On the measurement of occupational task content

Matthias Haslberger

Abstract

Which tasks workers perform on their jobs is critical for how technological change plays out in the labour market. This crucial insight sparked a large literature on routine-biased technological change (RBTC) which argues that occupations with a high share of repetitive and codifiable tasks are at risk of being automated. The RBTC hypothesis was formulated based on the US experience and stipulates that technological change leads to employment and wage polarisation; however, macro-comparative research on the issue is rare. My paper therefore adopts a broader scope, and in doing so makes the case for rethinking how we operationalise occupational task content to analyse occupational change.

Based on survey data from up to 27 European countries over the period from 2000 to 2015, I show that previous work on operationalising occupational task content suffers from conceptual and empirical problems. I proceed to construct alternative measures of routine task intensity and task complexity at the ISCO-88 2-digit level. Comparing them to existing operationalisations, I show that my indices lead to improvements in several critical areas. Firstly, my task dimensions capture the essence of the routine-biased and skill-biased technological change arguments and therefore have a straightforward theoretical interpretation. Secondly, the indices are operationalised in a way that corresponds to the underlying concepts of routine-intensity and complexity, leading to a better alignment of theory and measurement than before. Finally, instead of a one-size-fits-all approach, my indices leverage the richness of the underlying data by allowing researchers to analyse within-occupation change over time and country-differences. My paper can therefore contribute to a more sociologically informed understanding of the process of technological change. My indices will benefit both the sociological literature and the labour economics literature, as they will be helpful for investigating the nature of recent employment trends in OECD countries.
I. Introduction

Routine-biased technological change (RBTC) has been the dominant explanation in the literature on technological change and inequality in developed countries over the last 15 years or so. The RBTC hypothesis was formulated based on the US experience in a seminal paper by Autor, Levy, and Murnane (2003; hereafter ALM). The authors develop a set of measures of occupational task content which have since, by virtue of academic primogeniture, enjoyed a near-monopoly in the task literature. These measures, or a simpler summary index devised by (Autor, Katz, and Kearney 2006) and (Autor and Dorn 2013, hereafter AD), have been rarely scrutinised, with other authors contenting themselves to follow recommendations to use “off-the-shelf measures” (e.g., Goos, Manning, and Salomons 2014; Mahutga, Curran, and Roberts 2018).

More recently, however, scholars have begun subjecting these widely used measures to greater scrutiny. Invariably, they find conceptual, operational, and empirical problems with the standard indices (Fernández-Macías and Hurley 2017; Sebastian and Biagi 2018). My paper builds on and adds to this recent literature. I argue that in existing research on employment change within the task framework, the nature of occupational tasks is insufficiently theorised, and operationalised in an even less convincing fashion. It is therefore precisely this central building block of the task framework which requires theoretical and empirical refinement. The main contributions of my new measures are threefold: firstly, I propose two indices of routine task intensity and task complexity which correspond to prevalent theories of RBTC and SBTC. Secondly, my measures of these indices better align with the respective definitions, making sure we measure what we purport to be measuring. Finally, instead of a one-size-fits-all approach, my indices leverage the richness of the underlying data, affording researchers more flexibility by using different configurations. To this end, I first provide an overview and critique of the existing literature and operationalisations in section 2, followed by the theoretical case for an alternative approach in section 3. Section 4 describes the data and the strategy for constructing the new measures of occupational task content. Section 5 is devoted to quantifying the differences between the existing and new measures, and backs up the argument from section 3 empirically. Finally, section 6 concludes.

II. Overview and critique of existing operationalisations

What are routine tasks?

The task-based approach to occupational change came to prominence with the seminal paper by (Autor, Levy, and Murnane 2003) who proposed the first model of routine-biased technological change. This approach soon supplanted the older model of skill-biased technological change (SBTC) where researchers focused on worker characteristics such as average levels of education. In one of the seminal articles of the task literature, a task is defined as a “unit of work activity that produces output”, whereas “[a] skill is a worker’s endowment of capabilities for performing various tasks” (Acemoglu and Autor 2011, 1118). Essentially, the
RBTC argument predicts a reallocation of employment based on the task composition of occupations. The core of ALM’s argument is worth quoting in full and stipulates

“(1) that computer capital substitutes for workers in carrying out a limited and well-defined set of cognitive and manual activities, those that can be accomplished by following explicit rules (what we term “routine tasks”); and (2) that computer capital complements workers in carrying out problem-solving and complex communication activities (“nonroutine” tasks).” (ALM, p.1280).

Thus, production tasks are allocated to workers or capital (‘machines’, which in this context includes hardware and software) based on comparative advantage in performing the respective tasks, where machines have an advantage in performing routine tasks and workers have an advantage when it comes to non-routine tasks. Elsewhere in the paper, ALM argue that routine tasks “require [the] methodical repetition of an unwavering procedure” (ALM, 1283), thus introducing the element of repetitiveness. Other definitions require that routine tasks be “expressible in rules such that they are easily programmable and can be performed by computers at economically feasible costs” (Spitz-Oener 2006, 239) or define routine-intensity as “the extent to which an occupation is automatable or codifiable” (Caines, Hoffmann, and Kambourov 2017, 302). As Fernandez-Macias and Hurley (2017) point out, there are several more aspects to the term “routine” as it is commonly understood: most obviously, it refers to repetitiveness and standardisation – as do the definitions above, but not every repetitive task is necessarily routine, and subjectively “boring” tasks may be perceived as routine even if they are not particularly repetitive in the strict sense. Hence, conceptually, the existing literature focuses on codifiability and repetitiveness as the distinguishing features of routine tasks. Generally speaking, these definitions are technology-driven, in line with the RBTC theory which predicts that technology will replace workers in routine tasks. While the emphasis on codifiability is generally sensible, there is a danger of circularity: if routine tasks are defined as codifiable tasks that are being replaced by machines, technological change is by construction routine-biased.¹

Existing operationalisations of occupational task content

The RBTC hypothesis can be assessed empirically by quantifying which tasks make up an occupation and investigating whether the prediction of declining employment in more routine-intensive occupations is borne out. Broadly speaking, there are two popular approaches in the literature, roughly corresponding to the disciplinary boundary between economics and sociology. ALM identify five task dimensions in their empirical analysis: routine manual, routine cognitive, non-routine interactive, non-routine analytical, and non-routine manual tasks. (Autor, Katz, and Kearney 2006) have consolidated this approach into a framework comprising

¹ For example, (Acemoglu and Autor 2011, 1076) define routine tasks as those which are “sufficiently well understood that they can be fully specified as a series of instructions to be executed by a machine.”
three types of tasks which (Autor and Dorn 2013) have subsequently formalised. This framework, which classifies tasks as routine, abstract, or manual, has since become mainstream in the economics literature and is used in numerous studies by other authors (see Sebastian and Biagi 2018 for an overview). However, this dominance is not the result of rigorous debate how to best measure occupational task content, but mainly of convenience turned convention. A few more recent studies, more sociologically influenced, retain the basic logic of the task framework of focusing on task contents rather than worker characteristics, but choose different task dimensions, variables, data sources or units of analysis to operationalise their theories. (Fernández-Macías and Hurley 2017) are prominent proponents of one such approach. The following table contrasts the most influential operationalisations of the RBTC theory, including which other studies follow a similar approach as these landmark studies.
<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Autor, Levy and Murnane 2003</th>
<th>Fernandez-Macias and Hurley 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>routine manual; routine cognitive; non-routine interactive; non-routine analytic; non-routine manual.</td>
<td>Routine; cognitive (also: social interaction; trade intensity)</td>
</tr>
<tr>
<td>Variables</td>
<td>Routine manual: finger dexterity.</td>
<td>Routine: repetitive hand or arm movements; repetitive hand movements of less than 1 or 10 minutes; monotonous tasks; dealing with unforeseen problems.</td>
</tr>
<tr>
<td></td>
<td>Routine cognitive: adaptability to situations requiring the precise attainment of set limits, tolerances and standards.</td>
<td>Cognitive: complex tasks; use of computers at work; use of internet at work; number of years of formal education necessary.</td>
</tr>
<tr>
<td></td>
<td>Non-routine interactive: direction, control and planning.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-routine analytic: quantitative reasoning requirements.</td>
<td></td>
</tr>
<tr>
<td>Unit of analysis</td>
<td>3-digit census occupations (US)</td>
<td>“jobs”: 2-digit occupation, 2-digit industry cells (EU-28)</td>
</tr>
<tr>
<td>Variants also used by</td>
<td>Spitz-Oener 2006 (for Germany, dimensions only); Goos and Manning 2007 (for UK); Autor, Katz and Kearney 2008; Acemoglu and Autor 2011; Autor and Dorn 2013; Autor and Handel 2013 (dimensions only); Goos, Manning and Salomons 2014 (for 16 EU countries); Naticchioni, Ragusa and Massari 2014 (for 12 EU countries); Mahutga, Curran and Roberts 2018 (for LIS countries).</td>
<td>Eurofound 2014, 2017</td>
</tr>
</tbody>
</table>

Table 1: Source: Author's analysis of the literature quoted in the table. Studies following the ALM approach after Autor and Dorn 2013 (bold) generally employ their consolidated version of the ALM approach.
The overview in table 1 illustrates the proliferation of empirical studies based on the approaches developed by ALM and AD. Their methodologies have been adapted to a range of contexts beyond the United States, including studies of individual European countries and comparative studies. Some authors, like (Spitz-Oener 2006) and (Autor and Handel 2013), only adopt the analytical framework but use different data to calculate task intensity scores while others, such as (Goos, Manning, and Salomons 2014) embrace the approach wholesale. A few recent studies attempt to introduce alternative typologies and operationalisations into the task literature but by and large they have so far not been very successful in overcoming the predominance of ALM and AD, especially in economics. Probably the most promising alternative approach is that of (Fernández-Macías and Hurley 2017) which has also been used in a number of (Eurofound 2014, 2017) reports in which the same authors were involved. Concerned with employment in EU countries, it is devised with a comparative approach in mind. Further key differences from the AD approach are their use of workers’ self-assessment as the basis for measures of task content (similar to Spitz-Oener 2006) and the level of analysis which in their case is the ‘job’-level which refers to 2-digit occupations in 2-digit sectors. Another interesting approach is that of (Caines, Hoffmann, and Kambourov 2017), who follow AD in the way they measure routine but instead of abstract and manual tasks juxtapose task complexity. Also methodologically, by calculating a task complexity measure based on a principal component analysis, they depart from AD; however, their approach has not yet found many followers. This overview shows that the measurement of occupational task content indeed overwhelmingly follows the example of the first papers in the field. In the following section I will subject these approaches to a critical review and identify the manifold problems which make them dubious examples to follow. Based on these deliberations, I will then propose a novel way of operationalising occupational task content which improves upon existing approaches in important ways.

