Benign Effects of Automation: New Evidence From Patent Texts

Katja Mann
University of Bonn
katja.mann@uni-bonn.de

Lukas Püttmann
University of Bonn
lukas.puettmann@uni-bonn.de

31 December 2017

Abstract. We provide a new measure of automation based on patents and study its employment effects. Classifying all U.S. patents granted between 1976 and 2014 as automation or non-automation patents, we document a rise in the share of automation patents from 25 percent to 67 percent. We link patents to the industries of their use and, through local industry structure, to commuting zones. According to our estimates, advances in national automation technology have a positive influence on employment in local labor markets. Manufacturing employment declines, but this is more than compensated by service sector job growth. Commuting zones with more people working in routine occupations fare worse.

Keywords: automation, employment, labor demand, innovation, patents, local labor markets
JEL Codes: J23, O34, R23, C81

1 Introduction

What is the effect of automation technology on employment? The answer to this question is not obvious: While machines may replace workers, new jobs could also be created. For example, if self-driving vehicles become widely used, taxi and truck drivers might lose their jobs. Other sectors such as retail could, however, experience employment growth through lower transport costs.

To identify the employment effects of automation, this paper introduces a new indicator of automation technology. The large literature addressing this question has so far relied

• *We would like to thank Michael Böhm, Simon Jäger, Jürgen von Hagen and seminar participants at the University of Bonn, the University of California at Berkeley, the Institute of Labor Economics (IZA) and the Collegio Carlo Alberto for comments. We thank Christian Bayer, Hans-Martin von Gaudecker and Ralph Lütticke for computational help. Both authors thank the Deutsche Forschungsgemeinschaft for financial support and Püttmann additionally acknowledges funding by the Wissenschaftsförderung der Sparkassen-Finanzgruppe e. V.*

on indirect proxies of automation, such as routine task input (Autor, Katz, and Kearney, 2008, Autor, Levy, and Murnane, 2003, Goos and Manning, 2007, Autor and Dorn, 2013), investment in computer capital (Beaudry, Doms, and Lewis, 2010; Michaels, Natraj, and van Reenen, 2014) or investment in robots (Graetz and Michaels, 2015; Acemoglu and Restrepo, 2017). Many of these papers find evidence for job polarization, but the smaller literature on aggregate employment changes reports more ambiguous results. This may be due to difficulties in measuring automation comprehensively.

Our proposed automation indicator relies on patent grant texts. Patents are a natural candidate for measuring technological progress and frequently serve as proxies of innovation. However, few studies examine the consequences of technological progress through patents. Also, while patent meta-data such as citation counts or the identity of innovators is used regularly (Hall, Jaffe, and Trajtenberg, 2001; Acemoglu, Akcigit, and Celik, 2014; Bell, Chetty, Jaravel, Petkova, and Reenen, 2017), the actual patent texts have not been in the focus so far. We classify patents as automation patents if their texts describe physical inventions (such as robots) or immaterial or conceptual inventions (such as software), which carry out a process independently of human interference.

We extract the texts of all 5 million U.S. patents granted between 1976 and 2014 and train a machine learning algorithm on a sample of 560 manually classified patents to sort patents into automation and non-automation innovations. As a result, we document a strong rise in both the absolute and the relative number of automation patents. As a share of total patents, automation patents have increased from 25 percent in 1976 to 67 percent in 2014. Applying a probabilistic matching that is based on Canadian patents, we link patents to the 956 4-digit SIC industries where they are likely to be used. In this way, we quantify trends in newly available technology at the industry level.

Next, we compare the indicator to established measures of automation. The number of automation patents is positively correlated across industries both with investment in computer capital and with robots shipments. More automation patents have been granted in industries with a larger share of employment in routine occupations in 1960, a result that is in line with the literature on routine-biased technological change. Also, industries with more automation patents were characterized by a rise in non-routine cognitive and non-routine interactive task input and a fall in routine cognitive and routine manual task input.

To estimate the labor-market effects of automation, we transfer our industry-level data to U.S. commuting zones through industry-county employment counts. Commuting zones approximate local labor markets as workers tend to look for jobs within commuting distance from where they live. We obtain a panel dataset of new automation technology across 722 commuting zones over 39 years. Up to the late 1980s, there was a higher density of automation in the Great Lakes region, but automation technology has become less geographically concentrated over time.

Our empirical analysis benefits from the fact that we examine *local* economic outcomes which are impacted by, but unlikely to affect, the innovation activity of industries at the *national* level. Our key assumption is that commuting zone-specific developments in the medium-run do not affect automation innovation in industries that operate there. This is plausible for the following reasons: First, we separate the industries where patents originate from where they are used. Second, many patents belong to foreigners and universities who respond to other incentives than local firms. And third, local industries are small in comparison to national aggregate industries. Our approach thus follows Bartik (1991).

Our main econometric analysis is a fixed effects panel regression for five-year periods. Interpreting the automation index as a flow measure of technology, we assess the relationship between the sum of automation and changes in employment. While we find a positive effect of automation on total employment, this is driven by job growth in the service sector, which compensates for a fall in manufacturing employment. This result is robust to adding a variety of other economic and demographic controls and to weighting patents by the number of citations they received. We also consider separately patents belonging to specific groups of assignees: universities and public research institutes, foreigners and governments. All three should be less responsive to US labor market trends than US companies. Our results hold in the regressions for the subgroups of patentees as well as in an instrumental variable regression. Lastly, we find that automation is associated with more job creation in commuting zones where the share of routine occupations is low.

All in all, our study thus shows automation to be more beneficial for employment than some of the previous literature (Autor et al., 2015; Acemoglu and Restrepo, 2017), which might be due to our broader definition of automation. Our results are in line with Gregory, Salomons, and Zierahn (2016), who show that the detrimental substitution effect of automation on routine jobs is more than compensated by a positive labor demand effect due to larger product demand.

In the final part of our paper, we apply our indicator to replicate two central papers (Autor and Dorn, 2013 and Autor, Dorn, and Hanson, 2015) that study the influence of automation on labor markets using the routine task share of jobs. First, we show that non-college employment rose in commuting zones where more automation patents could be used and where more people worked in routine occupations. Second, we find that automation leads to rises in employment levels even when controlling for Chinese import competition, which stands in contrast to Autor et al. (2015). We provide further evidence that employment increases were driven to a larger extent by flows into the labor force than by a fall in unemployment.

There are strengths and weaknesses to our approach to quantifying automation technology. Text classification is an inherently imprecise activity and we introduce further inaccuracies through probabilistic matchings of patents to industries and commuting zones. Also, we make assumptions on the usefulness of patents and the way they are implemented. On the upside, we have to impose fewer ex-ante assumption on the nature of advances in automation technology, compared to the literature using routine task shares or computer and robot investment. Our indicator allows us to closely track the technology frontier, translating newly granted patents into a fine-grained industry- or commuting zone-level dataset. With these caveats in mind, we consider our indicator a complement to previous measures of automation.

2 New automation index

This section introduces the new automation index. We start by arguing why patents are a suitable data source for measuring technological progress and then define automation. We show how we construct the indicator and how the classification algorithm works. Then, we explain how to link patents to industries in which they are likely to be used. The resulting indicator traces the technology frontier across 956 industries and 39 years and displays plausible co-movement with existing indicators of automation such as computer investment, the number of robots used in production and the share of routine tasks across industries.

2.1 Patents as indicators of technological progress

The purpose of patents is to encourage innovation and technological progress by offering a temporary monopoly on an invention. Once granted, no one can re-engineer, create or sell the same object or idea. In return, the text of the patent is made publicly available. The language in the patent text is technical and highly standardized. Applicants have an incentive to provide exact and correct information about their innovation to obtain full protection of their ideas. Professional patent examiners judge a patent's claims and make changes where appropriate. In return for disclosing the content of the innovation to the public, an intellectual property right is granted for 20 years. To be patentable, an innovation must be *novel*, *non-obvious* and *useful*. The description must further be exact and detailed enough to allow for replication and it must name the invention's most important application. All these characteristics make patents a valuable data source.

Researchers in economics have made frequent use of patents, often in the form of the database established by Hall et al. (2001). Griliches (1990) provides an extensive survey of various issues related to using patents in economics. However, patents are so far usually interpreted as proxies for innovative activity, not as increments of technological progress whose effects can be studied (for an overview of the more recent literature, see Nagaoka et al., 2010). This is related to the fact that existing research almost exclusively uses patents' metadata, such as the location or affiliation of a patentee or a patent's importance.[1]

Magerman, Looy, and Song (2010) note that there is almost no research which uses the actual texts of the patent document, although this has been recommended as early as Griliches (1990). An exception is Bessen and Hunt (2007), who identify software patents by searching patent texts for keywords. Our approach differs as we do not specify a priori which words to search for, but use a state-of-the-art text classification algorithm. Also, we apply the derived measure to study the effects of technology on the labor market, whereas the goal of Bessen and Hunt (2007) is to characterize firms that file software patents.

[1]Patent citations, in particular, are widely applied as indicators of the value of an invention, for example by Bell et al. (2017).

In other areas of economics, text search has become common, with Gentzkow and Shapiro (2010) and Baker, Bloom, and Davis (2016) being prominent examples of papers that use newspaper articles. However, patent texts hold several advantages for researchers over other document collections: The precise technical language with a high degree of standardization, the incentive to deliver correct information, the additional check through the patent examiners' review and the public access to patent grant texts make patents well suited for text search analysis.

Patent text analysis is common in the private sector for prior art and freedom-to-operate searches by firms and lawyers. However, none of these providers – to the best of our knowledge – offers a comparison of technological trends over time, which leads us to develop our own approach.

2.2 Patent data

We obtain all 5 million utility patent documents granted in the United States from 1976 to 2014 from Google.[2] While Europe, Japan and increasingly China are also important patent legislations, of the roughly 10.9 million patents effective ("in force") worldwide in 2014, the largest fraction (about one fourth) had been granted in the United States (WIPO, 2016). In addition, the most important innovations are usually patented in all major patent legislations. These properties make U.S. patents a good proxy for the technological frontier in the United States and beyond. Also, given that this paper studies the effect of automation in the United States, U.S. patents are an obvious candidate for how available technology changes.

We only consider utility patents, which account for around 90 percent of all patents. Utility patents are "issued for the invention of a new and useful process, machine, manufacture, or composition of matter, or a new and useful improvement thereof" (USPTO, 2015). Other patent types are design, plant and reissue patents and do not track technology that we aim to measure. According to the United States Patent and Trademark Office (USPTO), in the period 1976-2014, 83 percent of all patents granted were owned by firms – mostly large multinational corporations. 15 percent of patents were owned by individuals and less than 2 percent by the U.S. government. About half of all patents are granted to foreign applicants, a share that has increased over time. During the period of our analysis, IBM, Canon and Samsung were the corporations with the largest number of patents granted (USPTO, 2014).

The patent grant document includes the title, patent number, name of the inventor, date, citations of other patents, legal information, drawings, abstract and a detailed description, as well as information on the technology class of the invention. Every patent is assigned one or more technology classification numbers by the patent examiner which describes technological and functional characteristics of a patent and on which we base our link from patents to industries. We exclude chemical and pharmaceutical patents from our

[2] google.com/googlebooks/uspto-patents.html

classification.[3] The overwhelming majority of these patents do not meet our definition of an automation patent (14 out of 560 manually classified patents were automation patents from those sectors), but including these patents might distor our classification.

2.3 Definition of automation

We define an automation patent to describe a *device that carries out a process independently*.[4] This broad definition captures technologies such as software, a robot used in a production or the self-driving vehicle mentioned in the introduction. The "device" can be a physical machine, a combination of machines, an algorithm or a computer program. The process it automates may be a production process, but also anything else where an input is altered to generate an output. An important element of the definition is the notion of independence: It works without human intervention, except at the start or for supervision. We require the automation innovation to be a reasonably complete process, product or machine. In addition, we require it to have an at least remotely-recognizable application. This excludes inventions that are minor parts of an automation innovation and highly abstract patents with no obvious application. We make no difference between process and product innovations, so an automation patent could describe either. Table 1 displays some examples of automation and non-automation patents.

Table 1: Examples of automation and non-automation patents

Patent title	**Patent number**	**Automation patent?**
"Automatic taco machine"	5531156	Yes
"Color measuring method and device"	6362849	Yes
"Coinfusion apparatus"	8857476	Yes
"Hair dye applicator "	6357449	Yes
"Hand-held scanner having adjustable light path"	5552597	No
"Bicycle frame with device cavity"	7878521	No
"Process for making pyridinethione salts"	4323683	No
"Golf ball"	4173345	No

Note: Authors' classifications according to manual coding guidelines. Click on the patent number for the weblink to the patent document.

2.4 Classification of patents

Based on the definition above, all patents can be classified as either automation or non-automation patents. We use an automated approach. To train a classification algorithm, we need reliable and objective classifications on which we can base the comparison. To this end, we manually classify 560 randomly drawn patents according to rules laid out in

[3]Excluded USPC technology numbers: 127, 252, 423, 424, 435, 436, 502, 510-585, 800, 930, 987.

