i.e.: more capital investment) which increases labour demand since more workers are needed to work with the machines.
- **Deepening of Automation:** What if automation strikes but does so with respect to activities that are already heavily automated? In that case, labour cannot be displaced but the productivity effects of this type of ‘intensive margin’ automation can increase labour demand. An example of this would be the introduction of tractors in agriculture, which replace horse-powered reapers and harvesters but not necessarily manual labour.
- **Reinstatement Effect:** Simply put, automation in some areas can open up the space for new, labour intensive tasks to be created. This is both because labour is needed to work adjacent to and service more sophisticated machines and because technology can create new classes of jobs entirely. The increasing specialisation of non-routine cognitive service sector employment is the most prominent recent example of this.

This framework for understanding the job creating dynamics of automation is relatively new and currently there is limited research available that explicitly quantifies these mechanisms. The main focus has been modelling that aims to pick up patterns of displacement via a process of ‘routine-biased technological change’. We review this evidence below and then put forward some modelling of employment growth in the post-2000 period. While this new modelling is not able to explicitly address the countervailing effects described above it does offer a more detailed analysis of the growth of non-routine occupations in the post-internet era.

4.4. Empirical Evidence from the ‘First Wave’

4.4.1. *Routine-Biased Technical Change?*

Employment Polarisation - The US Benchmark

Following the framework above, what can be characterized as a ‘first wave’ of post-WWII automation has been identified across a range of studies. Specifically, this first wave has been generally understood as a process of ‘routinization’ whereby jobs that (following the Autor et al. (2003) framework) involved either routine cognitive or routine manual tasks were subject to relatively stronger substitution effects arising from the introduction of new ICT-related technologies.

Figures 4.1-4.3 summarise the fundamental shift that has been documented in the literature. It summarises how occupational employment shares changed over the 1980s and 1990s. The occupations in Figure 4.1 are ranked according to ‘skill’ based on their average wage in 1979, that is, wages are used as a proxy of skill. Specifically, this figure then depicts a pattern of labour market polarisation whereby a segment of low-skill jobs grew strongly over time along with another segment of high-skill jobs. In combination, the two trends contributed to the phenomenon of a ‘disappearing middle’ in the employment distribution.

Two main patterns are evident from Figure 4.1-4.3. Developments in the 1980s and 1990s (Figure 4.1) took the form of declining employment shares in the lower middle of the skill distribution, defined roughly as the band between the 5th and 45th percentiles in the 1980s and the 15th-55th percentiles in the 1990s. Hence the lowest skill jobs (i.e.: up to the 15th percentile) actually grew again in the 1990s while the middle 45th-55th band was newly exposed to losses in their overall employment shares. This growth in low skill jobs from the 1990s onwards was mainly concentrated in service occupations involving non-routine manual tasks. Over both decades, the employment shares of high skill occupations beyond the 60th percentile increased, with a sharp ‘twist’ evident for those above the 80th percentile.

The 2000s and early 2010s saw a substantial shift in this pattern, as seen in Figure 4.2. This shows that there were large increases in low-skill jobs below the 20th percentile and a moderate decline in the middle from the 25th up to the 75th percentile. Further to this, there was limited growth at the top - the 1999-2007 period saw generally flat growth across percentiles with some growth re-emerging above the 80th percentile in 2007-2012. In order to show the contrast in these patterns over the decades we directly plot the different changes in 1979-1989 and 1999-2007 against each other in Figure 4.3. This decisively shows the overall polarisation of employment that occurred between 1979 and the late 2000s, specifically the loss in employment shares experienced by middle skill occupations where routine cognitive and routine manual work dominated.

The analogous changes in wages are summarised in Figure 4.4. This shows the change in wages by occupational skill percentile over all of the different intervals between 1979-2012. The patterns here are in line with the changes in employment shares seen in the earlier figures. Low and high-skill occupations experience higher wage growth in the 1980s and 1990s relative to the middle groups. However, in the 2000s there is a linear gradient in wage changes - the higher the level of skill then the higher is the increase in wages. Also notable is the fact that wage growth the 2000s and 2010s is of an order of magnitude lower than in the 1980s and 1990s. Indeed, wage growth is effectively flat across all skills levels in the 2007-2012 post-crisis period.

Non-Technological Explanations?

The patterns presented in these figures are consistent with a ‘routine-biased technical change’ hypothesis but the underlying issue in the research has been to distinguish a technology-based explanation from other potential reasons for the observed changes. Many economic changes occurred over this period - large increases in import penetration from low-wage countries such as China, shifts in institutional structures such as the minimum wage and unions, and changes arising from the contraction of government (for example, via privatization).

The research on the labour market has tackled this challenge with recourse to three strategies. Firstly, it has been consistently found (e.g.: Autor et al. (2003), David and Dorn (2013)) that the adoption of ICT has been correlated with measures of initial routine task intensity. That is, the more routine that an occupation is initially the greater the rate of ICT adoption in subsequent years. Secondly, measures of task structure (e.g.: the ‘amount’ of routine or non-routine tasks in a job) are correlated with employment and wage changes. Simply put, the nature of tasks at a finely measured level matters. This is an area where we contribute new evidence in this chapter via our analysis of the US OES data.

The third source of evidence for the routine-biased technical change argument is summarised in Figure 4.5. Derived from the data assembled by Goos et al. (2009), this shows the structural shifts in occupational employment across 16 European economies from 1993-2005 (the period for which consistent data is available). Indications of employment polarisation are apparent across the full range of countries. Declines in the share of middle-paying occupations are in the 10-15% range for countries such as France, Austria and the UK. Further to this, employment polarisation in these European economies is even more pronounced than in the US. The fact that this polarisation pattern is so common across countries underlines the routine-biased technical change hypothesis. Other explanations based on unions, wage-setting institutions or import penetration need to rely on having the same structure across all these countries to be convincing as a main explanation. Hence the pervasive nature of the structural change in occupation shares strongly suggests that a common technological shock related to routine task levels has been at play.

4.4.2. *Economic Challenges of Automation*

Worker Adjustment to Automation

The ‘technological optimist’ view on automation that we outlined in the framework section earlier makes the crucial point that, as well as displacing jobs, technology also creates them. The key question for policy is then how workers flow between different types of jobs over time. In particular, what happens to workers employed in routine jobs when those jobs disappear?

The evidence on adjustment is less well-developed than the broader research on polarisation discussed above but points to systematic problems in the transition of workers between different types of jobs. Cortes et al. (2014) conduct an analysis of employment transitions in the US labour force from 1976 onwards. Specifically, this work makes use of the longitudinal tracking component of the US Current Population Survey (CPS). This tracking component makes it feasible to observe the short-run (up to 15 months) history of workers in the periods after they exit from employment in routine occupations.

Table 4.2 reports data from Cortes et al. (2014) on some critical labour market transition rates. These rates compare trends in the early pre-polarisation period in the late 1970s to the post-crisis expansion from 2009-2012. The first panel (A) reports ‘return’ rates, that is, the rate at which workers flow back into the same type of occupational employment after a spell in unemployment. This shows very sharp declines in the return rate of routine manual and routine cognitive workers - approximately -9.3% and -7.9% respectively. These declines are 2.5-3.0 times larger than those experienced by non-routine workers. In short, in the current US labour market, once routine workers fall out of employment in their jobs it has become much less likely that they will be able to re-position themselves into the same type of work.

The second panel describes switching between different types of jobs. There is a -4-5% reduction in the transition rate of former non-routine manual and routine cognitive workers into routine employment of any type (columns (1) and (2)). More critically, the already low transition rates of routine workers into non-routine jobs falls by around 20-25% (columns (3) and (4), panel (A))⁷. The final panel then considers transitions from non-participation in the labour market back into employment. This shows that the route from non-participation into routine employment has contracted with, for example, the job finding rate into routine manual declining from 1.2% to 0.8%. Interestingly, the job finding rate for non-routine cognitive workers increased slightly but the most likely return destination for low-skill workers (non-routine manual employment) has been impeded (see the -0.3% decline).

In summary, these findings imply that it has gotten harder for former routine occupation workers to re-position themselves in the non-routine sector of newly created jobs. Note here that these declining transition rates have occurred even though the number of non-routine jobs has increased greatly in absolute terms.

Other work (Cortes (2016)) tracking workers over a number of years indicates that even if routine workers manage to stay in the routine job sector they are still subject greatly reduced wages (a -17% fall relative to nonroutine groups between 1976 and the mid-2000s) while other studies (Poletaev and Robinson (2008), Gathmann and Schönberg (2010)) have shown that there is significant task-specific ‘stickiness’ across occupations. Although this issue of limited occupational mobility has always been a feature of labour market dynamics the concern is that - based on evidence such as Cortes et al. (2014) - mobility has been declining in practical terms during the 2010s.

⁷ Specifically, an example of this calculation is the fall in the transition rate of former routine cognitive workers into non-routine cognitive employment from 14.8% to 11.7% in column (4).

The Role of the ‘Gig Economy’

The notion of a flexible ‘gig economy’ facilitated by decentralised online platforms has gained attention recently as highly visible companies such as Uber and Deliveroo have emerged, along with more specialised high skill operations such as Upwork and TaskRabbit. The flexibility of these new forms of employment opens up the question of whether they may be able to facilitate the adjustments needed to deal with the pressures of automation. But how common is gig economy work currently and exactly what structure has it taken?

The formal measurement of the gig economy is still in its early stages. One approach (taken by Katz and Krueger, 2016) is to focus on the prevalence of nonstandard or ‘contingent’ work. They conduct their own RAND-Princeton Survey in 2015 to follow up an earlier survey on contingent work in 2005 that has been discontinued. A similar format of questions and sampling scheme was used to ensure comparability across time. The definition of contingent work encompassed independent contractors, on-call workers, temp agencies and contract firms. The survey also explicitly asked respondents whether they sold goods or services directly to customers and whether they worked through an intermediary, including online platforms.

The results are reported in Table 4.3. The increase in the alternative work arrangements between 2005 and 2015 - a rise of 5.7% to 15.8% based on the authors preferred estimates (column (4)). Growth is strongest in proportional terms for workers in contract or ‘outsourcing’ firms, which increase their overall share 4.5 times over, coming from a low based of 0.6%. The figures for the growth of self-employment are also corroborated by direct administrative counts of the relevant tax returns.

One major implication of this growth is that it can account for all of the net employment growth in the US economy between 2005-2015. While the survey does not track whether the contingent jobs were mainly created after the recession it does suggest that traditional employment systematically lost ground to alternative work arrangements. This could have come about either through direct transitions (e.g.: former traditional workers got redeployed into contingent work) or turnover (new cohorts entering or re-entering the labour market took up this work).

More detailed breakdowns by Katz and Krueger (2016) indicate that the growth was driven by female workers, who increased their share from 38.6% to 50.8%. Multiple job holding also became more common among contingent workers, rising from 7.3% to 33%. There was also a mix of low and high skill workers across the arrangements with the low skill workers dominating in Temporary Help and On-Call work.