A critique of existing operationalisations of occupational task content

There are a host of conceptual and empirical problems with the measurements of occupational task content just presented which in my view have the potential to seriously undermine the inferences that can be made based on them. I am not the first to point out most of these issues – I take up several criticisms that have been voiced by comparative sociologists and economists – but to my knowledge they have not been addressed jointly in any previous research. For example, (Caines, Hoffmann, and Kambourov 2017) maintain that task characteristics such as complexity may be more important than routine-intensity and (Fernández-Macías and Hurley 2017) argue that the operationalisation of routine intensity in the influential papers of ALM and AD does not capture the essence of the concept as it is defined in these same papers. Furthermore, existing measures fail to account for potential differences across countries and time. Finally, several authors have commented on the relative merits of expert- and survey-based assessments of task content, and I argue that there are indeed strong arguments for preferring the latter. Taken together, this suggests that the measurement of occupational task content needs to be refined both conceptually and empirically. My argument in this section is

---

2 It is worth reiterating that AD is essentially a consolidated version of ALM: their three task categories are derived from the five categories in ALM and they use the same DOT variables as ALM.
largely a logical one; I will then below produce the empirical evidence to suggest that these problems can be addressed.

**Use of ill-defined auxiliary task dimensions**

The uniting feature of all approaches discussed above is their interest in occupational routine-intensity since this is the main variable of interest of for RBTC theory. However, all studies define one or more secondary axes of occupational tasks that are of interest. For example, in ALM that is the cognitive-manual axis and in (Fernández-Macías and Hurley 2017) simply a cognitive task dimension. In the papers following the AD approach, an abstract and a manual task dimension are also specified. (Caines, Hoffmann, and Kambourov 2017) juxtapose routine and complex tasks. While the choice of axes is predicated on the nature of the research question and hence there is no obvious right or wrong, I argue that, with some exceptions, these auxiliary axes do not serve a well-defined purpose. Only (Caines, Hoffmann, and Kambourov 2017) formulate a theory of task complexity in relation to routine-intensity and technological change. In most other analyses, for example the measure of cognitive task intensity does not add much of substantive interest; it serves a merely descriptive purpose. For instance, ALM expect routine occupations to decline regardless of whether they are cognitive or manual; they do not posit any independent relationship between technological change and cognitive and manual task inputs. Of course, there is no logical necessity that there should be such a relationship. Nevertheless, if the measure of routine-intensity is used to operationalise RBTC, whatever auxiliary measure is part of the analysis ideally should have some independent theoretical interpretation. In particular, since RBTC is an alternative to SBTC, it would be eminently helpful to have a measure for the latter that is constructed in a similar manner as the measure of RBTC.

**Use of variables that do not capture key concepts**

Important though the definition of task dimensions is, the choice of variables to operationalise these concepts is just as crucial. Hence, a central element of my critique is that the variables used to operationalise cognitive and manual routine tasks in ALM completely fail to capture key aspects of the notion of routine, most importantly, repetitiveness. Granted, their multifaceted nature complicates a meaningful operationalisation of routine tasks, yet existing measures fail to properly capture even the restricted definition they are based on. Routine tasks, we recall, are usually defined in the RBTC literature as tasks that are codifiable and/or repetitive. However, the measures which ALM and AD use do not properly capture these characteristics. For example, ALM measure high cognitive routine intensity with two items from the DOT: adaptability to situations requiring the precise attainment of set limits, tolerances and standards, and manual routine intensity with finger dexterity. While in the former a certain element of repetitiveness is discernible (the example tasks ALM list in their appendix 1 would certainly intuitively be considered routine, if not necessarily cognitive), the relationship between finger dexterity and codifiability and repetitiveness seems altogether questionable. In AD, the measure for routine tasks is the simple average of those two variables. Thus, even

---

3 It is also the case that these variables are invariably closely related to routine intensity itself; however, this is a problem that it seems impossible to circumvent without resorting to principal component analysis.
though repetitiveness is at the core of the concept, it does barely feature in the variables used to measure routine intensity in most research.

Fundamentally, the measures used in ALM and AD seem to have been chosen to identify occupations, not tasks. They are suited to pointing out the high routine-intensity of clerical and manufacturing occupations, but they do not get at the theoretical concept of routine-intensity. The item “adaptability to situations requiring the precise attainment of set limits, tolerances, or standards” (STS) as the sole measure of routine cognitive tasks is clearly geared towards low-level clerical jobs and jobs in the manufacturing sector (Autor, Levy, and Murnane 2003, 1323). The use of finger dexterity as a measure of manual routine intensity rests on the assumption that tasks which involve fine movements and coordination are repetitive and can be automated – a questionable assumption for example with regard to musicians and artisans who often require a great deal of finger dexterity but whose job tasks would generally not be considered repetitive and automatable. Overall, the variables used in ALM and AD appear to have been chosen not so much with the abstract concept of routine in mind as with a set of purportedly routine occupations.

Failure to account for change within occupations

The data used by AD, FMH, and most subsequent studies make it impossible to account for change over time within occupations. While the DOT was periodically updated, these revisions happened very infrequently and most published studies on RBTC measure occupational task content at one point in time and do not analyse within-occupation change. There are nevertheless some studies which suggest that within-occupation change cannot be ignored in the analysis of the effects of technological change. ALM analyse the “extensive” (between-occupation) as well as the “intensive” (within-occupation) margin of changes in job task requirements by using DOT data from 1977 and 1991. Within occupations, they find a pattern of change which cuts across the routine/non-routine divide: cognitive-analytic inputs have declined whereas interactive-manual task inputs have increased, independent of the level of routine inputs, even though the changes are relatively small compared to the extensive margin. (Spitz-Oener 2006), on the other hand, using the same analytical framework but data from a large German survey, finds that in Germany within-occupation changes account for most of the change in aggregate task requirements. Similarly, (Becker and Muendler 2015, 593) remark on the issue of task change within occupations that “the previous practice of mapping tasks to occupations in a time and sector invariant manner used to obfuscate this source of variation despite its dominance.” It is therefore clear that as the prevalence of occupations changes, so does their nature. A failure to account for this would result in underestimating the impact of technological change on the labour market.

---

4 While employment in clerical occupations declined substantially in the US, it was quite stable in most other countries. Since the RBTC hypothesis was originally formulated with the US experience in mind, this begs the question if not the observed decline of clerical occupations primed researchers to look for a measure of RTI that could accommodate this pattern.

5 (Sebastian and Biagi 2018) provide an overview of the years from which task data in various studies are taken.
Failure to account for differences between countries

Furthermore, all existing comparative studies fail to account for potential differences in task content between countries. Although (Eurofound 2014) rightly point out that the RBTC hypothesis focuses on attributes of job tasks which should be relatively similar across developed countries, the possibility that some occupations differ between countries as a result for example of regulatory requirements should not be dismissed. Indeed, it is not at all clear that the tasks that a worker in a particular occupation performs are the same across countries. For example, (Eurofound 2014) find that there are differences across countries regarding the demand for routine and cognitive tasks in a job, relative to that country’s overall task distribution. They argue that the degree of variation of task content is smaller than of skill and wage levels and hence country-specific wage and skill rankings are more important. True though this may be, it does not mean that country-specific task measures are unnecessary even if they do not come at the expense of the former.

However, many studies simply use the measures of ALM and AD outside the United States, such as (Goos and Manning 2007) and (Goos, Manning, and Salomons 2014), who use crosswalks to transfer the task scores from US census occupations to the SOC90 and ISCO-88 occupational classifications, respectively. (Fernández-Macías and Hurley 2017), while they do not use data from a country that is not part of the analysis, still only calculate one measure of task intensity for all countries and don’t exploit the fact that the EWCS would allow them to calculate country-specific task intensity scores. Yet ideally, country-specific measures of task content should be used for more detailed analyses. Where such measures are not possible due to data limitations, pooling the data across countries as in FMH would still reduce bias compared to imposing one country’s occupational task requirements on all others as in (Goos, Manning, and Salomons 2014).

