Mapping the industrial base for the new energy economy: predicting competitiveness in clean energy economic opportunities

Ishana Ratan, Jonas Goldman, Becky Han, Benjamin Bagozzi, Tim Sahay, and Bentley Allan

Abstract: Countries all over the world are designing green industrial policies to upgrade their economies and position their firms in global value chains. If successful, these strategies could continue to generate cost declines, locking in decarbonization pathways that draw on new energy technologies. However, there are two central risks. First, countries could make investments in sectors where they cannot be competitive. Second, poor quality information about the opportunity set and the specific nature of economic opportunities will lead many countries to compete in a small number of technologies. Instead, countries need to spread out and compete in their areas of strength but this requires granular tools to map capabilities and opportunities. In this paper, we seek to identify the priority sectors for investment in all countries. This in turn relies on answers to deeper questions. What are the countries' comparative advantages? What does the country have to contribute to net zero supply chains that it could transform into competitive advantages? We present a machine learning model that predicts country competitiveness across 10 clean energy technologies.

Keywords: industrial policy; industrial base; green comparative advantage; exports; machine learning; global value chains

Introduction

Countries all over the world are designing green industrial policies to upgrade their economies and position their firms in global value chains.[1] The success of China's industrial strategy for wind, solar, and batteries has spurred action in both developing and developed countries. Focused investment and coordination in China and elsewhere drove down the cost of these technologies to the point that they are now competitive with conventional energy incumbents. Now there is competition across a range of technology verticals including heat pumps, electrolyzers, permanent magnets, biofuels, nuclear, geothermal, and transmission equipment. If successful, these strategies could continue to generate cost declines, locking in decarbonization pathways that draw on new energy technologies.

However, there are two central risks. First, countries could make investments in sectors where they cannot be competitive. Countries pursuing green industrial policy must recognize that competitive advantage in high technology areas can be built over time, as China's example shows, but that it is not infinitely elastic in the short-term. Thus, industrial strategies are more likely to be successful if they build on existing strengths while charting pathways into new areas.

Second, poor quality information about the opportunity set and the specific nature of economic opportunities will lead many countries to compete in a small number of technologies. For example, in 2020 and 2021, many countries announced hydrogen or solar manufacturing strategies. This herding and strategic concentration can create real problems for developing countries, as overinvestment in coffee and other commodities did in the 1960s. Rising production pushed down prices, and exposed many countries' bets.

Instead, countries need to spread out and compete in their areas of strength. As Nahm has shown, clean energy technologies are built on complex global divisions of labour in which countries develop niche comparative advantages rooted in their economic institutions.[2] What this means is that countries' opportunities are likely to be far more specific than policy-makers may think, requiring granular tools to map capabilities and opportunities.

For both these reasons, countries need rigorous analytics and strategy to support industrial policy design. In this paper, we seek to help countries address a central problem: what are the **priority sectors** for investment? This in turn relies on answers to deeper questions. What are the countries' comparative advantages? What does the country have to contribute to net zero supply chains that it could transform into competitive advantages?

A number of studies have now set out to identify opportunities. The seminal work of Hausmann and Hidalgo on economic complexity argues that countries' should seek to develop capabilities in products that are related or adjacent in the product space to their current exports.[3] At the core of their approach is the idea that economies possess knowledge and skills that are not easy to acquire.

This is consistent with in-depth qualitative work on manufacturing and innovation clusters. Berger and her collaborators argue that successful industries emerge from dense interactions between the lab and the factory floor because only those interactions can build the knowledge needed to sustain advanced industries.[4]

Rosenow and Mealy, building on earlier work by Mealy and coauthors, create indices to assess countries decarbonization technology strengths and opportunities.[5] They follow

Hausmann and Hidalgo in arguing that countries have opportunities in areas where their current capabilities align with their current revealed comparative advantage profile.

In this paper, we build on the critical insights of this work to develop a machine learning model that predicts country competitiveness across 10 clean energy technologies. Consistent with the approaches above, we build the model on global export data because it provides the granularity necessary to deeply understand strengths and capabilities. Our model used three sets of independent variables: revealed comparative advantage in products in the supply and process chains for each technology, the co-exports for these products, and a slate of country characteristics (including GDP, manufacturing as a percentage of GDP, etc.).

We used a random forest model for classification and regression against the final product in each technology vertical. The strength of a random forest model is that unlike other strategies, it is broadly inductive and does not require us to make strong assumptions about what will predict competitiveness in final product.