The actual estimate of online-related ‘gig economy’ work as it is currently understood is derived from a question about direct selling to customers. In turn, workers answering ‘Yes’ to this question are then asked whether this direct selling involves an online intermediary. On this basis, Katz and Krueger (2016) estimate that 0.5% of all US workers are employed in this modern ‘gig’ format. This survey-based estimate is close to that obtained from other methods such as counts of Google searches (Hall and Krueger, 2015) and the frequency of platform-related bank deposits (Farrell and Greig, 2016).

Research by Costa et al. (2017) discusses UK gig economy work in the context of self-employment, the most likely place for these forms of work to show up in UK statistics. They note that the rise in self-employed workers in the UK (from about 9% in 2000 to 13% in 2016) was driven by the self-employed without workers, that is, solo operators. Earnings for this group are on average lower than other types of self-employed and actually fell 20% in real terms since 2007/08.

The current picture of self-employment and the gig economy therefore suggests that it may be acting as a ‘sink’ to absorb workers displaced from traditional employment by forces such as automation. In particular, gig economy work may be operating as a channel into non-routine manual employment in sectors such as transportation and home assistance jobs (e.g.: cleaning, repairs). However, this hypothesis is still waiting to be decisively tested.

In terms of platform-based ‘gig’ format work, the estimate of 0.5% for the US is actually quite large given the short time since the emergence of these forms of employment. The most likely scenario for the future is continued strong growth and the question is whether it will level off, perhaps as a result of regulatory action (for example, limitations on rideshare services). A point of interest at the moment is whether current strong employment conditions will translate into a rebound in traditional jobs or as an expansion of contingent work.

4.5. US Occupational Employment Trends in the 2000s and 2010

4.5.1. *Measuring Task Content*

In this section, we offer an analysis of US Occupational Employment Statistics (OES) data. As discussed, this data is novel due to the large range of detailed occupations covered. Begun in the late 1990s, the OES covers 700-800 occupations versus the 300-400 that have been most common in the literature on task-biased technological change.

A further advantage of the OES is that it can be directly matched to O*Net, the key database on task structure that has again been maintained since the late 1990s and is increasingly the benchmark source for standardized task-based skill measures in the literature. In the following analysis we use the classifications of occupational task groups devised by Acemoglu and Autor (2011) using O*Net. These groupings are comprehensive and split occupations across the full range of task groups, for example, including distinctions between four different classes of non-routine jobs. Most importantly, these measures are also continuous, that is, they give a ‘score’ describing how (say) routine manual an occupation is in terms of its underlying task structure.

This allows us to test whether relatively fine differences in the task content between jobs matters for the direction of change in employment and wages. The fact that the OES covers the post 1999 period up to 2016 also allows us to evaluate the most recent labour market trends. Finally, we supplement the O*Net-derived measures of task content with a ‘probability of automation’ measure devised by Frey and Osborne (2017). This measure is itself based on O*Net data with the added input of a qualitative expert analysis of ‘automation bottlenecks’ fed through a classifier algorithm.

4.5.2. *Employment, Wages and Task Content in the 2000s and 2010s*

Our basic modelling approach is to relate 10-year and 5-year changes in employment to occupation level measures of task content in the baseline year of 2000. That is, we are testing (for example) whether occupations that are highly intensive in Non-Routine Cognitive (NRC) Analytic task inputs grow faster than other occupations. This approach is common in the literature (e.g.: Autor et al., 2003) but the detailed nature of the OES data allows us to add the embellishment of extra controls in terms of higher order occupational structure.

In simple terms, we are able to include 2-digit and 3-digit occupational controls in our employment regressions, which themselves are defined in terms of 5-digit categories. The inclusion of these 2-digit and 3-digit controls has the effect of sweeping out the broad occupational differences in the data (for example, between routine manual and non-routine analytic) and forcing the empirical models to test for differences in employment patterns amongst narrower occupational groups. This is important insofar that it sheds light on whether small(er) differences in task content matter and, by implication, helps to back up the hypothesis that the changing structure of the labour market is driven by technological forces.

Figures 4.6 and 4.7 report some of detailed occupational evidence on employment growth, namely the ‘bottom 10’ and ‘top 10’ occupations based on 10-year changes. This shows some classic patterns consistent with ongoing routine-biased change. For example, data entry, correspondence clerk and telephone operator positions have declined sharply amongst the routine cognitive occupations.

Interestingly, textile jobs has suffered declines amongst the routine manual group. These occupations were previously notable for having resisted automation (due to the high level of manual dexterity required in sewing tasks) but some changes seem to have taken root.

The results of our regression modelling are reported in Table 4.4 where we study employment changes over the 10 years from 2000-2010. The first column shows that the various measures have a lot of ‘signal’ and help explain 10-year employment changes. The direction of these effects is also in the expected direction with low-skill routine categories showing negative effects and high skill non-routine groups recording positive effects. The size of these coefficients falls as 1-digit, 2-digit and 3-digit controls are included in the subsequent columns and the comparisons in task input scores between occupations becomes narrower.

The results in column (4) include 3-digit controls representing 89 occupations. The majority of variables are significant indicating that employment changes have been occurring along finely defined lines. The magnitude of these effects is meaningful - a one standard deviation shift in NRC Analytic task content is associated with a 5.7% increase employment over 10 years. As an indication, the comparison points for such an effect would be the difference between an occupation at around the 35th percentile of the NRC Analytic distribution versus one at around the 70th percentile. The Frey and Osborne (2017) measure is also significant and appears to effectively nest within it information from across the different task measures.

Table 4.5 then conducts the same set of tests for hourly wages at the occupation-level. The results here indicate that wage changes are flat with respect to task content once we control for 3- digit occupational categories. This is in line with previous evidence reviewed above (e.g.: Acemoglu and Autor (2011), which found that wage growth was muted across all skill groups in the 2000s.

We next go onto using the same testing framework on data for the period between 2010-2015, reported in Tables 4.6 and 4.7. The results here are substantially different. Both employment and wages are flat with respect to the individual task content measures and the Frey and Osborne (2017) index of automation. There is only limited evidence negative employment effects for Routine Manual workers once 3-digit controls are included while higher NRC Analytic task content is associated with a small wage premium.

This result does not necessarily imply that employment and wage changes have ‘slowed down’ with respect to task content in the 2010s. In Figure 4.8, we decompose the long period analysis into 1-year effects. These plots show in what years the overall effects seen in the earlier tables are concentrated. We do not show all the categories for reasons of brevity but the patterns shown are illustrative. Most importantly, it is evident that the main changes in employment according to task structure are concentrated in the recession years of 2009 and 2010. This is in line with other recent research, for example Cortes et al. (2014), who show that routine jobs disappear in recessions. The finding of a flat gradient with respect to task content in the 2010-2015 period is therefore most likely due to the fact there has been no sharp economic downturn (yet) in the 2010s.

4.6. Conclusion

The evidence presented in this chapter does not uncover any clear evidence of a turning point in employment patterns consistent with the beginning of a new wave of automation. We define this concept of a ‘new wave’ as a trend of new rates of employment loss in either high skill non-routine jobs or other types of jobs previously insulated from the main effects of automation.

Commentary on the prospect of a new wave of automation has recently been stressing the importance of considering ‘partial automation’. This is the argument that it is not occupations that get automated but rather tasks. Since the existing data analysis has mainly been conducted at the occupation level this means that there is a danger of over-estimating the consequences of automation (Bessen (2016)). That

is, when automation hits it does not automatically imply that whole occupations are displaced. This approach has been a focus of work that has sought to project the impacts of automation by using a task-level analysis. The OECD report by Arntz et al. (2016) conducts this analysis and estimates that, on average, 9% of jobs across the OECD face a high automation risk. This is clearly much lower than the 47% estimated by Frey and Osborne (2017) and a range of others using an occupation-based approach (Pajarinen et al. (2015), Bowles (2014)).

An additional important point about the prospect of large-scale automation is the timing of such a wave. As the work on routine-biased displacement has developed it has strongly suggested that the process of structural change has been concentrated during recessions, and our new evidence strongly supports this. Since we are currently in a more benign period of employment growth - albeit one which is characterised by more insecure jobs - the most decisive indication of a new wave of automation is unlikely to emerge until the next cyclical downturn.

Robots in the Labour Market

Recent research has begun to discuss the impact of investments in robotics on the economy in a range of countries. Note here that the investments in robotics here represent 'first generation' technologies from the 1980s and 1990s, so they precede the most recent advances surveyed (for example) by Brynjolfsson and McAfee (2014).

Robots at Work: Graetz and Michaels (2015) study industry-level data for 17 countries from 1993-2007. Their data is based on robot delivery numbers from the International Federation of Robotics (IFR). This data shows that preconditions for wider robot adoption have been in place: prices fell by approximately one-half from 1990-2005 and, similar to general ICT equipment, the effective price decline is larger when quality is taken into account. Robot density (defined as the number of robots per million hours worked) increased by 150% from 0.58 to 1.48 over the 1993-2007 period. Levels of density by 2007 were especially high in Germany (4.4), Italy (2.5) and Sweden (2.2). The US level in 2007 was 1.4 - about the upper middle across the 17 countries.

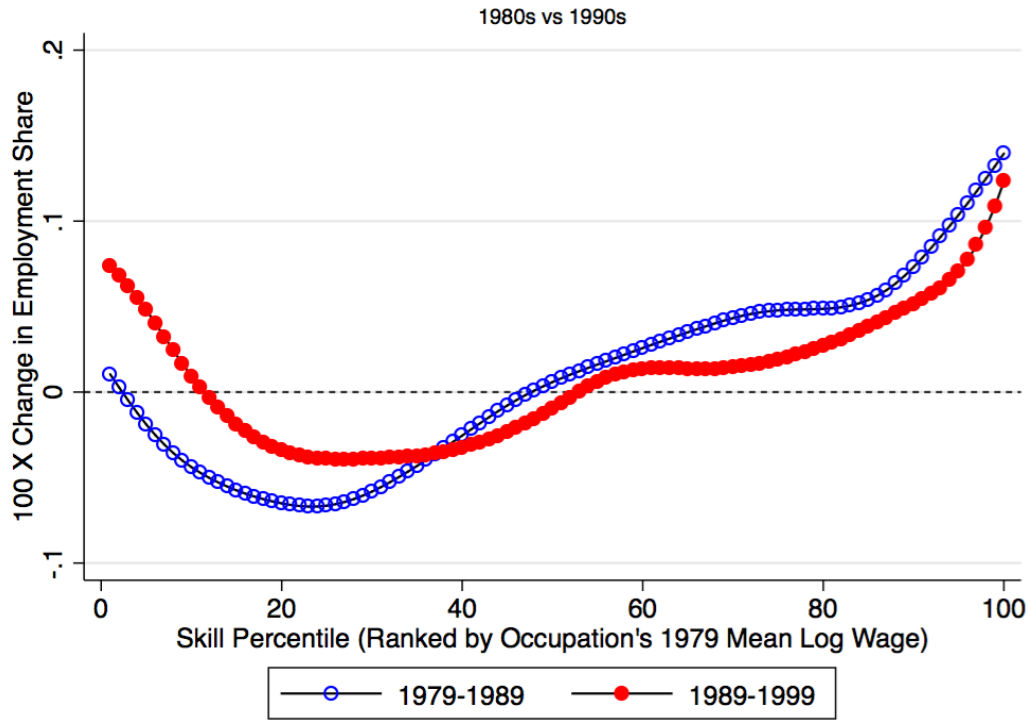
The analysis shows that increased robot density has a substantial association with labour productivity growth - about 0.36% points compared to the mean rate of 2.4%. These findings are supported by instrumental variables strategies (derived from measures of the scope of robot tasks) and some case studies of robots in practice. These case studies indicate very high returns to robot investment - around 25% - and the authors suggest some explanations for why robots have not been adopted more widely. These include: a still narrow range of replaceable tasks for robots to take up, limits to managerial resources, and a simple lack of information.