Use of expert-coded measures

Most task measures in the RBTC literature are derived from the Dictionary of Occupational Titles (DOT), in which expert coders assign scores on indicators that characterise occupations in the US. Of the studies reviewed above, only (Spitz-Oener 2006) and FMH use survey data to operationalise occupational task content. There are pros and cons to using expert-coded data as opposed to survey data, but on balance I argue that the former outweigh the latter. Survey data, by asking people what they actually do in their job, are conceptually closer to what the task-based approach is all about; furthermore, survey data can provide a sense of the variability of tasks within an occupation. (Spitz-Oener 2006) moreover points out that it is an established finding that experts tend to underestimate the true changes in task content.

On the other hand, surveys suffer from the problem that people may answer questions such as “Does your job require you to perform complex tasks?”, very differently not based on the actual tasks they perform but on their personal characteristics. While this could conceivably lead to problematic inconsistencies, research from the US using the DOT and the O*NET database which likewise relies on workers’ self-assessments, suggests that occupational task content is judged similarly by experts and workers (Caines, Hoffmann, and Kambourov 2017). Overall,

---

6 If that wasn’t the case, the question would remain who is closer to the truth, the experts or the workers.
the best way to find out what people do at work still seems to be to ask the workers themselves, provided the appropriate questions are asked in the survey.

In this section, I have subjected existing measures of occupational task content to a thorough critique mainly on conceptual and empirical grounds. While not every criticism applies to every measure that has been used in the literature, there is none that is immune. Therefore, conditional on the acceptance of at least the majority of my points, it follows that there is a need for a better way of measuring occupational task content. The remainder of this paper is an attempt to provide such a measure.

Towards better measures of task content

With the measures developed here, I intend to address as far as possible the criticisms I laid out above. In contrast to most of the existing literature, my task intensity indices focus on routine intensity and task complexity. While routine intensity remains of central theoretical importance but will be operationalised in a novel way, I devise an index of task complexity rather than cognitive intensity. With regard to technology, the routine index should capture to what extent the tasks that make up an occupation can be standardised, codified, and ultimately performed by machine labour (the RBTC hypothesis), whereas the complexity index should capture the degree to which the tasks required in an occupation are complemented by technology (the SBTC hypothesis).

Use of meaningful auxiliary task dimension

The first conceptual problem of much existing research on RBTC is the lack of a well-defined auxiliary dimension to the routine dimension. I propose a measure with the aim to enable an analysis of SBTC in parallel to RBTC. Following (Caines, Hoffmann, and Kambourov 2017), I call this the task complexity dimension. This approach is motivated by the fact that SBTC and RBTC are both attempts to make sense of the same underlying phenomenon, technological change, and to ascertain its nature and implications. More commonly, the SBTC hypothesis is assessed by looking at educational requirements of occupations. However, for consistency within a task-based framework, a measure based directly on self-reported tasks is a suitable alternative under the assumption that complex occupations on average require more education. Therefore, it would be beneficial to have measures that allow researchers to treat both as alternative hypotheses and investigate where their empirical implications overlap and differ.

By task complexity I mean the demand for higher-order skills such as effective communication, abstraction, and decision making (see also Caines et al. 2017). Occupations that comprise many tasks requiring these skills are likely to be highly skilled and less likely to be replaced by technology, for machines and AI cannot (yet) perform such tasks (Grace et al. 2018). On the contrary, these higher-order skills are prone to be complemented by modern technology: thanks

---

7 The precise content and label of these tasks varies: (Frey and Osborne 2017) identify “engineering bottlenecks” relating to perception and manipulation tasks, creative intelligence tasks, and social intelligence tasks, while (McKinsey Global Institute 2017) predicts that applying expertise, interacting with stakeholders, managing people, and unpredictable physical activities will see increased demand. Other authors such as (Bostrom 2014) and (Brynjolfsson and McAfee 2014) offer more or less similar assessments of the types of tasks that will remain immune to automation at least for some time.
to it, effective communicators can reach wider audiences, scientists have powerful tools at hand that facilitate abstraction and induction, vast amounts of easily accessible data enable more informed decision making, and so on. Hence, task complexity is suitable for measuring the prevalence of SBTC. Notably, routine task intensity and task complexity are not just two sides of the same coin, although they are of course related: there are a number of occupations that are relatively high-routine which also require performing a considerable amount of complex tasks, for example some health professionals and service-oriented occupations. (Caines, Hoffmann, and Kambourov 2017) argue that these occupations recruit from a different pool of workers than occupations that are in competition with computers. Thus, with these characteristics, task complexity is expected to be negatively correlated with routine intensity but is nevertheless analytically distinct.

Task complexity is furthermore different in theoretical and practical terms from the more familiar concept of cognitive intensity. ALM, who introduce the category of cognitive tasks, never properly define the term. They refer to nonroutine cognitive tasks as tasks “requiring flexibility, creativity, generalized problem-solving, and complex communications” but offer no definition of routine cognitive tasks (Autor, Levy, and Murnane 2003, 1284). Their description of nonroutine cognitive tasks comes quite close to my definition of complex tasks but their operationalisation does not actually measure the demand for these skills. Their indicator of routine cognitive tasks, measured as the demand for “adaptability to work requiring set limits, tolerances, or standards” (Autor, Levy, and Murnane 2003, 1293), has very little to do with task complexity as I – and indeed they – have defined it. Thus, rather than a poorly conceived auxiliary measure to augment the RTI index, task complexity is an independent measure which allows me to test the SBTC hypothesis alongside the RBTC hypothesis.

*Use of variables that capture key concepts*

The most common operationalisation of routine task intensity does not in fact measure RTI as it is defined in the same literature: the key concepts of repetitiveness and codifiability are insufficiently captured. Yet, as shown in table 1, numerous other influential studies either use essentially the same set of measurements as AD, slight variations thereof, or entirely different variables while relying on the same concepts, as Fernández-Macías and Hurley (2017) point out. Thus, there is a real need to better align the concept of routine-intensity and its measurement and place it on more solid empirical foundations. (Fernández-Macías and Hurley 2017) have taken an important step in this direction and my discussion of this issue relies heavily on their previous work; yet, I believe that their measure of routine-intensity can be improved further, even when using the same data source, the EWCS.

To better align the theoretical concept and its operationalisation, my routine index includes the following items: whether a job involves (1) repetitive arm or hand movements, (2) short repetitive tasks of less than 1 or 10 minutes, (3) monotonous tasks, and (4) meeting precise quality standards. Considering the discussion above, the four items included here are arguably

---

8 The two DOT items used to code nonroutine cognitive tasks are the requirement for direction, control, and planning activities and quantitative reasoning requirements. It is immediately apparent that this does not match very well with the definition given above.

9 These items correspond to questions number 30e, 48a, 48b, 53a, and 53d in the EWCS.
less occupation-specific than the ones used by ALM. At the same time, this measure affords the aspect of repetitiveness its due importance, which has been missing in most previous operationalisations even though it was at the core of the definition provided by ALM when they introduced the concept of routine tasks. Just looking at the choice of variables, the index in this form closely resembles the RTI measure of FMH. In my operationalisation of RTI, the key difference is the inclusion of item 4 (meeting precise quality standards) instead of “solving unforeseen problems on one’s own” which I show correctly belongs to the complexity index.

The task complexity index, we recall, aims to measure the demand for higher-order skills such as effective communication, abstraction, and decision making. It includes the following items from the EWCS: whether a job entails (1) working with computers, tablets, smartphones, etc., (2) solving unforeseen problems on one’s own, (3) complex tasks, and (4) learning new things.\(^\text{10}\) The task complexity index is where I depart further from (Fernández-Macías and Hurley 2017) who focus on cognitive intensity rather than task complexity as the second dimension of occupational task content. Consequently, the only overlapping question is item 3, whether a job involves complex tasks. Furthermore, while FMH include the years of formal education necessary to perform a job adequately, I maintain a consistent focus on self-reported job tasks. Instead, I include the additional question whether a job involves learning new things, since on-the-job learning is a key characteristic of complex jobs (Feng and Graetz 2015). The items that make up this index are the most appropriate questions from the EWCS to capture the demand for higher-order skills such as effective communication, abstraction, and decision making that have been complemented by technological advances (Caines, Hoffmann, and Kambourov 2017). Overall, using variables that truly capture the prevalence of routine and complex tasks is an important and long overdue advance in the comparative literature on RBTC which is finally possible with the EWCS data.

Account for change within occupations

With the EWCS, which is conducted quinquennially for the period from 2000 – 2015, it is possible to analyse change within occupations in most European countries.\(^\text{11}\) (Eurofound 2014) were the first to use the EWCS to construct indices of job task content but they only use the 2010 wave and ignore the temporal component of occupational change. In contrast, I am using data from the four most recent waves (3 – 6) to account for change within occupations over a period of 15 years. In principle, any data source with consistent occupation-level data for several points in time can be used to account for change over time within occupations. However, the DOT was only updated infrequently and was finally replaced by the O*NET database. It is therefore impossible to develop a time series of occupational task content based on the variables in ALM. Generally, regularly conducted surveys would appear to be the logical candidate for the analysis of occupational task requirements over time. For example, in Germany the Qualifications and Career Survey has been used for this purpose (Spitz-Oener 2006; Becker and Muendler 2015). However, with the interest in comparative analysis in mind, the EWCS is the most appropriate data source for my purposes.