[4]This is a standard definition that can be found in encyclopedias. For example, the Encyclopedia Britannica defines automata as "any of various mechanical objects that are relatively self-operating after they have been set in motion" and adds that "the term automaton is also applied to a class of electromechanical devices—either theoretical or real—that transform information from one form into another on the basis of predetermined instructions or procedures" (Encyclopædia Britannica, 2015).

manual coding guidelines.[5] Baker et al. (2016) proceed similarly when they manually classify newspaper articles to check the performance of their dictionary-based classification. We aim to minimize coding mistakes and biases by providing a structured classification process, by classifying patents in random order and by reviewing every classification by a second person.

The language in patent texts might have changed over time. But patents from the 1970s read very similar to those from the 2000s and important technological classes such as computers and robots are developed and patented throughout the sample period. The technical nature of the documents and the fact that legal terms change more slowly than other language also makes it less likely that there are short-lived trends that could pose a problem for a classification based on specific terms.

From our sample of patents, we extract word stems, called *tokens*, with the Porter2 stemming algorithm. This shortens "automation", "automated", "automatically", "automatable" to "automat". Table 2 summarizes these tokens. A typical title contains about 5 tokens, a typical abstract about 36 and the rest of the patent (the "body") about 500 to 600.

Table 2: Tokens in 560 manually classified patents

Part	All tokens	Unique tokens	Mean	Median
Title	2796	1301	4.99	5
Abstract	20781	3971	37.11	36
Body	339366	31499	606.01	506.5

Source: USPTO, Google and own calculations.

In principle, one could now record for all 5 million patents whether they contain one of the roughly 32,000 tokens that we can assign probabilities to. But to keep the computation-intensive data collection feasible and to remove noise features, we use the *mutual information criterion* to extract those tokens which are most informative about which class a patent belongs to. This is an established statistic for feature selection which prefers tokens that appear significantly more often in one of the classes and punishes tokens that appear rarely overall (Manning et al., 2009). We then pick the highest ranked (according to the mutual information criterion) 50 title tokens, 200 abstract tokens and 500 patent body tokens. The final search dictionary consists of 623 tokens.

Figure 1 visualizes the 150 tokens with the highest mutual information criterion. The most important token is unsurprisingly "automat". After that come "output", "execut", "inform", "input" and "detect". Some tokens are indicative of software, such as "microprocessor", "database", "comput", "program" or "transmiss". Others are more likely to appear in descriptions of physical machines, such as "motor", "move", "metal" or "apparatus". The last discernible group of tokens are action verbs that appear in descriptions of a wide range of independently operating devices, such as "distinguish", "command", "respons" or "perform".

[5] See: `http://lukaspuettmann.com/assets/pdf/manual_coding_guidelines.pdf`

Figure 1: Words that indicate an automation patent

Note: Token size is proportional to the value of the mutual information criterion in sample 560 classified patents. We show only the 150 highest ranked tokens excluding chemical and pharmaceutical words.
Source: USPTO, Google and own calculations.

Our algorithm emulates how a human being would have classified each patent. We apply the Naive Bayes algorithm which is a supervised learning method which is easy to interpret and which computationally scales well with large amounts of data. The "naive" assumption the probability of a token to appear in a document is independent from the appearance of other tokens. Despite its simplicity it has been shown to perform quite well (Domingos and Pazzani, 1997).[6] One reason for this that the low number of parameters it estimates make it unlikely to overfit (Murphy, 2012).

Manning et al. (2009) explain how this algorithm picks the class c for every document d with maximum a posteriori probability $P(c \mid d)$. In our analysis, the documents d correspond to patent grant texts and the two different classes are automation patents and non-automation patents. In the *Bernoulli* Naive Bayes that we use, every document d is represented by a vector e, where entry e_i ($i = 1, \dots, M$) is 1 if token i appears at least once in the document and 0 if it does not. Patent texts contain matter-of-fact language, where words are often repeated. So the occurrence of a word is more important than the frequency of its appearance and we therefore ignore how often a word appears in a document.

According to this language model, in any document in class c the token e_i occurs with conditional probability $P(e_i \mid c)$. Therefore, the probability of a document d to show up in class c is

$$P(d \mid c) = \prod_{1 \le i \le M} P(e_i \mid c), \tag{1}$$

[6] Gentzkow et al. (2017) also recommend this algorithm if the number of observed features (tokens) is much larger than the size of the training sample, as is the case in our analysis. Antweiler and Frank (2004) proceed similarly, as they manually classify 1000 messages and then use the Naive Bayes algorithm to generalize to over 1.5 million other messages.

and the conditional probability of document d to belong to class c is according to Bayes' rule[7]

$$P(c \mid d) \propto P(c) \prod_{1 \leq i \leq M} P(e_i \mid c). \tag{2}$$

We estimate the prior $\widehat{P}(c)$ as the relative frequency of documents in class c in the training set. This is $\widehat{P}(\text{autom}) = \frac{147}{483} = 0.304$, as about a third of eligible patents (i.e., after removing chemical and pharmaceutical patents) were manually labeled as automation patents. We then estimate the conditional probabilities of a certain token to occur in class c, $\widehat{P}(e_i \mid c)$ as

$$\widehat{P}(e_i \mid c) = \widehat{P}(i \mid c)e_i + (1 - \widehat{P}(i \mid c))(1 - e_i), \tag{3}$$

where $\widehat{P}(i \mid c)$ is the share of documents with token i in class c. In this way, we calculate posterior probabilities for all 5 million patents to belong to either class and assign each patent to the class with the higher posterior probability.

Table 3: Contingency table

		Computerized: No	Computerized: Yes	
Manual	No	323	88	411
	Yes	25	124	149
		348	212	560

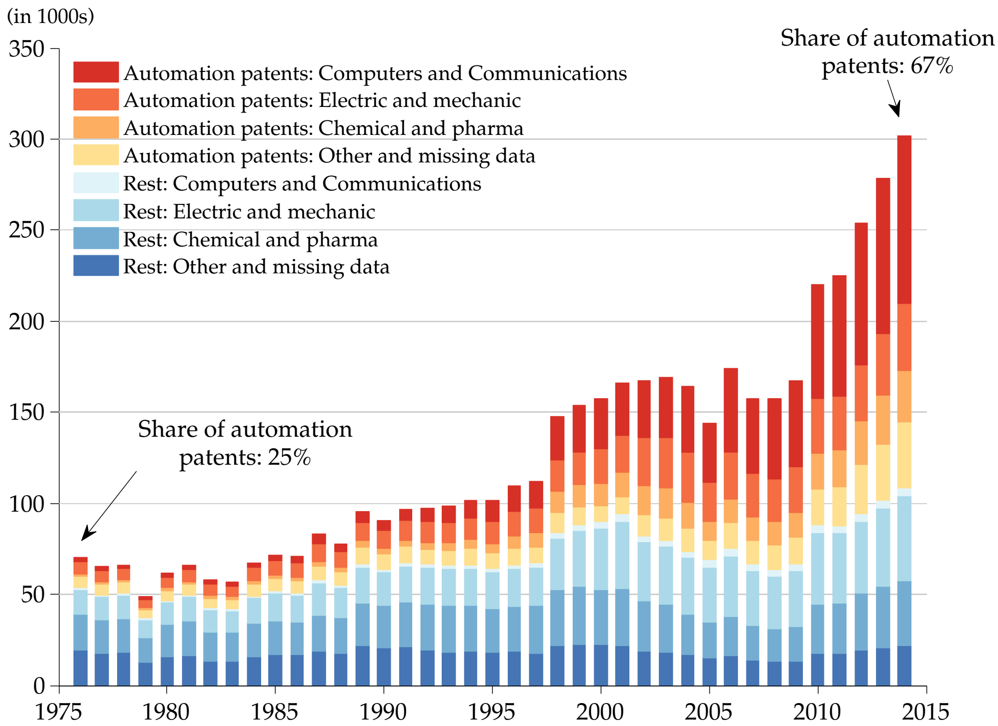
"No": not automation patent

Table 3 shows how human examiners and how the computer algorithm classified the set of manually investigated patents. Both the manual coding and the algorithmic classification judged around a quarter of patents to be automation patents. In 80 percent of cases ($= \frac{323+124}{560}$) both approaches agreed. The probability of a false positive (type I error) is 21 percent ($= \frac{88}{411}$). The probability of a false negative (type II error) is 17 percent ($= \frac{25}{149}$).

While some share of misclassified patents remains, as long as there is no underlying bias in the classication this should only add noise to our indicator series as we only aim to approximate trends in technology over time. Any noise should therefore push our empirical results towards zero, making it harder to detect an effect of automation.

A more precise classification might be possible when including patents' other observable characteristics such as their technological class (USPC and IPC numbers), grant years, the origins of inventors or the sector of firms. But we keep the classification into automation and non-automation separate from these observables to allow comparing automation trends across time and industries, without making these associations automatic.

Figure 2: Patents, 1976-2014



Note: See text for classification of automation patents and assignment of patents to categories.
Source: USPTO, Google, Hall, Jaffe, and Trajtenberg (2001) and own calculations.

2.5 Aggregate properties of the indicator

Figure 2 is a graphical representation of all 5 million patents granted in the United States between 1976 and 2014. We show patents by when they were granted, not when applied for, as inventions are unlikely to be shared before they are protected by a patent.

There has been a steady increase from 70,000 granted patents in 1976 to more than 300,000 patents in 2014. Over the whole period, we classify 2.2 million of these as automation patents. The red-shaded parts of the bars show the patents which we classified as automation patents and blue colors signal all other patents. We observe a sharp upward trend in automation patents from 16,000 in 1976 to 180,000 in 2014. The share of patents related to automation also increased, from 25 percent of patents in 1976 to 67 percent of patents in 2014. Table A1 in the Appendix provides the yearly numbers.

Figure 2 further shows broad categories of patents based on an aggregation method by Hall et al. (2001) which relies on the technological classification (USPC number) of patents.[8] Patents in the sub-category computers and communication have become much more frequent over the sample period and we mostly classify them as automation. Many of these are likely software patents. Electrical, electronic and mechanical patents also

[7] $P(c \mid d) = \frac{P(c)P(d|c)}{P(d)} \propto P(c)P(d \mid c)$.

[8] Note that this is a different classification than the one we will employ to match patents to the industries they are likely to be used in. See section 2.6.

contribute significantly to the stock of automation patents. Robots, for example, fall in this category. By design, most chemical and pharmaceutical patents are not classified as automation patents, but they make up a large portion of the non-automation patents.

The rise in the total number of patents granted is a potential concern for the interpretation of the time-dimension of patent texts. If the nature of patents had changed in parallel with the number, so if the increase in patents is due to something else than an increase in research productivity, the data might not be comparable across time. An increase in the number of automation patents would then not be interpretable as an increase in automation technology. Kortum and Lerner (1999) evaluate different possible explanations for why the number of patent grants has changed: increased patent protection due to patentee-friendly court rulings, regulatory capture by large firms that patent eagerly, new technology fields producing patentable inventions (e.g., information technology, biotechnology and financial intermediation) and more applied research. The authors refute all hypotheses except for the increase in research productivity. This result is in line with an OECD survey (OECD, 2004) in which 94 percent of surveyed firms responded that an increase in the number of inventions was an important or very important driver of their increased patenting activity (66 percent very important). In contrast, changes in patentability played only a minor role. We therefore conclude that the quality of patents granted has not changed over time and that we do not need to worry about any distortive effects of a change in grant numbers. As an additional check, we compute a deflated version of our indicator, for which we divide the number of automation patents in each industry and year by the total number of patents granted in that specific year relative to the number of patents granted in 1990. The resulting measure is an automation count in units of 1990 patents, which takes higher values for earlier years and lower values for later years than the original measure. Our empirical results in section 4 are insensitive to the time deflation.

2.6 From patents to industries

Various researchers have proposed matchings of patents to industries. Hall et al. (2001) identify firms filing for patents and Lybbert and Zolas (2014) propose an automated approach that compares descriptions of industries with descriptions of patents' technological classes. The OECD (2011) reviews these techniques in more detail and Griliches (1990) describes the difficulties in matching patents to industries.

However, we are interested in how automation technology affects labor markets. Therefore, we aim to find the industries where automation patents are *used*, not where they originate. These two need not be the same, so that the industry of the patentee is not necessarily the industry we want to assign the patent to. As an example, IBM owns many patents that are not used in the computer industry, but by companies in the manufacturing or in the retail sector. These patents are either sold or licensed out. Attributing them to the computer industry would overstate the automation intensity there, while understating it in the other sectors.

Table 4: Automation patents across industries of use

Industries	Manufacturing	Automation patents (1000s)	Share	SICs (1987)
Computers	✓	499	88%	357
Other electronics	✓	250	46%	36*
Measuring instruments; watches	✓	193	60%	38
Telephones and telegraphs	✓	185	68%	3661
Machines	✓	183	40%	35*
Hospitals		137	46%	8062
Househ. audio and video equip.	✓	104	69%	3651
Other services		118	47%	70-89*
Transportation equipment	✓	115	39%	37
Chemicals, rubber, plastics, oil	✓	101	18%	28, 30, 29
Utilities (transport, gas, sanitary)		57	44%	E
Fabricated metal products	✓	51	33%	34
Medical laboratories		37	64%	8071
Construction		34	24%	C
Printing publishing; paper	✓	34	32%	26, 27
Metal, stone, clay, glass, concrete	✓	29	22%	32, 33
Retail and wholesale trade		26	32%	G, F
Agriculture, forestry and fishing		24	33%	A
R&D, management	✓	23	64%	87
Miscellaneous manufacturing	✓	20	38%	39
Public administration; finance		20	47%	J, H
Food, tobacco	✓	19	24%	20, 21
Mining		16	37%	B
Apparel, wood, furniture	✓	15	17%	22-25, 31
total		2,290	46%	

Note: Patents are counted if they can be used in an industry, as described in text. Numbers are sums of patents 1976-2014. Shares are calculated by dividing automation patents by all patents in industry. An asterisk * indicates that some subindustries are shown separately.
Source: USPTO, Google, Silverman (2002) and own calculations.