Robots and Jobs: Acemoglu and Restrepo (2017) look at the relationship between robots, employment and wages in the US from 1990-2007. They also utilise data from the IFR but use geographical areas (commuting zones) as their unit of analysis.

Their modelling picks up some decisive co-movements between robot investment and outcomes. An extra robot per thousand workers is associated with a -0.34% lower employment-to-population ratio and 0.5% lower wages. These estimates incorporate the countervailing effects of adjustments across areas and are distinct from a range of other potential mechanisms (Chinese import penetration, offshoring, other types of ICT capital). However, an important extra point here is that robots may be part of an ensemble of labour saving technologies that are being deployed in parallel but are not measured directly in the data.

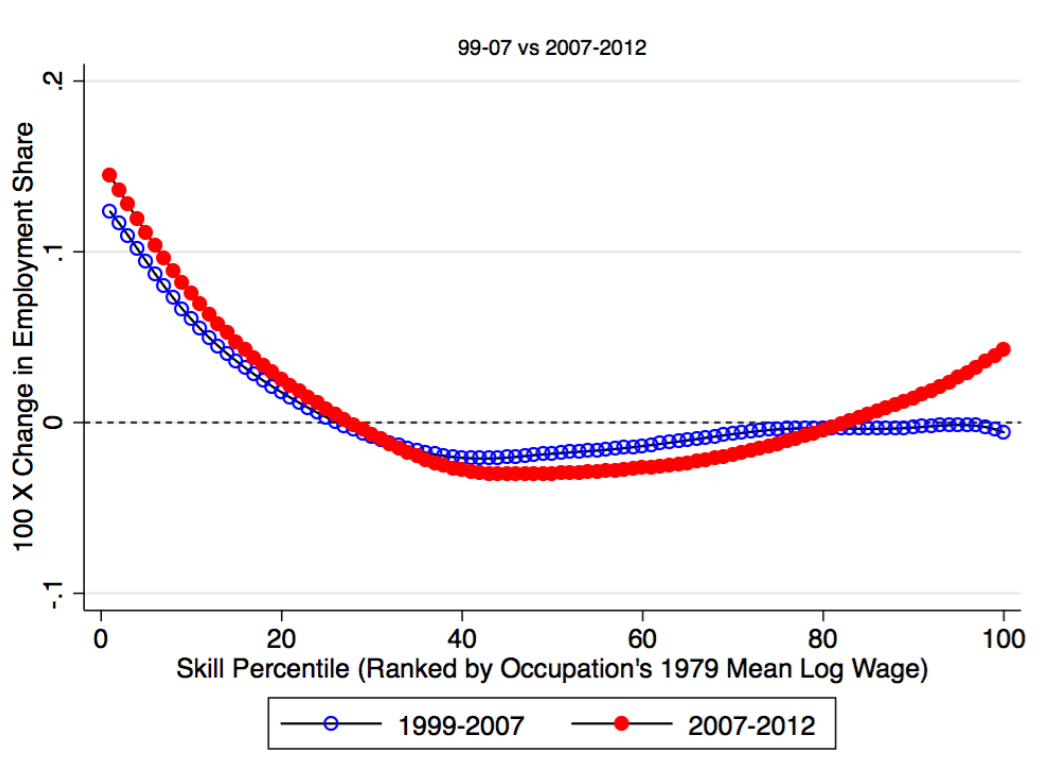
4.7. Tables and Figures

Figure 4.1: Employment Changes by Occupational Skill Percentile, 1980-1999



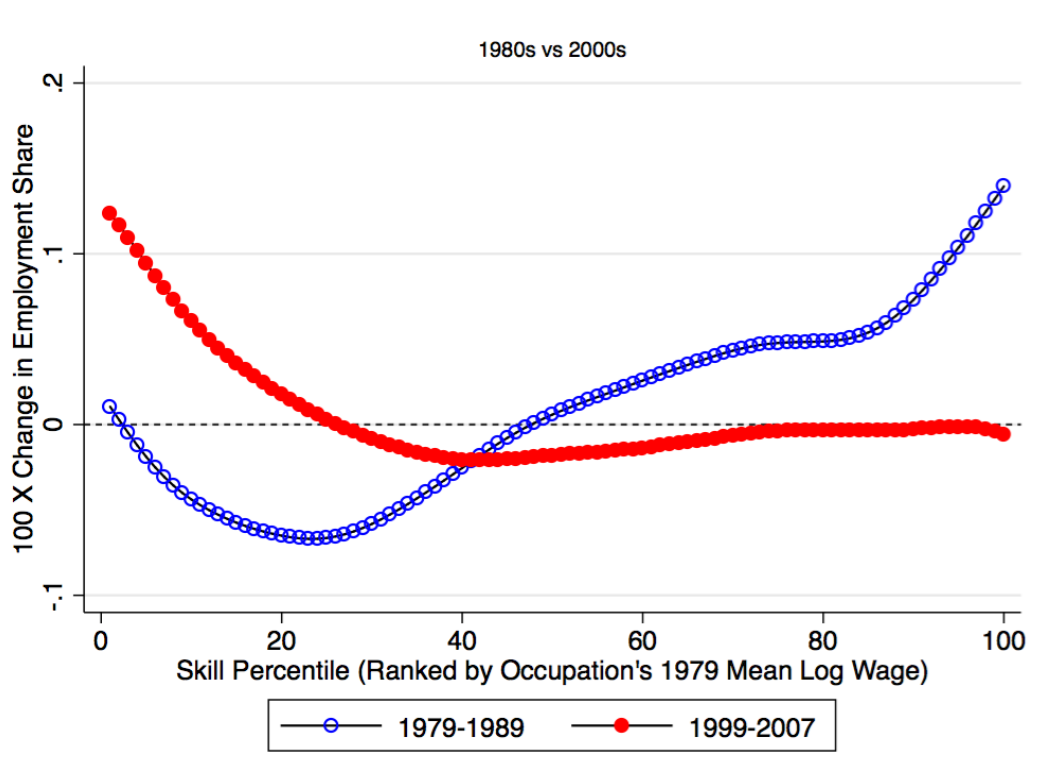
Notes: Data from Acemoglu and Autor (2011), based on US Census IPUMS data and American Community Survey (ACS). The horizontal axis represents the ranking of an occupation in the 1980 wage distribution.

Figure 4.2: Employment Changes by Occupational Skill Percentile, 1999-2012



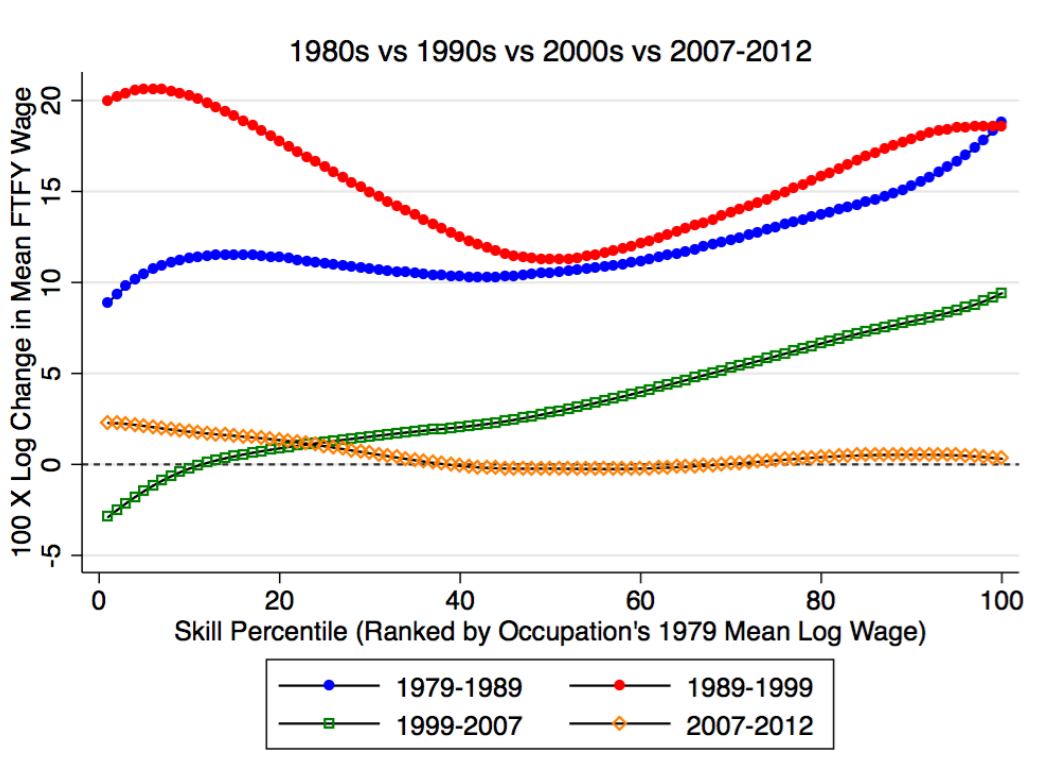
Notes: Data from Acemoglu and Autor (2011), based on US Census IPUMS data and American Community Survey (ACS). The horizontal axis represents the ranking of an occupation in the 1980 wage distribution.