\(^\text{10}\) These items correspond to questions number 30i, 53c, 53e, and 53f in the EWCS.
\(^\text{11}\) The first wave of the EWCS was conducted in 1990 but important countries and questions are only included in later waves from 2000 onwards.
Account for differences between countries

The EWCS data can not only be used to trace change within occupations, but also to analyse differences between countries. The gradually enlarged sample of the EWCS covered 35 European countries in its most recent wave, with a sample size between 1,000 and 3,300 individuals depending on country size. With these characteristics, a country-level analysis at the level of 2-digit occupations is feasible. To my knowledge, there is currently no study that allows for country differences in the routine intensity and complexity of occupations. Even though country differences in this domain are likely smaller than for example with regard to average occupational wages (Eurofound 2014), this represents an advance in the literature. In particular, leveraging country differences in task intensity measures will be useful in future comparative research seeking to establish a relationship between task change and employment change through regression analyses. The entirely European nature of the sample restricts the scope for such comparative analyses to exclude countries such as the US, but compared to the prevailing practice of using task content measures based on US data for all countries, my broader approach promises more robust insights (Mahutga, Curran, and Roberts 2018).

Use of survey data

Weighing the advantages and drawbacks of using survey data over expert-coded data, I have argued that the former predominate. Using the EWCS for my measures helps to realise these advantages. Instead of one number assigned to an occupation by an outsider, the score on each item is the result of many (often thousands of) practitioners of an occupation evaluating what they do on their job, and how they do it. An added benefit of this is that we get an indication of the degree of variation within a job. With this I follow the practice of a number of scholars mainly from Europe who have eschewed the American approach and have used survey data for the analysis of occupational task content all along.

III. Data and construction of the indices

This section describes in detail the main data source for this paper, the waves 3 – 6 of the European Working Conditions Survey (EWCS) which provides workplace information for the years 2000, 2005, 2010, and 2015. This means that instead of a single country-year, at the most general level my measures of task intensity are based on up to four waves of individual-level data from up to 35 countries. The first EWCS was launched in 1990 in the EU12 and was then gradually expanded. The survey is conducted quinquennially and in its latest edition included 35 countries (EU28, Norway, Switzerland, Albania, FYROM, Montenegro, Serbia, Turkey). For the sake of consistency, in most analyses below I restrict the sample to include only the EU-27 or EU-15. The target sample size is at least 1000 individuals in each country and wave,

---

12 Some authors argue that questions how people perform their jobs should be distinguished from questions about what they do, but in my view the two are so closely related that a neat distinction is not practical.
13 The correlation of the occupational scores in the EU-27 with the full EWCS sample is 0.999, dispelling concerns that the addition of generally less economically advanced countries would distort task intensity scores in later waves.
with only a few exceptions in small countries.\textsuperscript{14} The total N amounts thus to 151,100 individuals. The EWCS characterises the task profile of occupations through the surveyed workers’ answers to a set of relatively objective questions (for example, “Does your main paid job involve repetitive hand or arm movements?”) and a number of more subjective items (for example, “Does your main paid job involve complex tasks?”). The full wording of the questions and summary statistics can be found in the appendix.

For both routine task intensity and task complexity, I calculate three indices each: The overall index uses all available data and is pooled across countries and waves. In addition, I calculate a wave-specific version of the overall index, and a country-specific index with data for the respective country pooled over all available waves. I follow the approach in the majority of the literature and perform my analysis at the level of occupations, rather than occupation-sector cells (“jobs”) like in FMH. This entails a certain loss of precision compared to FMH but in my view is warranted for several reasons. On the one hand, only by restricting the analysis to the 27 2-digit ISCO-88 occupations is the analysis of country- and wave-specific scores feasible with the amount of data available. On the other hand, the method of FMH, using a sample of about 43,000 to populate 3,123 out of 3,784 hypothetical job cells, means that for a large number of small jobs, the task scores will be based on a single or uncomfortably few observations. FMH argue that this doesn’t affect their findings because these jobs only represent a tiny fraction of employment, but this begs the question, why bother in the first place? On the negative side, if occupational tasks differ between sectors, not accounting for sectoral differences may cause researchers to mistake changes in the sectoral composition of the economy for changes in occupational tasks. However, at present it is not possible to address all potential sources of variation at once, hence my decision to follow most of the previous literature and focus on the occupational level.

The indices are constructed by standardising the constituent variables to have a mean of 0 and a standard deviation of 1, and then first averaging across individual survey respondents and subsequently across 2-digit ISCO codes. This means that the indices are not technically bounded, although in practice all scores fall within a range from -1 to +1. Principal component analysis, which is sometimes used in the literature, is not useful in the present case because of the low number of items that make up each index. Thus, the routine index $r_{score2d}$ for occupation $o$ is calculated as:

$$r_{score2d, o} = \frac{\Sigma_{i \in I_o} (\Sigma_{j \in J} ewcs_{ji})}{I_o},$$

where $ewcs_{ji}$ is the value for individual $i$ on question $j$, $J$ is the set of items used to calculate $r_{score2d}$ and $I_o$ is the set of individuals with occupation $o$. Analogously, the complexity index $c_{score2d, o}$ is calculated as

$$c_{score2d, o} = \frac{\Sigma_{i \in I_o} (\Sigma_{l \in L} ewcs_{li})}{L_o},$$

\textsuperscript{14} The target sample size was 500 in Luxembourg in 2000, and 600 in Cyprus, Estonia, Luxembourg, Malta, and Slovenia in 2005. In all other country-years, the size of the target sample varies between 1,000 and 4,000.
where \( L \) denotes the set of indicators used to calculate the index. For the wave- and country-specific indices, the respective subscripts \( w \) and \( c \) have to be added to the formula. Given these indices, I classify occupations as either high-routine or low-routine, and high-complexity or low-complexity, depending on whether they are above or below the median of the respective measure. Thus, I obtain a 2x2 matrix of four types of occupations, as depicted in figure 1.

![Figure 1: Schematic view of the categorisation of occupations.](image)

While I cannot fully solve everything that is problematic about existing approaches, these indices represent a substantive improvement and can lead to more robust insights. Moreover, different versions can be used depending on the research question at hand. The table below provides an overview of the various versions of my index.

<table>
<thead>
<tr>
<th>Table 2: Different versions of the RTI index</th>
</tr>
</thead>
<tbody>
<tr>
<td>rscore2d(_sample)</td>
</tr>
<tr>
<td>rscore2d_w(_sample)</td>
</tr>
<tr>
<td>rscore2d_cs(_country)</td>
</tr>
</tbody>
</table>

*Table 2: An overview of the different versions of the RTI index. Equivalent versions exist of the complexity index (cscore). Extensions of the variable name (_sample; _country) differentiate different samples over which the index is calculated.*

**IV.  Comparing measures of task content**

After making the theoretical case for my novel approach to measuring occupational task content, it is now crucial to consider the empirical similarities with and differences from the “competitor” measures. Above all, it is important to show that my measures indeed produce more plausible routine-intensity scores and that differences over time and across countries warrant giving up the pooled dataset in favour of wave- or country-specific measures. As my analysis shows, there are indeed important differences, especially compared to the AD index, and some insightful similarities with the FMH indices. Countries’ average scores vary rather a
lot, casting doubt on the previous practice of extending one country’s measures to a range of potentially very different countries. Finally, while there is undoubtedly change over time, it is difficult to discern any consistent pattern of certain occupations becoming more or less routine or complex.

Comparison of my RTI and complexity indices

I plot the routine and complexity indices for the EU-27 at the 2-digit ISCO-88 level in figure 2 below. The figure shows a more or less linear increase of routine-intensity down the occupational hierarchy. At least based on this ordering of occupations, one of the key tenets of RBTC theory, that routine occupations cluster around the middle of the occupational distribution, is not borne out. On the one hand, clerical occupations, which are generally assumed to be one of the most routine occupations in the RBTC framework, do not appear to be particularly routine-intensive. According to my measure, they are near the median together with sales and service workers. On the other hand, elementary occupations exhibit relatively high routine intensity, as do crafts and manufacturing workers. As a result, the often-postulated pattern of medium-skill occupations being disproportionately high-routine is absent in my data. Below, I will contrast these findings with those of AD and argue why my index is far more plausible.

---

15 While at the 1-digit level, it is generally possible to speak of a hierarchy in terms of occupational prestige and average wage, this is questionable at the 2-digit level. For example, there is no substantive reason to place health and life science professionals (group 22) above teaching professionals (group 23).
Echoing the findings of previous research on routine tasks, cognitive tasks, and complexity, there is an inverse relationship between RTI and complexity. The Spearman correlation between the two indices is sizeable at -0.73, and the weighted Pearson correlation, at -0.66, is in the range reported in other studies for the correlation between RTI and cognitive task intensity. This inverse relationship between RTI and complexity entails that there are few occupations with both above-the-median RTI and complexity, or below-the-median RTI and complexity. In the former group, life science and health associate professionals (group 32), customer service clerks (group 42), and metal, machinery and related trades workers (group 72) form the group of occupations for which RBTC would predict declining employment shares while SBTC would offer the opposite prediction. In the latter group, models, salespersons and demonstrators (group 52) should grow if occupational change was strictly routine-biased and should decline if it was complementing high skill levels. The lines of best fit in figure 3 show furthermore that the relationship between routine-intensity and complexity is not strictly linear. A quadratic fit describes the relationship marginally better than a simple linear regression, with the lest complex occupations exhibiting above-average, but not very high routine-intensity. While it is thus clear that one is not simply the inverse of the other, the RBTC hypothesis would

---

**Figure 2**: This figure shows the routine and complexity scores at the 2-digit ISCO-88 level (full sample).

---

16 This shows that it remains a challenge to develop an index based on a priori considerations with dimensions that are at least approximately orthogonal at the occupational level and don’t explain the same underlying variation.