Linking patents to the industries of their use is difficult. If we wanted to measure the *actual* usage of a specific patent in a certain industry, we would need data on out-licensing. But this information is not available, as firms and research institutions have incentives to keep their licensing agreements private. Interpreting patents more indirectly as a proxy for automation technology rather than a direct measure, we can use information about the areas in which patents can *potentially* be applied. There have been attempts by Schmookler (1966) and Scherer (1984) to manually classify patents and link them to industries of use, but this would not be feasible for a large number of patents. Patent offices themselves usually do not provide information on the link of patents to industries. However, we benefit from an exception to this rule by the Canadian patent office. Between 1978 and 1993, Canadian patent officers assigned industries of use for all granted patents. Based on this information, Kortum and Putnam (1997) assembled the "Yale Technology Concordance", a way to link patents through their technological classification to the industries in which they are likely to be used. This is based on the assumption that the pattern linking patents' technological class to industries of use should be similar in Canada and the United States. We use the files provided by Silverman (2002), who calculates empirical frequencies for cross-overs from patent technology classes (IPCs) to 1987 SIC industries using 148,000 patents granted between 1990 and 1993.[9]

This allows for a probabilistic matching. We connect a patent to an industry with the probability of being used in that industry. So if patent A is used in two industries X and Y, then half the patent count is assigned to industry X and half to Y. However, patents are often assigned several IPC technology classifications. In that case, we divide each value for that patent by the number of its IPCs. So if patent A now is assigned another IPC number, then only a quarter of its value will now be attributed to industries X and Y each and the rest to industries in the new IPC. This fractional counting of patents ensures that more general patents that are assigned to several IPCs do not have get more weight than more specialized patents that are assigned to fewer IPCs.[10]

As a result, we obtain an annual dataset of new patents and new automation patents that can be used in 956 industries and over 39 years. Table 4 displays all automation patents by industries of use over the whole time period 1976-2014. (The totals differ slightly from Appendix Table A1 due to rounding errors and the probabilistic conversion to patent equivalents as described before.)

Out of a total of 2.3 million automation patents, 1.8 million (79 percent) are used in the manufacturing sector (division D in SIC 1987). Half a million automation patents could be used in the production of computers (SIC 357) which includes personal computers, mainframes, storage devices, terminals, billing machines, automatic teller machines and peripheral equipment such as printers, scanners, office equipments or typewriters. The production of electronic devices, sensors and communication equipment also received a large number of automation patents. Outside of the manufacturing sector, hospitals, utilities and medical laboratories were assigned a large number of automation patents. In large parts of the economy – such as agriculture, mining, public administration, finance or retail – only few automation patents were granted. We also calculate the share of patents used in an industry that we classify as automation. This ratio is high for the computer industry or communication-related industries and is low for the chemical industry or "Apparel, wood, furniture".

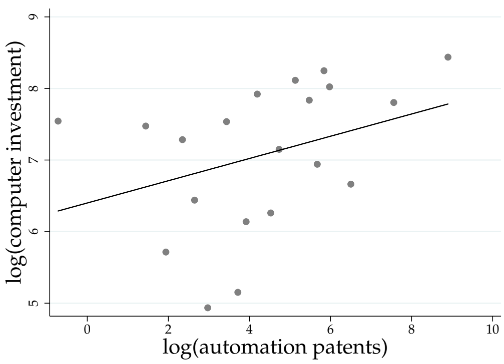
In our following empirical analysis, we interpret these indices as worker intensities by fully assigning all new (automation) patents in an industry to each person employed in that industry and year. This is equivalent to assuming that patents assigned to an industry will potentially be used by everyone working in that industry. If we considered our indicator narrowly as an exact measure of the use of patents in the production process, this would not be a realistic assumption. But to us, a patent is just one part of an innovation process that will produce many types of outputs. Being a measurable outcome of this process, patents serve as a proxy for it.In our regressions we will use the total number

[9] `http://www-2.rotman.utoronto.ca/~/silverman/ipcsic/documentation_IPC-SIC_concordance.htm`, accessed 25.10.2015. The fact that we use only data for 1990–1993 means that the matching should be most precise during this period, while becoming less exact the further away we move from this period. It helps that this periodis in the middle of our sample, but the fact that patents grow much more near in the later years is some cause for concern.

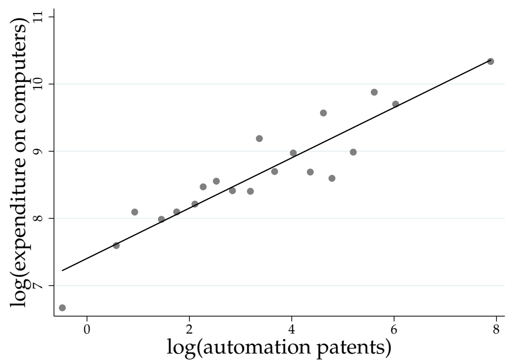
[10] This also enables us to interpret the resulting indicator as full patent equivalents which we will still refer to simply as "patents" in the remainder of the paper.

Figure 3: Comparison with other indicators of automation

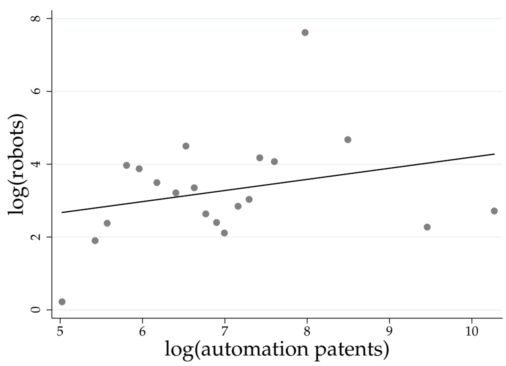
(a) Computer investment (NIPA), 1976-2001



(b) Computer investment (ASM), 2002-2014



(c) Robots (IFR), 2004-2014



Note: NIPA computer investment is the mean of 1976-2001 in millions of 1996 U.S. dollars, ASM computer investment is the mean of 2002-2014 in thousands of 2009 U.S. dollars. Robots is the mean number of robot shipments in the U.S. over 2003-2014 (U.S. data for 2003-2010 are imputed from North America data). Automation patents are counted for the same time period as the respective comparison data. All three figures show binscatters of log values.
Source: USPTO, Google, Silverman (2002), NIPA, ASM and IRF (2014).

of automation patents as our main explanatory variable, but we will also control for the amount of all other patents that can be used in an industry.

3 Comparison with previous automation proxies

Next, we analyze how our new industry measure of automation technology is related to established automation indicators. Previous proxies of automation differ from ours along two lines. First, they are indicators of realized automation in the production process, not indicators of automation technology. Second, most capture only one specific facet of automation technology, such as computers or robots, while our indicator incorporates both and even allows delineating it from other kinds of technological progress.

As a measure of computerization, studies use survey data of computer use at the workplace (Autor et al., 2003, Beaudry et al., 2010) or industry-level investment in computer capital (Autor et al., 2003, Michaels et al., 2014). Frey and Osborne (2017) manually assess the probability of computerization of a number of occupations. Akerman et al. (2015)

Table 5: Relationship between automation patents and other automation proxies

	contemp. automation	1st lag of automation	2nd lag of automation	3rd lag of automation
A. Time fixed effects				
computer invt (ASM)	0.391*** (0.0568)	0.393*** (0.0572)	0.395*** (0.0576)	0.398*** (0.0582)
computer invt (NIPA)	0.194** (0.0948)	0.193** (0.0953)	0.191* (0.0958)	0.189* (0.0963)
robot shipments (IRF)	0.0846 (0.308)	0.121 (0.323)	0.167 (0.342)	0.220 (0.365)
B. Time and industry fixed effects				
computer invt (ASM)	0.247** (0.125)	0.250** (0.125)	0.257** (0.124)	0.259** (0.128)
computer invt (NIPA)	0.336 (0.252)	0.322 (0.243)	0.272 (0.231)	0.250 (0.220)
robot shipments (IRF)	0.350*** (0.120)	0.335*** (0.107)	0.465*** (0.156)	3.095** (1.440)

Note: ASM: N = 2,524 (14-3 years with max 465 industries); NIPA: N = 1,380 (26-3 years with max 71 industries); IFR: N = 186 (11-3 years with max 24 industries). The table shows results of regressions of various automation proxies on the log of (one plus) the automation measure at the contemporaneous level and various lags. Each coefficient estimate represents a separate regression. Data are annual; industry fixed effects are at the most disaggregate level of industries, but at maximum at the 3-digit SIC level. Regressions include a constant. Industry-clustered standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

exploit a natural experiment, the introduction of broadband internet in Norway, to study employment effects of automation.

As a proxy for physical automation innovations, Graetz and Michaels (2015), Acemoglu and Restrepo (2017) and Dauth et al. (2017) count the number of robots used in production, a dataset assembled by the International Federation of Robotics. Lewis (2011) applies a more general understanding of automation by looking at adoption rates for new automation technologies, but with limited coverage of industries.

To show how our index relates to some of these measures, Figure 3 correlates automation patents with investment in computer capital and shipments of robots. We use two different data sources for investment in computer capital: The National Income and Product Accounts (NIPA), which provides annual data until 2001 for 71 2- and 3-digit SIC industries and the Annual Survey of Manufactures (ASM), which is available annually from 2002 onwards and for 465 4-digit SIC industries, the majority of them being manufacturing industries. As a measure of robots, we use the dataset on robot shipments by the International Federation of Robotics, which is provided at an annual frequency for North America starting from 2004 for 24 SIC industries. All correlations are highly positive, which indicates that our automation measure captures both advances in robotics and in software, which are then translated into production and trade of computers and robots.

This positive relationship holds even in a panel regression that controls for time- and/or

Table 6: Automation and industry task input

		Outcome: Within-industry change in task input		
		1970-1980	1980-1990	1990-98
Δ Non-routine analytic	Auto Technology	-0.012	0.033***	0.011
		(0.011)	(0.005)	(0.014)
	Constant	0.068***	0.110***	0.139***
		(0.011)	(0.014)	(0.019)
	R^2	0.004	0.019	0.001
Δ Non-routine interactive	Auto Technology	0.017*	0.062***	0.007
		(0.010)	(0.008)	(0.018)
	Constant	0.131***	0.206***	0.279***
		(0.017)	(0.030)	(0.036)
	R^2	0.004	0.016	0.000
Δ Routine cognitive	Auto Technology	-0.032**	-0.066***	-0.031***
		(0.016)	(0.011)	(0.011)
	Constant	-0.081***	-0.185***	-0.254***
		(0.022)	(0.024)	(0.038)
	R^2	0.008	0.027	0.003
Δ Routine manual	Auto Technology	-0.010***	-0.022***	-0.003
		(0.003)	(0.004)	(0.004)
	Constant	0.002	-0.058***	-0.095***
		(0.007)	(0.009)	(0.011)
	R^2	0.008	0.021	0.000

Note: The table presents separate OLS regressions for the subperiods 1970-1980, 1980-1990 and 1990-1998, always using as explanatory variable the average change of new automation patents between 1976 and 1998 (divided by 1000). The dependent variable is the change in industry-level task input as calculated by Autor et al. (2003). Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

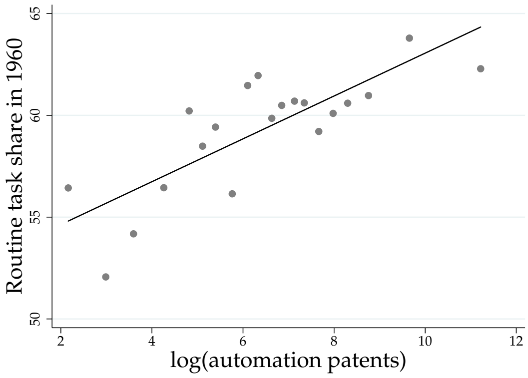
industry-specific effects. Table 5 shows further that the correlations are significant at various lags of automation, accounting for the fact that it may take some time to translate a patented innovation into the actual production of this technology. Although the results across the three variables are not directly comparable due to different time periods and industries covered, the link with computer investment might be slightly stronger than that with robot shipments.

Another way to contextualize our indicator is to evaluate how it relates to the nature of jobs. A large strand of literature, pioneered by Autor et al. (2003), analyzes the labor market effects of automation based on the assumption that automated machines are good at carrying out repeated tasks and fail at complex intellectual or manual tasks. For each occupation, they calculate what share of a job comprises routine (manual or cognitive) tasks. The resulting routine-task index thus measures the outcome of automation given specific – theory- and data-supported – assumptions. Weighing the index by occupation-specific employment, Autor et al. (2003) further create a routine task intensity measure across 140 industries from 1960 to 1998, based on which they show that changes in routine-task intensity are predicted by investment in computer capital: The share of non-routine tasks increases, whereas that of routine tasks decreases as a result of computer investment.

