
This version is available at https://strathprints.strath.ac.uk/64600/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

The Strathprints institutional repository (https://strathprints.strath.ac.uk) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde’s intellectual output.
Probing “green” industry enterprises in the UK: A new identification approach

Philip Shapira a,b,⁎, Abdullah Gök a, Evgeny Klochikhin a, Marianne Sensier c

a Manchester Institute of Innovation Research, Manchester Business School, University of Manchester, Manchester M13 9PL, UK
b School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332-03345, USA
c Economics, School of Social Sciences, University of Manchester, Manchester M13 9PL, UK

A R T I C L E   I N F O

Article history:
Received 6 June 2013
Received in revised form 21 October 2013
Accepted 26 October 2013
Available online 5 December 2013

Keywords:
Green industry
Measurement
Definition of industries

A B S T R A C T

There is a growing interest in innovation for sustainability and the development of green industries and green jobs. But how can green industries, green manufacturing jobs, and green goods innovation be measured? This paper probes current and recent attempts to define and measure these categories, with a focus on studies in the UK. We review the methods, estimates and trends contained in these studies. While these efforts have value, they also raise significant conceptual and measurement issues. The paper discusses a series of these issues and considers strategies to further refine the categorization and detection of green sector enterprises. A new identification approach is put forward using search term combinations and text mining to discern green goods sector companies. This method is tested through a search of small and medium-size green goods enterprises in the UK. Findings from our search approach are presented, along with a discussion of advantages and limitations.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

There is growing policy interest in the development of green industries to address challenges of environmental sustainability. At the same time, it is also suggested that stimulating green industries is an avenue for generating jobs and new business development, particularly among small and midsize enterprises (SMEs), and for rectifying regional imbalances. We are particularly interested in the business and economic justification for policy interest in green industries. However, we do not take it as axiomatic that industries grow just because they are greener. Rather, as part of a larger project on “Sustaining Growth for Innovative New Enterprises” (39), we seek to understand the performance and strategies of SMEs in emerging green goods sectors and to generate evidence about how and why such companies stay in business and grow.

One problem that immediately arises when one attempts to investigate green industries is that it is not straightforward to define and identify the industries and companies that comprise this sector. It is precisely this problem that we focus upon in this paper, where – dissatisfied with currently available methods – we propose a new approach to identify green enterprises. We concentrate on what we term the green goods sector (GGS) — comprising companies in a range of industries that produce or market manufactured items that have environmental or natural resource benefits when used by other businesses, organizations or households.

The paper proceeds as follows. After a concise review of the literature which highlights the potential of green sectors and green SMEs in fostering economic development, we consider existing definitions and methods of data collection on green industries. We note a series of measurement issues associated with current approaches. We then present a new method for identifying green firms. We illustrate the method by applying it to a search of green goods SMEs in the United
Kingdom (UK). We present the results in a section on findings, and also compare our strategy with results from searches using a standard industrial classification method. The concluding section discusses the strengths and limitations of our new identification method and considers opportunities for further application and improvement.

2. Literature overview

Efforts to foster “green” technologies and sectors have been launched in many countries and regions. For example, the European Commission launched a program dedicated to eco-innovation through the Executive Agency for Competitiveness and Innovation in 2008.1 Here the objective is “to boost Europe’s environmental and competitive standing by supporting innovative solutions that protect the environment while creating a larger market for ‘green’ technologies, management methods, products and services.” As Kemp and Oltra [24] highlight “policy is crucial for giving environmental benefits a value in the marketplace through the use of regulation, taxes and tradable emission rights.” Suppliers and customers of business also need to grasp the value of eco-innovation in order to stimulate both supply and demand for green technologies that have utility in reducing harmful environmental consequences and conserving scarce natural resources.

Although eco-innovation can raise concerns about potential negative consequences on business and employment, proponents of eco-innovation maintain that such policies will generally result in positive economic growth outcomes as well as beneficial environmental effects. There is indeed an empirical literature that tests the hypothesis [36,37] that environmental policies can foster competitiveness by inducing technological innovation. Böhringer et al. [7] analyze a panel of German manufacturing sectors and find a positive impact of environmental investment on production growth. Constantini and Mazzanti [13], in an analysis of sectors across 15 European Union (EU) countries, report that high tech sector exports have responded positively to energy and environmental taxation, although they suggest further research on the effect of environmental policies in inducing firms in specific green technology sectors to increase their innovative efforts (an issue that our broader study is also now considering). Using German innovation survey firm data, Rennings and Rammer [38] find that innovations stimulated by environmental regulation increased sales but when looking at different sectors within the green industry they report mixed results on profitability. Firms in recycling and waste management benefit, in terms of higher profit margins, from regulations but those in water management experience lower profitability, as costs for eco-innovation cannot be fully passed on through prices in this sector. Yet, while there has been a growth of research studies on eco-innovation and sustainability [8,28], there is also recognition that the linkages between eco-innovation, business development and jobs are still not fully explicated, including understanding the combined effects of mixes of policies related to eco-innovation, enterprise promotion, and sustainability [3].

The need for continued work, particularly at disaggregated levels, on the linkages between eco-innovation and business, is emphasized by changes in the context for green technology development. In recent years, interest in eco-innovation has been influenced not only by environmental and sustainability concerns but also by debate, in multiple countries around the world, on the importance of economic rebalancing following the financial sector crises of the late 2000s and their ongoing aftershocks [18,29,40]. While there is much variation as to how the objectives and processes of rebalancing are described, common themes can be discerned. In the UK, there has been heightened discourse about strengthening manufacturing and shifting towards low-carbon and greener production and consumption so as to counterbalance an overreliance on financial services, foster regional growth outside of the financial capital of London, promote exports, and ensure a more resilient path for economic recovery and sustainable growth [20]. Similarly, in the United States, there has been dialog about how to rebalance the economy, strengthen manufacturing, and using clean production technologies to foster domestic growth and exports [9,33,34]. The Obama administration launched several key initiatives including the Advanced Manufacturing Partnership, the Clean Energy Manufacturing Initiative, and a proposed National Network for Manufacturing Innovation [1,14]. China has also emphasized (in its latest 12th Five Year Plan and in recent governmental statements) the need for rebalancing strategies for regional and social development and to promote environmentally sustainable internal growth ([11]; see also [17]).