17 The example of subsistence agricultural and fishery workers illustrates that it matters why a job is or isn’t routine-intensive. Clearly, one would not expect subsistence agriculture to be a growing field in advanced economies. Nevertheless, because of the low degree of specialisation and division of tasks, it is also classified as a low-routine occupation whereas market-oriented skilled agricultural and fishery workers are above the median on the RTI mesure.
lead us to expect a more pronounced inverse-U shaped relationship. However, the high correlation is problematic for research aiming to distinguish RBTC from SBTC, a problem which plagues all widely used operationalisations of the RBTC hypothesis that try to address this issue. Before looking at country differences and changes over time, I will now first briefly situate my index in relation to the most prominent alternatives, the operationalisations of AD and FMH.

Figure 3: This figure plots the routine and complexity scores at the 2-digit ISCO-88 level, and their relationship with each other. The quadratic line of best fit describes the relationship better than a simple linear fit; however, the quadratic term is not statistically significant in the regression. Occupations are weighted by their average employment share in 2015 in the EU-27 countries.

Comparison with the AD routine index

As stated earlier, the RTI measure that was formalised by AD has become the standard “off-the-shelf” measure for routine task intensity in the US and beyond. Hence, the first test for my index is how it measures up to this widely used standard measure.\textsuperscript{18} While AD work at the level

\textsuperscript{18} Here a reminder is appropriate that AD calculate an index of predominance, by subtracting the measures for abstract and manual task inputs from routine task inputs. This means that there is no separate index of cognitive or abstract task requirements that I could compare the complexity index to – the only possible comparison is that of my index of the prevalence of routine tasks with AD’s measure of the predominance of routine tasks.
of US census occupations, (Goos, Manning, and Salomons 2014) take their measure and map it onto ISCO-88, thus making the index applicable outside the US.\textsuperscript{19} In line with the classical RBTC theory, this measure identifies middling occupations such as clerks and craftsmen as particularly routine intensive, whereas in my framework, better occupations are generally less routine intensive.\textsuperscript{20} The differences between the two measures become visible when we plot the two indices in figure 4.\textsuperscript{21} In contrast to the figures above which showed the absolute value of the scores, I now plot employment-weighted percentiles. Hence, if the two measures were equivalent, the markers for the occupations would be lined up on a straight 45-degree line through the origin. Instead, the markers are dispersed widely over the plot region and the line of best fit from a weighted linear regression meets the y-axis nowhere near the origin. This, and the relatively low adjusted R\textsuperscript{2} of 0.35 between two measures that purportedly capture the same concept implies that my operationalisation is substantively different from the approach adopted in previous research.

A look at the outliers is instructive to assess why the indices differ so much. Office clerks (group 41) have the highest RTI score of all occupations in Goos et al. but are just below the median according to my measure. This illustrates the aforementioned point of the in my view unrealistic characterisation of clerical occupations as far surpassing any other occupation in routine intensity. While secretaries, finance clerks, or librarians (examples of 3-digit occupations subsumed under office clerks) undoubtedly perform a fair share of routine tasks, it seems implausible that their job tasks are vastly more routine-intensive than those of printing machine operators, mechanical equipment assemblers, or weavers, to name just a few examples. The time dimension likely plays a role here, since by 2015 these occupations undoubtedly had become less routine compared to 1977, the year from which the task data in AD are taken (Autor, Levy, and Murnane 2003, 1313). Similarly, it seems rather implausible to assign drivers and mobile plant operators (group 83) one of the lowest routine intensity scores, third only to managers of small enterprises and teaching professionals. A position close to the median, as reported by my measure, is eminently more plausible. While driving and operating moving equipment thus far eludes complete automation, mainly because of the lack of a controlled environment, much of the time workers in these occupations still perform comparatively repetitive tasks.

Overall, my measure tends to assign relatively higher routine-intensity scores to occupations in major groups 7, 8, and 9 (plant and machine operators and assemblers, craft and related trades occupations, and elementary occupations). At the same time, managers, professionals, and clerks tend to receive lower RTI scores than we would expect if both measures were perfectly

\textsuperscript{19} In addition to the suboptimal measures used by AD, mapping the 318 US census occupations onto 27 ISCO-88 2-digit occupations may have introduced further measurement error.

\textsuperscript{20} In the ISCO-88 framework, occupational groups reflect similarity in the skill level and skill specialisation of the tasks and duties that make up a job. Class schemata which assign individuals to a social class based on the mode of regulation of their employment result in a substantively similar ranking with professionals at the top and unskilled and manual workers at the bottom (Goldthorpe & McKnight 2004). This, and the fact that wages follow a similar gradient, indicates that one can really speak of a hierarchy, at least at the 1-digit level.

\textsuperscript{21} In figure 4 I merged ISCO group 62 (subsistence agricultural and fishery workers) with group 61 (market-oriented skilled agricultural and fishery workers) because the former is not coded in all countries and tends to have very small sample sizes. The same approach is adopted in figure 5.
correlated. Most of this accords with the classical RBTC hypothesis, but the finding that elementary occupations – which include many low-skilled service jobs – are relatively high-routine contradicts the notion that displaced routine workers would move into such occupations (Autor and Dorn 2013). Another interesting difference is the reversed order of occupational groups 51 (personal and protective services workers) and 52 (models, salespersons and demonstrators). While according to AD the former are near the 40th percentile and the latter near the 60th percentile, the opposite is true in my routine index. In light of recent research noting the incipient displacement of restaurant services workers (a 3-digit occupation accounting for a substantial part of employment in group 51) by automated ordering machines and the like (Lordan and Neumark 2018), again, my measure seems more plausible.

Thus, the two routine measures have slightly different implications for employment change. Both theories are unanimous that professional occupations should experience employment growth; however, differences are visible above all regarding clerical occupations and some crafts and manufacturing occupations. What is more, some of these occupations with large discrepancies account for substantial portions of total employment. It will be interesting to see how these differences play out empirically in further research.

Figure 4: Comparison of my EWCS-based RTI index and the DOT-based index of AD as reported in Mahutga et al. To plot the occupations, I calculate their average percentile ranks on the RTI measures and weigh them by their employment share in the EU-15 countries in 2000. For two measures that purportedly capture the same concept, an adjusted $R^2$ of 0.35 and the associated rank-order correlation of 0.74 is quite low.
Comparison with the FMH routine and cognitive indices

My indices are more closely related to the routine and cognitive indices that are employed by (Fernández-Macías and Hurley 2017) and in previous Eurofound publications by the same authors (Eurofound 2014, 2017). For one, like them, I use the EWCS as the main data source, although I use waves 3 – 6 (quinquennially from 2000 to 2015) whereas they only use wave 5 (2010). More importantly, while FMH differentiate jobs by occupation and sector of employment, I look predominantly at the occupational level.22 In turn, this allows me to calculate country- and wave-specific scores for which FMH lack sufficient data. In terms of the construction of the variables, again, my index is similar although there are some important differences.

In my operationalisation of RTI, the key difference is the inclusion of the item “meeting precise quality standards” which I include instead of “solving unforeseen problems on one’s own” which I contend is better suited as a component of the complexity index. FMH argue explicitly against the inclusion of a quality control variable, on the grounds that this assigns relatively high routine scores to most higher-skilled occupations which often include monitoring tasks (Eurofound 2014, 48). They are correct, however, there is a crucial difference between enforcing quality standards and being forced to meet them. Presumably, if the end product can be standardised the production process can be expressed as a series of discrete steps. Therefore, conceptually, the requirement to meet precise quality standards fits with my definition of routine which focuses on repetitiveness and codifiability.

The left-hand panel of figure 5 compares the percentile rankings according to the two routine-intensity measures. While FMH argue in quite strong terms against some key findings of the earlier literature, my findings qualify their stance. When it comes to the routine intensity of service occupations, FMH argue that they are in fact quite routine-intensive, which would contradict the RBTC theory. However, salespersons, as well as elementary occupations (which are often service occupations), all rank lower on the routine axis according to my measure compared to FMH. Conversely, blue-collar jobs in ISCO-88 major groups 7 and 8 are closer to the upper end of the RTI scale based on my index. Overall, this suggests that FMH may be going a bit too far when they claim that routine tasks are most frequent at the bottom of the “skills-wage-cognitive tasks continuum” (Fernández-Macías and Hurley 2017, 575). While they are right to point out the higher routine-intensity of elementary occupations compared to previous research, they underestimate the extent to which crafts and manufacturing occupations are routine.

---

22 On the one hand, the analysis at the job level adds an important level of nuance since it is the case that in some cases occupational task contents differ between sectors. Hence, without accounting for the sector of employment, differences in the sectoral composition of the economy could lead to a false positive finding of country differences in occupational task content. On the other hand, the practice of FMH to use some 43,000 observations to fill a job matrix with 3784 theoretical (and 3123 actual) job cells entails that task scores for a large number of small jobs will be based on very few observations. In my view, therefore, the jobs approach is only feasible with more data or with sectoral data at the 1-digit sector level. For some later analyses I intend to produce an index at that level of detail (27 ISCO occupations and 21 ISIC industries = 567 job cells). For the time being, however, I restrict the analysis to the 27 ISCO-88 occupations.
Figure 5: The panel on the left compares my RTI index with the one in FMH, whereas the panel on the right compares my complexity index with the cognitive index proposed by FMH. Like above, average percentile scores for my indices are calculated from data for the EU-15 in 2000; the percentile scores for the FMH indices are taken from Table 1 of their 2017 paper and are weighted by employment in the EU-15 in 1997. Compared to the AD index, the FMH indices resemble my indices much more closely, especially their cognitive index.