Figure 4 plots the routine task share of industries in 1960 against new automation techno-

Figure 4: Automation patents and routine labor

(a) Routine task share



(b) Per decade

	Correlation
1976-1985	0.35
1986-1995	0.34
1996-2005	0.32
2006-2014	0.28

Note: Binscatter of log of total number of automation patents in industries against routine task input share in 1960 across 258 SIC 3-digit industries, 1976-2014.
Source: Autor et al. (2003) and see text.

logy patented between 1976 and 2014.[11] The relationship between automation patents and the routine-task index is positive: The larger the routine task share of an industry in 1960, the more automation technology was subsequently invented, patented and potentially used in that industry in the following decades. Our indicator thus seems to be capturing the same phenomenon as described by the literature on routine-biased technological change. The correlation is strongest in the 1970s to 1980s and declines over time. We interpret this as a sign that the nature of automation technology may have changed: While in the 1970s until 1990s, automation technology mostly replaced routine tasks, it nowadays spreads into other tasks. This could be because many routine jobs have already been replaced by automation, so that additional research in this area is less demanded and less profitable. Another possible explanation is that recent advances in the automation technology frontier affect non-routine workers by being able to replace more complex intellectual or manual tasks. (The self-driving vehicle comes to mind.)

To explore this finding further, we examine the effects of technological change separately for routine manual, routine cognitive, non-routine analytic and non-routine interactive tasks. We regress changes in industry task input within each decade on our measure of new automation technology. This is a replication of a regression analysis by Autor et al. (2003), but we replace investment in computer capital with our index. To stay as close to the analysis of Autor et al. (2003) as possible we calculate the left-hand side variable separately for 1970-1980, 1980-1990 and 1990-1998 whereas on the right-hand side, we use the mean of new automation patents over the whole time period from 1976 to 1998.[12]

Table 6 shows that *more* automation patents were granted in industries where routine

[11] Data on routine-task intensities at the industry level is obtained from David Autor's website `economics.mit.edu/faculty/dautor` (accessed 14.07.2015). Their dataset is for U.S. Census industries which we translate into SIC industries using a concordance scheme of the U.S. Census Bureau.

[12] Results are very similar when we use the whole period that our indicator covers, 1976-2014. Alternatively, we can count only automation patents of the decade for which the change in task input is calculated. The results stay qualitatively the same. Regression outputs are available from the authors upon request.

cognitive and routine manual task inputs *declined* and where the share of non-routine analytic and non-routine interactive tasks *increased*. It is noteworthy that for all four task inputs the effect is strongest in 1980-1990. This differs from Autor et al. (2003) who found that for routine tasks the effect had monotonically increased over time.

4 Labor market effects of automation technology

In this section, we first motivate our unit of analysis, local labor markets, before explaining how we translate our index from industries to U.S. commuting zones. We show graphically how automation across commuting zones changed over time. Then, we apply the derived measure in our econometric analysis of employment effects. In the regression set-up, we rely on fixed effects five-year overlapping time periods, which we explain in detail before discussing the results. We run regressions for the full sample and separately for manufacturing and non-manufacturing employment.

4.1 Commuting zones as level of analysis

We study the effects of automation on employment at the level of U.S. commuting zones. Tolbert and Sizer (1996) have grouped all counties of the U.S. mainland into 722 commuting zones which each exhibit strong commuting ties within, but weak commuting ties between one another. These regions are meant to approximate local labor markets. In response to a shock to labor demand, most adjustments in the short- and medium-run will take place within the local labor market (Blanchard and Katz, 1992, Moretti, 2011). Workers, when laid off, usually first look for a new job within the same commuting zone. This is particularly true for low-skill workers, who are likely to be affected the most by automation (Notowidigdo, 2011). Therefore, studying the effects of automation on employment on the level of commuting zones gives us a more complete picture of the employment effects of automation than an industry-level analysis, which would neglect worker flows from one industry to another. This is of particular relevance because of the substantial shift of employment from manufacturing to services in the sample period.

We use employment data by the *County Business Patterns* (CBP) to convert patents per industry to worker patent automation intensities on a commuting zone level.[13] To create the commuting zone measure of automation, we first take (one plus) the natural logarithm of industry-level automation patents in order to account for the different levels of patenting across industries: In some industries the pace of technological progress is too fast for patents to be a feasible way to protect innovations, while in others, inventors have strategic reasons not to file for a patent. We then divide the employment-weighted sum of

[13] In this dataset, employment numbers are reported by county and 4-digit SIC (6-digit NAICS) industry. In contrast to Census data, which is sometimes used for commuting zone analysis, CBP provides annual data for the whole period of analysis. Agriculture (SIC < 1000) and public administration (SIC > 9000) are excluded from CBP. To avoid imprecision due to SIC-NAICS correspondences and missing CBP employment data for some particular industries, we aggregate employment and the automation index on the 3-digit SIC level before matching.

automation patents by total employment in the commuting zone. The resulting measure is

$$\underbrace{\text{autoint}_{c,t}}_{\text{automation intensity}} = \frac{\sum_i \ln(1 + \text{automation patents}_{i,t}) L_{i,c,t}}{L_{c,t}}, \tag{4}$$

where L is employment, i stands for industry, c for commuting zone and t for time period.

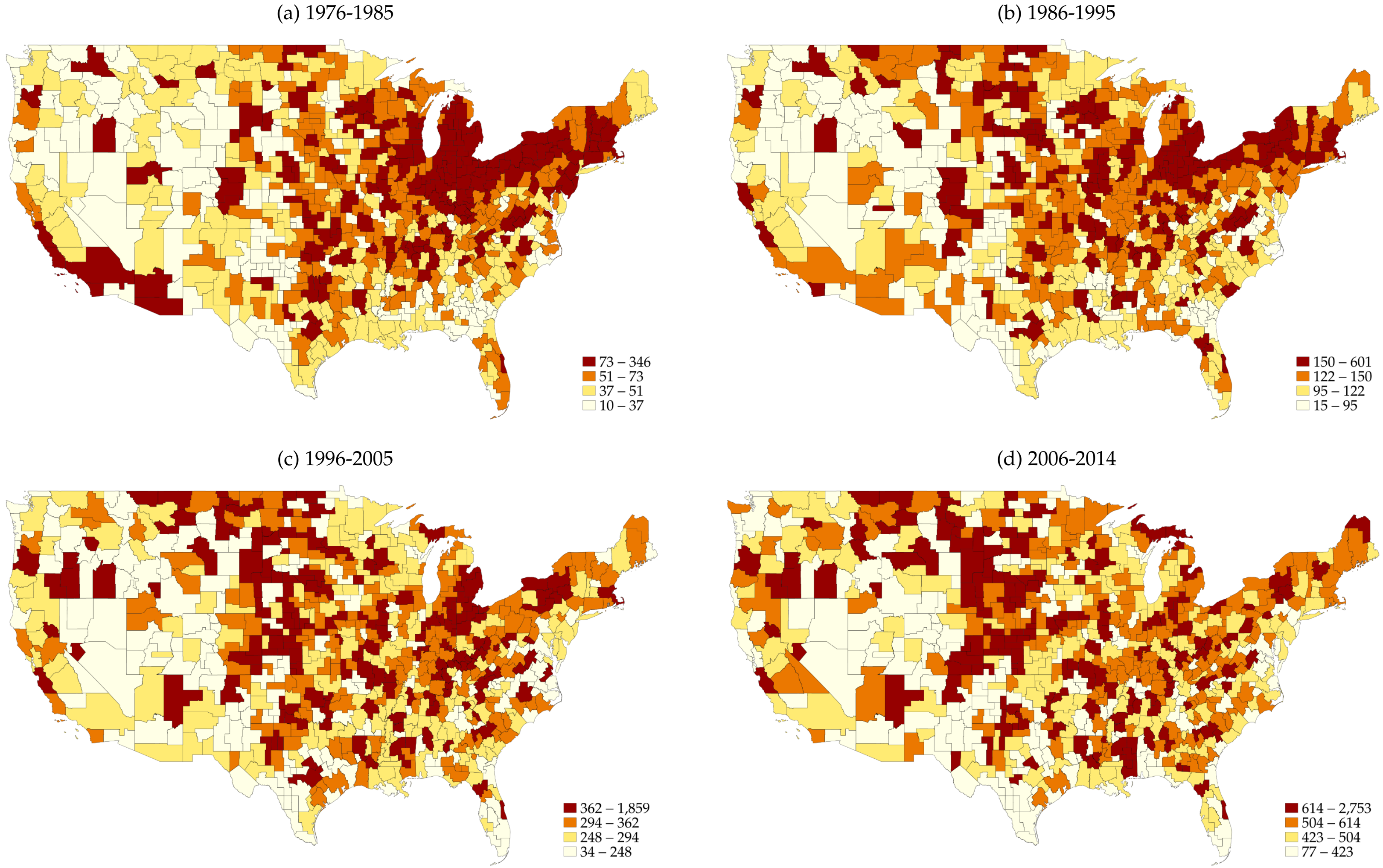
Figure 5 shows the number of automation patents per worker across U.S. commuting zones in four subperiods: 1976-85, 1986-95, 1996-2005 and 2006-14. The colors represent four quartiles of the distribution of automation intensity (in levels) in these subperiods: dark red color signals the 25 percent of commuting zones with the most patents, white color signals the 25 percent with the least patents. The map thus indicates which commuting zones have a high or low share of patents *relative* to the rest of the United States in the specific sub-period.[14]

There are pronounced regional patterns in the dispersion of available automation technology. Between 1976 and 1995, the region around the Great Lakes had a large automation patent intensity relative to the rest of the United States. This stems from the conjunction of both a high number of patents in manufacturing industries and a large share of industrial employment in this area. Starting in the mid-1990s, many commuting zones in this region move to a lower quartile as the number of manufacturing employees decreased relative to the number of employees in sectors with fewer patents. But our map of automation density is not simply a reflection of the manufacturing share. In a particular the Southern United States

The commuting zones with the highest automation intensities are more dispersed in the 1990s and 2000s. Commuting zones in Montana, North and South Dakota and Nebraska attract many automation patents per employee. The Rocky Mountain region has a low share of patents throughout the whole sample period. The map therefore reveals substantial geographic variation over time, which we exploit in the regression analysis.

[14] As the legend shows, the absolute number of patents has increased across all quartiles. An individual commuting zone may thus have had its absolute number of patents increase constantly over time, but change from dark red to white because the index increased relatively more slowly than in other commuting zones.

Figure 5: Intensity of automation patents across commuting zones, 1976-2014



Note: Shows averages of the number of national automation patents that can be used by a single worker.
Source: USPTO, Google, Silverman (2002), CBP and own calculations.

4.2 Empirical strategy

Our dependent variable is the five-year change in the employment-to-population ratio L_c/pop_c in commuting zone c:

$$\Delta \frac{L_{c,t}}{\text{pop}_{c,t}} = \frac{L_{c,t}}{\text{pop}_{c,t}} - \frac{L_{c,t-5}}{\text{pop}_{c,t-5}},$$

where in contrast to automation, we observe employment directly at the commuting-zone level. We choose a medium term period as new patents might start to be used by firms only with some lags.[15] This also holds the additional benefit of smoothing out business cycle effects.

The main explanatory variable is the five-year sum of the automation intensity in a commuting zone: $\sum_{s=0}^{4} \text{autoint}_{c,t-s}$. By using sums, we interpret patents as a flow measure of technology and therefore, the five-year sum of new patents is the five-year difference in the stock of patents.

In our econometric analysis we ask the following question: What is the impact of newly available nationwide automation technology on changes in the employment structure at the local level? In order to answer this question causally, we need to argue convincingly that our automation measure is exogenous to employment changes. The main potential source of endogeneity is that in their research activity, firms may be reacting to local developments, for example changes in labor costs, regulations or demand, thus introducing a reverse causality bias. There are several reasons why this is less of a concern for us:

Automation by industry of use: Assigning patents to the industries where they are likely to be used, not filed, weakens the danger of reverse causality: The research effort of a firm in one industry is less directly linked to employment trends in another industry than, for example, data on actual investment in automation technology. Additionally, many patents are granted to universities, research institutes or individuals that might follow other objectives than profit maximization, for example intellectual curiosity or an interest in advancing science. These sources of innovation are of relevance, as in year 2000 about 7000 patent licenses to firms were issued by U.S. universities and U.S. public research institutions (OECD, 2003). Further, around half of the patents granted by the USPTO are filed by foreign applicants. This reduces the potential for a feedback from industry wage structure to innovative activity, as a patent from, for example, a manufacturer in Japan is less likely to respond to employment conditions in the manufacturing industry in the United States.

National innovation, local effects: We measure innovations at the level of national industries, whereas we observe employment changes locally. Our constructed commuting zone automation measure is thus a proxy for unobserved locally applicable innovation in the spirit of Bartik (1991), as recently explained by Goldsmith-Pinkham, Sorkin, and Swift (2017). A national industry is unlikely to react to local employment trends in its research

[15] Results are robust to changing the length of a period.

activity unless the following conditions hold: First, the specific commuting zone is of key importance to the industry (by hosting a large share of industry employment) and second, the industry is represented strongly in the commuting zone, so that industry trends will translate directly into commuting zone employment trends. These conditions do not drive our findings: In our sample, only two commuting zones are above the 25 percent double threshold (CZ 35002 in Arizona and CZ 37601 in Nevada, in both of which mining is dominant) and only 34 commuting zones are above the 10 percent double threshold. Excluding these does not significantly change the results.