Figure 4.3: Employment Changes by Occupational Skill Percentile, 1980-2012



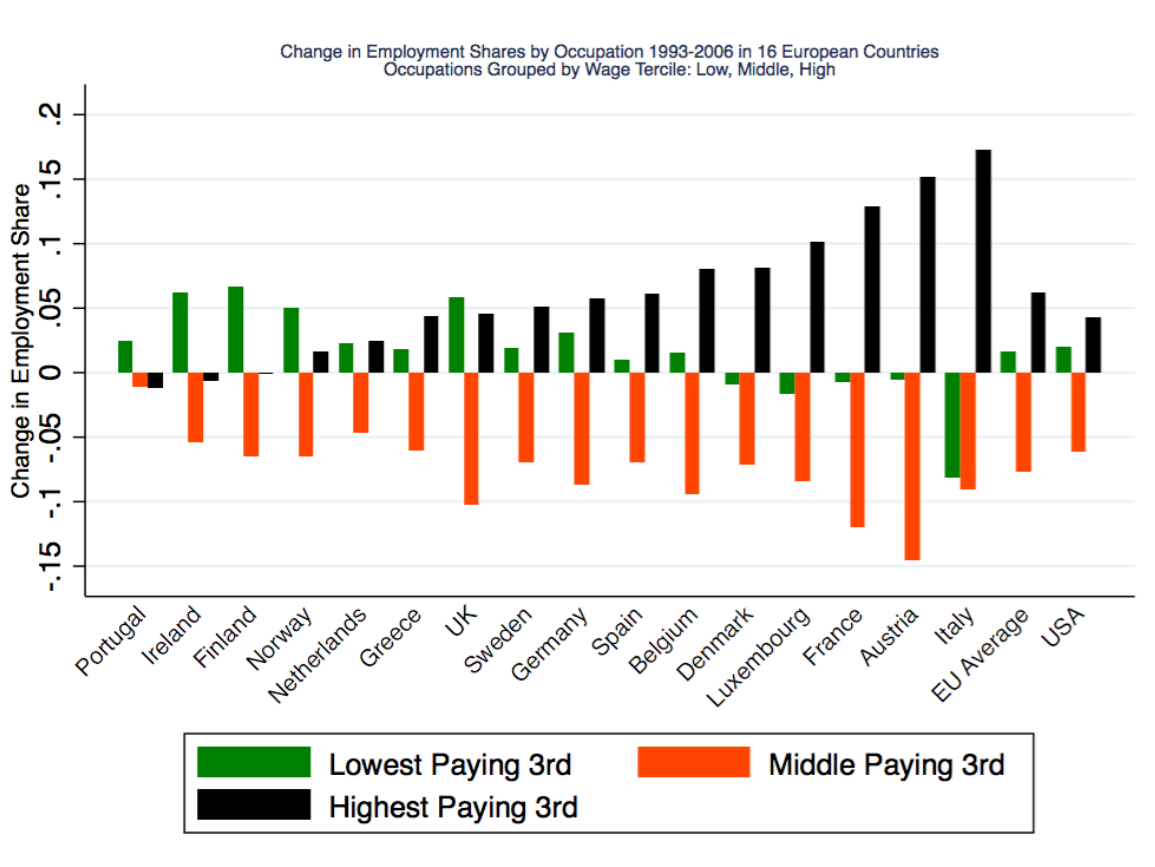
Notes: Data from Acemoglu and Autor (2011), based on US Census IPUMS data and American Community Survey (ACS). The horizontal axis represents the ranking of an occupation in the 1980 wage distribution.

Figure 4.4: Wages Changes by Occupational Skill Percentile, 1980-2012



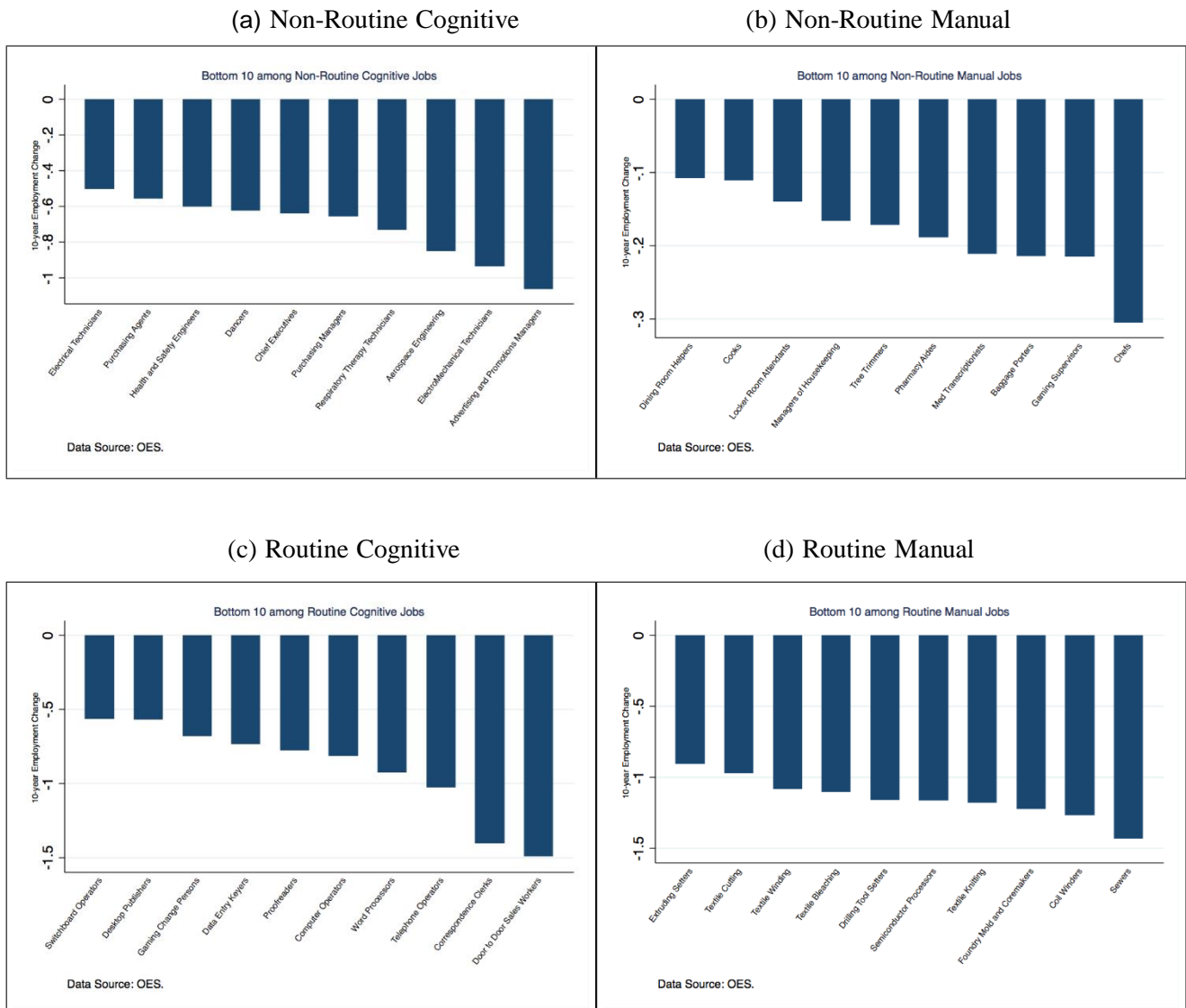
Notes: Data from Acemoglu and Autor (2011), based on US Census IPUMS data and American Community Survey (ACS). The horizontal axis represents the ranking of an occupation in the 1980 wage distribution.

Figure 4.5: Employment Changes by Occupational Group, Europe, 1993-2005



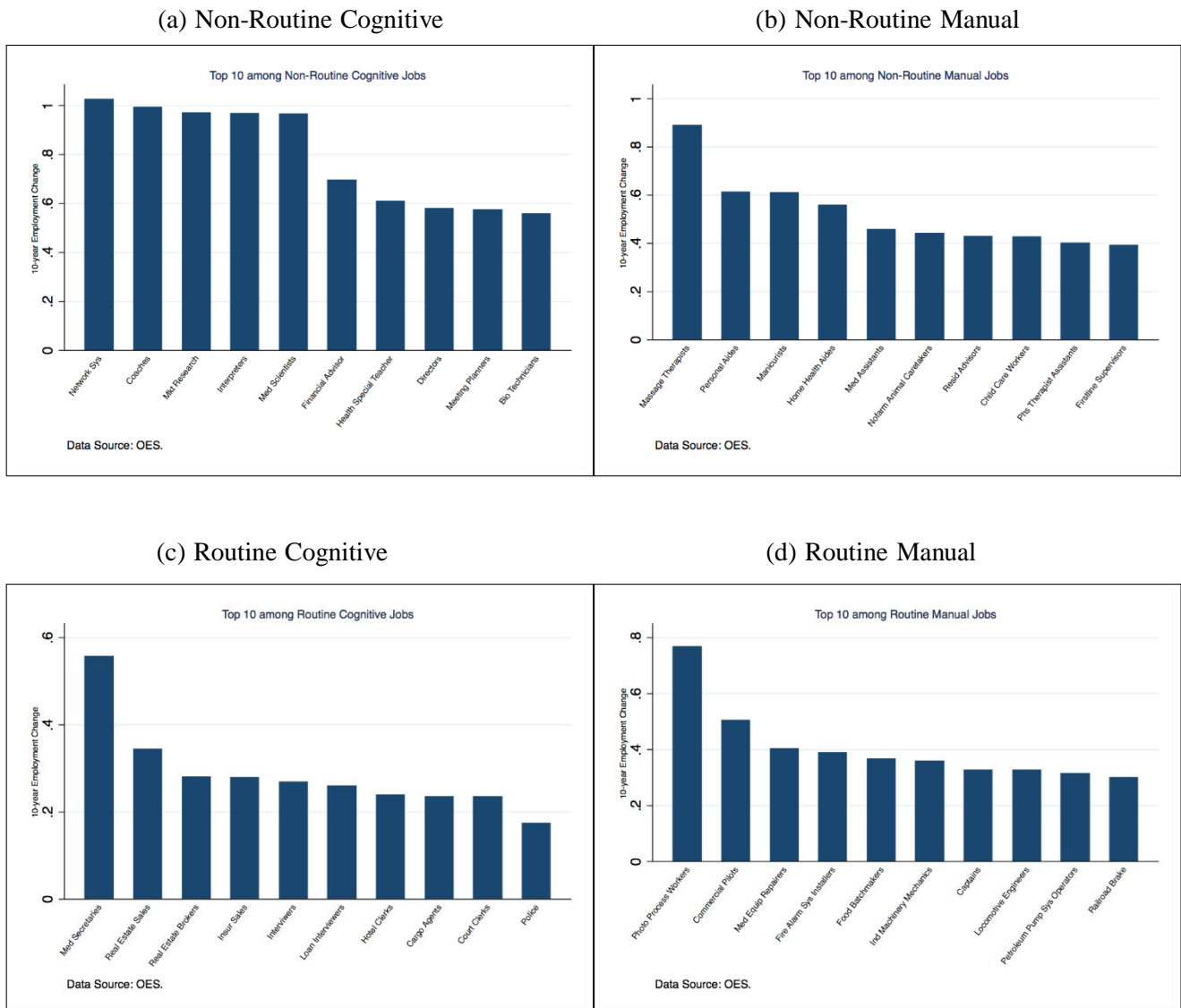
Notes: Data Goos, Manning and Salomons (2014).

Figure 4.6: Declining Occupations - OES, 2000-2016



Notes: These figures show the ten occupations with highest rate of employment decline over the 2000-2016 period in the OES panel. The occupational classification from Cortes et al. (2014) is used to define the occupational groups. The y-axis shows log employment change over the long difference. We condition on the set of all occupations that are at the 25th percentile or higher in employment share distribution in 2000.