We present results for 10 clean energy supply chains including the top predictors in each technology. We find that advanced economies including China, Japan, Korea, Germany, and the United States are competitive across many verticals. But there are many opportunities for emerging markets and developing economies (EMDE) with India, Malaysia, Thailand, Philippines, Turkey, and Mexico exhibiting particular strength.

The model can support countries' strategic focus by highlighting granular areas of strength. Further it encourages countries to diversify their investments and compete in a range of sectors. But it also provides a non-mechanical view on how countries should think about industrial strategy. A simple, low-risk strategy to compete in a given vertical is to build the industrial base in the underlying areas necessary for that technology. This finding is consistent with seminal and recent findings in the literature on industrial strategy that highlight the importance of investments in the upstream industrial base.[6] Korea, for example, transformed itself into a modern manufacturing and chemicals powerhouse by strongly subsidizing upstream metal and chemical production and allowing the benefits to flow downstream.

A key finding of our model is that five sets of capabilities are critical for predicting country competitiveness: electronics, machinery, mining and metals, industrial materials, and chemicals. Moreover, each technology has a particular profile corresponding to the kind of industrial base needed to be competitive. For example, solar competitiveness is driven primarily by strength in electronics and industrial materials (e.g. glass), while battery competitiveness is driven by strength in metals and mining and chemicals.

The model

[Explain ML via RF]

As in Rosenow and Mealy, we begin by calculating countries' revealed comparative advantage in all the products along the supply chain. These continuous RCA scores in product components are the key independent variables. But in a critical difference between our model and others', we also include RCA scores in what we call the process chain in our list of product components.

The process chain includes all the technologies that are needed to make each part in the supply chain. For example, in batteries we include the HS code for cathode active material and the copper foil, but also the mixers used to transform that cathode active material into a slurry that is then sprayed on the foil. In this way, we sought to ensure that our model captured not just exports in the supply chain, but in the capital goods needed to produce the technologies.

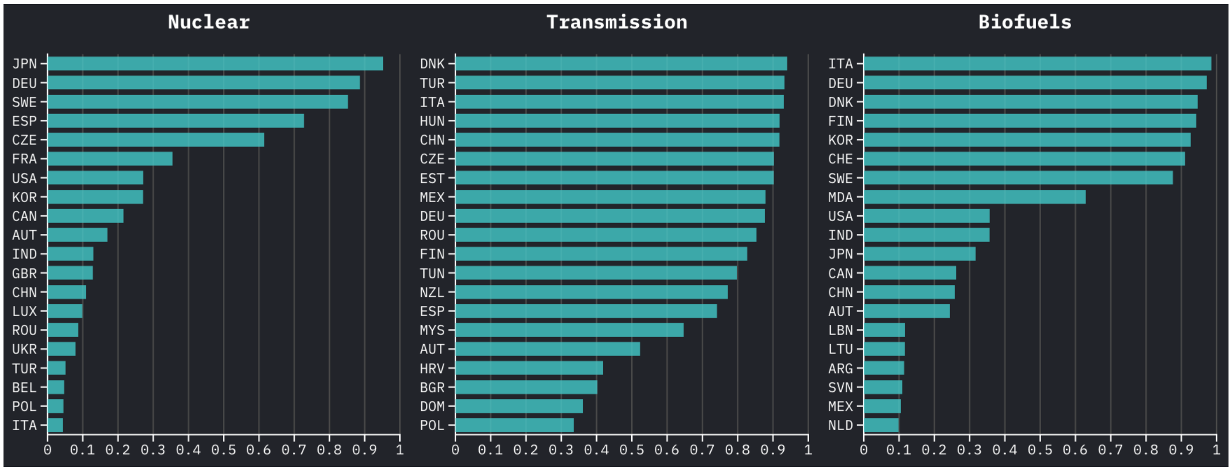
Just as capabilities in "related" products should predict competitiveness so too should capabilities in production techniques. Again, going back to Amsden, Berger and others, the idea is that knowledge is forged in interaction between the lab, capital goods manufacturers, and the factory floor. Engineers and scientists across these domains must interact and work together to increase innovation and efficiency. We feel our model captures this basic fact about successful manufacturing in a way other models do not.

The machine learning model allows us to include over one hundred of these predictors along the supply chain and among related components and identifies which are most influential in determining whether a country has revealed comparative advantage in a final product. Another advantage of setting up the model in this way is that it enables us to identify the specific products that are driving predictive competitiveness for each country over time. That is, we can identify each country's strengths in each sector. To do so, we calculated the predictors (SHAP values) driving competitiveness at the level of the technology, and for each individual country. We focus on the predictors with the highest SHAP values, standardized using the absolute value of z-scores, to identify which variables most strongly shape competitiveness.