My complexity index has its counterpart in the cognitive task index of FMH, with which it is compared in the right-hand panel of figure 5. Again, there is some overlap due to the limited number of questions on task content in the EWCs: both measures ask for a self-assessment whether a person’s job involves complex tasks. However, in contrast to FMH, I include questions whether a job involves “solving unforeseen problem on one’s own” and “learning new things” in the complexity index. In my view, these two questions are a more direct proxy for complexity than the “number of years of formal education necessary to perform the job adequately” which FMH include in their measure of cognitive intensity. Using worker characteristics like average years of education is moreover a departure from the task-based framework that FMH purport to follow. Furthermore, rather than two separate questions whether a job involves the use of computers and the use of internet at work, I include one question which ask for the use of “computers, tablets, smartphones, etc.”. This serves to avoid an undue emphasis on office jobs since it is in practice unlikely that a job involves the use of computers but not the internet, or vice versa.

Given these substantial differences in the construction of the two indices, it is somewhat surprising that most occupations rank very similarly – in fact, the differences between complexity and cognitive intensity, two related but distinct concepts, are less pronounced than between the two alternative measures of routine intensity. The only larger discrepancies are between life science and health associate professionals and teaching professionals who rank in
the 4th quintile of FMH’s cognitive index but in the 5th quintile of my complexity index, and vice versa. Thus, even though the concept and operationalisation are different, the practical implications of operating with task complexity rather than cognitive intensity are likely to be small. Yet, it is important to be scientifically precise and therefore, if possible, use the most appropriate concept available which I believe is task complexity rather than cognitive intensity.

Overall, it is apparent that the routine intensity and task complexity indices as I have operationalised them do not simply replicate previous research. Moreover, sound theoretical arguments exist for the different choices that I made. The table below displays the rank order correlation between the various indices discussed and shows that despite all differences between the measures, the ordinal ranking of occupations based on their demand for routine, complex, or cognitive tasks is fairly similar. While some degree of similarity is to be expected (and to be desired, for an entirely unrelated alternative measure would not be credible), in the case of the complexity and cognitive indices it is striking how very different questions yield and almost identical ordering of occupations. The correlation table furthermore shows that the correlation between the RTI and cognitive indices of FMH is -0.86, and thus substantially stronger than the -0.73 between my indices. Therefore, if my measures are said to be so similar as to represent two sides of the same coin, this applies a fortiori to the FMH and Eurofound measures. The AD index, calculated as a measure of predominance of routine tasks, exhibits comparatively lower correlations with my and FMH’s indices which measure the prevalence of the respective tasks.

To strengthen confidence in my findings, I conducted a battery of robustness checks such as using the EU-27 countries instead of the EU-15 for the comparisons and using average occupational employment shares from different waves of the EWCS. None of this affects the conclusions presented here, as can be seen from the supplementary results in the appendix.

| Table 2: Rank order correlation between different indices |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Index            | MH RTI | AD RTI | FMH RTI | MH Comp | FMH Cog |
| MH RTI           | 1      | .74    | .94     | -.73    | -.76    |
| AD RTI           | 1      | .69    | -.83    | -.40    | -.47    |
| FMH RTI          | 1      |       |         | .98     |         |
| MH Comp          |       | 1      | -.40    |         |         |
| FMH Cog          |       |       | .98     |         |         |

Table 3: Rank order correlations between the various indices, calculated from percentile scores based on employment shares in the EU-15 in 2000 in the case of my indices and the AD RTI index, and taken from Table 1 in FMH in the case of the FMH RTI and cognitive indices.

Task content across countries

My index is the first to account for potential country differences by calculating country-specific versions of the routine and complexity indices. The sample sizes of the EWCS, with at least 1,000 respondents per country and wave (except for a few exceptions mentioned earlier), make this possible. Indeed, it is surprising that authors who have previously used the EWCS for task-related research have not considered country differences in occupational task content a matter of interest, although the fact that they only use one wave of the EWCS is probably part of the explanation (Eurofound 2014, 2017; Fernández-Macías and Hurley 2017). Figure 6 below
shows for each 2-digit occupation the range of routine and complexity scores in the EU-15 countries. Country-specific RTI scores are depicted in the panel on the left and complexity scores on the right side. The plots show that there are only few countries where individual occupations fall outside the range circumscribed by 1.5 times the interquartile range. Luxembourg and Greece, the two countries which account for two thirds of the outliers, are in many ways atypical economies. Furthermore, the interquartile range which circumscribes the observations between the 25th and 75th percentile is noticeably narrower for the RTI index than for the complexity index.

We see that while large occupations such as teaching professionals typically have fairly narrow boxes and few outliers, for some occupations there is large variation or extreme outliers. The EWCS data of course do not allow me to ascertain whether these differences in perception have any substantive basis or are due to, for example, cultural differences. Regardless, it suggests that country-specific data, where available, are to be preferred unless small sample sizes make country-specific scores susceptible to outliers. While extreme outliers are in most cases the result of very small sample sizes, another likely contributing factor are differences in the sectoral composition of these occupations. There is an interesting geographic pattern in that the Nordic countries and Germany have the lowest overall routine-intensity, with Central and Western European countries in the middle and Southern Europe exhibiting the highest routine intensity. However, no comparable pattern is discernible for average task complexity. The North-South gradient of the routine-intensity measure is interesting insofar as it broadly corresponds to levels of economic development: work in more affluent countries is less routine intensive. This illustrates the importance of supplementing the investigation at the occupational level with an analysis at the job level, albeit for data reasons with broader sectors than in FMH.
Figure 6: This figure shows the range of country-specific task intensity scores in the EU-15.

Task content over time

It is widely accepted that not only the frequency of occupations changes in response to technological advances, but also the tasks they entail (Spitz-Oener 2006; Becker and Muendler 2015). However, in the literature on RBTC, this facet of technological has received scant attention. Most previous studies use data on occupational task content from only one year, thus making any investigation of task content over time impossible. Moreover, as described above, the few studies that do look at within-occupation change over time find contradictory results. My index makes it possible to analyse within-occupation changes over four points in time over a 15-year period from 2000 until 2015 in a large sample of European countries. With this, my approach covers a wider geographical area and a more recent time period than other studies. I find that there was no general trend towards less routine-intensive work and a small increase in overall occupational complexity. There was, however, a mean-preserving convergence of RTI scores, with high-routine occupations becoming less routine and low-routine occupations becoming more routine, and a divergence of complexity scores, with increases in complex occupations larger than the concomitant declines in simple occupations. At the level of individual occupations, wave-on-wave increases in complexity are associated with a decline in routine-intensity.
**Routine intensity over time**

If there is a trend to reallocate employment to less routine-intensive occupations, this constitutes *prima facie* grounds to expect the same pattern to hold within occupations. Presumably, in occupations that are under pressure from routine-biased technological change, those who remain employed in these occupations will end up performing more non-routine tasks. At the same time, in the occupations which are complemented by RBTC there is no reason to expect the share of routine tasks to increase. While this is not a necessary consequence of the change between occupations, the forces that drive the latter can be expected to also effect the former.

![Routine intensity scores by wave in EU-15 countries](image)

*Figure 7: This figure shows a preliminary analysis of changes of routine-intensity over time.*

Contrary to theoretical expectations, there seems to be no clear-cut trend for work to become more non-routine. At the aggregate level, EWCS respondents in 2000 in fact reported on average the lowest routine intensity of all four waves, followed by a jump upwards in 2005. Over the following 10 years, routine-intensity declined back towards 2000 levels. Figure 8 shows that there is no clear picture at the 2-digit level either. Occupational scores vary widely between waves, and it is not apparent that a majority of occupations have consistently become more or less routine-intensive over the years. In some cases, such as group 11 (legislators and senior officials) part of the explanation is probably the comparatively small sample size at the

23 In fact, the only occupations which have consistently become more routine-intensive are group 22 (life sciences and health professionals) and group 52 (models, salespersons and demonstrators). There is not a single occupation with a lower RTI score in each subsequent wave.
individual level, but this cannot explain the variation in larger occupations such as general managers, clerks, or machine operators.  

Table 4 shows that there has in fact been some convergence: low-routine occupations (defined as having median or lower routine-intensity in a given wave) have seen their routine score increase successively with each wave, while the routine scores of occupations with values above the median have declined slightly. Overall, however, routine-intensity has remained remarkably stable. This is an interesting, if somewhat counterintuitive, finding. Low-routine occupations, which tend to be high-value-added occupations, would have been expected to further reduce their routine task component. However, computerisation and automation are not new phenomena and one possibility, compatible with previous research and economic reasoning, is that routine tasks have first been eliminated in high-value-added occupations but that this process had been largely completed by 2000. In the absence of further innovations, this may have led to such tasks creeping back in, for example due to tighter regulations at the workplace.  

Firm-level research over a longer period would be necessary to determine the veracity of this conjecture, especially in light of contrary findings for the US (Hershbein and Kahn 2018).

**Complexity over time**

Persistent SBTC would suggest a sustained increase in the complexity of occupational tasks. As above, the reasoning that leads us to expect simple occupations to shrink at the expense of complex ones also implies an increase in the demand for complex tasks within occupations. This is what ALM referred to as the intensive margin of technological change (Autor, Levy, and Murnane 2003). Moreover, we expect the demand for complex tasks to increase within occupations regardless of their position in the complexity distribution. In figure 9, we see again a pattern that is not easily interpretable, with no clear tendency in either direction. Interestingly, the complexity scores of small occupations such as groups 11 and 92 are much more stable than their RTI scores; the outlier with unusually large changes is group 13 (general managers) which has become substantially more complex.