Alongside broader macro-economic, fiscal, and trade policies, it is evident that economic rebalancing requires enhanced and renewed enterprise growth and development particularly in sectors that can most readily meet rebalancing objectives. One of the domains which could be especially appropriate in this regard is the production of green goods — comprising manufactured items which help others to conserve natural resources or meet carbon reduction and other environmental goals. In this process, jobs associated with manufacturing may be sustained or expanded, and opportunities developed to increase manufactured exports and services associated with green goods production and use [21]. On a broader scale, the fostering of green economies, sectors, and industries has been highlighted as an imperative (to address sustainability challenges) and an opportunity (to reinvigorate economic growth) for both developed and developing economies [32,43].

3. Defining and measuring green sectors

While fostering a green economy has risen up as a priority target for policy makers throughout the world, the tasks of defining and identifying green industries, green companies, green products and green jobs are neither simple nor trivial. As a recent ILO [21] study points out, these categories do not always overlap. For example, energy conserving products may be marketed to consumers as green but may not necessarily be made of materials or processes that are themselves green. Similarly, green jobs in numerous occupations, for example, waste reduction engineering, may not be located in industries or companies designated as green. In this section, we consider a subset of this larger question of how greenness is demarcated: we focus on the specific problem of how green industrial sectors and green companies are defined and pinpointed. The reason for this is that if we (and others) wish to track whether

---

green industries and companies are growing in terms of sales, jobs, or other metrics, we need to start with a workable and reasonably robust definition so that such industries and companies can be identified.

A starting point for this definitional investigation is to consider the green definitions currently in use by various governmental bodies and other relevant organizations. To take three examples:

- The UK Government states that a green economy is "...one in which value and growth are maximized across the whole economy, while natural assets are managed sustainably." This economy is based upon a "thriving low-carbon and environmental goods and services sector" leading to reduced environmental damage and increased energy security, resource efficiency and climate change resilience [19].

- The United Nations Industrial Development Organization (UNIDO) indicates that green industry is "industrial production and development that does not come at the expense of the health of natural systems or lead to adverse human health outcomes." At the industrial level, this involves "greening" existing industries as well as creating new "green" sectors [44].

- The US Bureau of Labor Statistics defines a green job as a job that either "...provides goods or services that benefit the environment or conserve natural resources" or makes an establishment's processes of production "...more environmentally friendly or use fewer natural resources" [4].

These green economic definitions share some common concepts, such as comprising activities that foster resource efficiency and benefit the environment. However, it is difficult to operationalize and measure these green economic concepts, given the lack of readily available data on the "greenness" of individual firms' outputs and the effects those outputs might have when deployed by businesses, consumers and other actors in the economy. To address this, approaches have been developed which use aggregated information, for example at the level of standard industrial or occupational classifications (SICs or SOCs), as proxies for industries that most contribute to green outputs and green jobs. For example, in the US, the Bureau of Labor Statistics (BLS) designated 333 (subsequently revised to 325) green industry sectors—equivalent to just over one-quarter of the 1192 detailed industries that exist in the North American Industrial Classification System (NAICS). BLS uses an approach that identifies industries which produce green goods and services or which deploy environmentally-friendly production processes. The BLS list of included green goods and services sectors was based on a review of other existing definitions and literature, consultation with industry groups and government agencies, and other stakeholder and public input [4,5,12]. This process led to the exclusion of certain sectors from the BLS green goods and services definition such as the processing of organic agriculture products (determined to have no direct environmental benefit compared with the processing of non-organic food) and transporting or selling green goods (also judged to have no direct environmental benefit compared with selling any other good).

BLS categorized its selected industries into five broad groupings: (1) energy from renewable sources, including wind, biomass, geothermal, solar, ocean, hydropower, and landfill gas and municipal solid waste; (2) energy efficiency, including energy-efficient equipment, appliances, buildings, and vehicles, as well as products and services that improve the energy efficiency of buildings and the efficiency of energy storage and distribution, such as Smart Grid technologies; (3) pollution reduction and removal, greenhouse gas reduction, and recycling and reuse; (4) natural resources conservation, including organic agriculture and sustainable forestry; land management; soil, water, or wildlife conservation; and storm-water management; and (5) environmental compliance, education and training, and public awareness. BLS then undertook surveys (in 2010 and 2011) of firms (businesses with employees, excluding the self-employed) in the selected NAICS industrial classes to determine the proportion of revenues and jobs attributed to various types of green goods and services and weighted the results to obtain estimates for each selected industry. With this method, BLS calculated that the US had 3.4 million green goods and services jobs in 2011, equivalent to 2.6% of all US employment [6].