Fixed industry structure: We fix the employment structure in equation (4) to the beginning of each five-year period. This means that in the following five years we assign all patents to a commuting zone according to the initial employment share of relevant industries in this commuting zone. Our indicator thus does not pick up employment changes that happen within the five-year period. A downside of keeping the employment structure fixed is that we potentially do not count all those patents which workers in a commuting zone can use, but might over-represent declining and under-represent growing industries.[16]

Additionally, in Secion 4.6 we exploit information on the owners of patents in order to identify innovations that more likely result from research effort that is unrelated to trends in US labor markets. We show that our baseline regression results hold when focusing only on patents held by foreigners, governments or universities and public research institutes, or when using these as instruments for the patents held by US companies.

4.3 Regression set-up

We consider changes in overlapping five-year time periods and the sample therefore comprises 34 consecutive five-year periods across 722 commuting zones.[17]

The estimation equation takes the form

$$\begin{aligned}\Delta \frac{L_{c,t}}{pop_{c,t}} = \alpha_k + \gamma_t + \beta_1 \sum_{s=0}^{4} \text{autoint}_{c,t-s} + \beta_2 \sum_{s=0}^{4} \text{non-autoint}_{c,t-s} + \beta_3 \text{routine}_{c,t-5} + \\ \beta_4 \left(\sum_{s=0}^{4} \text{autoint}_{c,t-s} \times \text{routine}_{c,t-5} \right) + X'_{c,t-5}\,\beta_5 + \varepsilon_{c,t,t-5}\,, \end{aligned} \tag{5}$$

where γ_t are time fixed effects and α_k are state fixed effects. $X_{c,t-5}$ are additional control variables. The main variable of interest *autoint* is automation intensity, *non-autoint* is the intensity of any non-automation patents and *routine* is the routine task share which we describe below. To construct the left-hand side variable, we take county level population data from the *Census Population and Housing Unit Estimates* and county-level employment

[16]The results are however robust to using an adaptive industry structure.

[17]The overlapping data structure generates serial correlation. We correct the standard errors by using the Driscoll and Kraay (1998) estimator, which corrects both for serial and spacial correlation. An alternative would be to use non-overlapping time periods. But not only would this mean losing a considerable amount of observations (and thus precision), but it would also require us to choose cut-off points for the five-year intervals, which would always be to some extent arbitrary. As shown in the appendix, all main results go through using this more standard estimation procedure instead.

Table 7: Summary statistics of main variables in baseline regression

Variable	**Mean**	**Overall Std. Dev.**	**Between Std. Dev.**	**Within Std. Dev.**	**Min**	**Max**
Δ emp/pop	1.19	2.71	0.710	2.62	-9.40	13.2
Δ manu emp/pop	-0.342	1.08	0.457	0.977	-5.35	4.33
Δ non-manu emp/pop	1.53	2.19	0.542	2.12	-8.63	12.9
autoint	16.4	3.02	1.23	2.76	7.63	28.6
non-autoint	18.8	1.89	1.38	1.29	8.88	26.7
routine	34.4	5.32	4.25	3.20	8.51	56.3

Note: Variables are as defined in the text.

data from CBP. Because the CBP omits employment in some SIC industries for certain years, there are a few large jumps in the outcome variable, which we exclude from the analysis by dropping data below the 1^{th} and beyond the 99^{th} percentile in each year.[18]

In addition to commuting zone intensities of automation patents, we include intensities of non-automation patents (*non-autoint*) in the regression, computed analogously to equation (4). This variable controls for the effect of technological change other than in automation technology. Given that some industries generally patent more, it is likely that the number of automation patents and non-automation patents granted annually are correlated across industries and commuting zones. At the same time, non-automation inventions may also have an independent effect on employment. In particular, they may be interpreted as an indicator for local growth potential, which we might otherwise suspect to be accountable for correlations between automation and employment: If growing industries increase their workforce as well as invest more in R&D, this should be reflected by the coefficient on *non-autoint*.

As described in Section 3, an often-used measure of susceptibility to automation is the routine-task index by Autor et al. (2003). The different construction of this measure from ours creates the opportunity to explore how the effects of these two are related and to ask the question: How does the effect of automation depend on the routine task share of a commuting zone? We therefore include the initial ($t - 5$) routine task share (*routine*) in the regression as well as an interaction term between this measure and the variable for automation intensity.

We further include the initial share of manufacturing employment in total employment (CBP) to capture structural change in the economy. Automation patents occur to a larger extent in the manufacturing sector than in the service sector, so an increase in the automation index may parallel a decline in the manufacturing industry for other reasons, such as the cheap import of manufactured goods from abroad or changes in the demand for goods. If not included as a control, any effect stemming from non-automation-related structural change might be attributed to automation technology.

[18] For details, see `census.gov/program-surveys/cbp/technical-documentation`. The number of commuting zones in each year falls to 708.

Similar to Acemoglu and Restrepo (2017), our set-up also includes the log of initial commuting zone population because employment in larger and smaller commuting zones – in particular when interpreting this as a proxy for urban vs. rural areas – might react differently to automation. We also control for the share of non-white citizens in the commuting zone population and for the (log of) per capita level of personal income. Data on the demographic variables are taken from the *Census Population and Housing Unit Estimates*, data on income come from the Bureau of Economic Analysis' *Regional Economic Information System* (REIS), which exploits county-level data from administrative records and censuses.

Table 7 summarizes the main variables of interest. Employment per population grew on average over the sample period.[19] Employment changes were negative on average for the manufacturing sector and positive for the non-manufacturing sector with more within and across variation for the latter.[20] Our automation intensity measure *autoint* takes the value 16.4 on average across years and commuting zones. This value is equivalent to a commuting zone with a flat industry structure (i.e., all 377 SIC 3-digit industries having the same employment share) where 25 new automation patents are granted every year in all industries. Because patents are skewed across industries, this number will be larger for most industries.

4.4 Estimation results: Total employment

Table 8 presents the baseline results. Throughout almost all specifications, *autoint* has a significantly positive coefficient in the range of around 0.10 to 0.23 percentage points. So new automation technology per worker is significantly related to employment gains in the same commuting zone. This result is robust to controlling for several economic and demographic variables.

Column (1) shows the positive association between automation and employment when no further controls but time and industry fixed effects are included. The relationship becomes more pronounced when we control for other non-automation patents in column (2). Columns (3) shows our preferred regression specification. The coefficient on *autoint* in column (3) can be interpreted such that a one-unit increase in the automation intensity leads to a 0.178 percentage point increase in the employment-to-population ratio. As laid out in Table 7, this is about one sixth of the average five-year increase across all observations. The within-year interquartile range of *autoint* lies between 1.23 and 2.15, so a one-unit increase is well within the range of variation of the sample. In terms of the actual number of new patents that this implies, a one-unit increase in *autoint* around its mean is equivalent to the number of new automation patents in a commuting zone with a

[19]This is mainly driven by increases in female labor market participation, which rose from 47 percent in 1976 to 57 percent in 2014, peaking at 60 percent in 1999. (See the BLS series LNS11300000, LNS11300001 and LNS11300002.) Male participation rates fell quite monotonously from 78 percent in 1976 to 69 percent in 2014. We take care of these structural long-run changes in the labor market not related to automation through time fixed effects.

[20]We will use "non-manufacturing" and "services" interchangeably, but "non-manufacturing" also includes mining and construction.

Table 8: Labor market effects of automation, five-year overlapping time periods

	Outcome: Employment-to-population				
	(1)	(2)	(3)	(4)	(5)
autoint	0.105***	0.222***	0.178**	0.144	0.563**
	(0.0363)	(0.0783)	(0.0853)	(0.0886)	(0.214)
non-autoint		-0.120	-0.0245	0.0249	-0.0170
		(0.0997)	(0.0931)	(0.0920)	(0.0989)
manufacturing			-1.782*	-1.211	-1.177
			(1.016)	(1.082)	(1.121)
population			0.0875	0.0745	0.0525
			(0.114)	(0.108)	(0.102)
income			-1.319***	-1.284***	-1.232***
			(0.351)	(0.347)	(0.338)
non-white			-1.222***	-1.256***	-1.383***
			(0.259)	(0.273)	(0.283)
routine				-0.0257	0.143*
				(0.0161)	(0.0787)
autoint $\times$ routine					-0.0109**
					(0.00468)
Observations	24,064	24,064	24,064	24,064	24,064
R^2	0.42	0.42	0.43	0.43	0.43

Note: The table presents fixed effects regressions using five-year changes in employment as percent of commuting zone population as the dependent variable. *autoint* and *non-autoint* are five-year sums of new automation and non-automation technology. *routine* is the initial percentage of routine tasks in commuting zone employment. The initial manufacturing share, the log of initial commuting zone population, the log of initial per capita income and the initial share of non-white citizens in the population are further controls . All regressions include state and year fixed effects and a constant. Driscoll-Kraay standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

flat industry structure rising from 23 to 29 per year.

A particularly interesting result is how automation technology interacts with the routine task share. In the setup with both variables in column (4), the coefficients on automation and on routine-intensity become insignificant. This is likely due to the fact that the variables measure overlapping concepts, as argued in Section 3. However, both coefficients are significant when we include the interaction between the two variables. The negative coefficient on the interaction shows that the magnitude of the effect of automation on employment varies with the level of the routine task share: In commuting zones with more routine labor, automation technology has a less positive effect. The total effect of automation in column (5) turns negative for commuting zones with a routine task share larger or equal to 54.5 percent. The mean of *routine* is 34.4 and in only 0.1 percent of all observations it exceeds 54.5 percent. So, the total effect of automation is positive in the overwhelming majority of commuting zones.

Non-automation patents are not associated with changes in employment. This might be driven by the nature of these innovations. Many non-automation patents are chemical or pharmaceutical and some are patents without any clear applications. In contrast, automation patents are required by our definition to have at least a distantly recognizable application.

The initial manufacturing share has a mildly significant negative coefficient in our baseline setup of column (3), which might capture the part of the secular trend from manufacturing to services that takes place in the five-year periods we study. The population size is not significantly related to employment changes. A higher per capita income negatively predicts employment changes across all specifications. The employment level is generally higher in commuting zones with a higher per capita income. This could be a sign of convergence in employment shares across commuting zones, but could also reflect a reversely causal effect: as personal income is composed to a large extent of labor income, there could be slower employment growth in commuting zones with a higher wage level, because it is more costly to create jobs. A higher share of the non-white population is negatively associated with employment changes.

Our findings thus paint a more positive picture of the net employment effects of automation than Autor et al. (2015), Graetz and Michaels (2015) and Acemoglu and Restrepo (2017), who found negative or insignificant effects of automation on jobs.[21] It is, however, in line with the findings by Gregory et al. (2016), who show that next to a substitution effect on routine-task jobs, automation lower the production costs. Declining goods prices boost product demand, and so new (non-routine) jobs are created. The positive product demand effect trumps the negative substitution effect. Both the positive level effect of automation and the negative coefficient on the interaction term with the routine task share in our regressions support this explanation. By using a broader measure of automation, we can thus extend the knowledge on its employment effects beyond the findings of a literature that focuses on specific types of automation.

4.5 Estimation results: Sectoral employment

We further study the effect of automation on different types of employment separately. Table 9 shows pointedly different effects of automation technology on manufacturing and non-manufacturing employment.

Panel A consistently shows that manufacturing employment falls when the automation intensity increases. The effect is significant in our preferred specification (3) and when adding the routine task share in column (4). In contrast to the total US population, the group of manufacturing workers experiences job losses - even when controlling for the initial manufacturing share, which itself has a significantly negative effect. The negative employment effect of automation is more pronounced in commuting zones with a higher routine task share, as the interaction term shows. It turns positive only for commuting zones with a routine task share below 20.9 percent. This is only the case for 115 out of 24,058 observations. Panel B paints a very different picture. In non-manufacturing industries, automation has a very robust job-creating effect. The coefficients are twice as large as in Table 8. Non-manufacturing occupations are clear beneficiaries from automation in terms of employment numbers. In contrast to Panel A, the routine task share in the commuting

[21] Section 4.5 sheds light on why this is the case.

Table 9: Labor market effects of automation for manufacturing and non-manufacturing employment, fixed employment structure

	(1)	(2)	(3)	(4)	(5)
A. Outcome: Manufacturing employment-to-population					
autoint	-0.0169	-0.0480	-0.173***	-0.200***	0.144
	(0.0176)	(0.0665)	(0.0300)	(0.0300)	(0.0911)
non-autoint		0.0317	0.235***	0.275***	0.240***
		(0.0747)	(0.0299)	(0.0296)	(0.0218)
manufacturing			-2.581***	-2.142***	-2.127***
			(0.587)	(0.617)	(0.656)
population			-0.0335**	-0.0437***	-0.0608***
			(0.0133)	(0.0128)	(0.0149)
income			-0.739***	-0.712***	-0.668***
			(0.206)	(0.206)	(0.201)
non-white			-0.122	-0.150	-0.259
			(0.238)	(0.232)	(0.214)
routine				-0.0200***	0.119**
				(0.00247)	(0.0437)
autoint × routine					-0.00898***
					(0.00243)
Observations	24,058	24,058	24,058	24,058	24,058
R^2	0.21	0.21	0.25	0.25	0.26
B. Outcome: Non-manufacturing employment-to-population					
autoint	0.113***	0.278***	0.372***	0.370***	0.420***
	(0.0344)	(0.0984)	(0.0768)	(0.0799)	(0.147)
non-autoint		-0.169	-0.293***	-0.290***	-0.296***
		(0.112)	(0.0870)	(0.0840)	(0.0894)
manufacturing			0.852	0.883	0.887
			(0.728)	(0.719)	(0.726)
population			0.118	0.117	0.115
			(0.109)	(0.103)	(0.101)
income			-0.612**	-0.610**	-0.604*
			(0.291)	(0.299)	(0.298)
non-white			-1.105***	-1.107***	-1.122***
			(0.178)	(0.188)	(0.194)
routine				-0.00136	0.0186
				(0.0173)	(0.0384)
autoint × routine					-0.00129
					(0.00256)
Observations	24,067	24,067	24,067	24,067	24,067
R^2	0.38	0.39	0.39	0.39	0.39

Note: The table presents fixed effects regressions for five-year changes in manufacturing employment-to-population and non-manufacturing employment-to-population. The control variables are defined as in Table 8. All regressions include state and year fixed effects and a constant. Driscoll-Kraay standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

zone does not play a significant role for the size of the automation effect.