Table 4 provides more insights. There is indeed some limited evidence for upskilling, as the weighted average of the complexity score increased moderately by about .02 points. More prominent in the data than the overall increase in task complexity is the apparent divergence, in contrast to the convergence exhibited by the RTI index. While simple occupations have become moderately less complex with every successive wave, complex occupations have become substantially more complex, especially between 2005 and 2010. Consequently, this interval also accounts for most of the increase in overall complexity. Since this interval contains the core period of the financial crisis, the pattern appears to conform to arguments that low-skilled employment is the first to be shed during a crisis, leaving behind a more skilled core working population. In this case it would be incorrect to attribute the task changes to technological

---

24 Another finding supports my earlier argument that occupational (or job) scores based on insufficient sample sizes are unlikely to be meaningful measures of the tasks of an average worker.

25 With each wave, a higher percentage of respondents gave a positive answer to the question, “does your main job involve meeting precise quality standards?” Hence, this may indeed explain part of the increase in routine task intensity through a regulatory channel in high-value-added occupations.
change, but the fact that most of the change is concentrated in high-complexity occupations suggests otherwise.

Figure 8: This figure shows the complexity scores by wave.

Table 4: The trajectory of the RTI and complexity measures, 2000 – 2015

<table>
<thead>
<tr>
<th>Measure/occupations</th>
<th>2000</th>
<th>2005</th>
<th>2010</th>
<th>2015</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routine task intensity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All occupations</td>
<td>-.0089547</td>
<td>-.0022459</td>
<td>-.0031969</td>
<td>-.0066702</td>
<td>+</td>
</tr>
<tr>
<td>Low-routine occupations</td>
<td>-.1648208</td>
<td>-.1465081</td>
<td>-.1322215</td>
<td>-.1184708</td>
<td>++</td>
</tr>
<tr>
<td>High-routine occupations</td>
<td>.1702327</td>
<td>.1573867</td>
<td>.159215</td>
<td>.1524918</td>
<td>-</td>
</tr>
<tr>
<td>Task complexity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All occupations</td>
<td>.009769</td>
<td>.0084734</td>
<td>.0226035</td>
<td>.02974</td>
<td>+</td>
</tr>
<tr>
<td>Simple occupations</td>
<td>-.2690487</td>
<td>-.2784464</td>
<td>-.2868856</td>
<td>-.2899873</td>
<td>-</td>
</tr>
<tr>
<td>Complex occupations</td>
<td>.3156649</td>
<td>.3090257</td>
<td>.3460918</td>
<td>.3497246</td>
<td>++</td>
</tr>
</tbody>
</table>

Table 4: The trajectory of my RTI and complexity measures, 2000 - 2015. The table shows that while there has been convergence in terms of routine-intensity, task complexity has diverged. Data for EU-15, occupations weighted by employment share; low-routine and simple occupations are those that score at the median or below, high-routine and complex occupations those that score above the median of the respective measure.

Finally, it is important to consider to what extent changes in RTI and task complexity are related. To this end, I calculate the wave-on-wave changes in RTI and task complexity for each occupation and regress the former on the latter. This relationship is depicted in figure 10, which
shows a moderate negative relationship between changes in the two task dimensions. More precisely, an increase of the complexity measure by 0.1 is associated with a decrease in routine intensity by roughly 0.03 points. Controlling for survey wave does not change the coefficient on complexity, while the coefficient on wave is negative but statistically insignificant. Of course, there is no reason to assume that changes in complexity cause changes in routine-intensity; rather, it stands to reason that an omitted variable – technological change – affects both simultaneously. The absence of an effect of the wave on changes in RTI or complexity means that there is no tendency for change to accelerate or slow down(?) As even high-skill jobs become increasingly automatable (Acemoglu and Restrepo 2018), it is likely that after a period of relative stability, in the foreseeable future more pronounced within-occupation changes will take place.

Figure 9: This figure depicts the unconditional relationship between wave-on-wave changes in RTI and task complexity. There is a statistically significant relationship, with a 0.1 point increase in task complexity being associated with a 0.03 point reduction in routine-task intensity. The relationship is robust to controlling for wave, suggesting that the association between changes in RTI and complexity has not changed between 2000 and 2015.
Conclusion

The literature on occupational task content has long relied on just a few off-the-shelf measures without giving much thought to how key concepts are operationalised. Hence it came to be that cognitive tasks were defined without a clear purpose, routine tasks operationalised with unsuitable variables, and diversity across countries and over time ignored. Yet, for a meaningful analysis of the effects of technological change, conceptually and empirically solid measures are crucial. The measures proposed in this article address the most serious problems with existing operationalisations and, with their various configurations and subsamples, offer researchers a flexible tool for the analysis of technological change and the labour market. Most importantly, both the routine-intensity and the complexity index have a clear theoretical interpretation and the variables used to operationalise them capture the essence of the concepts. Furthermore, with my country- and wave-specific indices, a much more detailed analysis of the impact of RBTC and SBTC on employment change becomes possible for the first time. This represents a significant improvement over the measures developed by AD and used in most of the task literature, and goes beyond the important contribution of FMH who also set out to improve the measurement of occupational task content.

In detail, my analyses show that in my sample a number of occupations, above all clerical and sales occupations, are less routine-intensive than the AD index suggests. On the other hand, crafts, manufacturing, and elementary occupations tend to have higher RTI scores. Task complexity, as a measure of the task characteristics that are affected by SBTC, closely resembles the cognitive dimension in FMH, even though there are important conceptual and empirical differences. All this implies that my measures are a viable alternative to the prevalent operationalisations. I furthermore shed light on the extent of country differences, arguing that they call into question the practice of applying task measures from one country in very different contexts. Finally, I show that routine-intensity and task complexity have changed over time. While occupations’ RTI scores converged slightly without the mean changing noticeably, overall task complexity has gone up and simple and complex occupations have actually further diverged.

The measures proposed here will be useful for future research on employment trends, especially when it comes to answering the important question, which is the nature of recent technological change? In my ongoing research project, I will pursue this question further by analysing employment change in a large sample of OECD countries in relation to occupational wages and task characteristics. My goal is to determine whether the RBTC hypothesis as postulated by AD, (Goos and Manning 2007; Goos, Manning, and Salomons 2014) and others holds up, or has to be qualified as suggested by FMH and Eurofound (2014, 2017). The findings of this study place both within the realm of possibility, with key tenets of both strands of the literature called into question.
Bibliography


Appendix

Appendix A: ISCO-88 codes

Below are the ISCO-88 codes at two-digit level.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Legislators and senior officials</td>
</tr>
<tr>
<td>12</td>
<td>Corporate managers</td>
</tr>
<tr>
<td>13</td>
<td>General managers</td>
</tr>
<tr>
<td>21</td>
<td>Physical, mathematical and engineering science professionals</td>
</tr>
<tr>
<td>22</td>
<td>Life science and health professionals</td>
</tr>
<tr>
<td>23</td>
<td>Teaching professionals</td>
</tr>
<tr>
<td>24</td>
<td>Other professionals</td>
</tr>
<tr>
<td>31</td>
<td>Physical and engineering science associate professionals</td>
</tr>
<tr>
<td>32</td>
<td>Life science and health associate professionals</td>
</tr>
<tr>
<td>33</td>
<td>Teaching associate professionals</td>
</tr>
<tr>
<td>34</td>
<td>Other associate professionals</td>
</tr>
<tr>
<td>41</td>
<td>Office clerks</td>
</tr>
<tr>
<td>42</td>
<td>Customer services clerks</td>
</tr>
<tr>
<td>51</td>
<td>Personal and protective services workers</td>
</tr>
<tr>
<td>52</td>
<td>Models, salespersons and demonstrators</td>
</tr>
<tr>
<td>61</td>
<td>Market-oriented skilled agricultural and fishery workers</td>
</tr>
<tr>
<td>62</td>
<td>Subsistence agricultural and fishery workers</td>
</tr>
<tr>
<td>71</td>
<td>Extraction and building trades workers</td>
</tr>
<tr>
<td>72</td>
<td>Metal, machinery and related trades workers</td>
</tr>
<tr>
<td>73</td>
<td>Precision, handicraft, printing and related trades workers</td>
</tr>
<tr>
<td>74</td>
<td>Other craft and related trades workers</td>
</tr>
<tr>
<td>81</td>
<td>Stationary-plant and related operators</td>
</tr>
<tr>
<td>82</td>
<td>Machine operators and assemblers</td>
</tr>
<tr>
<td>83</td>
<td>Drivers and mobile-plant operators</td>
</tr>
<tr>
<td>91</td>
<td>Sales and services elementary occupations</td>
</tr>
<tr>
<td>92</td>
<td>Agricultural, fishery and related labourers</td>
</tr>
<tr>
<td>93</td>
<td>Labourers in mining, construction, manufacturing and transport</td>
</tr>
</tbody>
</table>
Appendix B: Details on the construction of the indices

Survey questions

The questions that are part of the index are a mix of questions on how the work is organised and general questions about the nature of the work. The precise set of questions comprising each index is as follows:

Routine intensity index

- Does your main job involve repetitive hand or arm movements? (Scale from 1: “all of the time” to 7: “never”)
- Does your job involve short repetitive tasks of less than 1 minute? (Scale from 1:”yes” to 2: “no”)
- Does your job involve short repetitive tasks of less than 10 minutes? (Scale from 1:”yes” to 2: “no”)
- Does your main paid job involve monotonous tasks? (Scale from 1:”yes” to 2: “no”)
- Does your main paid job involve meeting precise quality standards? (Scale from 1:”yes” to 2: “no”)