In another US study, the Brookings Institute estimated the number of jobs in what was termed the "clean economy." A definition similar to BLS was used, whereby the clean economy comprised establishments and associated jobs producing or adding value to goods and services with an environmental benefit [30]. A two-part method was used to operationalize this definition. First, a list of industrial classifications considered to be part of the clean economy was compiled, drawing on classifications developed by other researchers. This list was developed using an eight-digit industrial classification system provided by Dun and Bradstreet (D&B). Companies in these industries were identified by D&B. Second, more than 60 lists of clean economy companies identified by other organizations were added in. After removing duplicates, the companies were checked again against D&B's information on their detailed industrial activities, with adjustments made for large companies (allocating a percentage of employees based on the estimated share of clean products to all products) and for small non-standalone establishments. The resulting establishment and employment estimates were re-aggregated at the level of five high-level categories (similar to those used by BLS) and 39 finer segments. Overall, Brookings estimated that 2.7 million workers were employed in the clean economy in the US in 2010, which is similar to the BLS estimate. The largest segments in Brookings' estimate of the US clean economy were waste management, public mass transit, and resources conservation, followed by energy-saving building materials, regulation and compliance, professional environmental services, organic food and farming, and recycling.

A variation on the definition of green goods and services is provided by Eurostat [16] in conjunction with the Organisation for Economic Cooperation and Development (OECD). The Eurostat/OECD approach differs in that it focuses on activities "...to measure, prevent, limit, minimise or correct environmental damage to water, air, soil, as well as problems

---

2 In March 2013, the BLS program to measure employment in green goods and services was discontinued under federal government sequestration spending cuts, with no indication as to whether the program might be restarted at a later date.
related to waste, noise and eco-systems. This includes cleaner technologies, goods and services that reduce environmental risk and minimise pollution and resource use” (Eurostat [16]). The main purpose of the technology, good or service must be environmental for it to be included. A company is included if 50% of its sales are in the sector, but only the sales value that can be attributed to a product group is included.

In the United Kingdom, attempts have also been made to measure green goods and services activities. A study by Innovas Solutions Ltd [23] for the UK Department for Business, Enterprise and Regulatory Reform expanded the classification of the environmental goods and services sector to include low carbon activities. This study defined three broad (“Level 1”) categories of low carbon and environmental goods and services (LCEGS): environment (including waste management, recycling, pollution control, and environmental consultancy); renewable energy (such as geothermal, hydro, and wave energy products and services); and emerging low carbon (comprising goods and services which reduce emissions from transport and construction, nuclear energy, energy management, carbon capture, and carbon finance). These are consistent with aggregations of SIC codes but are not limited to them. Level 1 is then successively subdivided, going down to the most detailed Level 5, which comprises 2490 industry sub-sectors within 23 sub-categories. Some sectors which potentially overlap are subdivided to avoid double-counting. For example biofuels appear under “alternative fuels” and “alternative fuels for vehicles” as they can be used in heat generating equipment (first category) or to power cars and trucks (second category). Companies are included if 20% or more of their sales are supplied into the LCEGS sector (but only LCEGS sales are counted when aggregating up).

The Innovas/kMatrix study uses a “bottom-up” approach — collecting and cross-referencing data from multiple sources, including data reported by companies for regulatory purposes and available in proprietary business data bases, surveys of firms, and other business directories. Data triangulation is used along with other econometric techniques to verify multiple data sources. Employment in different sectors is estimated on a pro rata basis rather than on the measurable environmental content of the job itself. Thus, if only part of a firm’s output goes into an environmental goods sector, just that proportion will be measured, allowing supply chain relationships to be assessed. However, this depends on whether and how firms report their green goods outputs. The Innovas/kMatrix methodology is broad in coverage. It is structured around selected sectors, but these are not pre-restricted to those defined by SIC classifications (as is the case for BLS). Innovas/kMatrix includes environmental sector activities such as land management and noise management, which some other definitions of green goods and services exclude. Clean technologies and processes developed and used by companies in sectors outside of the three selected categories fall outside of the scope of the study. The Innovas/kMatrix method is not easily verified and cannot readily be replicated by others given the significant proprietary data required. However, the successor government agency commissioned three further reports on the LCEGS sector using this approach [25–27]. For the UK, these studies estimate LCEGS employment and sales by industry categories, regional distributions, and exports and imports. LCEGS sales are also estimated for other countries so as to benchmark the UK and also indicate global LCEGS trade opportunities. Among the seven leading countries, the US tops the list in terms of total LCEGS sales, followed in second place by China, with the UK taking 6th place (Table 1). Annual growth rates from 2007/2008 to 2010/2011 are available. This covers the global economic downturn, following the 2007–2008 financial crisis. Significantly, with the exception of one year for the US (2009/2010), all countries saw growth in LCEGS sales, with the UK achieving the largest growth rate of more than 4% annually. At least in part, this growth was boosted by the measures that many governments took in recent years to provide support to green economy sectors as a counter-cyclical public strategy to promote economic investment and demand.

There are other methods to identify firms in the green goods sector. For example, enterprises involved in inventing new green technologies and processes can be identified through applications for “green” patents. Green patent classifications have been formulated by various patent offices such as the UK Intellectual Property Office’s Green Channel [22] or the United States Patent and Trademark Office’s Green Technology Pilot Program (although the latter is now discontinued, see [45]). In these programs, firms can explicitly apply for a green patent, within designated patent classes, subject to review by patent examiners. The European Patent Office [15] has also developed a specific patent tag (Y02) for climate-change mitigation technology. Meanwhile, in 2010, the World Intellectual Property Organization (WIPO) published a listing of WIPO international patent classifications (the “IPC Green Inventory”) judged by experts to be related to environmentally sound technologies. Seven top-level groupings are identified: alternative energy production, transportation, energy conservation, waste management, agriculture and forestry, administrative, regulatory or design aspects, and nuclear power generation (see [46,47]). Similarly, Thomson Reuters [41] has identified green technology codes that can be applied to its Derwent World Patents Index patent database. These codes focus on green technologies in chemistry, life sciences, and engineering organized under six headings: transportation, power sources, green fuels, environmental awareness, pollution, and recycling. Using patent data allows the use of accessible secondary data sources, but there are significant limitations. Excluded are incumbent and new green technologies where patents applications are not applicable or sought (e.g. trade secrets or obvious processes and practices) or where firms, for strategic or competitive reasons, chose to avoid green designations to make their patent applications less visible. The patent classification system can be idiosyncratic and typically reflects mature technologies (since it takes time for new technologies and patent classes to be recognized). Furthermore, a patent application (or grant) does not necessarily mean that this “green invention” is actually being used by a company in its business or product lines.