Related to this, the coefficient on the routine task share also reveals strong differences between manufacturing and non-manufacturing employment. Commuting zones with a lot of routine labor lose more manufacturing jobs, but this is not the case for non-manufacturing employment. This is likely due to the larger share of routine tasks in the manufacturing than in the service sector. These findings may explain why Acemoglu and Restrepo (2017), in their analysis of the impact of robot use on employment, found automation to be harmful for employment and why Graetz and Michaels (2015), using the same dataset, found evidence for skill polarizing effects of robots: Robots are mainly used in the manufacturing sector and indeed 19 out of the 24 industries covered by IRF robot data are manufacturing industries. Other types of automation innovations, in particular those that can be used in the non-manufacturing sector, may have a more positive effect on employment than industrial robots. Indeed, Acemoglu and Restrepo (2017) show that the effect of robots is less negative or even positive in non-manufacturing industries. They also find that computer usage tends to increase the demand for labor.

We add to the existing literature by documenting different effects of automation on manufacturing and non-manufacturing employment: Next to a polarization in skills and tasks, automation has lead to a sectoral shift. Manufacturing sector jobs win, while non-manufacturing jobs lose from automation.

The results presented in this and the previous section are robust to weighing patents by how often they have been cited. Patent citations are sometimes used as an indicator of the value of an invention and therefore, giving stronger weight to highly cited patents might paint a more realistic picture of the degree to which a patent is used in the production process. In Tables A4 and A5 we replicate the regressions presented in Tables 8 and 9 using a citations-weighted measure of automation, which we explain further in the Appendix. While our sample is thus shortened by several years, we still find a mildly positive effect of automation for total employment and a pronounced disparity between manufacturing and non-manufacturing.

4.6 Effects of automation by assignees

Patents contain information on who owns (or "is assigned") a patent. This information is valuable, because it hints on how closely a patentee's research activities are linked to developments in US labor markets. Innovation activity by entities that do not have business interests in US markets is less likely to be influenced by developments on US labor markets. By focusing on new automation technologies that are originating from such groups, we therefore get a cleaner identification.

To classify the patents, we use data by Lai, D'Amour, Yu, Sun, Doolin, and Fleming (2011), who extract the names of assignees from 1976 until 2012 and provide a host of other information about patents and their owners. We focus on patents held by three groups of assignees, who we believe to be less directly responsive to US labor market trends than US companies: foreigners (these can be companies, individuals or public entities), government bodies (US or foreign) and universities and public research institutes.[22]

Research by foreigners can be assumed to respond to developments in their home country rather than in the United States, as long as the following two conditions are met: The company does not operate on a large scale in the United States, and the domestic labor market trends are not linked to US trends. We do not observe if these conditions hold, so the group of foreigners is the most endogenous of the three. Universities and public research institutes conduct more basic research than corporations, so for them, the immediate applicability or profit maximization might only be a distant motivation. Government patents are also unlikely to be motivated by labor market developments, but should rather respond to military buildups, the needs of certain ministries or cycles in budgetary planning.

Table 10: Assignee summary statistics, 1976-2012

Assignee	Patents (1000s)	Automat (1000s)	Share	Cit.	Cit. (weighted)	Excl.	Length
US firm	1875.7	948.3	51%	12.2	1.24	14%	1012.3
foreigners	1827.8	777.1	43%	7.1	0.78	12%	831.5
universities	115.1	67.0	58%	10.4	1.02	41%	1435.8
governments	44.8	19.1	43%	8.6	0.74	17%	700.8
missing	609.9	187.5	31%	9.7	0.91	9%	653.8

Note: "Automat" are automation patents as described in text. "Cit." are the average number of citations, "Cit. (weighted)" are the number of citations after removing time-subclassification (HJT) means, where subgroups correspond to those of Table A2. "Excl." is the share of excluded patents due to being pharmaceutical and chemical patents. "Length" is the average number of lines in a patent document.
Source: Lai et al. (2011) and own calculations.

Table 10 shows summary statistics for patents by the different groups of assignees. US firms are the largest group with around 1.9 million patents. The second largest group are foreigners, who hold 1.8 million patents. Based on the classification by Lai et al. (2011), we identify 45 thousand patents that are assigned to governments. The most important assignees in this category are the US Navy with 10,922 patents, the US Army with 6,217 patents, the US Department of Energy with 4,416 patents, the US Air Force with 3448 patents and NASA with 2,823 patents. The largest foreign government institutions owning US patents are French nuclear energy and aviation commissions and the British and Canadian defense ministries. To identify patents assigned to universities, we inspected the 10,000 assignees with the most patents and determined whether they are an university or a public research institute. There are 581 such entities holding a total of 115 thousand patents. The most productive are the University of California (5,400 patents), the Industrial Research Institute of Taiwan (4,289 patents), the Massachusetts Institute of Technology (3897 patents), the Electronics and Telecommunications Research Institute from South Korea (3,606 patents) and the French Institute of Petroleum (2,471 patents). For the remaining 610 thousand patents, we do not know the assignee, as this information is missing in Lai et al. (2011). A casual inspection of these patents suggests that most of these also belong to US firms or individuals.

[22] These groups are mostly mutually exclusive, but we count foreign governments (a small group) in both the "foreign" and the "governments" category and foreign universities also show up in the foreigners category.

Table 11: SIC-level correlation of patents in assignee subcategories with US companies

Assignee	Patents		Automation	
	year	year & SIC	year	year & SIC
foreigners	0.33	0.33	0.94	0.95
universities	0.35	0.36	0.88	0.88
governments	-0.45	-0.43	0.02	0.04

Note: Numbers show correlations of subcategories with the categories of US firms and missing assignees. "year" indicates that year trends are taken out, "year & SIC" indicates that year and industry trends are taken out.

The automation patents assigned to foreigners, universities or governments may be of a different nature than those held by US firms – not just for their less direct link to economic developments in the United States, but for reasons related to their applicability. We might see different effects of automation on employment if they were not representative of the technology frontier in automation. Table 10 shows that patents held by US firms are characterized by a larger share of automation patents and are more widely cited than those held by other patentees. However, automation patents are highly correlated across groups at the industry level, as Table 11 shows. Automation innovations by governmental, foreign and university patentees seem to be applicable in similar industries as automation innovations patented by US firms or individuals. This is not the case when considering all patents. So while it is reasonable to assume that patented automation technology is similar across assignee groups, this is not the case for technology in general.

Indeed, the types of patented innovations differ across technology subgroups. As Table A2 shows, US firms hold a particularly high share of "Communication & Computer" patents, which contain a large number of automation patents. Foreigners hold fewer pharmaceutical patents, but many mechanical patents and their patents are cited least often. The column "Cit. (weighted)" in Table 10 shows that this holds even after controlling for time and subgroup fixed effects. Universities hold many chemical and pharmaceutical patents and few in the "Communication & Computer" category. These patents are also particularly lengthy. In contrast, governments hold many patents on electric and electronic innovations, and the corresponding patent texts are shorter than those from other assignees.

We replicate our empirical analysis from the previous section in two ways. First, we repeat the panel data regressions of Table 8 and Table 9, but for *autoint* and *non-autoint* we use the intensities computed from either only university patents, foreign patents or government patents. Second, we use all three automation sub-indicators as instrumental variables for possibly more endogenous category of US companies and non-identified assignees. The purpose of this exercise is to extract only the component of automation that is unrelated to US labor market developments. As we only have assignee data until 2012, we limit our analysis to the period 1976 to 2012.

For university patents, we document positive net effects of automation on employment. The same holds when using all three groups of automation patents as instruments in column (4). It is striking that again none of the effects of automation on total employment

Table 12: Labor market effects of automation, various assignee groups

	Outcome: Employment-to-population			
	(1) university	(2) foreign	(3) gov't	(4) IV
autoint	0.410*	0.153	-0.108	0.128*
	(0.217)	(0.128)	(0.223)	(0.0717)
non-autoint	-0.332	0.0145	0.379	0.0344
	(0.238)	(0.144)	(0.252)	(0.0756)
manufacturing	-0.769	-2.017	-2.061*	-1.961***
	(1.217)	(1.203)	(1.058)	(0.377)
population	0.121	0.110	0.113	0.119***
	(0.114)	(0.112)	(0.116)	(0.0232)
income	-1.225***	-1.393***	-1.358***	-1.358***
	(0.342)	(0.369)	(0.370)	(0.192)
non-white	-1.277***	-1.256***	-1.301***	-1.255***
	(0.233)	(0.256)	(0.250)	(0.255)
Observations	22,648	22,648	22,648	22,648
R^2	0.41	0.42	0.42	0.42

Note: All columns replicate column (3) of Table 8. In columns (1) - (3), the full automation measure is replaced by automation by universities, foreigners and governments, respectively. The non-automation measure is constructed accordingly. The last column represents an IV regression, where university, foreign and government (automation) patents are used as instruments for the remaining (automation) patents. Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

is negative. The size of the coefficient in Table 8 lies in the middle of the new estimates. Table 13 reports separate results for manufacturing and non-manufacturing employment. We find negative effects of automation on manufacturing employment for all assignee groups apart from university patents. All types of patented automation technology lead a rise in non-manufacturing employment. The magnitude of the coefficients again frame the previous estimates. The findings strongly support the results from our baseline analysis and thus show that the earlier findings were likely not biased by endogeneity of the regressors.

While having roughly the same effects on employment, we can detect slight differences between the patent assignee categories. Automation technology patented by universities and public research institutes has the most strongly positive effects on employment and even the manufacturing sector does not lose from this type of technology. The negative employment effects of automation on the manufacturing sector are strongest when we consider only government patents. Why could this be the case? Universities hold many chemical and pharmaceutical patents, while governments patent many electrical and mechanical patents (Table A2). But as explained before, we exclude most chemical and pharmaceutical patents and the classification algorithm further extracts only a relevant subset of patents. As Table A3 shows, the makeup of the final automation patents does not differ much between those two groups of assignees. Pharmaceutical patents make for 4 percent of university automation patents and 1 percent of government university patent. A more likely explanation is that the innovations by universities and governments differ

Table 13: Labor market effects of automation for manufacturing and non-manufacturing employment, various assignee groups

	(1) university	(2) foreign	(3) gov't	(4) IV
A. Outcome: Manufacturing employment-to-population				
autoint	-0.120	-0.208***	-0.435***	-0.216***
	(0.114)	(0.0314)	(0.128)	(0.0329)
non-autoint	0.157	0.286***	0.518***	0.303***
	(0.145)	(0.0375)	(0.169)	(0.0331)
manufacturing	-1.796**	-2.827***	-2.441***	-2.807***
	(0.672)	(0.652)	(0.693)	(0.171)
population	-0.0399***	-0.0429***	-0.0419***	-0.0321***
	(0.0137)	(0.0138)	(0.0140)	(0.00941)
income	-0.807***	-0.793***	-0.862***	-0.724***
	(0.213)	(0.234)	(0.213)	(0.0746)
non-white	-0.287	-0.130	-0.310	-0.0937
	(0.264)	(0.257)	(0.270)	(0.125)
Observations	22,642	22,642	22,642	22,642
R^2	0.24	0.25	0.25	0.25
B. Outcome: Non-manufacturing employment-to-population				
autoint	0.518***	0.380***	0.354***	0.374***
	(0.170)	(0.112)	(0.125)	(0.0598)
non-autoint	-0.479**	-0.304**	-0.175	-0.314***
	(0.204)	(0.135)	(0.148)	(0.0637)
manufacturing	0.912	0.897	0.337	0.963***
	(0.849)	(0.842)	(0.605)	(0.310)
population	0.157	0.150	0.150	0.147***
	(0.108)	(0.108)	(0.110)	(0.0200)
income	-0.431	-0.611*	-0.506	-0.661***
	(0.304)	(0.311)	(0.314)	(0.167)
non-white	-0.989***	-1.143***	-0.978***	-1.188***
	(0.138)	(0.131)	(0.146)	(0.207)
Observations	22,650	22,650	22,650	22,650
R^2	0.37	0.37	0.37	0.37

Note: All columns replicate column (3) of Table 9. In columns (1) - (3), the full automation measure is replaced by automation by universities, foreigners and governments, respectively. The non-automation measure is constructed accordingly. The last column represents an IV regression, where university, foreign and government (automation) patents are jointly used as instruments for the remaining (automation) patents. Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

along other dimensions that we do not measure.