Results

The model presents three kinds of results. First, a ranking of all countries by predictive competitiveness across the 10 technologies. Second, a list of the products driving competitiveness in each sector and in each country. Third, a simple model of the underlying global industrial base for green technology.

Figure 1. Top 20 countries by predicted competitiveness across all opportunity areas.



[Alt. Some version of Figure 3. in Rosenow and Mealy. 10 plots in one figure?]

The model performs well in that it identifies countries that currently produce and export the technologies. In solar, for example, all the leading producers appear in the top 10 countries. In batteries,

Second, the model provides a granular assessment of what drives competitiveness for each technology and each country. At the level of technology, we get a sense of which products are driving predicted competitiveness. These lists reveal that each technology has a unique set of competencies that predict competitiveness. The model also identifies key predictors for each country. Thus, we can decompose Japan's strength in solar to see which parts of its industrial base are driving its predicted competitiveness.

APPENDIX. Top 20 predictors for 3 opportunity areas with full HS code descriptions

[*do we want a table or bar graph? (e.g. bar corresponds to our SHAP measure)

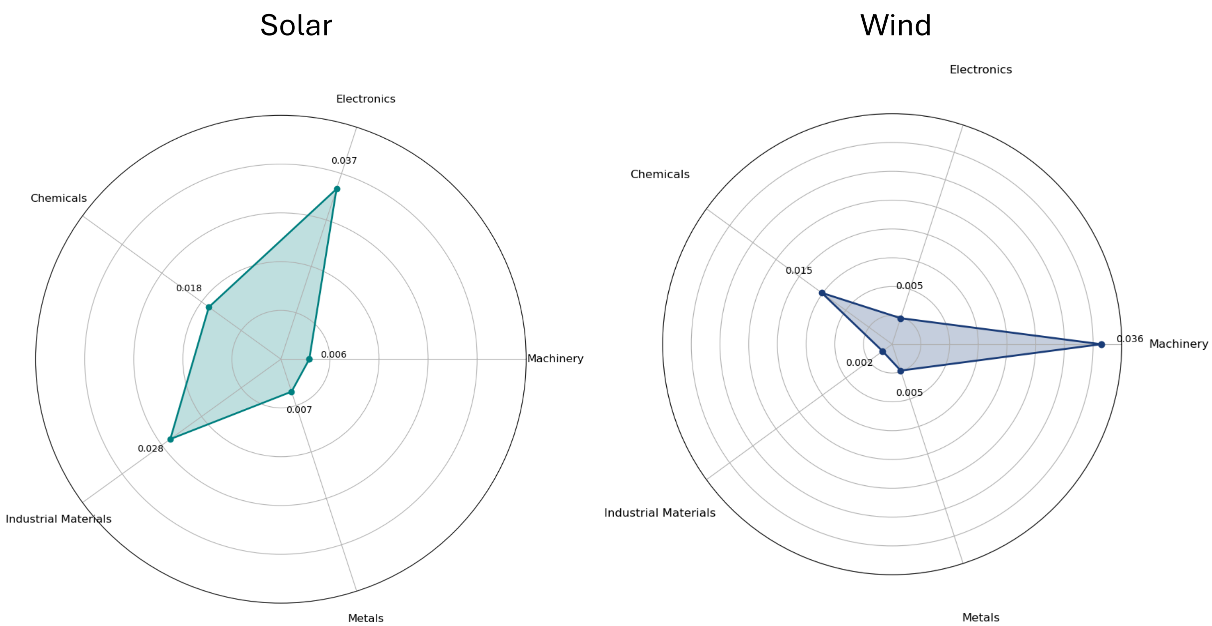
*table may be easier to include HS descriptions]

This suggests that strength in the immediate supply and process chain is not a necessary condition for success in each technology. The implication is that competitiveness emerges from something deeper. To make sense of this intuition, we classified the leading predictors via SHAP values according to the groups of HS code chapters. Categorize... We sum the probabilities within each categories.

The classification and analysis of key predictors shows that clean energy industrial capabilities cluster into five key capability areas: electronics, machinery, industrial materials, mining and metals, and chemicals. Interestingly, automotive and textile capabilities are almost irrelevant. We expected capabilities in these manufacturing sectors to be predictive, especially since textiles has emerged as a critical bridge to higher value manufacturing.