Task complexity index

- Does your main paid job involve working with computers, laptops, smartphones etc.? (Scale from 1: “all of the time” to 7: “never”)
- Does your main paid job involve complex tasks? (Scale from 1:”yes” to 2: “no”)
- Does your main paid job involve solving unforeseen problems on your own? (Scale from 1:”yes” to 2: “no”)
- Does your main paid job involve learning new things? (Scale from 1:”yes” to 2: “no”)

Correlations

Since the items in the questionnaire have different scales, I normalise the answers to have mean 0 and a standard deviation of 1. Where appropriate, I reverse code the answers. The correlations of the components of the indices are displayed in table B1, showing that the indices are internally consistent measures. Since the correlations are calculated at the individual level, they are not very high but it is visible that they form two clusters, corresponding to the indices. All components of the RTI and complexity indices are positively and significantly correlated with the other components of the index they are part of, with correlation coefficients usually in the range between 0.2 and 0.5. On the other hand, they tend to be weakly negatively correlated with the components of the other index. The only outlier is the quality standards item in the routine index which is positively correlated with all components of both indices (highlighted in yellow in the table below). This suggests that there is some ambiguity whether the item belongs in the index. Yet, for the conceptual reasons outlined above, I keep the question as a component of the RTI index. If the RTI index were to be calculated without the quality standards item, the two occupations which are otherwise classified as low-routine simple occupations (groups 52 and 62) would be routine simple occupations for which the implications of SBTC and RBTC would be identical. Furthermore, my index also shows is that the item “dealing with unforeseen

26 There is a wider debate to be had whether the components of a composite indicator should be highly correlated so as to reflect the same latent concept, or whether a useful composite indicator should comprise different aspects (McGillivray 1991; Kovacevic 2011).
problems” which FMH include in their RTI index in fact belongs in the complexity index, since it is quite strongly positively correlated with the items in that index while the correlations with most RTI items are weakly negative.

Table B1: Correlation coefficients of items constituting the RTI and complexity indices

<table>
<thead>
<tr>
<th></th>
<th>Rep. Arm/ hand</th>
<th>1 min rep. tasks</th>
<th>10 min rep. tasks</th>
<th>Monotonous tasks</th>
<th>Quality standards</th>
<th>Computer use</th>
<th>Complex tasks</th>
<th>Unforeseen problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm/hand movements</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 min rep. tasks</td>
<td>0.301***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 min rep. tasks</td>
<td>0.337***</td>
<td>0.489***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monotonous tasks</td>
<td>0.250***</td>
<td>0.210***</td>
<td>0.215***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality standards</td>
<td>0.131***</td>
<td>0.0935***</td>
<td>0.113***</td>
<td>0.0769***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer use</td>
<td>-0.0886***</td>
<td>-0.0574***</td>
<td>-0.0536***</td>
<td>-0.0895***</td>
<td>0.0590***</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex tasks</td>
<td>-0.0500***</td>
<td>-0.0147***</td>
<td>-0.00614***</td>
<td>-0.0192***</td>
<td>0.155***</td>
<td>0.175***</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Unforeseen problems</td>
<td>-0.0202***</td>
<td>-0.0162***</td>
<td>0.00861***</td>
<td>-</td>
<td>0.208***</td>
<td>0.262***</td>
<td>0.285***</td>
<td>1</td>
</tr>
<tr>
<td>Learning new things</td>
<td>-0.0553***</td>
<td>-0.0193***</td>
<td>-0.00140</td>
<td>-0.0912***</td>
<td>0.209***</td>
<td>0.303***</td>
<td>0.300***</td>
<td>0.424***</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01, ***p < 0.001

The table provides support for the argument that my indices are internally consistent measures of what they purport to be measuring. The only borderline case is the quality standards item which could be a component of either index. For all other items, the smallest absolute value of the correlation with a component of the own index substantially exceeds the maximal absolute value of the correlation with a component of the other index. This indicates that although they are correlated, the two indices nevertheless measure different underlying concepts.

Weights

I experimented with using the post-stratification weights provided in the EWCS dataset for calculating the indices but can confirm that the weights do not affect the occupational task intensity scores. Weights are relevant, and I indeed use them, for the analysis of occupational employment shares, as some occupations are over- or undersampled in national samples. However, for the calculation of my indices, weights would only be important if high- or low-routine or -complexity individuals would be systematically over- or undersampled within an occupation. A comparison of a version of the index with and without weights (r = 0.9993) confirms that this is not the case. Hence, I do not use weights to calculate the routine and complexity scores, but I include weights in the comparisons with other indices to account for the varying employment shares of occupations.

Distribution of individual responses

A final note concerns the distribution of individuals’ responses to the survey questions included in the index. With most of the constituent questions being binary variables, and the answers being normalised to have a mean of 0 and a standard deviation of 1 in the relevant sample, the value for each individual depends on the overall distribution: the fewer individuals give the same answer, the further away from 0 will the value be. However, there is in all cases a substantial number of people answering either way, so that most values range between -1.5 and -0.5 and between 0.5 and 1.5. Since individuals’ routine and complexity scores are the simple
average of the questions making up the indicator, and it is rare for survey respondents to answer all questions affirmatively or negatively (the constituent questions are positively correlated, but not overly highly so), the individual-level scores in each occupation still have high standard deviations around 0.6.

The wide dispersion of individual answers does not change if I define the analytical groups more restrictively, such as occupation-sector combinations where one would expect tasks to be more similar than in broad occupations spanning all sectors. Partly, this is certainly due to measurement error which may be induced by survey respondents misunderstanding the question, and various other complications. However, it also illustrates that the tasks workers perform on their jobs vary considerably, and more than one expert-assigned value would make one believe. Yet, when it comes to analysing the data, I nevertheless have to reduce this complexity to a single number. At the same time, conventional measures of uncertainty offer not much insight, as they essentially span the entire range of the data. This is why I do not plot confidence intervals in the graphs illustrating change over time and country differences.
Appendix C: Detailed overview of studies of RBTC

Table A1: The operationalisation of RBTC in key papers – detailed table

<table>
<thead>
<tr>
<th>Autor, Levy and Murnane 2003</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dimensions</strong></td>
<td>routine manual; routine cognitive; non-routine interactive; non-routine analytic; non-routine manual.</td>
</tr>
</tbody>
</table>
| **Variables** | Routine manual: finger dexterity.  
Routine cognitive: adaptability to situations requiring the precise attainment of set limits, tolerances and standards.  
Non-routine interactive: direction, control and planning.  
Non-routine analytic: quantitative reasoning requirements.  
| **Data source** | Dictionary of Occupational Titles (US; 1977 and 1991 versions) |
| **Unit of analysis** | 3-digit census occupations (US) |
| **Also used by** | Spitz-Oener 2006 (for Germany, dimensions only); Goos and Manning 2007 (for UK); Autor, Katz and Kearney 2008; Acemoglu and Autor 2011. |

<table>
<thead>
<tr>
<th>Autor and Dorn 2013</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dimensions</strong></td>
<td>Routine; abstract; manual</td>
</tr>
</tbody>
</table>
| **Variables** | Routine: finger dexterity; adaptability to situations requiring the precise attainment of set limits, tolerances and standards.  
Abstract: direction, control and planning; quantitative reasoning requirements.  
| **Data source** | Dictionary of Occupational Titles (US; 1977 version) |
| **Unit of analysis** | Not specified (US) |
Also used by
Autor and Handel 2013 (dimensions only); Goos, Manning and Salomons 2014 (for 16 EU countries); Naticchioni, Ragusa and Massari 2014 (for 12 EU countries); Mahutga, Curran and Roberts 2018 (for LIS countries).

**Fernandez-Macias and Hurley 2017**

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Routine; cognitive; social interaction; trade intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables</strong></td>
<td>Routine: repetitive hand or arm movements; repetitive hand movements of less than 1 or 10 minutes; monotonous tasks; dealing with unforeseen problems. Cognitive: complex tasks; use of computers at work; use of internet at work; number of years of formal education necessary. Social interaction: direct interaction with non-colleagues; pace of work determined by demands from customers. Trade intensity: domestic value-added of exports; gross value added of imports relative to gross output.</td>
</tr>
<tr>
<td><strong>Data source</strong></td>
<td>European Working Conditions Survey; 2010 wave</td>
</tr>
<tr>
<td><strong>Unit of analysis</strong></td>
<td>“jobs”: 2-digit occupation, 2-digit industry cells (EU-28)</td>
</tr>
<tr>
<td><strong>Also used by</strong></td>
<td>Eurofound 2014, 2017</td>
</tr>
</tbody>
</table>

**Caines, Hoffmann and Kambourov 2017**

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Routine, complex</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables</strong></td>
<td>Routine: as in Autor and Dorn (2013) Complex: PCA of 35 O*NET descriptors assumed to be associated with task complexity</td>
</tr>
<tr>
<td><strong>Data source</strong></td>
<td>Dictionary of Occupational Titles; Occupational Information Network (O*NET)</td>
</tr>
<tr>
<td><strong>Unit of analysis</strong></td>
<td>3-digit census occupations (US)</td>
</tr>
<tr>
<td><strong>Also used by</strong></td>
<td>---</td>
</tr>
</tbody>
</table>

Table 5: Source: Author’s analysis of the literature quoted in the table.