Figure 4.7: Fastest Growing Occupations - OES, 2000-2016

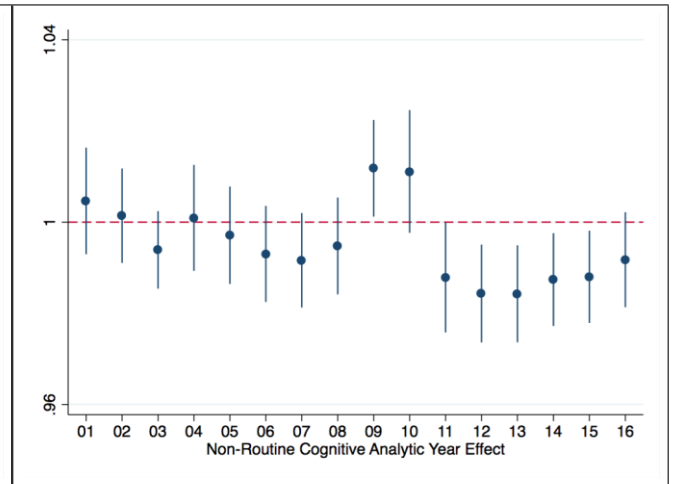
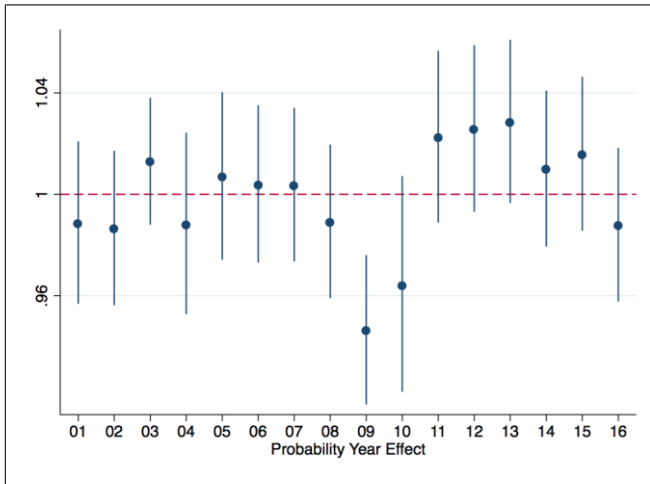


Notes: These figures show the Top Ten occupations with highest rate of employment growth over the 2000-2016 period in the OES panel. The occupational classification from Cortes et al. (2014) is used to define the occupational groups. The y-axis shows log employment change over the long difference. We condition on the set of all occupations that are at the 25th percentile or higher in employment share distribution in 2000.

Figure 4.8: Yearly Employment Effects - OES, 2000-2016

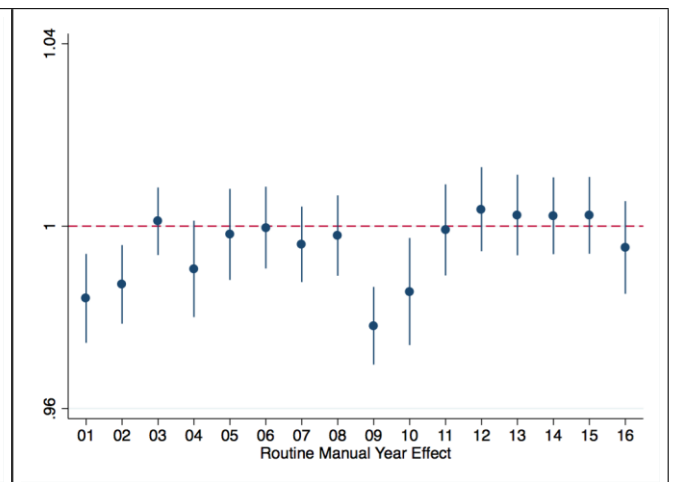
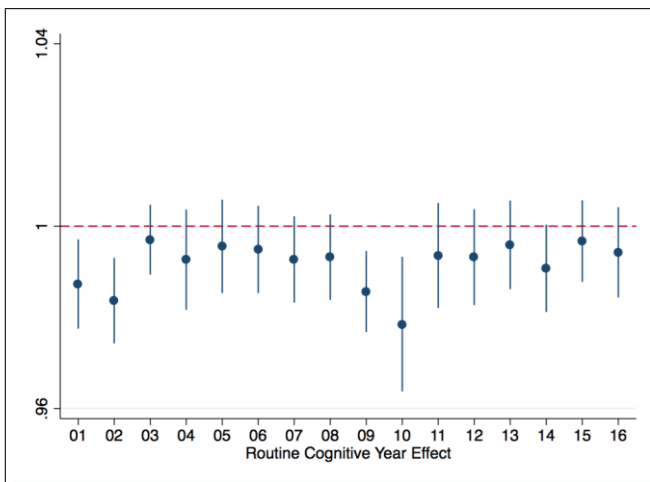
(a) Frey-Osborne Probability Measure

(b) NR Cognitive - Analytic



(c) Routine Cognitive

(d) Routine Manual



Notes: These figures shows the annual interaction effects from regressions of 1-year employment changes where we interact occupational task categories with the year effects. The Frey-Osborne measure is the ‘probability of automation’ variable reported by occupation in their data appendix. The regressions are run on the US Occupational Employment Statistics (OES) panel.

Table 4.1: Basic Task Model - Routine versus Non-routine Tasks

	Routine	Non-Routine
Analytic/Interactive	Record-keeping calculation Repetitive Customer Service	Forming/Testing Hypotheses Medical Diagnosis Legal Writing Persuading/Selling
Manual	Picking or sorting Repetitive assembly	Janitorial Services Truck Driving

Edited version of Autor, Levy and Murnane's (2003) task model description.

Table 4.2: Changes in US Job Finding Rates, 1976-2012

	(1) URM-ERM	(2) URC-ERC	(3) UNRM-ENRM	(4) UNRC-ENRC
A: 'Return' Job Finding Rates				
Pre-Polarisation Expansion	23.1	15.8	14.0	14.8
2009-2012 Expansion	13.9	7.9	11.3	11.7
Difference	-9.3	-7.9	-2.7	-3.1
	NRM-ER	RC-ER	RM-ENR	RC-ENR
B: 'Switching' Job Finding Rates				
Pre-Polarisation Expansion	9.7	10.4	4.4	6.5
2009-2012 Expansion	5.8	5.0	3.3	5.0
Difference	-4.0	-5.3	-1.0	-1.5
	NLF-ERM	NLF-ERC	NLF-ENRM	NLF-ENRC
C: Non-Participation to Employment				
Pre-Polarisation Expansion	1.2	1.6	1.6	1.0
2009-2012 Expansion	0.8	1.2	1.3	1.2
Difference	-0.4	-0.4	-0.3	0.2

Notes: Information reported from Cortes et al (2017), as based on analysis of Current Population Survey (CPS) data. The letter U stands for unemployed while E is employed, RM is routine manual, RC is routine cognitive, NRM is non-routine manual and NRC is non-routine cognitive. NLF represents Not in the Labour Force. Pre-polarisation period defined as two expansions in the 1976-1981 period.

Table 4.3: Estimates of US Alternative Work Arrangements, 2005-2015

	CPS-based Feb-95	CPS-based Feb-05	RAND Survey Weight(1)	RAND Survey Weight(2)
All	9.3	10.1	17.2	15.8
Independent Contractors	6.4	6.9	9.6	8.4
On-Call Workers	1.6	1.7	2.8	2.6
Temp Agencies	1.0	0.9	1.6	1.6
Contract Firms	0.5	0.6	3.3	3.1
Observations	61,752	63,437	2,194	2,194

Notes: Information reported from Katz and Krueger (2015). ‘Rand Survey’ represents data from a survey conducted by the authors to reflect BLS methodology from the 1995 CPS. The second weight in column (4) reduces the weight on self-employed respondents. The numbers reflect percentages of the total workforce in the named arrangements.

Table 4.4: Employment Changes 2000-2010, US OES Panel

	Baseline	Digit 1	Digit 2	Digit 3
(A) Net Measure				
Probability	-0.395*** (0.045)	-0.179*** (0.051)	-0.167*** (0.055)	-0.135** (0.063)
(B) Non-Routine High-Skill				
NRC Analytic	0.129*** (0.017)	0.079*** (0.023)	0.065*** (0.022)	0.057** (0.023)
NRC Interpersonal	0.100*** (0.015)	0.035** (0.016)	0.058*** (0.017)	0.037* (0.021)
(C) Non-Routine Low-Skill				
NRM Interpersonal	0.119*** (0.014)	0.052*** (0.016)	0.075*** (0.019)	0.063*** (0.020)
NRM Physical	-0.073*** (0.015)	0.035* (0.019)	-0.006 (0.024)	0.017 (0.028)
(D) Routine				
R Cognitive	-0.052*** (0.018)	-0.010 (0.017)	-0.009 (0.019)	-0.012 (0.019)
R Manual	-0.118*** (0.014)	-0.041** (0.021)	-0.053** (0.023)	-0.052** (0.024)
Number of Categories	n/a	5	22	89
Observation	636	636	636	636

Notes: Standard errors clustered by SOC occupation in parentheses. *, **, and *** represent 10 %, 5 % and 1% significance level, respectively. NRC refers to “Non-Routine Cognitive”; NRM refers to “Non-Routine Manual”; R refers to “Routine”. The Net Measure is the probability of automation as calculated by Frey-Osborne (2015). Data is a panel constructed from the US Occupational Employment Statistics (OES).

Table 4.5: Hourly Wage Changes 2000-2010, US OES Panel

	Baseline	Digit 1	Digit 2	Digit 3
(A) Net Measure				
Probability	-0.082*** (0.009)	-0.049*** (0.011)	-0.048*** (0.012)	-0.031* (0.016)
(B) Non-Routine High-Skill				
NRC Analytic	0.027*** (0.004)	0.011** (0.005)	0.016*** (0.005)	0.010* (0.006)
NRC Interpersonal	0.021*** (0.004)	0.011*** (0.004)	0.003 (0.004)	-0.002 (0.005)
(C) Non-Routine Low-Skill				
NRM Interpersonal	0.018*** (0.003)	0.012*** (0.004)	0.007 (0.005)	0.002 (0.006)
NRM Physical	-0.022*** (0.003)	-0.006 (0.004)	-0.008 (0.005)	-0.002 (0.006)
(D) Routine				
R Cognitive	-0.006* (0.004)	-0.001 (0.004)	0.000 (0.004)	0.004 (0.004)
R Manual	-0.020*** (0.003)	-0.006 (0.004)	-0.004 (0.004)	0.002 (0.004)
Number of Categories	0	5	22	89
Observation	615	615	615	615

Notes: Standard errors clustered by SOC occupation in parentheses. *, **, and *** represent 10 %, 5 % and 1% significance level, respectively. NRC refers to “Non- Routine Cognitive”; NRM refers to “Non-Routine Manual”; R refers to “Routine”. The Net Measure is the probability of automation as calculated by Frey-Osborne (2015). Data is a panel constructed from the US Occupational Employment Statistics (OES).