Moreover, each technology has a particular profile such that solar competitiveness is driven by strengths in electronics and industrial materials, batteries in chemicals mining and metals, and so on (Figure 4).

Figure 2. Radar plots for all the techs



Note: In this HS code system, chapter 84 refers to machinery, especially precision components and capital goods, chapter 85 corresponds to electronic devices and computing components, chapters 68-70 are industrial materials, and so on.

This is an intuitive but powerful finding. Solar panels are complex electronic devices attached to glass. Wind turbines integrate a range of precision components into an efficient moving machine. Batteries require complex metallurgical and chemical expertise to perfect metals performance. Electrolyzers bring together complex chemical polymers with precision machines. The value-add for heat pumps is concentrated in heat exchangers and other HVAC components in the machinery sections of the HS code list. Rare earth magnets require mining and chemical refining, but principally they are electromagnets. Nuclear installations are primarily complex machines with millions of machined components that use a complex mineral fuel. Biofuels competitiveness is predicted by strength in engines and engine components, in the machinery chapters of the HS code.

Figure 3. Radar plots for 9 countries generally

[]

The result is a parsimonious mapping of the global industrial base for clean energy technology that outlines the capabilities needed to compete in each technology. With this characterization of the industrial base in hand, the industrial base of each country can also be mapped in a simple way. This model gives countries a sense of their current strengths and a tool that enables them to map progress over time at both a high level and at a very granular level.

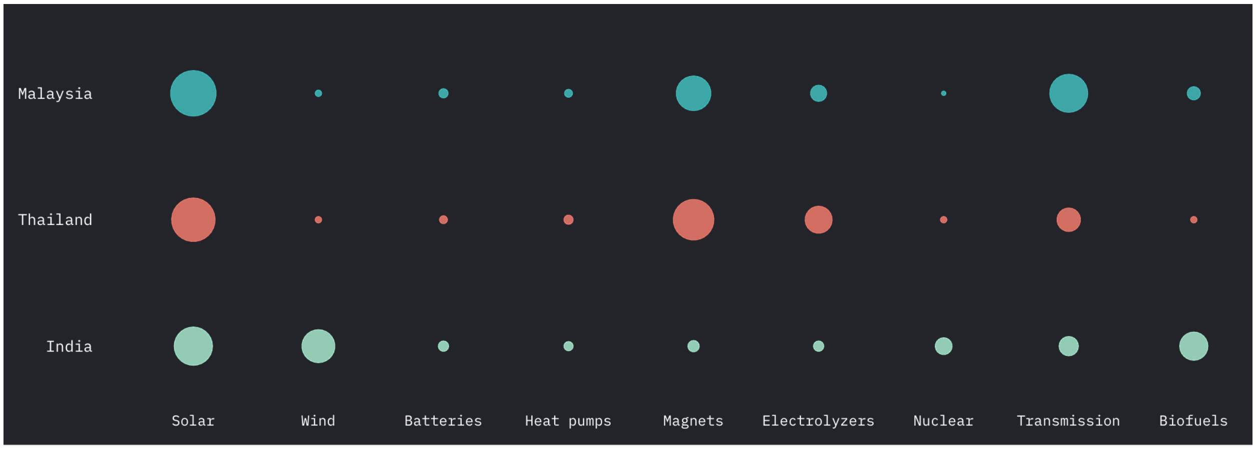
Discussion

The rankings of predicted competitiveness highlight the strength of the world's industrial giants: China, Japan, Korea, and Germany across a wide range of opportunity areas (Table 1). The United States has strengths in batteries, nuclear, and biofuels, but for its size it lags far behind its peers. Denmark, Italy, Spain, and France also emerge as strong competitors alongside a range of smaller European countries. The industrial base for clean energy is highly concentrated in Asia and Europe.

However, many emerging market and developing economies also demonstrate broad strengths. India ranks ninth on average overall scores with good competitiveness in solar, wind, magnets, nuclear, and biofuels. This suggests that its recent manufacturing push has been more successful than many realize and that U.S. investments in India may be well-placed to generate long-term supply chain diversification.[7] Outside of the top 20, Thailand, Turkey, and Mexico all have strong overall rankings.

But the point of this model is not to rank overall strengths, but to allow countries to identify areas of competitive advantage so that they can make smart industrial strategy investments. Consider Figure 5, which summarizes competitiveness across all areas for a number of countries. Here the lesson is not that Malaysia, the top ranked EMDE country, is a powerhouse, but that it has particular strengths in solar, magnets, and transmission.