5 Reassessing the literature

With our new dataset we revisit findings from two important papers of the literature on the local labor market effects of automation. We investigate whether our measure of automation predicts different effects for the growth of non-college service sector jobs (Autor and Dorn, 2013) and how the effects of automation compare with those from China import competition (Autor et al., 2015).[23] Apart from gaining additional insights through our new indicator, this allows comparing our results to the findings from the literature using the established routine-share measure.

5.1 Revisiting Autor and Dorn (2013): The non-college service sector and employment polarization

Autor and Dorn (2013) address the issue why there has been an increasing polarization in both employment and wages in 1980-2005. They focus on non-college service sector jobs (e.g., cleaners or security guards), which have grown more rapidly than other less-educated and low-paying occupations (such as factory work) and which have experienced wage increases. The authors hypothesize that this is due, among other things, to an increase in automation technology: Automation has reduced the demand for routine manual tasks, while increasing the demand for non-routine manual tasks, thus benefiting non-college service sector jobs at the expense of non-college production jobs.

In their empirical analysis, Autor and Dorn (2013) use the routine-task share as a proxy for automation and show that in commuting zones where initially more people worked in routine occupations, there was a larger increase in non-college service employment. In Table 14, column (1), we reproduce their finding to the letter.

Table 14: Automation and non-college service employment, 1980-2005

	Outcome: 10 × annual change in share of non-college employment in service occupations			
	(1)	(2)	(3)	(4)
routine	0.105***		0.105***	-0.336
	(0.0320)		(0.0284)	(0.230)
autoint		-0.00100	-0.000990	-0.00533**
		(0.000688)	(0.000645)	(0.00227)
routine × autoint				0.0139*
				(0.00695)
Constant	-0.00632	0.0568***	0.0241	0.161**
	(0.0104)	(0.0210)	(0.0202)	(0.0740)
R^2	0.179	0.171	0.185	0.188

Note: 2,166 observations (3 time periods × 722 commuting zones); robust standard errors in parentheses; all models include state fixed-effects and period fixed effects and are weighted by start of period commuting zone share of national population.
*** p<0.01, ** p<0.05, * p<0.1
Source: Own calculations following Autor and Dorn (2013), Table 5.

[23]Data and replication files for both papers are from David Dorn's website, `ddorn.net/data` (accessed 10.02.2017).

We then add *autoint*, our new automation intensity measure. The interaction term in column (4) between *autoint* and *routine* is positive and significant: Non-college service jobs rise in commuting zones with a high routine-task share initially *and* where many new automation patents could be used. This is consistent with the model presented by Autor and Dorn (2013) and highlights an important piece of evidence: the presence of those routine jobs that can be easily automated is necessary for the shift of low-skilled employment into the service sector, not the availability of automation technology by itself.

However, the total effect of automation changes from negative to positive only at a routine-task share of 0.38, a number reached by just 2 out of 2,166 observations and the coefficient on *autoint* in columns (2) and (3) is insignificant. So although we found in Section 4.5 that automation creates non-manufacturing jobs, the rise in non-college service jobs depends crucially on the mix between automation and the existence of routine jobs.

5.2 Revisiting Autor, Dorn, and Hanson (2015): Employment effects and relation to exposure to Chinese trade competition

Since the 1990's, there has been a strong rise in trade between the United States and China. A number of papers, such as Autor et al. (2013), Acemoglu et al. (2016) and Pierce and Schott (2016), argue that Chinese import competition is responsible for employment losses in those regions where firms reside that are most exposed to it. Autor et al. (2015) investigate whether this "China shock" or automation has a larger impact on U.S. labor markets. They find that while import competition reduces employment in local labor markets, automation – as measured by the routine task share – is not related to employment changes.

We revisit this finding with our dataset. Table 15 replicates the baseline analysis of Autor et al. (2015), Table 1, in which the authors regress 10-year equivalent changes in the employment-to-population ratio, unemployment-to-population ratio and non-participation rate among working age adults. The two main variables of interest are the contemporaneous change in Chinese import exposure per worker and the start-of-decade employment share in routine occupations, both of which are being instrumented.[24]

Columns (1) and (4) of Table 15 are exact replications of columns (1) and (3) of Autor et al. (2015), one containing only the initial routine share, the other one both the routine share and the China shock as explanatory variables. In columns (2) and (5), we replace the routine share by our commuting zone automation intensity. While the coefficient on the routine share is always insignificant, our automation measure has a significantly positive effect on the employment share and a significantly negative effects on both share of unemployed workers and the share of workers that are not in the labor force. This even holds when including both *autoint* and the routine task share. Automation patents have positive effects by reducing the unemployment rate and the number of people outside of

[24]The instrument for the trade variable is imports from China to other advanced economies. For the initial routine task share, Autor et al. (2015) use its 1950 value in all states but the one that contains the commuting zone, weighted by 1950 employment shares. They argue that in this way, they can isolate the stable, long-run differences in the production structure across commuting zones.

Table 15: Labor market effects of automation patents, routine employment share and exposure to Chinese import competition, 1990-2007

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Outcome: Share of employed in workage population					
routine	-0.0481		-0.0369	-0.207		-0.185
	(0.224)		(0.233)	(0.254)		(0.260)
autoint		0.215***	0.206***		0.331***	0.297***
		(0.0670)	(0.0748)		(0.0757)	(0.0792)
Δ (Imports from China to US)/Worker				-0.831***	-0.832***	-0.942***
				(0.215)	(0.181)	(0.221)
	B. Outcome: Share of unemployed in workage population					
routine	-0.0144		-0.0247	-0.00513		-0.0104
	(0.0616)		(0.0653)	(0.0702)		(0.0728)
autoint		-0.0579**	-0.0645**		-0.0926***	-0.0914***
		(0.0255)	(0.0282)		(0.0222)	(0.0285)
Δ (Imports from China to US)/Worker				0.186***	0.249***	0.221***
				(0.0527)	(0.0676)	(0.0612)
	C. Outcome: Share of not in labor force in workage population					
routine	0.0624		0.0616	0.213		0.195
	(0.172)		(0.178)	(0.194)		(0.197)
autoint		-0.158***	-0.141**		-0.239***	-0.206***
		(0.0538)	(0.0608)		(0.0667)	(0.0672)
Δ (Imports from China to US)/Worker				0.645***	0.583***	0.721***
				(0.188)	(0.155)	(0.190)

Note: The table is based on Autor et al. (2015), Table 1, juxtaposing the effect of Chinese import competition and routine biased technological change on 10-year equivalent changes in the employment status of the working-age population. N = 1444 (2 time periods 1990-2000, 2000-2007, 722 commuting zones). All regressions control for the start of period levels of share of employment in manufacturing, share of population that is college educated, share of population that is foreign born, employment rate among females and Census division dummies. Robust standard errors in parentheses are clustered on state. Models are weighted by start of period commuting zone share of national population. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the labor force, with a larger effect on the latter group.

An additional finding is that while the effect of the routine task share stays insignificant when including the China shock in column (5), the estimates become even more strongly positive when using our automation indicator. The coefficient on the China shock change little when using *autoint* (column (5)) instead of the *routine* (column (4)). This lends further support to the findings of Autor et al. (2015) on the detrimental effect of Chinese import competition, while automation is playing a more positive role now.

6 Conclusion

This paper makes two contributions: First, it provides a new indicator of automation by applying a text classification algorithm to the universe of U.S. patents granted since 1976. Linking patents to their industry of use and, ultimately, to commuting zones, we construct geographical intensities of newly available automation technology. The second contribution is a fresh assessment of the labor market effects of automation. In an econometric analysis, we show that in commuting zones where more newly-invented automation technology becomes available, the employment-to-population ratio increases. At the same time, there is a shift from routine manufacturing jobs towards non-routine service sector jobs. These results hold when we study only patents by universities, governments or foreigners, which are likely less responsive to developments in US labor markets than domestic firms.

While rising employment ratios in response to automation technology are good news, the benefits of automation may be unevenly distributed. We hope that future research will provide more insights in this respect. A more general contribution of this paper is that it pioneers a way of extracting trends in innovation which can also be used to study the effects of other technologies on the economy.

References

Acemoglu, D., U. Akcigit, and M. A. Celik (2014). Young, restless and creative: Openness to disruption and creative innovations. NBER Working Paper No. 19894.

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2016). Import competition and the great US employment sag of the 2000s. *Journal of Labor Economics 34*(1), 141–198.

Acemoglu, D. and P. Restrepo (2017). Robots and jobs: Evidence from US labor markets. NBER Working Paper No. 23285.

Akerman, A., I. Gaarder, and M. Mogstad (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics 130*(4), 1781–1824.

Antweiler, W. and M. Z. Frank (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance 59*(3), 1259–1294.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review 103*(5), 1553–1597.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review 103*(6), 2121–2168.

Autor, D. H., D. Dorn, and G. H. Hanson (2015). Untangling trade and technology: Evidence from local labor markets. *The Economic Journal 125*(584), 621–646.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in US wage inequality: Revising the revisionists. *The Review of Economics and Statistics 90*(2), 300–323.

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

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Beaudry, P., M. Doms, and E. Lewis (2010). Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas. *Journal of Political Economy 118*(5), 988–1036.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. V. Reenen (2017). Who becomes an inventor in America? The importance of exposure to innovation. NBER Working Paper No. 24062.

Bessen, J. and R. Hunt (2007). An empirical look at software patents. *Journal of Economics and Management Strategy 16*(1), 157–189.

Blanchard, O. J. and L. F. Katz (1992). Regional evolutions. *Brookings Papers on Economic Activity 1992*(1), 1–75.

Dauth, W., S. Findeisen, J. Südekum, and N. Woessner (2017). German robots - the impact of industrial robots on workers. CEPR Discussion Paper 12306.

Domingos, P. and M. Pazzani (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning 29*(2), 103–130.

Driscoll, J. C. and A. C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics 80*(4), 549–560.

Encyclopædia Britannica (2015). *Automaton*.

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

Gentzkow, M., B. T. Kelly, and M. Taddy (2017). Text as data. *Journal of Economic Literature*. (Forthcoming).

Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica 78*(1), 35–71.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2017). Bartik instruments: What, when, why and how. Unpublished manuscript.

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

Graetz, G. and G. Michaels (2015). Robots at work. CEPR Discussion Paper, 1335.

Gregory, T., A. Salomons, and U. Zierahn (2016). Racing with or against the machine? Evidence from Europe.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature 28*, 1661–1707.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). The NBER patent citation data file: Lessons, insights and methodologial tools. NBER Working Paper No. 8498.

IRF (2014). World robotics: Industrial robots. Technical report, International Federation of Robotics.

Kortum, S. and J. Lerner (1999). What is behind the recent surge in patenting? *Research Policy 28*(1), 1–22.

Kortum, S. and J. Putnam (1997). Assigning patents to industries: Tests of the Yale technology concordance. *Economic Systems Research 9*(2), 161–176.

Lai, R., A. D'Amour, A. Yu, Y. Sun, D. M. Doolin, and L. Fleming (2011). Disambiguation and coauthorship networks of the U.S. patent inventor database (1975 -2010). Fung Institute.

Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. *The Quarterly Journal of Economics 126*(2), 1029–1069.

Lybbert, T. J. and N. J. Zolas (2014). Getting patents and economic data to speak to each other: An 'algorithmic links with probabilities' approach for joint analyses of patenting and economic activity. *Research Policy 43*(3), 530–542.

Magerman, T., B. V. Looy, and X. Song (2010). Exploring the feasibility and accuracy of latent semantic analysis based on text mining techniques to detect similarity between patent documents and scientific publication. *Scientometrics 82*(2), 289–306.

Manning, C. D., P. Raghavan, and H. Schütze (2009). *An Introduction to Information Retrieval*. Cambridge University Press.

Michaels, G., A. Natraj, and J. van Reenen (2014). Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years. *The Review of Economics and Statistics 96*(1), 60–77.

Moretti, E. (2011). Local labor markets. *Handbook of labor economics 4*, 1237–1313.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Nagaoka, S., K. Motohashi, and A. Goto (2010). Patent statistics as an innovation indicator. In B. H. Hall and N. Rosenberg (Eds.), *Handbook of Economics of Innovation*, pp. 1083–1127. Amsterdam: North Holland.

Notowidigdo, M. J. (2011). The incidence of local labor demand shocks. NBER Working Paper No. 17167.

OECD (2003). *Turning Science Into Business: Patenting and Licensing at Public Research Organisations*. OECD Publishing.

OECD (2004). Understanding business patenting and licensing: Results of a survey. In *Patents, Innovation and Economic Performance. OECD Conference Proceedings*, pp. 89–110.

OECD (2011). *OECD Guide to Measuring the Information Society 2011*. OECD Publishing.

Pierce, J. R. and P. K. Schott (2016). The surprisingly swift decline of US manufacturing employment. *American Economic Review 106*(7), 1632–1662.

Scherer, F. M. (1984). Using linked patent and R&D data to measure interindustry technology flows. In Z. Griliches (Ed.), *R&D, Patents, and Productivity*. University of Chicago Press.

Scherer, F. M. and D. Harhoff (2000). Technology policy for a world of skew-distributed outcomes. *Research Policy 29*(4-5), 559–566.

Schmookler, J. (1966). *Invention and Economic Growth*. Harvard University Press.

Silverman, B. S. (2002). *Technological Resources and the Logic of Corporate Diversification*. Routledge.