Table 4.6: Employment Changes 2010-2015

	Baseline	Digit 1	Digit 2	Digit 3
(A) Net Measure				
Probability	-0.044** (0.021)	-0.030 (0.026)	-0.023 (0.031)	0.009 (0.042)
(B) Non-Routine High-Skill				
NRC Analytic	0.001 (0.009)	-0.012 (0.011)	-0.010 (0.012)	-0.017 (0.015)
NRC Interpersonal	0.001 (0.009)	-0.005 (0.010)	-0.006 (0.012)	-0.014 (0.014)
(C) Non-Routine Low-Skill				
NRM Physical	0.001 (0.008)	0.015 (0.011)	0.002 (0.013)	-0.011 (0.015)
NRM Interpersonal	-0.002 (0.007)	-0.010 (0.008)	-0.009 (0.012)	-0.008 (0.015)
(D) Routine				
R Cognitive	-0.023*** (0.008)	-0.021*** (0.008)	-0.017* (0.009)	-0.007 (0.009)
R Manual	-0.014** (0.007)	-0.014 (0.010)	-0.023** (0.010)	-0.028** (0.012)
Number of Categories	0	5	22	89
Observation	668	668	668	668

Notes: Standard errors clustered by SOC occupation in parentheses. *, **, and *** represent 10 %, 5 % and 1% significance level, respectively. NRC refers to “Non- Routine Cognitive”; NRM refers to “Non-Routine Manual”; R refers to “Routine”. The Net Measure is the probability of automation as calculated by Frey-Osborne (2015). Data is a panel constructed from the US Occupational Employment Statistics (OES).

Table 4.7: Hourly Wage Changes 2010-2015

	Baseline	Digit 1	Digit 2	Digit 3
(A) Net Measure				
Probability	0.002 (0.006)	0.009 (0.007)	0.009 (0.008)	-0.001 (0.010)
(B) Non-Routine High-Skill				
NRC Analytic	0.004* (0.002)	0.004 (0.003)	0.006** (0.003)	0.007** (0.003)
NRC Interpersonal	0.000 (0.002)	-0.001 (0.002)	-0.001 (0.003)	0.001 (0.003)
(C) Non-Routine Low-Skill				
NRM Physical	-0.001 (0.002)	0.000 (0.002)	-0.002 (0.003)	-0.004 (0.004)
NRM Interpersonal	-0.001 (0.002)	-0.002 (0.002)	-0.003 (0.003)	0.000 (0.003)
(D) Routine				
R Cognitive	0.002 (0.002)	0.003 (0.002)	0.002 (0.002)	-0.002 (0.003)
R Manual	-0.001 (0.002)	-0.000 (0.002)	-0.002 (0.003)	-0.005 (0.003)
Number of Categories	n/a	5	22	89
Observation	651	651	651	651

Notes: Standard errors clustered by SOC occupation in parentheses. *, **, and *** represent 10 %, 5 % and 1% significance level, respectively. NRC refers to “Non-Routine Cognitive”; NRM refers to “Non-Routine Manual”; R refers to “Routine”. The Net Measure is the probability of automation as calculated by Frey-Osborne (2015). Data is a panel constructed from the US Occupational Employment Statistics (OES).

5. Bibliography

- Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, Volume 4, pages 1043–1171. Elsevier.
- Acemoglu, D. and Restrepo, P. (2017). Robots and Jobs: Evidence from US Labor Markets. Working Paper.
- Acemoglu, D. and Restrepo, P. (2018). Artificial Intelligence, Automation and Work. Technical Report, National Bureau of Economic Research.
- Ahlfeldt, G., Koutroumpis, P., and Valletti, T. (2017). Speed 2.0: Evaluating Access to Universal Digital Highways. *Journal of the European Economic Association*, 15(3): 586–625.
- Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *Quarterly Journal of Economics*, 130(4): 1781–1824.
- Arntz, M., Gregory, T., and Zierahn, U. (2016). The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. *OECD Social, Employment, and Migration Working Papers*.
- Atasoy, H. (2013). The Effects of Broadband Internet Expansion on Labor Market Outcomes. *ILR Review*, 66(2): 315–345.
- Autor, D. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3): 3–30.
- Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4): 1279–1333.
- Bauernschuster, S., Falck, O., and Woessmann, L. (2014). Surfing alone? The Internet and Social Capital: Evidence from an Unforeseeable Technological Mistake. *Journal of Public Economics*, 117(C): 73–89.
- Bellou, A. (2015). The Impact of Internet Diffusion on Marriage Rates: Evidence from the Broadband Market. *Journal of Population Economics*, 28(2): 265–297.
- Bergeaud, A., Cette, G., and Lecat, R. (2016). Productivity Trends in Advanced Countries between 1890 and 2012. *Review of Income and Wealth*, 62(3): 420–444.
- Bertschek, I., Briglauer, W., Hüschelrath, K., Kauf, B., and Niebel, T. (2015). The Economic Impacts of Broadband Internet: A Survey. *Review of Network Economics*, 14(4): 201–227.
- Bertschek, I., Cerquera, D., and Klein, G. J. (2013). More Bits – More Bucks? Measuring the Impact of Broadband Internet on Firm Performance. *Information Economics and Policy*, 25(3): 190–203.
- Bertschek, I. and Niebel, T. (2016). Mobile and More Productive? Firm-Level Evidence on the Productivity Effects of Mobile Internet Use. *Telecommunications Policy*, 40(9): 888–898.
- Bessen, J. E. (2016). How Computer Automation Affects Occupations: Technology, Jobs, and Skills. Working Paper.

- Bhuller, M., Havnes, T., Leuven, E., and Mogstad, M. (2013). Broadband Internet: An Information Superhighway to Sex Crime? *Review of Economic Studies*, 8(4): 1237–1266.
- Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2017). Are Ideas Getting Harder to Find? Technical Report, National Bureau of Economic Research.
- Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 81(4): 1347–1393.
- Bowen, H. R. (1966). Report of the National Commission on Technology, Automation, and Economic Progress: Volume 1. *Washington: US Government Printing Office*.
- Bowles, J. (2014). The Computerisation of European Jobs – Who Will Win and Who Will Lose from the Impact of New Technology onto old Areas of Employment. *Bruegel blog*, 17.
- Bresnahan, T. F. and Trajtenberg, M. (1995). General Purpose Technologies ‘Engines of Growth’? *Journal of Econometrics*, 65(1): 83–108.
- Brynjolfsson, E. and McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. WW Norton & Company.
- Brynjolfsson, E. and Oh, J. (2012). The Attention Economy: Measuring the Value of Free Digital Services on the Internet. Thirty-third International Conference on Information Systems.
- Brynjolfsson, E., Rock, D., and Syverson, C. (2017). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. Technical Report, National Bureau of Economic Research.
- Byrne, D., Oliner, S., and Sichel, D. (2013). Is the Information Technology Revolution Over?
- Byrne, D. M. and Corrado, C. A. (2017). ICT Prices and ICT Services: What do they tell us about Productivity and Technology?
- Campante, F. R., Durante, R., and Sobbrío, F. (2013). Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation. NBER Working Paper Num. 19029.
- Cheshire, P., Hilber, C., Montebruno, P., and Sanchis-Guarner, R. (2017). Using Micro-Geography Data to Identify Town-Centre Space in Great Britain. SERC discussion papers, Spatial Economics Research Centre, LSE.
- Cockburn, I. M., Henderson, R., and Stern, S. (2018). *The Impact of Artificial Intelligence on Innovation. Economics of Artificial Intelligence*. University of Chicago Press.
- Cohen, P., Hahn, R., Hall, J., Levitt, S., and Metcalfe, R. (2016). Using Big Data to Estimate Consumer Surplus: The Case of Uber. Technical report, National Bureau of Economic Research.
- Commission, LSE. Growth. (2013). Investing for Prosperity: Skills, Infrastructure and Innovation. Centre for Economic Performance, London School of Economics and Political Science.
- Cortes, G. M. (2016). Where Have the Middle-Wage Workers Gone? A Study of Polarization Using

- Panel Data. *Journal of Labor Economics*, 34(1): 63–105.
- Cortes, G. M., Jaimovich, N., Nekarda, C. J., and Siu, H. E. (2014). The micro and macro of disappearing routine jobs: A flows approach. Technical Report, National Bureau of Economic Research.
- Costa, R., Machin, S., et al. (2017). Real Wages and Living Standards in the UK. *CEP Election Analysis*, (36).
- CTO (2012). The Socio-Economic Impact of Broadband in Sub-Saharan Africa: The Satellite Advantage. Commonwealth Telecommunications Organisation.
- Autor, D. and Dorn, D. (2013). The Growth of Low-skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5): 1553–97.
- David, P. A. (1990). The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox. *The American Economic Review*, 80(2): 355–361.
- DCMS (2013). UK Broadband Impact Study Impact Report. Technical Report, SQW for UK Department for Digital, Culture, Media and Sport.
- Dechezleprêtre, A., Martin, R., and Mohnen, M. (2014). Knowledge Spillovers from Clean and Dirty Technologies. Working Paper.
- Dutz, M. A., Orszag, J. M., and Willig, R. D. (2012). The Liftoff of Consumer Benefits from the Broadband Revolution. *Review of Network Economics*, 11(4).
- EIU (2013). Redefining the Digital Divide. Technical Report, Economist Intelligence Unit.
- EIU (2014). The Future of Broadband in South-East Asia. Technical Report, Economist Intelligence Unit.
- Faber, B., Sanchis-Guarner, R., and Weinhardt, F. (2015). ICT and Education: Evidence from Student Home Addresses. NBER Working Papers 21306, National Bureau of Economic Research.
- Falck, O., Gold, R., and Heblich, S. (2014). E-Lectons: Voting Behavior and the Internet. *American Economic Review*, 107(7): 2238–2265.
- Farrell, D. and Greig, F. (2016). Paychecks, Paydays, and the Online Platform Economy: Big Data on Income Volatility. *JP Morgan Chase Institute*.
- Forman, C., Goldfarb, A., and Greenstein, S. (2012). The Internet and Local Wages: A Puzzle. *American Economic Review*, 102(1): 556–75.
- Frey, C. B. and Osborne, M. A. (2017). The Future of Employment: How Susceptible Are Jobs to Computerisation? *Technological Forecasting and Social Change*, 114: 254–280.
- Gathmann, C. and Schönberg, U. (2010). How General is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, 28(1): 1–49.
- Gavazza, A., Nardotto, M., and Valletti, T. M. (2017). Internet and Politics: Evidence from UK Local Elections and Local Government Policies. CESifo Working Paper Series 6659, CESifo Group

Munich.

Goolsbee, A. and Klenow, P. J. (2006). Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet. *American Economic Review*, 96(2): 108–113.

Goos, M., Manning, A., and Salomons, A. (2009). Job Polarization in Europe. *American Economic Review*, 99(2): 58–63.