A further approach has been to use survey-based methods to identify green goods firms in a population or sample.
Enterprise surveys involve time, burden and cost, and tend to suffer from low response rates unless replies are government-mandated. Survey methods were used by BLS to obtain more detailed information about firms within the selected green SIC classes. (In this case, while commissioned by a government agency, response was not mandatory. BLS applied weighting and other statistical methods to adjust for non-response.) Other less obtrusive methods include identifying green firms through membership of specific networks or associations related to green or environmental goods and services (for example, the Prince Charles’ MayDay Network in the UK) or by sampling firms who have won “green awards” (such as the Green Business Award). Such approaches are obviously partial in that many firms producing green goods and services do not belong to green trade associations (and some that do may do so for public relations reasons), while only a small number of firms are selected for prizes or awards.

All of the methods discussed in this section have advantages and limitations. For example, using only selected industrial classifications to define green goods and services is straightforward (once those classifications have been demarcated) and allows the use of available business data reported by SIC categories. Subsequent surveys of firms within the selected green SIC classes can provide additional new data (as with BLS), although this involves additional cost. However, while certain industrial classifications appear to wholly capture green sector activities, such as NAICS code 221119 (other electric power generation) covering solar, wind, geothermal and biomass, the disadvantage of a SIC filter is that many included industrial classifications contain a mix of green and non-green firms and activities, while there are likely to be numerous green firms and jobs in SIC categories not included within the initial green filter. A further challenge of SIC-based approaches, which is recognized by statistical agencies (e.g. [16]), is that the purpose of the products falling within SIC classes categorized as green is not necessarily clear, for example, whether the product seeks to reduce environmental pollution or carbon emissions. It is also not necessarily clear whether a product that is classed within a green sector is actually “green” and, if it is, whether it is an adaptation of an existing technology, or whether is a new innovation.

The deficiencies of using a purely SIC-based selection method have been addressed by combination methods that mix selected green industry classifications with other lists of green companies (some of whom will be from other SIC categories), as illustrated by the UK Innovas/kMatrix studies. This requires using either lists produced by other organizations or undertaking a special new screening of potential green products and services companies. However, the selection criteria used to include (or exclude) companies are not always transparent and the approach is not easily checked or replicated by other researchers. Moreover, if selection is restricted to companies already identified within pre-designated green technology categories, then green companies that are listed outside of these categories will be overlooked.

4. A search-based method for identifying green goods enterprises

In order to address the issues raised by current SIC-based approaches to identifying green firms, and lacking the resources to undertake our own large-scale field survey of enterprises, we have sought to develop a new method. We denote this as a search-based method, because we formulate and apply a comprehensive set of key search terms that describe a set of green technologies, products and processes. These search terms are not dependent on SIC codes, and thus can be applied across a range of sectors, although in the example we present in this section, we have fine-tuned and deployed the approach to identify green goods (rather than green services) firms. The green goods enterprise method that we have developed can be driven by information derived from textual searches of business databases (which is the case we discuss here) but is also applicable to searching other sources, such as enterprise web sites. Additionally, a search term approach is not only replicable by others, but can also be readily refined to improve accuracy and updated to incorporate new technological developments. The method also allows structuring the search approach within categories of green applications (such as carbon reduction or waste minimization). Previously, similar search-based methods have been successfully applied with respect to other emerging technologies, nanotechnology in particular (cf. [2,35]).

The starting point for our green goods search-based method was the development of key search terms. After investigating several options, we used as a base the list of expert-validated green technology terms developed for the UK by Innovas Solutions Ltd [23]. As discussed earlier, this offered a typology of green goods and services comprised of 23 categories in three broad areas: environmental, renewable energy, and emerging low carbon. We focused on the terms related to green technology manufacturing firms (hence in our own application did not include terms related to environmental consultancy or carbon finance, although these could be added if needed). We
then drew on other sources to enrich the search approach. We added terms in general categories related to green manufacturing technologies and also introduced specific terms included in the Derwent Manual Codes for Green Technology, including biological treatment, electrochemical processes, and batteries. We selected terms that had relevance to green goods sectors and which could be disambiguated by adding Boolean operators such as “AND” and “OR” in order to optimize precision and recall.5 Our final search approach for green goods sectors comprised four major categories (general, environmental, energy, and emerging low carbon) and 26 sub-classifications (see Appendix A).

The underlying research need for developing these search terms was to identify small and mid-size enterprises (SMEs) in green goods industries so as to examine the factors that cause some of these firms to grow particularly rapidly over time [39]. SMEs are usually defined as companies which are not members of a larger group and which fall below certain thresholds by employment size, turnover, and/or assets. In Europe, including the UK, companies with not more than 250 employees are typically defined as SMEs by employment size. In the United States, small businesses are generally defined by federal agencies as those with fewer than 500 employees. The justification for our focus on SMEs draws on the rationale that with growing global needs and demands to address sustainability challenges through novel solutions and systems, there will likely be multiple opportunities for the establishment of new startups and the growth of existing SMEs in green goods sectors as drivers of innovation, components of eco-innovation networks and clusters, and partners in the supply chains of larger incumbent firms [31]. Classifying and identifying green goods SMEs is a foundation step in analyzing their formation and growth trajectories and how they might be influenced by policy. For the purposes of developing and piloting our green goods search term approach, we focus on the United Kingdom (UK). In concurrent work, these green goods search terms have been used to identify SMEs in the United States [42]. We are also applying this search approach to identify Chinese green goods SMEs. (It would also be possible to apply these terms to identify the green goods activities of larger enterprises, although in our current work we are focusing on SMEs.)

To identify a panel of relevant green goods SMEs in the UK, we used FAME — an enterprise database which contains legal, organizational, financial, business line, and other information on millions of UK companies ([10]). FAME allows us to search not only for companies by year of incorporation, employment size, or industrial sector but also in the text of fields related to trade description and business lines. We applied the goods search terms to these fields for companies that were SMEs when incorporated within our target years, then undertook manual reviews (by two separate coders) of the companies so identified using the FAME record and web searches.

The major steps within this search and review process were as follows. First, we identified companies that matched the broad eligibility criteria of our study. In this case, we searched for companies incorporated between 1995 and 2007. Since we were interested in the growth of SMEs, we added the criteria that the company had to be of small or medium size (using the definition of under 500 employees) during its early years. FAME did not report data on prior employment earlier than 2002, so we applied the size criteria to any year between 2002 and 2007. Second, within this group of companies we searched for any of our key terms within the “Trade Description”, “UK SIC code” and “Overview” fields available in FAME. We note that these fields are self-reported by the companies themselves. Each of our 26 sub-category searches was undertaken sequentially. Third, using manual review and in some cases industrial classifications, we eliminated firms that were clearly services firms (given our interest in green goods producers). For the remaining firms, further manual review using the FAME record and web sources was undertaken to classify firms by their relevance to green goods production. A four-point coding system (very relevant; relevant; somewhat relevant; not relevant) was applied. The “relevance” criterion was defined empirically by the participating researchers based on the “Overview” and “Trade description” of the firms’ activities as well as the firms’ web sites. In order to increase the robustness of the process and include an acceptable validation mechanism, each record was coded independently by two data reviewers and then control-validated by the overseeing senior researcher to reconcile any differences and assign a final relevance score (1 to 4). This exercise resulted in a population of about 500 green goods companies, excluding services companies and companies where we could not determine from secondary evidence that they undertook manufacturing and production activities. From this set, we selected a sample of 304 companies that appeared most relevant to green goods production using the top two coding categories “1” (very relevant) and “2” (relevant) (Table 2). The next section discusses in detail the outcomes from our green goods search process.

5. Search outcomes

The search process identified a broad range of enterprises reporting that they were engaged in green goods manufacturing.

Table 2

Distribution of UK green goods companies using green goods search approach.
Source: Application of green goods search approach (Appendix A) to FAME business database [10].

<table>
<thead>
<tr>
<th>Sector</th>
<th>All identified enterprises</th>
<th>Green goods enterprise sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Pollution control</td>
<td>42</td>
<td>8.4</td>
</tr>
<tr>
<td>Building technologies</td>
<td>113</td>
<td>22.6</td>
</tr>
<tr>
<td>Battery</td>
<td>16</td>
<td>3.2</td>
</tr>
<tr>
<td>Alternative vehicle/fuel</td>
<td>141</td>
<td>28.2</td>
</tr>
<tr>
<td>Renewables and water treatment</td>
<td>103</td>
<td>20.6</td>
</tr>
<tr>
<td>Other green goods</td>
<td>85</td>
<td>17.0</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>100.0</td>
</tr>
</tbody>
</table>

a All enterprises identified by the green goods search approach (April-May 2012).

b Enterprises classified in top two classes (very relevant and relevant), see text.
By broad categories, the leading sectors for the 304 sample UK enterprises were green building technologies (24%), alternative vehicles and fuels (24%), renewables and water treatment (20%), pollution control (8%), and batteries (4%) (Table 3). The underlying data set records enterprises by units of legal business incorporation. Of our UK green goods sample, almost half (46.7%) were independent entities, about one half (49.0%) were subsidiaries of parent companies (enterprises owned by other firms), and the balance (4.3%) comprised holding companies (enterprises that owned other firms). Almost one-third (31.6%) of the enterprises were subsidiaries of non-UK companies (Table 3). This mix of ownership forms reflects the nature of the UK small and mid-size enterprise sector, where many SMEs are ultimately owned, or have been taken over, by larger parent companies and where there is a relatively high degree of foreign ownership.

By employment size, small enterprises with 50 or fewer employees comprised 38.2% of the UK green goods sample, while 38.5% were enterprises with 51–250 employees (Table 3). Some 21 enterprises (6.9% of the sample) employed more than 251 employees in 2010. (In our selection process, enterprises were included if they had up to 500 employees between 2002 and 2007, with the possibility that some might have grown beyond this by 2010). There were 50 enterprises (16.4%) with missing data in FAME. These are probably very small firms exempt from returning accounts data, hence are not fully captured in FAME.

We examined the sales and employment growth rates of our enterprise sample by sector. By change in nominal turnover, we see that not all green sectors grew over the last decade or so. The fastest employment growth, particularly for 2008–2011 period covering the recession. The fastest employment growth rates in alternative vehicles/fuel and pollution control. However, over this period, the battery sector saw negative growth rates in alternative vehicles/fuel and pollution control. Over this period, the battery sector saw negative growth rates in alternative vehicles/fuel and pollution control. By employment size, small enterprises with 50 or fewer employees comprised 38.2% of the UK green goods sample, while 38.5% were enterprises with 51–250 employees (Table 3).

Table 4
Source: UK green goods enterprise sample (see Table 2). Averages over annual nominal turnover and employment growth rates in specified years, from FAME.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Average annual growth rate (percent) over period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollution control</td>
<td>39.7</td>
</tr>
<tr>
<td>Building technologies</td>
<td>8.9</td>
</tr>
<tr>
<td>Battery</td>
<td>10.2</td>
</tr>
<tr>
<td>Alternative vehicle/fuel</td>
<td>34.8</td>
</tr>
<tr>
<td>Renewables and water treatment</td>
<td>24.5</td>
</tr>
<tr>
<td>Other green goods</td>
<td>17.3</td>
</tr>
<tr>
<td>Total</td>
<td>21.2</td>
</tr>
</tbody>
</table>


One of the hypothesized advantages of our search approach method is that it finds green goods companies that strict SIC green technology allocation methods would overlook. Our comparison with SIC classifications suggests that this seems to be the case. There are limitations and cautions associated with this comparison. We find that firms often operate in more than one SIC code (although FAME should prioritize the primary SIC code, this allocation may not be precise). Our manual inspections also indicate many errors in SIC code allocations, where the descriptions of what companies (in their trade descriptions and on their web sites) do not or no longer match their allocated SIC designations. We find companies in manufacturing that appear to do no manufacturing, with some companies listed in services sectors that do produce manufactured goods. Additionally, SIC codes tend to be based on historic classifications of technologies, meaning that some new green technologies are hard to allocate to current SIC classes. A caution related to our search term approach is that firms may modify how they describe their products to present them as environmentally friendly given the general growth of environmental concerns.

With these caveats acknowledged, we compared the distribution of our UK green goods SME sample with that of all enterprises in the FAME database, using SIC codes (Fig. 1). We observe that some of the most dominant sectors in the overall distribution of all firms are not as important in the green goods firms’ selection for two reasons. First, as we define our selection as goods producing (i.e. manufacturing) firms, the relative importance of service sectors in our selection such as “office administrative, office support and other business activities” (overall: 10.1%, our sample: 7.6%), “activities of head
Fig. 1. Relative importance of sectors in the population and green goods sample.

Sources: Authors’ analysis of UK green goods enterprise sample and all enterprises in FAME.
offices; management consultancy activities” (overall: 6.1%, our sample: 1.0%) and “computer programming, consultancy and related services” (overall: 4.6%, our sample: 0.3%) is relatively low compared to their overall relative importance. The fact that the proportion of these service SIC codes is not zero in our manufacturing sample selection is due to the above discussed data integrity issues related to SIC descriptors. Overall, while only 0.46% of the all firms in the FAME database are green goods SMEs, this ratio increases to almost one quarter in electrical equipment manufacturing sector and over one-fifth in motor vehicle manufacturing.

6. Conclusions

This paper details a method for identifying green goods enterprises using a search term approach rather than relying on the conventional method of pre-selecting and designating SIC classes as green. Building on prior work by others, we developed a broadly-based series of green goods search terms and applied this to the FAME database of enterprises in the UK, restricting the selection to enterprises that were established between 1995 and 2007 and which were small or medium sized (under 500 employees) during the period 2002–2007. The enterprises that were identified in this search process were then reviewed manually to ensure that they undertook activities relevant to green goods production. Using the associated business data set, we presented descriptive statistics by sector, ownership, and sales and employment growth for the sample of selected green goods enterprises. Descriptive business data was available for most (although not all) of the selected enterprises.

Given that firms vary in the ways in which they report their lines of business and products, we anticipate that it is appropriate to use an approach that searches broadly but then in a second round excludes those enterprises that fit less well. There are some limitations to our approach. The precision of the search results depends not only on the accuracy of the terms included but also on how well firms describe their activities when providing the company reports used by business databases. In our approach, sample precision is improved by manual review of all identified firms. In future iterations, it is likely that search precision could be improved by refining the search terms. There is also a caveat inherent in the underlying database. While we searched for firms that were incorporated in the period 1995–2007, some of these firms may have been established earlier but were re-registered as a new company. Further investigation of these firms could identify this, if it was desired to exclude such firms. A limitation of our method is that it needs access to an underlying source of data that includes searchable text fields that reasonably describe what companies are doing. In the UK, such data is available in proprietary but still accessible business databases. Comparable data is also available in the US and other developed economies. It may be less available in developing economies, although there is the option (not discussed in this paper) to apply the search terms to enterprise web sites. Of course, when working in countries where the language is other than English, there would be a need to translate and refine the search terms in the appropriate home language.

While there are limitations to the search term identification approach presented in this paper, there are advantages in that it captures relevant companies notwithstanding SIC designations. The method can be used for international comparisons. In our own further research, we are analyzing companies identified through a green goods identification search approach not only in the UK but also in the United States and China to examine factors associated with growth. A benefit of this method is that it can readily incorporate search terms for new technologies (such as green goods technologies) that are not captured by specific SIC classifications. The approach allows structuring and analysis within categories of green applications and allows terms that describe new green innovations to be introduced. Broad indications can be gleaned from some of the search terms as to whether green goods are comprised of novel materials or systems. However, there are clear limitations here as in many cases both adapted and new technologies may be captured by the same search terms. Subsequent review of additional information (for example, from web sources) about the company and its products would help if it was desired to distinguish types of adaption, novelty, and applications of specific green goods technologies. Nonetheless, a benefit of the approach is that it provides researchers control over defining sectors of interest, which is an important feature where there are contested definitions of what is a green technology. For instance, some argue that nuclear power and biofuels are neither green nor environmentally friendly. Our method allows transparent and flexible treatment of such issues, and contested terms can be explicitly included or excluded.

The approach in this paper also has broader implications for technology forecasting and the analysis of social change. In an era not only of rapid technological change and the emergence of new crosscutting technologies but also of large-scale data, there are new opportunities and needs to analyze structured and unstructured text to develop intelligence about technological and societal trends. Conventional data sources, with their historic industrial classifications and time lags in publication, can now be greatly supplemented by the use of search methods that can address new technologies and web-based and other online sources. The method in this paper can readily be updated, focused or expanded, and it can be adapted to identify and track companies with business lines in other technologies and fields on a real-time basis. Adaptations are possible for bibliometric, patent, and social media searches. Finally, there are also parallel opportunities to build on the approach with other more modeling strategies (including topic modeling) to identify and cluster emerging technological fields.

Acknowledgments

Research for this working paper was supported by the Project on Sustaining Growth for Innovative New Enterprises, sponsored by the UK Economic and Social Research Council (grant reference: ES/J008303/1). The opinions and conclusions presented in the paper are those of the authors and should not be attributed to the project sponsor. The authors gratefully acknowledge input and feedback from other project colleagues, including Jan Youtie and Sanjay Arora.
## Appendix A. Green goods sectors — search terms.

<table>
<thead>
<tr>
<th>Search number, sector and sub-sector</th>
<th>Search terms</th>
<th>FAME enterprises (1995–2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No. Categories 1 &amp; 2</td>
</tr>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(sustainab* OR (green good*) OR (green technolog*) OR (green innov*) OR (eco<em>innov</em>) OR (green manufac*) OR (green prod*) OR pollut* OR (ecolabel) OR (environ* product declarat*) OR (EPD AND environ*) OR (environ* prefer* product*) OR (environ* label*))</td>
<td>93 30</td>
</tr>
<tr>
<td>2 All-purpose</td>
<td>([natur* environ*] OR (environ* friend*) OR (environment* conserv*) OR biocompat*) OR biodivers* OR filter* OR filtra* OR (synth* gas*) OR regenerat* OR recircul* OR gasification OR gasifier OR (fluidized clean gas) OR (gas cleaning))</td>
<td>125 48</td>
</tr>
<tr>
<td>3 Biological treatment</td>
<td>([biogas* OR bioreact* OR polyolef* OR biopolymer* OR disinfect* OR biofilm* OR biosens* OR biosolid* OR caprolact* OR ([ultraviol* OR UV) AND (radiat* OR sol*)]) AND (bioremed* OR biorecov* OR (biolog* treat*) OR biodegrad*))</td>
<td>0 0</td>
</tr>
<tr>
<td>4 Air pollution</td>
<td>([(air* contr*) OR (particul* contr*) OR (air* qual*]) AND (pollut*))</td>
<td>14 7</td>
</tr>
<tr>
<td>5 Environmental monitoring and analysis</td>
<td>([environ* monitor*) AND (((environ* AND instrument*) OR (environ* AND analys*)) OR (life<em>cycle analysis) OR (life cycle analys</em>)))</td>
<td>23 5</td>
</tr>
<tr>
<td>6 Marine pollution control</td>
<td>([marin* control*) AND pollut*)</td>
<td>2 1</td>
</tr>
<tr>
<td>7 Noise &amp; vibration control</td>
<td>([(nois* abat*) OR (nois* reduc*) OR (nois* lessen*)])</td>
<td>5 2</td>
</tr>
<tr>
<td>8 Contaminated land reclamation &amp; remediation</td>
<td>(land AND (reclam* OR remediat* OR contamin*))</td>
<td>116 9</td>
</tr>
<tr>
<td>9 Waste management</td>
<td>(wast* OR sewag* OR inciner*)</td>
<td>112 13</td>
</tr>
<tr>
<td>10 Water supply and waste treatment</td>
<td>([(slurr* OR sludg* OR (aque* solution*) OR wastewat* OR effluent* OR sediment* OR 70coul* OR (wind power*) OR (wind energ*) OR (wind farm*) OR (turbin* AND wind*))</td>
<td>60 23</td>
</tr>
<tr>
<td>11 Recovery and recycling</td>
<td>(recycl* OR compost* OR (stock process*) OR (coal combust*) OR remanufactur* OR (coal AND (PCC) OR (circular fluid bed combust*) OR (combust* AND CFBC))</td>
<td>110 48</td>
</tr>
<tr>
<td><strong>Renewable energy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 All-purpose</td>
<td>(renewabl* AND (energ* OR electric*))</td>
<td>20 17</td>
</tr>
<tr>
<td>13 Wave &amp; tidal</td>
<td>(((two basin schem*) OR (wave* ener*) OR (tid* ener*) AND (electric*)))</td>
<td>13 10</td>
</tr>
<tr>
<td>14 Biomass</td>
<td>(biomass* OR (enzymat* hydrolys*) OR (bio<em>bas</em> product*))</td>
<td>58 23</td>
</tr>
<tr>
<td>15 Wind</td>
<td>(wind power*) OR (wind ener*) OR (wind farm*) OR (turbin* AND wind*))</td>
<td>40 35</td>
</tr>
<tr>
<td>16 Geothermal</td>
<td>(((whole system*) AND geotherm*) OR geotherm* OR 5oexchang*)</td>
<td>10 7</td>
</tr>
<tr>
<td>17 Photovoltaic &amp; solar</td>
<td>(solar* AND (ener* OR (linear fresnel sys*) OR electric* OR cell*) OR (heat* OR cool* OR photovolt* OR PV OR cdtre OR (cadmium tellurid*) OR PVC-U OR photoelect* OR photactiv* OR (sol<em>gel</em> process*) OR (evacuat* tub*) OR (flat plate collect*) OR (roof integrat* system*))</td>
<td>102 50</td>
</tr>
<tr>
<td><strong>Emerging low carbon</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 All-purpose</td>
<td>(((low carbon) OR (zero carbon) OR (no carbon) OR (0 carbon) OR (loww<em>carbon) OR (zero</em>carbon) OR (no(carbon))</td>
<td>63 12</td>
</tr>
<tr>
<td>19 Alternative fuel vehicle</td>
<td>(((electric* vehic*) OR (hybrid vehic*) OR (electric* motor*) OR (hybrid motor*) OR (electric* driv*) OR (electric* car*) OR (hybrid car*) OR (electric* machin*)) OR (electric* auto*) OR (hybrid auto*) OR (yaw* cat* sens*))</td>
<td>449 145</td>
</tr>
<tr>
<td>20 Alternative fuels</td>
<td>(((alternat* fuel*) OR (mainstream* fuel*) OR (fuel cell*)) OR (nuclear powe*) OR (nuclear stat*) OR (nuclear plant*) OR (nuclear ener*) OR (nuclear AND electric*) OR (nuclear fuel*) OR (fuel* process*) OR (porous* struct*) OR (porous* substrat*) OR (solid* oxid* fuel*) OR (Fischer<em>Tropsch</em>) OR (refus* deriv* fuel*) OR (fuel* AND biotech* AND (ethanol* OR hydrogen*)) OR (biofuel*) OR (synthetic fuel) OR (combined heat and power) OR (synth* gas*) OR (sysgyas))</td>
<td>299 57</td>
</tr>
<tr>
<td>21 Electrochemical processes</td>
<td>((electrochem* cell*) OR (electrochem* fuel*) OR (membran* electrod*) OR (ion* exchang* membran*) OR (ion<em>exchang</em> membran*) OR (electrolyt* cell*) OR (catalyt* convers*) OR (solid* separat*) OR (membran* separat*) OR (ion* exchang* resinf*) OR (ion<em>exchang</em> resinf*) OR (proton* exchang* membra*) OR (proton<em>exchang</em> membra*) OR (catalyt* reduc*) OR (electrod* membra*) OR (therm* engin*))</td>
<td>85 27</td>
</tr>
<tr>
<td>22 Battery</td>
<td>((batter* OR accumul*) AND (charge* OR rechar*) OR turbocharg* OR (high capacit*) OR (rapid charg*) OR (long life) OR (Ultra* OR solar* OR (no lead) OR (no mercury) OR (no cadmium) OR (lithium* ion*) OR (lithium* ion*) OR (Li*ion)))</td>
<td>23 16</td>
</tr>
<tr>
<td>23 Additional energy sources</td>
<td>(((addition* energ* sourc*) OR (addition* sourc* of ener*))</td>
<td>14 8</td>
</tr>
<tr>
<td>24 Carbon capture &amp; storage</td>
<td>(((carbon AND captu*) OR (carbon AND stor*) OR (carbon dioxid*) OR CO2)</td>
<td>61 21</td>
</tr>
<tr>
<td>25 Energy management</td>
<td>(((ener* sav*) OR (ener* effic*) OR (energ<em>effic</em>) OR (energ.sav*) OR (light* emit* diod*)) OR LED OR (organic LED) OR OLED OR CFL OR (compact fluorescent*) OR (energ* conserve*))</td>
<td>66 39</td>
</tr>
<tr>
<td>26 Building technologies</td>
<td>(((build* OR construct*) AND (insula* OR (heat* retent*) OR (heat* exchang*) OR (heat* pump*)) OR (therm* exchange) OR (therm* decompos*) OR (therm* energ*) OR (therm* communic*) OR (thermoplast* OR thermocoup* OR (heat* recover*)))</td>
<td>183 53</td>
</tr>
</tbody>
</table>
References


[28] PCAST, Report to the President on Capturing Domestic Competitive Advantage in Advanced Manufacturing, President’s Council of Advisors on Science and Technology, Executive Office of the President, Washington, DC, April 2012.


policy evaluation. He holds a Ph.D. in City and Regional Planning and an MA in Economics from the University of California, Berkeley.

**Dr Abdullah Gök** is a Research Fellow at the Manchester Institute of Innovation Research (MIoIR), Manchester Business School, University of Manchester. His research spans i) behavioral additionality, ii) the concepts, methods and findings of evaluation of science and innovation policies and iii) the use of advanced and innovative methods to address a variety of micro- and macro-level research questions in innovation studies. He holds a BSc in Economics and an MSc in Science and Technology Policy Studies from METU, Turkey, and a PhD in Science and Innovation Policy from the University of Manchester.

**Evgeny Klochikhin** is a Researcher at the American Institutes for Research in Washington, D.C., USA, with expertise and experience in innovation studies, science policy, international relations, and world development. His current work applies advanced methods of computational linguistics and natural language processing to the study of science policy and organization. He is involved in several research projects, including the Sloan Foundation-funded project on Examining the Outcomes of Investments in Caltech Entrepreneurs and the ESRC-funded project on Sustaining Growth for Innovative New Enterprises. He holds a PhD in Innovation Policy and Management from the University of Manchester and a BA in Area Studies and an MA in International Relations from the Moscow State Institute of International Relations (MGIMO-University) of the Russian Ministry of Foreign Affairs.

**Dr Marianne Sensier** is a Research Fellow in Economics at the University of Manchester. At Manchester, she has worked on a range of research projects and was a Hallsworth Fellow at the Centre for Growth and Business Cycle Research. She is currently working on two ESRC projects at the Manchester Institute of Innovation Research: Sustaining Growth for Innovative New Enterprises, and Emerging Technologies, Trajectories and Implications of Next Generation Innovation Systems Development. Additionally, she is investigating the resilience of regions in Europe with an ESPON research project based at Cardiff University. She obtained a BA in Economics with Econometrics and a PhD in Econometrics from the University of Sheffield, and is a member of the Euro Area Business Cycle Network (EABCN).