Tolbert, C. M. and M. Sizer (1996). U.S. commuting zones and local labor areas: A 1990 update. Economic Research Service Staff Paper No. 9614.

USPTO (2014). Extended year set. All technologies (utility patents) report. Technical report, United States Patent and Trademark Office.

USPTO (2015). Types of patents. Technical report, United States Patent and Trademark Office.

WIPO (2016). WIPO IP Facts and Figures 2016. Technical Report No. 943E/16.

A Additional tables

Table A1: Yearly automation and non-automation patents

	#A	#P			#A	#P			#A	#P	
1976	16279	70194	(25%)	*1989*	27928	95565	(35%)	*2002*	77267	167400	(54%)
1977	15433	65215	(26%)	*1990*	25925	90421	(34%)	*2003*	82017	169077	(56%)
1978	15412	66087	(26%)	*1991*	28037	96561	(35%)	*2004*	84372	164384	(58%)
1979	11721	48840	(28%)	*1992*	29165	97472	(36%)	*2005*	69602	143891	(54%)
1980	14937	61815	(28%)	*1993*	30439	98385	(38%)	*2006*	91201	173822	(59%)
1981	15885	65770	(28%)	*1994*	33699	101695	(39%)	*2007*	83196	157331	(60%)
1982	15092	57877	(31%)	*1995*	35135	101431	(41%)	*2008*	86705	157788	(62%)
1983	14546	56863	(31%)	*1996*	40411	109654	(44%)	*2009*	92843	167463	(62%)
1984	17665	67212	(31%)	*1997*	40217	112019	(44%)	*2010*	121163	219835	(62%)
1985	19415	71668	(32%)	*1998*	57293	147577	(46%)	*2011*	126328	224871	(63%)
1986	19515	70867	(32%)	*1999*	58464	153591	(45%)	*2012*	147550	253633	(65%)
1987	24359	82963	(34%)	*2000*	61273	157595	(45%)	*2013*	163112	278507	(66%)
1988	22006	77938	(33%)	*2001*	64796	166158	(46%)	*2014*	178422	301643	(67%)
								total	**2158825**	**4971078**	**(43%)**

Note: #A: number of automation patents as classified by own algorithm; the patent totals #P are reported as counted by us in the patent files. The USPTO reports slightly different numbers for total patent counts on its website, but the difference is below 0.5% in all years.
Source: USPTO, Google and own calculations.

Table A2: Assignee's patents across technological categories, 1976-2012

Assignee	Patents (1000s)	Chem-ical	Comm., Comput.	Drugs, Med.	Electr., Electron.	Mech-anical	Miss-ing	Oth-ers
US firm	1875.7	17%	22%	11%	15%	13%	10%	12%
foreigners	1827.8	16%	20%	7%	19%	18%	10%	9%
universities	115.1	23%	12%	31%	17%	6%	6%	5%
governments	44.8	21%	14%	11%	21%	14%	11%	8%
missing	609.9	11%	7%	11%	9%	21%	11%	29%

Note: Technological classifications are based on USPC numbers and aggregated using the scheme by Hall et al. (2001).
Source: Lai et al. (2011), Hall et al. (2001) and own calculations.

Table A3: Share of automation patents after excluding patents

Assignee	Patents (1000s)	Chem-ical	Comm., Comput.	Drugs, Med.	Electr., Electron.	Mech-anical	Miss-ing	Oth-ers
US firm	1875.7	2%	21%	2%	8%	5%	6%	3%
foreigners	1827.8	1%	17%	1%	7%	7%	6%	2%
universities	115.1	2%	11%	4%	10%	2%	4%	2%
governments	44.8	2%	11%	1%	10%	4%	6%	2%
missing	609.9	1%	6%	2%	4%	6%	4%	5%

Note: Technological classifications are based on USPC numbers and aggregated using the scheme by Hall et al. (2001). This table excludes all patents based on the selected pharmaceutical and chemical industries as explained in text.
Source: Lai et al. (2011), Hall et al. (2001) and own calculations.

B Further robustness checks

B.1 Patent citations

Not all patents are of the same importance. Scherer and Harhoff (2000) show that the returns on innovation are highly concentrated, with the 10 percent most valuable patents accounting for around 80 percent of realized value. While Griliches (1990) argues that using a large number of patents partly addresses this concern, we can count how often a patent was cited by other patents as an indicator of its value. We use the patent citations files by Lai et al. (2011) until 2009. The number of citations per patents follow a well-known hump-shape, as newer patents are cited less frequently, but the propensity to cite has risen. Also, some industries (such as pharmaceutical and chemical patents) cite many more patents than others (such as electronics). To control for this, we demean citations across years and the broad technology classes defined by Hall et al. (2001). This is the "fixed effect" method proposed by Hall et al. (2001).

We then weight patents by how often they were cited and replicate our analysis. The analysis shows similar results: Manufacturing employment falls and service employment rises when more (citation-weighted) automation patents become available. The baseline effect on all employment becomes insignificant in this specification, but the interaction between automation and routine task share is still significant.

Table A4: Labor market effects of automation of citations-weighted patents

	Outcome: Employment-to-population				
	(1)	(2)	(3)	(4)	(5)
autoint	0.0896**	0.177*	0.0337	-0.0226	0.456**
	(0.0339)	(0.0990)	(0.0748)	(0.0834)	(0.212)
non-autoint		-0.0917	0.104	0.182*	0.106
		(0.124)	(0.0885)	(0.0949)	(0.107)
manufacturing			-2.391**	-1.458	-1.264
			(0.887)	(1.080)	(1.171)
population			0.192**	0.171**	0.146*
			(0.0846)	(0.0800)	(0.0771)
income			-1.337***	-1.285***	-1.222***
			(0.389)	(0.380)	(0.378)
non-white			-1.374***	-1.420***	-1.559***
			(0.129)	(0.136)	(0.123)
routine				-0.0417**	0.142*
				(0.0152)	(0.0772)
autoint $\times$ routine					-0.0117**
					(0.00424)
Observations	20,524	20,524	20,524	20,524	20,524
R^2	0.32	0.32	0.34	0.34	0.34

Note: The table replicates the regressions of Table 8 but using citations-weighted five-year sums of new automation and non-automation technology. We only include observations until 2009. Citations are adjusted with the Hall et al. (2001) fixed effect method. Driscoll-Kraay standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5: Labor market effects of citations-weighted automation patents for manufacturing and non-manufacturing employment

	(1)	(2)	(3)	(4)	(5)
A. Outcome: Manufacturing employment-to-population					
autoint	-0.0198	-0.110	-0.262***	-0.292***	0.0443
	(0.0186)	(0.0762)	(0.0373)	(0.0425)	(0.136)
non-autoint		0.0948	0.332***	0.375***	0.323***
		(0.0884)	(0.0418)	(0.0490)	(0.0488)
manufacturing			-2.880***	-2.374***	-2.237***
			(0.536)	(0.590)	(0.613)
population			-0.0310**	-0.0425***	-0.0595***
			(0.0151)	(0.0146)	(0.0169)
income			-0.792***	-0.765***	-0.721***
			(0.169)	(0.162)	(0.169)
non-white			-0.167	-0.192	-0.287
			(0.224)	(0.216)	(0.175)
routine				-0.0228***	0.106*
				(0.00440)	(0.0592)
autoint $\times$ routine					-0.00826**
					(0.00320)
Observations	20,520	20,520	20,520	20,520	20,520
R^2	0.19	0.19	0.24	0.24	0.26
B. Outcome: Non-manufacturing employment-to-population					
autoint	0.103***	0.295***	0.315***	0.297***	0.405**
	(0.0313)	(0.0722)	(0.0615)	(0.0705)	(0.157)
non-autoint		-0.202**	-0.258***	-0.233***	-0.250***
		(0.0832)	(0.0741)	(0.0799)	(0.0882)
manufacturing			0.519	0.818	0.861
			(0.673)	(0.776)	(0.829)
population			0.218**	0.211**	0.205**
			(0.0832)	(0.0782)	(0.0757)
income			-0.580*	-0.563	-0.549
			(0.332)	(0.334)	(0.336)
non-white			-1.275***	-1.289***	-1.321***
			(0.110)	(0.105)	(0.0997)
routine				-0.0132	0.0281
				(0.0169)	(0.0413)
autoint $\times$ routine					-0.00263
					(0.00296)
Observations	20,529	20,529	20,529	20,529	20,529
R^2	0.26	0.26	0.28	0.28	0.28

Note: The table presents fixed effects regressions for five-year changes in manufacturing employment-to-population and non-manufacturing employment-to-population. Automation and non-automation are citations-weighted. We only include observations until 2009. Citations are adjusted with the Hall et al. (2001) fixed effect method. The other control variables are defined as in Table 9. All regressions include state and year fixed effects and a constant. Driscoll-Kraay standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.2 Non-overlapping five-year periods

Table A6: Labor market effects of automation, five-year non-overlapping time periods

	Outcome: Employment-to-population				
	(1)	(2)	(3)	(4)	(5)
autoint	0.154***	0.324***	0.258**	0.246*	0.611***
	(0.0334)	(0.0892)	(0.126)	(0.134)	(0.162)
non-autoint		-0.173**	-0.0776	-0.0610	-0.0825
		(0.0740)	(0.125)	(0.137)	(0.133)
manufacturing			-1.191*	-1.031*	-0.886
			(0.616)	(0.601)	(0.595)
population			0.107***	0.102***	0.0908***
			(0.0256)	(0.0236)	(0.0241)
income			-0.644***	-0.627***	-0.601***
			(0.228)	(0.224)	(0.223)
non-white			-1.215***	-1.232***	-1.281***
			(0.447)	(0.444)	(0.427)
routine				-0.00751	0.132**
				(0.0136)	(0.0507)
autoint*routine					-0.00969***
					(0.00356)
Observations	5,663	5,663	5,663	5,663	5,663
R^2	0.40	0.40	0.41	0.41	0.41

Note: The table presents fixed effects panel data regressions using non-overlapping five-year equivalent changes in employment as percent of commuting zone population as the dependent variable. *autoint* and *non-autoint* are five-year sums of new automation technology and non-automation technology, as defined in the text. *routine* is the initial percentage of routine tasks in commuting zone employment. Further controls are the initial manufacturing employment share, the log of the initial commuting zone employment, the log of initial per capita income and the initial share of non-white citizens in the population. All regressions include state and year fixed effects and a constant. Standard errors clustered at the state level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

As an alternative to the five-year overlapping regressions presented in the main part of the paper, we show regression results for non-overlapping periods. These are 1977-1981, 1982-1986, 1987-1991, 1992-1996, 1997-2001, 2002-2006, 2007-2011 and 2012-2014, for which we compute five-year equivalents for the last period that covers only three years. The panel therefore comprises 8 time periods and 708 commuting zones. The results are similar to those presented in the main text. The coefficients in Table A6 are slightly larger and more significant than those presented in Table 8. The effects of automation for the two employment groups of Table A7 are also each slightly more positive than those of Table 9, but the finding of the contrary effect of automation is strongly supported.

Table A7: Labor market effects of automation for manufacturing and non-manufacturing employment

	(1)	(2)	(3)	(4)	(5)
A. Outcome: Manufacturing employment-to-population					
autoint	0.00382	-0.0365*	-0.110***	-0.137***	0.255***
	(0.0102)	(0.0216)	(0.0269)	(0.0297)	(0.0653)
non-autoint		0.0409**	0.164***	0.205***	0.179***
		(0.0187)	(0.0285)	(0.0328)	(0.0390)
manufacturing			-1.588***	-1.209***	-1.055***
			(0.222)	(0.223)	(0.231)
population			-0.00100	-0.0118	-0.0239*
			(0.0117)	(0.0128)	(0.0125)
income			-0.710***	-0.666***	-0.645***
			(0.139)	(0.143)	(0.140)
non-white			-0.149	-0.191	-0.237
			(0.198)	(0.208)	(0.190)
routine				-0.0181***	0.133***
				(0.00402)	(0.0262)
auto*routine					-0.0104***
					(0.00184)
Observations	5,660	5,660	5,660	5,660	5,660
R^2	0.21	0.21	0.23	0.24	0.26
B. Outcome: Non-manufacturing employment-to-population					
autoint	0.137***	0.363***	0.368***	0.382***	0.260
	(0.0317)	(0.0778)	(0.113)	(0.124)	(0.160)
non-autoint		-0.230***	-0.253**	-0.274**	-0.267**
		(0.0612)	(0.104)	(0.121)	(0.123)
manufacturing			0.321	0.115	0.0700
			(0.425)	(0.419)	(0.424)
population			0.105***	0.111***	0.114***
			(0.0246)	(0.0207)	(0.0210)
routine				0.00957	-0.0371
				(0.0138)	(0.0390)
autoint*routine					0.00324
					(0.00239)
income			0.0393	0.0176	0.0101
			(0.172)	(0.178)	(0.176)
non-white			-1.032***	-1.013***	-0.992***
			(0.229)	(0.219)	(0.219)
Observations	5,662	5,662	5,662	5,662	5,662
R^2	0.36	0.36	0.37	0.37	0.37

Note: The table presents fixed effects panel data regressions for non-overlapping five-year equivalent changes in manufacturing employment-to-population and non-manufacturing employment-to-population. See Table 8 for variable definitions. All regressions include state and year fixed effects and a constant. Standard errors clustered at the state level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.