Gordon, R. J. (2012). Is US Economic Growth Over? Faltering Innovation Confronts the Six Headwinds. Technical Report, National Bureau of Economic Research.

Gordon, R. J. (2015). Secular Stagnation: A Supply-Side View. *American Economic Review*, 105(5): 54–59.

Graetz, G. and Michaels, G. (2015). Robots at work. *Review of Economics and Statistics*. Forthcoming.

Greenstein, S. and McDevitt, R. C. (2011). The Broadband Bonus: Estimating Broadband Internet's Economic Value. *Telecommunications Policy*, 35(7): 617–632.

Griliches, Z. (1981). Market Value, R&D, and Patents. *Economics letters*, 7(2): 183–187.

Griliches, Z. (1991). The Search for R&D Spillovers. Technical report, National Bureau of Economic Research.

Hall, J. V. and Krueger, A. B. (2015). An Analysis of the Labor Market for Uber's Driver-Partners in the United States. *ILR Review*.

Helpman, E. (1998). *General Purpose Technologies and Economic Growth*. MIT press.

Hilbert, M. (2016). The Bad News is that the Digital Access Divide is here to Stay: Domestically Installed Bandwidths among 172 Countries for 1986–2014. *Telecommunications Policy*, 40(6): 567–581.

Katz, L. F. and Krueger, A. B. (2016). The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015. Technical Report, National Bureau of Economic Research.

NESTA (2015). Exploring the Costs and Benefits of Fibre to the Home (FTTH) in the UK. Technical Report, NESTA Innovation Foundation.

Nevo, A., Turner, J. L., and Williams, J. W. (2016). Usage-Based Pricing and Demand for Residential Broadband. *Econometrica*, 84(2): 411–443.

OECD (2008). Broadband and the Economy. OECD Digital Economy Papers 146, Organisation for Economic Cooperation and Development, OECD Publishing, Paris.

OECD (2014). OECD Insights: International Cables, Gateways, Backhaul and International Exchange. Organisation for Economic Cooperation and Development.

OFCOM (2011). International Communications Market Report 2011. Report.

- OFCOM (2013). The Availability of Communications Services in the UK. Technical Report 146, OfCom UK.
- Pajarinen, M., Rouvinen, P., Ekeland, A., et al. (2015). Computerization threatens one-third of Finnish and Norwegian employment. *ETLA Brief*, 34: 1–8.
- Poletaev, M. and Robinson, C. (2008). Human capital specificity: evidence from the dictionary of occupational titles and displaced worker surveys, 1984–2000. *Journal of Labor Economics*, 26(3): 387–420.
- Qiang, C. Z.-W., Rossotto, C. M., and Kimura, K. (2009). Economic Impacts of Broadband. *Information and Communications for Development 2009: Extending Reach and Increasing Impact*, 3:35–50.
- Rosston, G. L., Savage, S. J., and Waldman, D. M. (2010). Household Demand for Broadband Internet in 2010. *The BE Journal of Economic Analysis & Policy*, 10(1).
- Segal, B. (2015). A Short History of Internet Protocols at CERN. Mimeo.
- Stefano, T. D., Kneller, R., and Timmis, J. (2014). The (Fuzzy) Digital Divide: The Effect of Broadband Internet Use on UK Firm Performance. Discussion Papers 14/06, University of Nottingham, School of Economics.
- Syverson, C. (2017). Challenges to Mismeasurement Explanations for the US Productivity Slowdown. *Journal of Economic Perspectives*, 31(2): 165–86.
- TC (2012). Submarine Telecoms Industry Report. Terabit Consulting.
- Van Ark, B. and Jäger, K. (2017). Recent trends in Europe’s output and productivity growth performance at the sector level, 2002-2015. *International Productivity Monitor*, (33):8–23.
- World Bank (2014). Broadband Networks in the Middle East and North Africa: Accelerating High-Speed Internet Access. Technical report.

Appendices

Appendix A

A.1 Calculating Patent Rank

We start with the value of an individual innovation V_i . This will be a combination of the private value of the innovation PV_i and the external value EV_i that it generates. The external value is a fraction of the (social) values of all innovations that are inspired by i ; i.e.

$$V_i = PV_i + \sum_j \phi_{Sij} V_j \quad (\text{A.1})$$

ϕ_{Sij} is the fraction of innovation areas j 's value that can be attributed to area i . We will explore various assumptions for j . Initially we assume that

$$\phi_{Sij} = \sigma \times \frac{1}{c_j}$$

where c_j is the number of (backward) citations innovation j is making and σ is a scale factor which we can interpret as the marginal productivity of the spillover stock.¹ We set it to 0.2 on the basis of results in Aghion et al (2014). We can motivate dividing by c_j by the assumption that i has more influence on j if it is only one of very few citations made by j . Equation A.1 represents a system of linear equations that we can solve for the individual values V_i in principle. In practice this will be difficult by brute force as we are dealing with millions of innovations. However, it is easy to show that as long as the ϕ_{Sij} are smaller than one, we can solve this recursively, starting with $V_i(1) = PV_i$.

How do we know PV_i ? We start with some back of the envelope figures derived from stock market value regressions of innovating firms as discussed in Appendix B; i.e. the private value of an innovation/patent is derived from the stock market response to a successful innovation. Initially, we assume a uniform average value. However, there is clearly room to develop more sophisticated and specific measures of innovation value. From this we get that the average value of an innovation

¹ This can be made equivalent to page rank by setting $PV_i = \frac{1}{\text{Number of Patents}} \alpha$ and $\sigma = 1 - \alpha$. V_i would then

become the probability of that a random patent document surfer following up citations links with probability $1 - \alpha$ and with probability α randomly picks an innovation would come across innovation is \$0.25 million.

A.1.1 From Patent Rank to Promising Technology Areas

How can the patent rank help us figure out which technology areas might be more promising than others for government intervention? We will examine the extent of external value generated in different technology areas by looking at the average external value of an area

$$\frac{1}{N(A_r)} \sum_{i \in A_r} (EV_i)$$

where A_r is the set of innovations in technology area r and $EV_i = V_i - PV_i$. To understand which technology area is more promising for a given country or region we repeat the calculation suggested by equation A.1 by only taking into account private values that occur in a specific area; i.e.

$$PV_{ic} = PV_i \times I[i \in A_c]$$

where A_c is the set of innovations that are undertaken by innovators in country c .

A.2 Computing the Private Value of an Innovation

What is the private value of an innovation and are there differences between technologies according to their scientific area? We build on the framework first proposed by Griliches (1981), suggesting to derive the value of an innovation on the basis of changes to a firm's stock market value upon getting or applying for a patent. According to this we describe a firm stock market value at time as

$$V_{it} = m_{it} \times (A_{it} + K_{it})$$

where A_{it} are the firm's conventional assets, K_{it} are intangible assets from innovative activities and v_{it} is the market's valuation. We decompose m_{it} as

$$m_{it} = \exp(\alpha_i + \alpha_t + \epsilon_{it})$$

i.e. a firm and time fixed effects and an idiosyncratic shock. K_{it} is the composite of various knowledge assets, which we derive by perpetual inventory from the flow variables. In our benchmark regression we consider two: R&D investments and patent applications. We consider extended specifications where we split up patent applications by technology type (clean and dirty). We also consider backward citations for different technology types separately. Our basic regression equation becomes

$$\ln Q_{it} = \ln \left[1 + \beta_{RD} \frac{RDK_{it}}{A_{it}} + \beta_{PAT} \frac{PATK_{it}}{A_{it}} \right] + \alpha_i + \alpha_t + \epsilon_{it} \quad (\text{A.2})$$

where RDK_{it} is the knowledge stock from doing R&D whereas $PATK_{it}$ is the knowledge stock from being successful - as measured by patenting - at innovating. We compute all knowledge stocks by applying a perpetual inventory method to the respective flows; e.g.

$$RDK_{it} = RD_{it} + (1 - \delta) RDK_{it-1}$$

where RD_{it} is R&D spending in year t and δ is a geometric depreciation rate.²

We implement A.2 by differencing (to get rid of firm fixed effects α_i) and nonlinear least squares. In extended specifications we distinguish between different technology areas of patenting separately. We also consider that knowledge stocks might develop heterogeneously depending on where innovations draw their inspiration from. For that we include terms that take into account how many backward citations a firm makes.

² We experiment with various rates. In our main results we set this to 20%.

Appendix B

B.1 Modelling Framework

In the following we outline the background of the empirical modelling conducted in the body of the report. Note here that in order to head off the proliferation of notation we keep to a generic notation across different models and use stylisations such as β to represent parameters for the main variables of interest in an equation, X_{it} for vectors of controls, and E_{it} for error terms. In addition, we will also move generically between

B.1.1 Modelling the Diffusion of Internet Technologies.

To begin, we make a distinction between ‘adoption’ and ‘diffusion’. We define adoption as the individual act of acquiring a technology at the firm or household level. This individual act will depend on the incentives faced by the decision-maker, their resources and their environmental constraints (eg: in our case the availability of ADSL technology via an activated exchange is such a constraint). In contrast, diffusion describes the spread of an innovation across a group of firms or households.

The specific form of diffusion that we consider is the roll-out of ADSL across areas by BT. We frame this roll-out process as ‘diffusion’ because the switching on process was a prerequisite for individual adoption of better internet technologies across areas. Hence, the roll-out decision governed the pattern of group-level diffusion.

Our approach for modelling the spread of ADSL technology involves two types of empirical models. The first model looks at the timing of the roll-out in area i relative to baseline area characteristics:

$$\ln(RDays_i) = \alpha + X'_{i(t=2001)}\theta + \gamma_j + \epsilon_i \quad (\text{B.1})$$

where $\ln(RDays_i)$ represents days until ADSL is rolled out, and $X_{i(t=2001)}$ is a set of area characteristics measured in the baseline period of 2001, and γ_j represents region fixed effects. The constant is α and the error term is E_{it} . Within the data that we consider the area unit i is the

catchment area around the postcode of BT's telephone exchange - approximately 3,900 areas in total.

Hence the conceptual idea underpinning this first model is to study how the different characteristics of areas at the baseline point are correlated with the roll-out date. That is, what factors (statistically) shaped the path of BT's roll-out? Practically, the main exercise that we run is to look at the estimated coefficients on different covariates in the X vector as well as test out omitted variable relationships when we add and subtract covariates from the full regression. For example, this latter 'adding and subtracting' approach lets us examine how much of the roll-out variation is soaked up key variables such as population density.

The second empirical model we use relates to the analysis of internet technologies at the LAD-level, which defines 326 areas across England. We do this because a broader range of information is available at the LAD-level. In this case, in order to measure the presence of ADSL we aggregate up from the Catchment Area level to LAD, weighting the Catchment Areas by population.