Figure 4. Technology competitiveness in selected countries, 2023



Thailand might easily conclude that it has strengths in electrolyzers and transmission it might like to boost. The model provides tools for thinking through how to do this. One could easily interpret the results in a highly mechanical way. For example, to be competitive in electrolyzers, a country could aim to develop germanium oxides & zirconium dioxide exports—the top predictor for electrolyzer competitiveness. Or, to compete in batteries, a country could aim to build an export sector in flat-rolled steel coated with chromium, the top predictor there. However, this would be applying a false level of precision to the model and misinterpret what it is telling us.

Instead, we can have confidence in the simpler, but more elegant finding that the knowledge and capabilities required to succeed in industrial development differs across technology areas. To compete, countries need to build upstream and intermediate capabilities in electronics, chemicals, machinery, industrial materials, and mining and metals. A comprehensive approach to building these capacities in focused areas is more likely to succeed.

The model is subject to further limitations. The dependent variable is export in the final product, so it does not tell us anything about how much value-added is happening in the economy. Re-exports could confound some of the results. We expect this is a limited problem, but it may help explain the overperformance of some countries. Cross-referencing the results with granular production or value-added data at the country level, not possible in a universal study like this, would be advisable before designing industrial strategy on the results.

Conclusion

The goal of green industrial policy is to secure a country's position in global clean energy value chains that supports technological development and the creation of innovative, competitive ecosystems at home. To do that, countries need to build on their strengths without remaining bound by the status quo.

The model presented here aims to balance future orientation with a realistic assessment of existing strengths. It helps countries identify which sectors countries can compete in based on existing capabilities. But beyond specific opportunities, it also identifies which parts of the industrial base drive predicted competitiveness. Thus, it can help identify where investments in the industrial base might be possible and beneficial to improve long-term export capability. Thus, it provides an in-depth picture of what countries are currently good at and how this positions them to produce green technologies in adjacent, emerging sectors.

It does this by presenting a simple mapping of the five core capabilities needed to compete in clean energy technologies. The industrial base revealed here consists of capabilities in electronics, machinery, mining and metals, industrial materials, and chemicals. This finding appears to support new work on industrial policy that suggests that states who invest in core, upstream capabilities are most likely to succeed. Further work could be done to understand the nature of this industrial base and its importance to technology and development.

The detail in country-level advice could also be improved by running the model at each stage of the supply chain, so that competitiveness in key components could also be ranked. This is important because the more specific industrial policy targeting is, the easier it is to build supportive ecosystems.

Another important extension is to understand the policy determinants of competitiveness. Which industrial and trade policies are most likely to improve competitiveness? Understanding the effectiveness of various policy tools will be critical for countries success.

---

[1] Meckling et al; Allan et al, GIP; Allan and Nahm; Juhasz, Lane, and Rodrik.
[2] Nahm 2021.
[3] César A. Hidalgo, Bailey Klinger, A-L. Barabási, and Ricardo Hausmann. "The product space conditions the development of nations." *Science* 317, no. 5837 (2007): 482-487.
[4] Berger 2013; Breznitz 2021.
[5] Rosenow and Mealy; Mealy and Tietelbaum 2022

[6] A.H. Amsden, *Asia's Next Giant* (New York: Oxford University Press), 1989; Ernest Liu, "Industrial Policies in Production Networks," *Quarterly Journal of Economics* (2019); Nathan Lane, "Manufacturing revolutions: Industrial policy and industrialization in South Korea," *The Quarterly Journal of Economics* (2025).

[7] U.S. International Development Finance Corporation, "DFC Announces Approval to Provide up to $500 Million of Debt Financing for First Solar's Vertically-Integrated Thin Film Solar Manufacturing Facility in India." https://www.dfc.gov/media/press-releases/dfc-announces-approval-provide-500-million-debt-financing-first-solars

APPENDICES

Table 1. Model calibration

Technology	Final product (HS code with description)	Market size (2023)
Solar		
Wind		
Batteries		
Electrolyzers		
Heat pumps		
Magnets		
Biofuels		
Geothermal		
Nuclear		
Transmission		

Table 2. Industrial base categories

Technology	HS Code Chapters
Electronics	85
Machinery	84, 90-96
Industrial materials	68-70
Mining and Metals	25-27, 71-83
Chemicals	28-40
Agriculture and resources	1-24, 41-49
Automotive and transport	86-89
Textiles	50-67