This results in a population weighted measure of ADSL penetration at the LAD level. The measure of ADSL penetration evolves on a year-to-year basis so we run a panel regression over the full period of the ADSL diffusion from 2000-2007. We define the following regression model where i now indexes LAD-level units rather than catchment areas:

$$PRate_{it} = \alpha + X'_{i(t=2001)}\theta + \tau_{jt} + \epsilon_{it} \quad (\text{B.2})$$

where $PRate_{it}$ is the ADSL penetration rate in area i at time t and, as before, a set of area characteristics measured in the baseline period of 2001¹, with α as the constant and ϵ_{it} as the error term. The special point of interest in this model is actually τ_{jt} , which represents region-year fixed effects. The inclusion of these region-year effects lets us test how the spread of ADSL across LADs in the same region is related to the various X_t -vector characteristics. As a specific example, with the τ_{jt} terms included we can test whether factors such as LAD-level average educational qualifications continue to matter once the general rise in ADSL penetration in a region-year cell is controlled for.

B.2 Modelling Effects on Outcomes

B.2.1 Short-Run Outcomes

We use a panel fixed effects approach to model the short-run relationship between ADSL penetration rates and wage outcomes. By 'short-run' we mean the correlation between annual levels of

¹ We use baseline characteristics measured specifically in 2001 since this is year that the UK Census was taken and because this data source offers the of flexibility in terms of variable scope and different levels of aggregation.

ADSL penetration and wages - the immediate effects that ADSL may have had as it diffused. The empirical model is as follows:

$$\ln(Wage_{it}) = \alpha_i + \beta ADSL_{i,t-1} + X'_{i,t-1} \theta + \tau_j t + \epsilon_{it} \quad (B.3)$$

where $\ln(Wage_{it})$ is a log wage measure for area i in time t , $ADSL_{i,t-1}$ is the lagged area-level ADSL penetration rate, X_{it} is a vector of covariates that can either be fixed or time-varying and $\tau_j t$ represents region-year fixed effects. In this case we have an area-specific constant α_i and the error term is ϵ_{it} . The simple logic of this approach is see if any effects of ADSL on wages prevail after controlling for the X variables and various fixed effects. We also take the approach of formulating this equation in terms of first differences as an alternative specification for tracking out potential effects in changes.

B.2.2 Long-Run Outcomes - Decennial Regressions, 1991-2011.

In order to study the impact of ADSL diffusion on economic outcomes we utilise the UK Census to provide information for the periods 'before' and 'after' ADSL. The evidence on the roll-out of ADSL indicates that the process was staggered over a window of approximately 7 years. In turn, this means that some areas received access to ADSL substantially earlier than other areas. This early access may have given these areas a strategic advantage in exploiting the economic benefits of ADSL.

We operationalise this idea by using the LAD-level penetration rate of ADSL in 2001 as an indicator of early roll-out. Using occupational shares as our main outcome of interest we can then formulate a long difference model define over the 10-year post-ADSL period between 2001 and 2011:

$$\Delta_{10}(Occ_{i(t=2011)}) = \alpha + \beta ADSL_{i(t=2001)} + X'_{i(t=2001)} \theta + \gamma_j + \epsilon_{it} \quad (B.4)$$

where $\Delta_{10}(Occ_{i(t=2011)})$ is the 10-year change in an occupation's share in total employment between 2001-2011, $ADSL_{2001_{i(t=2001)}}$ is the penetration rate of ADSL in 2001, γ_j represents region fixed effects, and $X_{i(t=2001)}$ is a set of LAD-level characteristics measured in the baseline period of 2001. Generically, α is the constant while ϵ_{it} is an error term.

This approach effectively tests whether there is a gradient of occupational change associated with the initial placement of ADSL, as measured by the parameter β . One modelling issue is that $ADSL_{it}$ could be proxying for other area characteristics that could be generating the growth of particular occupations, for example, levels of education or economic activity. We deal with this by including a vector $X_{i(t=2001)}$ of controls defined in the baseline year of 2001. Despite these controls, we cannot rule out that any growth trend associated $ADSL_{2001_{ii}}$ could still be the result of the remaining area-level unobservables. In particular, could be correlated with systematic

trends in growth or decline at the LAD level. Hence, as an extra measure we study the pattern of occupational growth over the previous decade:

$$\Delta_{10}(Occ_{i(t=2001)}) = \alpha + \beta ADSL_{i(t=2001)} + X'_{i(t=1991)}\theta + \gamma_j + \epsilon_{it} \quad (B.5)$$

The key difference with respect to our previous regression is that, firstly, we are studying the changes $\Delta_{10}(Occ_{i(t=2001)})$ that occur in the pre-ADSL decade as our dependent variable and using the baseline characteristics in 1991 $X_{i(t=1991)}$ as our area controls. The inclusion of $ADSL_{i(t=2001)}$ in the model can then be interpreted as a test of whether systematic economic trends existed in relation to ADSL before the actual roll-out. The framework we outline here can also be re-formulated as a difference-in-difference model that tests for a trend shift over two periods and we comment on this in the discussion of results.

To conclude, we note two modifications of this basic framework also canvassed in the results. Firstly, in the testing of wage effects we have access (via the ASHE dataset) to annual measures over the 2001-2011 period. This lets us define both long difference measures of wages (e.g.: a 10-year difference $\Delta_1 W_{it}$) and ‘smoothed’ short differences. These smoothed short differences are, for example, the average one-year $\Delta_1 W_{it}$ changes in wages over the 10-year period. The premise of this short difference approach is that there may be information in the adjustment process during the 2001-2011 period that is lost in the long difference specification.

Secondly, following (Forman et al., 2012) we can interact the ADSL measure $ADSL_{i(t=2001)}$ with some of our baseline area characteristics $X_{i(t=2001)}$. These interactions capture the idea that some type of areas may have had the capacity to benefit more from the roll-out of internet technology relative to others. We comment further on these heterogeneous effects in the discussion of results in the body of the report.

Appendix C

C.1 Data Appendix

Broadband and rollout patterns data

The broadband data used in this report mostly comes from data collected for the paper [Faber et al. \(2015\)](#). Telephone exchange - postcode mapping and telephone exchange date of upgrade information was collected from SamKnows (<https://samknows.com/>) and OFCOM. From the postcode level dataset we aggregated the information to the telephone exchange catchment area level (for the rollout analysis) and to the Local Authority Level (LAD) level (for the economic analysis).

For the determinants of rollout analysis we aggregated data from very small geographies (output areas, over 165,000 in England) to the catchment area level, weighting the variables appropriately when a given OA belonged to two distinct catchment areas. The information on the determinants of the rollout are obtained combining data on geographical factors (using data from different Ordnance Survey product which can be accessed here <https://www.ordnancesurvey.co.uk/opendatadownload/products.html>) and 2001 census characteristics (accessed via <https://census.ukdataservice.ac.uk/>). More information on the definition and data sources of some of this data can be found in [Cheshire et al. \(2017\)](#). Some of this data was also aggregated up to the LAD level to carry out the economic impact analysis.

Information on available broadband speeds at the postcode level and mobile broadband technologies at the local county areas were obtained from OFCOM (<https://www.ofcom.org.uk/research-and-data/data/map-data> and the Infrastructure and Connected Nations Reports).

Local Authority Level socio-economic data

The main source of data for the Local Authority Level socio-economic variables is NOMIS (www.nomisweb.co.uk), a source provided by the Office for National Statistics, which contains the most detailed and up-to-date UK labour market statistics from official sources. All data are collected at the Local Authority District (LAD) level defined prior to April 2015 and restricted to being in

England. This results in 326 LADs.

The data on earnings and hours worked come from the workplace analysis of the Annual Survey of Hours and Earnings and it covers the years from 1997 to 2017¹. This data set provides information about earnings of employees who are working in an area, who are on adult rates and whose pay for the survey pay-period was not affected by absence. ASHE is based on a 1% sample of employee jobs taken from HM Revenue & Customs PAYE records. This dataset does not cover the self-employed nor does it cover employees not paid during the reference period. It provides information on the overall annual, monthly and weekly pays as well as the average number of hours worked in each LAD, for all individuals, male workers, female workers, full-time workers and part-time workers. Note that all earnings are an estimate and, therefore, are rounded to the closest round number.

The occupations data come from the 1991, 2001 and 2011 Censuses which classify the workplace population in England by occupation². All usual residents aged 16 to 74 in employment in the area the week before the census are computed. The occupations used in the analysis are based on the minor groups classification which distinguishes between 9 different levels of occupation, from managers to elementary occupations. It is important to note that the 1991 figures are the result of a 10% sample whereas 2001 and 2011 cover the whole population. In the three censuses, the Standard Occupational Classification (SOC) codes are differently defined although they somewhat comparable when it comes to the minor groups³. More specifically, the 1991 Census uses the SOC90 classification, the 2001 Census uses the SOC2000 classification and the 2011 Census uses the SOC2010.

The data on gross disposable income is extracted from the Regional Accounts for the period from 1997 to 2015 and it is computed at current basic prices and in pounds⁴. Gross disposable household income (GDHI) is the amount of money that all of the individuals in the household sector have available for spending or saving after income distribution measures (for example, taxes, social contributions and benefits) have taken effect. GDHI is a concept which is seen to reflect the ‘material welfare’ of the household sector. It should be noted that regional GDHI estimates relate to totals for all individuals within the household sector for a LAD rather than to an average household or family unit. The household sector comprises all individuals in an economy, including people living in traditional households as well as those living in institutions such as retirement homes and prisons.

We use the Labour Force Survey (LFS) and the Annual Population Survey (APS) to construct

¹ For more information on ASHE visit <https://www.nomisweb.co.uk/articles/299.aspx>

² Visit <https://www.ons.gov.uk/census/2001censusandearlier> for more information on the 1991 and 2001 Censuses and <https://www.ons.gov.uk/census/2011census/2011ukcensuses> for more information on the 2011 Census.

³ For equivalences between SOC classifications of different years visit <https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsof>

⁴ Visit <https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome> for more information on the Regional Accounts Dataset.

the time varying control variables. The Labour Force Survey (LFS) is a study of the employment circumstances of the UK population⁵. It is the largest household study in the UK and provides the official measures of employment and unemployment along with other characteristics like education. The APS comprises key variables from the Labour Force Survey (LFS) and thus, can be used in combination with it⁶. The common topics between these two datasets include issues like education, employment, health and ethnicity. In NOMIS, the LFS is only available for the period 1999 to 2003 but, given the comparability between LFS and APS, the latter can be used to enlarge the period of analysis up to 2016. The variables extracted from these data sources are the LAD yearly averages of the unemployment rates, the employment rates, the activity rates and the percentage employed in each national vocational qualification (NVQ) level.

Finally, we also use data on the Annual ONS Population Estimates for the period 1991 to 2017⁷. In particular, these estimates provide information on the overall population in each Local Authority District. Moreover, it also provides an age break down of the population figures for young people — i.e. 0 to 15 years old, working age population — i.e. 16 to 64 years old and retiring age people — i.e. more than 64 years old.

⁵ Visit <https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/labourforcesurveylfs> for information on the labour force survey in the UK.

⁶ Go to <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/annualpopulationsurveyapsqmi> for more information on the Annual Population Survey.

⁷ For more information on the ONS population estimates visit <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates>