Winds of Change: Estimating Learning by Doing without Cost or Input Data *

Thomas R. Covert †
Richard L. Sweeney ‡

January 2024

Abstract

We measure how much learning has reduced costs in the wind turbine industry. As in many industrial settings, we do not observe costs, so we infer them using standard demand system data. To measure wind farm developer preferences, we embed a simple but physically realistic model of how wind turbine characteristics, like rotor size, relate to power production, into a standard discrete choice demand system. Next, we use a standard oligopoly model to invert these preferences and recover manufacturing costs, and their dependence on cumulative manufacturing experience. Because current sales increase future experience, manufacturers have dynamic incentives when setting prices. We account for these dynamic markdowns using methods developed in Berry and Pakes (2000), which allow us to control for dynamics without computing the equilibrium of a dynamic game. We find that a doubling of manufacturing experience reduces manufacturing costs by 14 to 29 percent. Only 1 to 2 percent of experience spills over to other turbine models produced by the same firm, and spillovers to turbines produced by other firms are on the order 0.1 to 0.6 percent. Though inter-firm spillovers are small, in aggregate, they are responsible for significant cost reductions over time. These results are consistent with policymaker motivation for generously subsidizing the industry.

Keywords: Innovation, Renewable Energy, Learning

*Both authors declare they have no interests, financial or otherwise, that relate to the research described in this paper, nor do they have any current ties, directly or indirectly to the energy industry. Financial support for data acquisition from Boston College, the University of Chicago's Becker Friedman Institute, and the NBER-Sloan conference on the Economics of Innovation in the Energy Sector is gratefully acknowledged. We also thank Jackson Dorsey, Gautam Gowrisankaran, Ashley Langer and Jing Li for helpful comments. Devin McNulty, Grace Park and Miriam Goldgeil provided excellent research assistance.

†University of Chicago Booth School of Business and NBER, Thomas.Covert@chicagobooth.edu.
‡Boston College, sweeneri@bc.edu.

1 Introduction

Researchers have been trying to measure the relationship between production costs and production experience, commonly referred to as "learning by doing" (LBD), for nearly a century (Wright, 1936). The primary challenge in this endeavor is the fact that although experience is usually easy to quantify, manufacturing costs are generally not recorded in research datasets. In lieu of this, recent empirical literature has made progress by studying special settings where detailed data on production *inputs* is available.[1] While analyses like this can, in principle, measure the true underlying relationship between manufacturing costs and experience, this kind of detailed data is rarely available. Moreover, given the rarity of such data, it is uncommon for it to be available for two or more firms in the same industry, making it impossible to measure spillovers. Given the importance of LBD for growth (Arrow, 1962), industrial and trade policy (Dasgupta and Stiglitz, 1988), and environmental policy (Jaffe et al., 2005), we need tools to estimate LBD, even when unit level costs or inputs are not available, as is frequently the case.

This paper presents a new method for measuring learning by doing using information that is typically available for econometric analysis of static demand models. Like much of the existing industrial organization literature that estimates such models, our approach assumes that manufacturers set prices to maximize some notion of present and future profits, and inverts this relationship to infer costs. However, unlike most existing applications, we propose a method which explicitly accounts for the dynamic incentives firms have to learn faster through their pricing choices (Besanko et al., 2014; Benkard, 2004). Although our method directly accounts for these dynamic incentives, it does not require their explicit computation, in contrast to existing full solution methods for dynamic games. Because our method does not require data beyond standard demand estimation datasets, it can also easily accomodate specifications that allow for spillovers across firms.

We use this method to measure the effect of manufacturing experience on costs in the global wind turbine industry over the past two decades. While the wind industry is now large, covering more than ten percent of global power generation capacity, at the start of our sample it was less than one percent. Much of this growth was fueled by generous government subsidies for wind farm project developers, with the explicit goal of instigating future cost declines in the wind turbine manufacturing sector through learning and knowledge spillovers. Existing research has quantified the *static* benefits of these policies, in the form

[1] Benkard (2000) estimates learning and forgetting in the commercial aircraft industry, where he observes the labor inputs into each aircraft produced at a manufacturer that decided to exit the market. Thornton and Thompson (2001) observe labor inputs for World War II boat building, and Levitt et al. (2013) observe manhours and output by shift at a major automobile assembly plant.

of reduced CO_2 local pollutant emissions from new power plant construction, and concluded that they are smaller than their associated costs (Van Benthem et al., 2008; Abrell et al., 2019; Greenstone and Nath, 2020). Thus, the extent to which renewable energy policy to date was welfare enahcing may depend on whether or not renewable energy device manufacturing exhibits significant learning economies. This paper provides direct evidence of these learning economies in the wind turbine industry.

We measure the evolution of manufacturing costs using data on the global wind turbine industry which covers the near-universe of wind turbine manufacturers and wind plants constructed using their turbines, between 2000-2019. Our data provides information about the choice set each plant faced — engineering estimates of the output each available turbine model would generate at each plant location — as well as the specific turbine models each plant chose to install. Unlike other discrete choice settings, where there is a published average price or MSRP, wind turbines are heavy industrial goods, and are acquired through informal procurement processes that do not generate public transaction records, so we have no prices. In light of this, we model each wind plant's turbine selection problem as a procurement scoring auction, and show how data on turbine characteristics, site wind speeds, and turbine choices, as well as standard discrete choice modelling tools, identify the otherwise latent "bids" that turbine manufacturers submit in this process.

Having characterized a notion of turbine prices as well as wind farm preferences over turbine characteristics, we specify and estimate a model of optimal turbine pricing in the presence of the dynamic incentives implied by learning-by-doing. Our model assumes that firms maximize the sum of current period profits from selling turbines and a continuation value which may depend on the sales of the firm's own turbines and the sales of other firm's turbines. Normally, models like this in the structural industrial organization literature assume that the continuation value satisfies the Bellman equation for an equilibrium of a dynamic game, and papers that use these models either fully solve the Bellman equation in a nested estimation procedure (e.g., Rust (1987)) or rely on a stationarity assumption to apply conditional choice probability approaches (Bajari et al., 2007). Neither approach is feasible in our setting, as the number of firms and turbines are large, implying a computationally large state space, the controls (prices) are set in a continuous, not discrete fashion, and a learning environment is, by definition, nonstationary, ruling out CCP estimators.

Instead, we adopt a procedure suggested in Berry and Pakes (2000) which characterizes a first-order condition for dynamic oligopoly problems with continuous controls in the spirit of the rational expectations and Euler equations frameworks. The Berry and Pakes (2000) insight is that when firms have optimally set a continuous control, like a bid in a procurement auction, the dynamic component of their decision problem can be expressed by a rational expectations term plus a shock. The rational expectations term is a function of the state transition probability distribution and subsequent realized observable terms. These objects are much easier to compute than the full equilibrium structure of a dynamic game, and scale easily with large numbers of firms, products and other state variables. We use this method to separately characterize the static manufacturing costs and dynamic pricing incentives that drive bids in the procurement auction. In implementing this idea, we also propose a new approach to handling endogeneity problems in rational expectations models by employing higher-order moment methods from the classical measurement error literature (Lewbel, 1997).

Using this framework, we document considerable learning by doing in wind turbine manufacturing and broad support for experience spillovers within and across firms. A doubling of manufacturing experience reduces manufacturing costs by 25 to 33 percent, an effect which is similar to the results on aircraft manufacturing in Benkard (2000). Additionally, spillovers appear to be important, though small on the margin. Approximately 1 to 2 percent of experience spills over to other turbine models produced by the same firm, and spillovers from turbines produced by other firms are on the order 0.1 to 0.2 percent. Though the marginal effects of spillovers are small, aggregate experience within a firm is often 2 orders of magnitude or more larger than turbine-specific experience, and aggregate experience outside the firm is another order of magnitude larger. Thus, spillovers have generated significant cost reductions over time. As an example, we show that when the Chinese wind industry began in the late 2000's, Chinese manufacturers' entered at a cost structure much more commensurate with established western firms than their limited own manufacturing experience would suggest.

The small scale of within- and across-firm spillovers that we estimate imply that for mature turbines, most experience capital is turbine specific. As a result, older turbines tend to have significant cost advantages over newer turbines. For example, our estimates suggest that a brand new 100 meter turbine has costs that are almost 6 times larger than a "mature" 90 meter turbine in its tenth year of production. The theoretical maximum performance of the larger turbine is only 24% higher than the smaller turbine, so it is unlikely that the firm could initially sell the larger turbine at a price high enough to offset the additional costs. However, a firm that manages to sell enough of the larger turbine eventually can make it at costs that are substantially lower. In the example above, after 6 years of sales, the 100 meter turbine has costs that are only 11% larger than contemporaneous costs of the smaller turbine. Thus, our results support the idea that new turbine introductions require a meaningful "experience investment" before they can generate positive operating profits.

In addition to measuring learning in an economically large and increasingly policy relevant industry, this paper also provides a novel and generic tool for measuring the effects of learning by doing, as well as any other dynamic component of a first-order optimality condition, in settings where only demand side data are available. The paper closest to ours in this dimension, and indeed a key inspiration for our functional form assumptions, is Irwin and Klenow (1994), which studied learning by doing and spillovers in the global semiconductor industry in the 1970s-1980s. Like this paper, Irwin and Klenow (1994) was only able to observe standard demand estimation data, like prices and quantities. However, to handle the dynamic incentives that semiconductor manufacturers may have faced, Irwin and Klenow (1994) assumed that Cournot quantity choices were first-order optimal with respect to a standard Euler condition, an approach that is not necessarily consistent with most modern models of dynamic oligopoly behavior. Our method of using the Berry and Pakes (2000) approach to a learning setting thus complements the original idea in Irwin and Klenow (1994), bringing its insights to the modern structural IO research paradigm.

Our paper also contributes to the broader learning-by-doing and innovation literatures with a specific focus on energy and environmental economics (Acemoglu et al., 2019). Newell et al. (1999) demonstrate that energy price shocks cause manufacturers to adopt more energy saving technology, consistent with an "induced innovation" hypothesis. Using data similar to ours, Knittel (2011) estimates considerable improvements in the production capabilities of car manufacturers, despite being seemingly constrained by fuel economy standards.

2 Background

In this section, we provide additional background on some key physical concepts and industry features that we leverage in estimation. To summarize, wind turbine production is quadratic in the size of the device, while the costs manufacturing a turbine are cubic in size. The industry is highly concentrated, and, while all the major players are active globally, "home" region preferences generate even more concentration at the market level. All manufacturers produce multiple turbine sizes at the same point in time, a fact we will later leverage in our estimation strategy.

2.1 Wind Power Basics

A wind turbine consists of a rotor with three long blades connected to a gearbox and generator atop a large tower. As wind passes through the blades, the rotor spins a drive shaft connected through a series of gears to a generator that converts this kinetic energy to electrical energy. The amount of energy such a device can capture is given by "Betz" law:

$$\text{Power}\ \ Q = C_p \left[\frac{1}{2}\pi r^2\right] \left[dv^3\right] \tag{1}$$

where C_p is the "power coefficient", or the ratio of the power flowing through the device that is captured, and d is the density of the air the turbine is exposed to. Betz (1926) demonstrated that the theoretical limit on C_p is $C_p^{\max} = \frac{16}{27} \approx 0.593$.

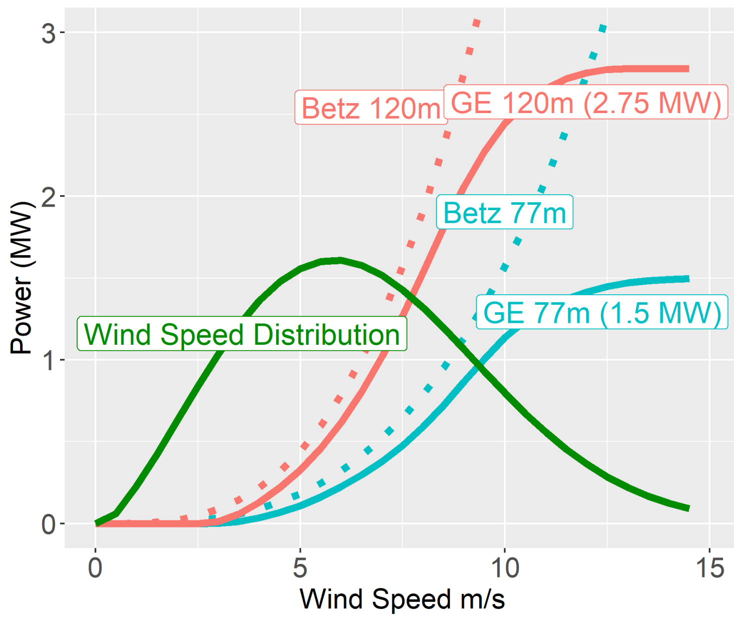
In practice, commercial turbines produce less than this optimum for two reasons. First, real world turbine blade designs never quite achieve the theoretical maximum at any speed (although some are remarkably close for a wide range of v's). Second, and more importantly, generators, which convert captured power into electricity, have a maximum capacity which typically chokes off the devices power at high wind speeds. This is partially to avoid extreme stress on the device, but also largely for economic reasons: if generator costs are increasing in size, and high wind speeds occur infrequently, then it doesn't make sense to pay for generating capacity that will be rarely used.

Figure 1 presents the Betz frontier and power curves for two different sized General Electric wind turbines. Power curves are functions, provided by the manufacturer to prospective buyers, which map wind speeds into the device's output. They are the empirical analogue to Betz' law. The figure also includes the probability density function of wind speed for a typical location. Wind speeds are well approximated by a Weibull distribution, with means between 6 and 8 meters per second (m/s). In this example, both turbines closely match the Betz frontier, until they achieve their "rated power", denoted in megawatts (MW).

One immediate implication of Betz law is that the wind turbine production function is characterized by increasing returns to size. This means that, all else equal, *better* wind turbines, are generally *bigger* wind turbines. This physical relationship has underpinned much of the rapid growth the industry has experienced in recent years. Since 2000, wind turbine rotors have more than doubled in size, while expected output per turbine has increased nearly fourfold (Figure 2).

Given that the underlying technology demonstrates increasing returns to scale (rotor size), it is natural to question why wind turbines weren't initially larger, and why they aren't even bigger now. It turns out that the constraints on rotor size are also physical in nature. Galileo's "square-cube" law, adapted to this setting, states that the volume of materials necessary to produce a rotor of radius r should be proportional to r^3. Thus, while large turbines produce more power than smaller turbines do, their costs of manufacturing are more than proportionately larger (for the same material and design). To profitably make bigger turbines, firms must either develop new turbine designs or they must employ newer

Figure 1: Example Power Curves



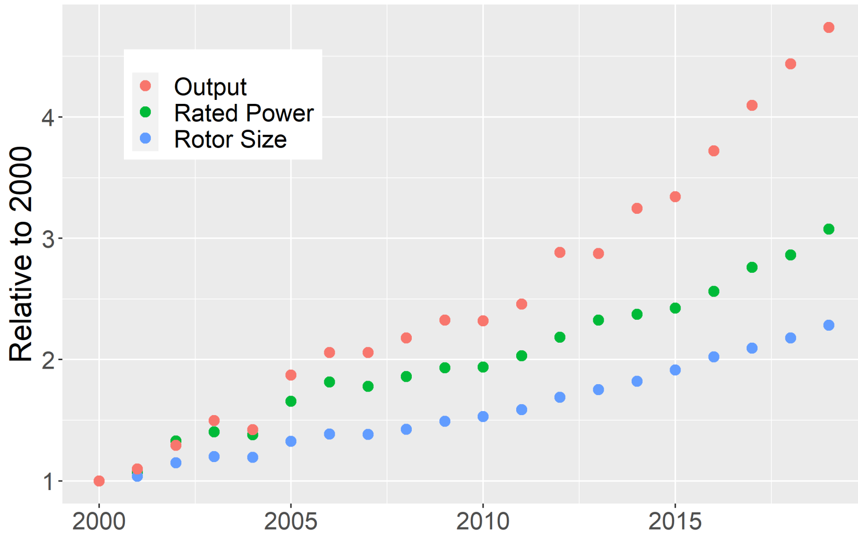
manufacturing materials in existing designs (or both). Newer designs or materials allow firms to increase rotor size r with a quadratic benefit in power generation at a less than cubic increase in costs. Innovation in the wind turbine industry, as reflected in the arrival of ever bigger turbines, is grounded in these engineering investments.

Beyond a simple physical observation, this cubic mass relationship has also been demonstrated in real world turbines. In the 2000s, the DOE conducted a series of studies on the limits to manufacturing large turbines. As part of this, they produced a software tool which estimates the mass of materials necessary to produce a turbine given user-entered characteristics. Figure 3 plots the log mass vs rotor size for all turbines in our sample, showing that estimated turbine mass indeed grows faster than quadratically in rotor size.

2.2 Industry Economics

The wind turbine market is highly concentrated. Table 1 presents sales for the ten largest firms. The four largest firms have over half of global turbine sales, and the top ten have nearly 80 percent. Although the industry is global, in that the top firms supply every market, sales are regionally concentrated. This is at least partially due to home market bias (Coşar et al., 2015). Figure 4 presents the distribution of sales by region, for the top four firms in each region, confirming that General Electric does most of its sales in the US market, Enercon, Siemens Gamesa, and Vestas do most of their sales the European market, and the

Figure 2: Global Turbine Size and Output Trends



Global average installed rotor size, rated power and *predicted* output (author's calculations), relative to their year 2000 values. Source: BNEF.

two major Chinese manufacturers, Goldwind and Guodian UP, sell primarily in Asia.

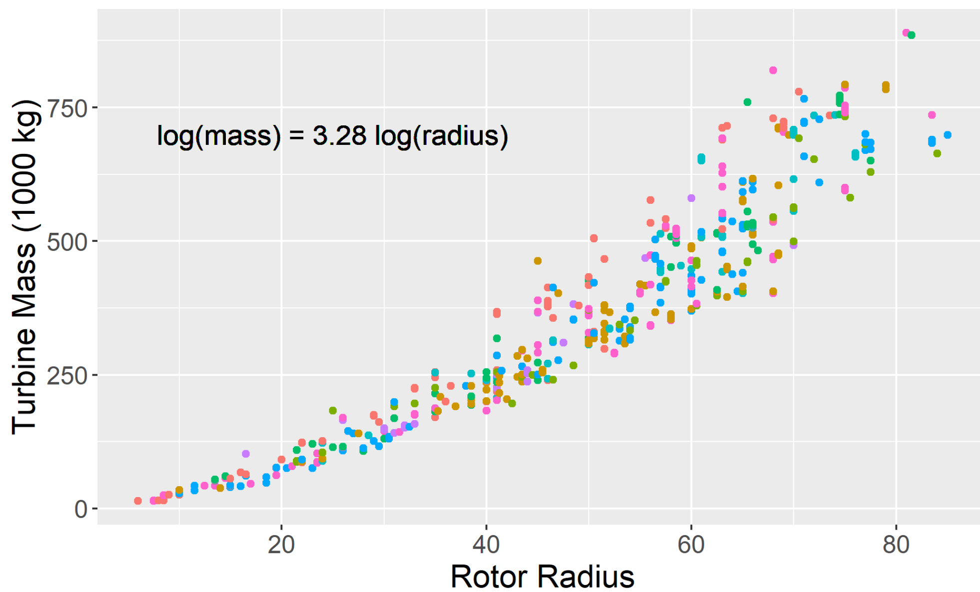
Table 1: Sales by Firm

Firm	Country	Models	Capacity	Projects	Turbines
Vestas	Germany	41	101045	4290	50246
General Electric	United States	28	68336	1434	37111
Gamesa	Spain	28	49022	1571	31260
Goldwind	China	23	47858	912	29418
Enercon	Germany	29	42983	3976	24098
Siemens	Germany	18	28913	661	13680
Nordex	Germany	21	22995	1217	9464
Guodian UP	China	11	18728	378	11256
Senvion	Germany	13	14702	1085	6919
Suzlon	India	13	14679	1060	8476

This table presents the total number of turbines sold, total capacity and number of wind farms for the ten largest firms (by turbine sales) in the BNEF data.

At any given point in time, manufacturers produce multiple turbines. As new turbine models are introduced, old models are discontinued. Figure 5 plots the annual mix of turbine sizes offered, scaled by sales in that year, for the six largest western manufacturers. For all firms, the mix of rotor sizes increases over time, with the largest rotors in 2018 being 50%-100% bigger than the largest rotors in 2000, when our data begins. Additionally, Figure

Figure 3: Turbine Mass vs Size (NREL)



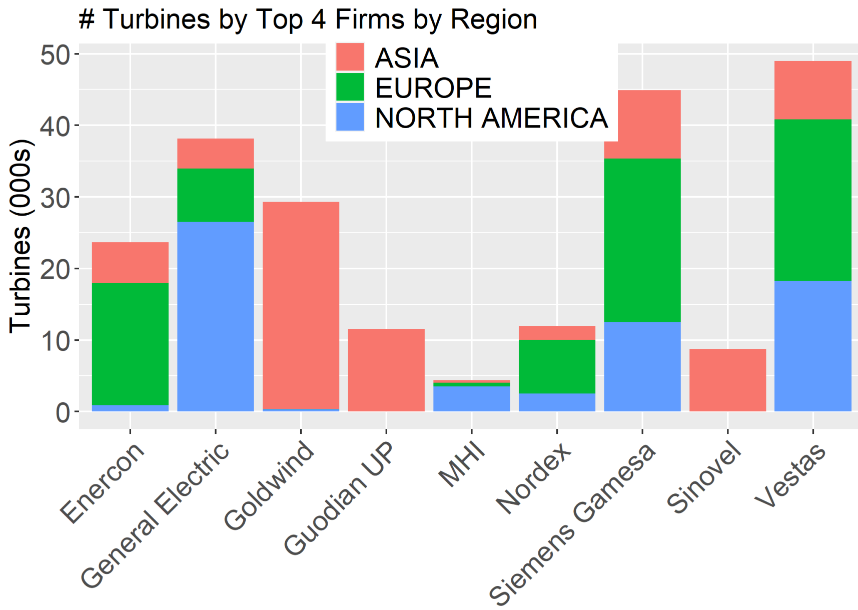
5 also shows that when new, larger turbines are introduced, they do not immediately take over all of the market share from existing smaller turbines, despite being more productive. Instead, these larger turbines gradually gain market share over time, which provides suggestive evidence that the relative cost difference must be declining over time, consistent with learning.

3 Data

The primary data come from a proprietary list of global wind farms maintained by Bloomberg New Energy Finance (BNEF). For almost all projects, this list includes the exact location of the wind farm (geocoded), its capacity in turbines and output, the date proposed and the date commissioned (if ever). Importantly, for most commissioned wind farms, BNEF also records the exact turbine model installed. We match these turbines to a detailed database of turbine characteristics from The Wind Power (TWP), an energy marketing consultancy. TWP's turbine database includes information about the rotor diameter and rated power of each turbine, along with dozens of other technical specifications. TWP also maintains a nearly comprehensive list of turbine power curves (as presented in Figure 1).

We supplement this database with project-specific information on wind speeds and power prices. We purchased information on the distribution of wind speeds from Vaisala, a commercial vendor widely used in the industry for siting purposes. For each project, we obtain

Figure 4: Sales by Region



the parameters of a Weibull distribution of wind speeds. Integrating a turbine's power curve over the distribution of wind speeds allows us to estimate the expected output from any given location-turbine combination, including those not selected in the data.

To estimate the value of this output at a given site, we employ data from several sources, of varying degrees of specificity depending on the region and time period. For all non-regulated US projects, EIA Form 860 provides average annual revenue from resale sales. Many countries outside the US support wind farms with feed-in-tariffs, which we obtained by year from the OECD and BNEF. In recent years, turbine contracts have become increasingly awarded via auction. BNEF maintains a database of all wind farm auctions, as well as the winning projects, their bids and the award price ($/MWh). Where none of these prices are available, we use the average wholesale price in a country-year, or country-state-year. Figure 6 plots average wind-specific power prices over time for selected large markets.

For analysis, we make a number of sample restrictions. Starting with the full BNEF database, which contains projects going back as far as the 1980s, and many not yet built, we exclude projects built prior to 2000 or not yet commissioned. We exclude small projects, of less than a megawatt. Of this sample, we exclude some projects where BNEF either does not have turbine information or we were unable to locate the turbine specified in the TWP database. Finally, we exclude offshore projects.

Figure 5: Sales by Turbine Size, Top Western Firms

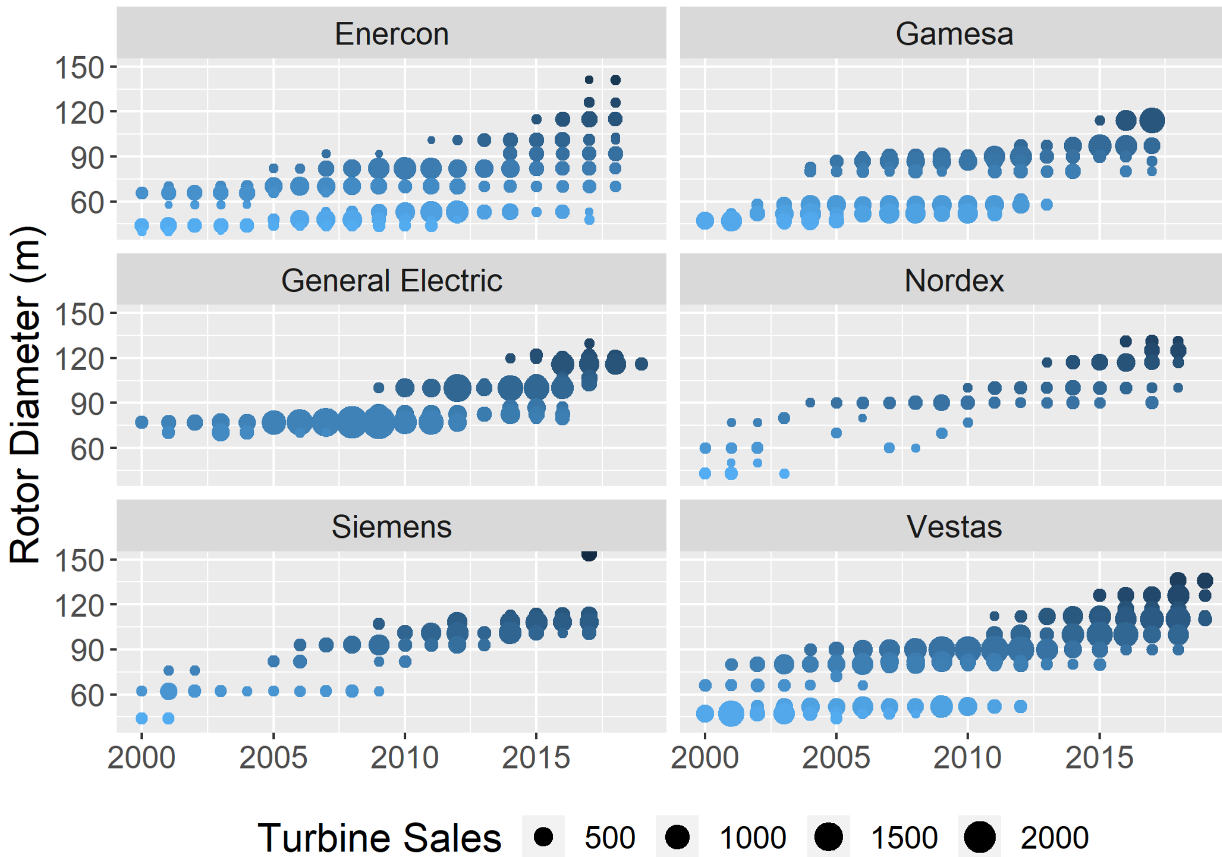


Table 2: Sample Construction

Group	N	mean
All Projects	22,186	1.00
Completed, Post 2000	21,316	0.96
Capacity 1 MW	20,514	0.92
Turbine match	20,025	0.90
Onshore	19,795	0.89

Source: BNEF Project Database

Table 3: Sample Summary

Variable	mean	sd	max	median	min
Capacity (MW)	27.57	40.56	644.40	12.50	1.00
# Turbines	15.47	23.19	460.00	7.00	1.00
Rated Power (MW)	2.01	0.86	12.00	2.00	0.06
Wind speed (m/s)	6.61	1.00	15.86	6.52	2.36
Rotor Radius (M)	43.87	12.33	110.00	43.50	7.50
Est. Turbine Output (kW/hour)	700.19	423.18	4,419.43	637.12	5.80

Figure 6: Average Wind Output Price by Country

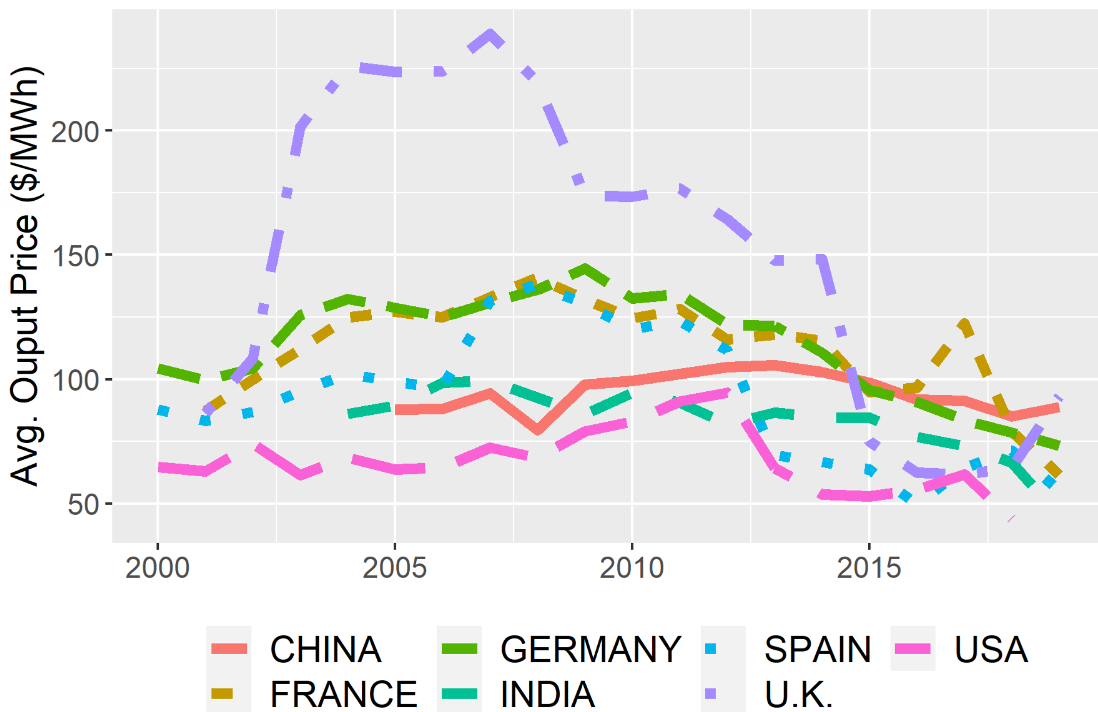


Table 4: Demand Estimation Observations by Country

Year	AU-DK-PT	CHINA	DEU	FRA	ITA	SPAIN	SWE	U.K.	U.S.A.	Excluded	Share
2000	25	3	188	9	11	43	14	7	8	62	0.83
2001	12	0	311	3	19	43	11	7	37	77	0.85
2002	21	0	341	7	10	42	15	9	13	86	0.84
2003	37	6	282	9	31	59	30	7	39	116	0.81
2004	41	2	238	16	24	93	22	5	19	117	0.80
2005	71	16	202	40	17	62	15	16	29	162	0.74
2006	37	25	222	73	26	89	26	22	39	243	0.70
2007	31	48	189	78	34	96	43	27	58	226	0.73
2008	34	94	131	102	37	96	36	33	97	266	0.71
2009	39	190	214	97	56	86	47	35	107	289	0.75
2010	42	223	164	100	43	47	64	30	61	427	0.64
2011	20	311	185	79	31	35	48	33	87	544	0.60
2012	25	224	122	51	26	29	32	32	144	400	0.63
2013	39	206	138	49	14	11	19	67	12	214	0.72
2014	34	230	318	80	6	0	25	45	50	407	0.66
2015	35	343	344	65	11	2	18	31	61	452	0.67
2016	44	218	412	95	16	3	17	40	72	425	0.68
2017	25	170	511	123	11	2	8	58	62	368	0.72
2018	31	139	226	88	22	22	14	21	54	397	0.61
2019	9	117	107	95	18	71	15	14	40	298	0.62
2020	7	280	150	74	7	56	13	6	0	358	0.62

4 Recovering Turbine Bids

In the typical dataset used for estimating marginal costs of producing consumer products, researchers have access to both aggregate quantities and posted (assumed common) prices. For large capital goods like wind turbines, there are no posted prices. Turbines are procured in a confidential, developer-specific process. Once completed, the terms agreed upon are rarely disclosed. Thus, before estimating manufacturing costs using the inversion of a demand system, we must recover the prices developers face when chosing which turbine to install. To do this, we develop a model of turbine procurement auctions, and show how the parameters of this model represent the bids manufacturers submit.

4.1 A model of turbine procurement

We begin by describing a procurement model with purely static incentives (i.e., no learning-by-doing incentives, yet). At time t, each firm f in the set of manufacturers F has a portfolio of turbines $j \in K_{ft}$. During t, each project i in a set of wind farm projects W_t requests bids for each turbine offered by each firm. If project i selects turbine j, i earns discounted expected revenues $R_{ijt} + \epsilon_{ijt}$, where ϵ is an iid shock to developer profits, known only by the developer.

The revenues that turbine j generates at site i are determined by the known wind speed distribution $F_i(v)$ for site i, the known power curve $P_j(v)$ which maps wind speeds v into megawatts of output for turbine j, and p_i^{output}, the output price site i receives for its production. We assume that each turbine will generate revenues according to these terms for 20 years, so that the discounted expected revenues R_{ijt} are:

$$R_{ij} = \underbrace{8760}_{\text{hours in a year}} \times \underbrace{p_i^{\text{output}}}_{\text{price per megawatt hour}} \times \underbrace{\int P_j(v) dF_i(v) dv}_{\text{expected megawatt hours generated}} \times \underbrace{\sum_{t=1}^{20} \delta^{t-1}}_{\text{discounted value for 20 years}}$$

Project i will buy n_i turbines, regardless of which turbine j it selects.[2] However, because R_{ijt} varies with j, i's revenues depend on what turbine model it selects.

We assume that i conducts a second price procurement auction to determine which turbine it buys, and what price it pays. In this mechanism, each turbine manufacturer submits a (potentially) project-specific bid b_{ijt} for each turbine j in its portfolio. Project i selects the turbine that delivers the highest net surplus, which we define as $R_{ijt} - b_{ijt} + \epsilon_{ijt}$. If turbine j from firm f wins the auction, the payment $\widetilde{b}_{ijkt}$ from i to f is designed to make i indifferent between choosing j and paying this price, and choosing the next highest net surplus turbine, k, and paying the actual bid offered by k's manufacturer. That is, $\widetilde{b}_{ijkt}$ satisfies

$$R_{ijt} - \widetilde{b}_{ijkt} + \epsilon_{ijt} = R_{ikt} - b_{ikt} + \epsilon_{ikt}$$

so that the payment is

$$\widetilde{b}_{ijkt} = \underbrace{b_{ikt}}_{\text{bid of second best turbine proposal}} + \underbrace{(R_{ijt} + \epsilon_{ijt})}_{\text{revenues from best turbine proposal}} - \underbrace{(R_{ikt} + \epsilon_{ikt})}_{\text{revenues from second best turbine proposal}}$$

Thus, though the probability that a firm wins a sale does depend on its bid, its payment, conditional on winning, does not, as a result of these second-price rules.

Asker and Cantillon (2008) show that the unique dominant strategy equilibrium for this game is for each firm to bid exactly its opportunity cost of delivering each turbine, and,

[2] This means we are ruling out situations where a developer is choosing between three 1 MW turbines or two 1.5 MW turbines, and plans to keep output fixed.

as a result, site developers pick the turbine which maximizes social surplus, the sum of developer and manufacturer profits. This implies that, conditional on the opportunity cost of supply, bids will not be project-specific in equilibrium. Based on this result, we assume that manufacturing costs are constant within a given year, so that if two projects i and i' are built in the same year, $b_{ijt} = b_{i'jt}$.[3] In a slight abuse of notation, we define the (common) bid for turbine j in year t as b_{jt}.

For tractability, we make three additional assumptions. First, we assume that the project-by-turbine revenue shocks are distributed as type-1 extreme value. Second, we assume that every project which solicits bids from manufacturers chooses a turbine (e.g., there is no outside option). Finally, we assume that developer's consider bids on all turbines offered by all active manufacturers at a given point in time (e.g., that choice sets are common). These additional assumptions imply that the probability firm f wins the site i auction with turbine j at time t is:

$$s_{ijt} = \frac{\exp(R_{ijt} - b_{jt})}{\sum_{f' \in F} \sum_{l \in J_{f'}} \exp(R_{ilt} - b_{lt})}$$

With this notation, we can define the static expected profits for firm f at time t. Let c_{jt} be the *marginal*, not opportunity, cost of delivering turbine j at time t.[4] Given our previous definition for f's revenues when it sells turbine j, its static profits from a successful sale to site i are:

$$\pi_{ijt} = R_{ijt} - b_{jt} + \epsilon_{ijt} - \max_{k \in \prod_{\phi \neq f} K_{\phi,t}} (R_{ikt} - b_{kt} + \epsilon_{ikt}) + (b_{jt} - c_{jt})$$

Because developers always choose the turbine which generates the highest net surplus, we can write f's total realized static profits from offering its turbines to site i as:

$$\pi_{ift} = \max_{j \in \prod_{\phi} K_{\phi,t}} (R_{ijt} - b_{jt} + \epsilon_{ijt}) - \max_{k \in \prod_{\phi \neq f} K_{\phi,t}} (R_{ikt} - b_{kt} + \epsilon_{ikt}) + \sum_{j \in K_{ft}} \mathbb{I}\left[i \text{ chooses } j\right] (b_{jt} - c_{jt})$$

Note that this expression includes the possibility that firm f does not make a sale: when i chooses a turbine *outside* of K_{ft}, the first expression above is identical to the second, so the entire expression is zero. Thanks to the type-1 extreme value assumption on the ϵ's, the

[3] When we introduce dynamic pricing incentives in Section 5, we'll discuss what additional assumptions we require in order to maintain this assumption that bids do not vary across projects.

[4] When we later introduce dynamic pricing incentives, marginal and opportunity costs will differ.

firm's expected profits at site i are:

$$\mathbb{E}\pi_{ift} = \underbrace{\log \sum_{j \in \prod_\phi K_{\phi,t}} \exp\left(R_{ijt} - b_{jt}\right)}_{S_{it}} - \underbrace{\log \sum_{k \in \prod_{\phi \neq f} K_{\phi,t}} \exp\left(R_{ikt} - b_{kt}\right)}_{S_{it}^{-f}} + \sum_{j \in K_{ft}} s_{ijt}(b_{jt} - c_{jt})$$

In words, this is the logit inclusive value of the wind developer's net surplus among all turbines (S_{it}), minus the logit inclusive value of the wind developer's net surplus among all turbines excluding those from firm f (which we call S_{it}^{-f}), plus firm f's expected markup over its marginal costs.

4.1.1 Empirical implementation

We use this structure to recover the unobserved procurement auction bids by modeling the developer's discrete choice problem, allowing for a slightly richer notion of developer preferences. Project developer i in period t chooses a utility-maximizing turbine j given overall revenues R_{ijt}, revenues that come specifically from high wind speeds R_{ijt}^H, bids b_{jt}, and an additional set of characteristics X_{ijt} which capture other costs and benefits of a given turbine-site combination:

$$u_{ijt} = \alpha_0 R_{ijt} + \alpha_H R_{ijt}^H + \sum_c \alpha_c \mathbb{I}[c(i) = c] R_{ijt} + X_{ijt}\beta^D - b_{jt} + \epsilon_{ij} \tag{2}$$

The variable R_{ijt} represents all wind revenues, while the variable R_{ijt}^H represents the subset of those revenues which accrue at high wind speeds.[5] We allow developers in different countries c to have different preferences over the revenues generated, to account for differences in discount rates, revenue volatility or uncertainty, and curtailment policies. The parameter α_0 represents the (common) marginal utility project developers get from turbines which produce more revenue, and α_H represents the differential marginal value of output at high wind speeds. The vector of parameters $\{\alpha_c\}$ represents the country-specific marginal utility for revenues.

[5] We define "high" wind speeds as those greater than 7.5 m/s. We allow developers to have distinct preferences over wind production revenues that accrue at high wind speeds for several reasons. First, episodes of high wind speeds often coincide with electricity grid transmission constraints, meaning that the high production associated with high wind speeds might either be priced at (lower) congested prices, or may require curtailment Aldy et al. (2019). Second, high wind speeds often occur in short bursts of time, and because wind turbines have lots of physical inertia, they do not immediately spin faster when the wind is blowing faster. This means that power production which accrues from high wind speeds may be lower than what we calculate using a wind speed distribution and a power curve. Third, our estimates of the distribution of wind speeds are necessarily less precise at high wind speeds, since they are observed less frequently than moderate speeds.

The vector X_{ijt} captures two sources of preference heterogeneity across projects. First, although we assume that all turbines are available to all projects, in practice there are physical constraints which limit the suitability of larger turbines at especially windy sites. Because there are no documented hard and fast rules which indicate which turbines are allowed at which sites, we instead include dummy variables that indicate whether a site's measured wind speed distribution lies outside of the turbine's recommended range, and interact these variables with the turbine's rated power capacity to account for scale differences in revenues across different capacity turbines. We also control for well-documented home-market bias in the wind turbine industry (Coşar et al., 2015). To do this, we include dummy variables which are equal to 1 if the project's country is the same as the country where the turbine manufacturer's headquarters lie, and also interact these variables with rated power capacity. Finally, we included a dummy for whether a firm has a manufacturing facility in the country.

Our assumption that manufacturer turbine bids are constant within each year means that we can estimate those bids as turbine-year fixed effects. Because our demand system does not have an outside option, what our estimates recover is actually the difference in bids between turbine j and the bid for a reference turbine, $\hat{b}_{jt} = b_{jt} - b_{0t}$. We define the reference turbine as the best selling turbine manufactured by Vestas in each year. With data on site-by-turbine expected revenues, other turbine characteristics, and each site's turbine choice, we estimate the b's, the α's and the vector β^D using maximum likelihood.

4.2 Bid estimation results

Table 5 presents the estimated preferences from four separate demand specifications. We select Germany as the base country, since it has the most projects and very rich output price variation. The first row indicates that German developers value a discounted expected dollar of revenue at 92 cents, and we cannot reject one. The second row indicates that revenues during high wind hours are much less valuable than low wind, consistent with congestion and curtailment being a factor. The next section of coefficients indicates significant heterogeneity in estimated revenue sensitivity across countries. This suggests heterogeneity in capital costs (via discount rates), policy and price uncertainty, as well as likely measurement error in our estimated output price measures and currency conversions.

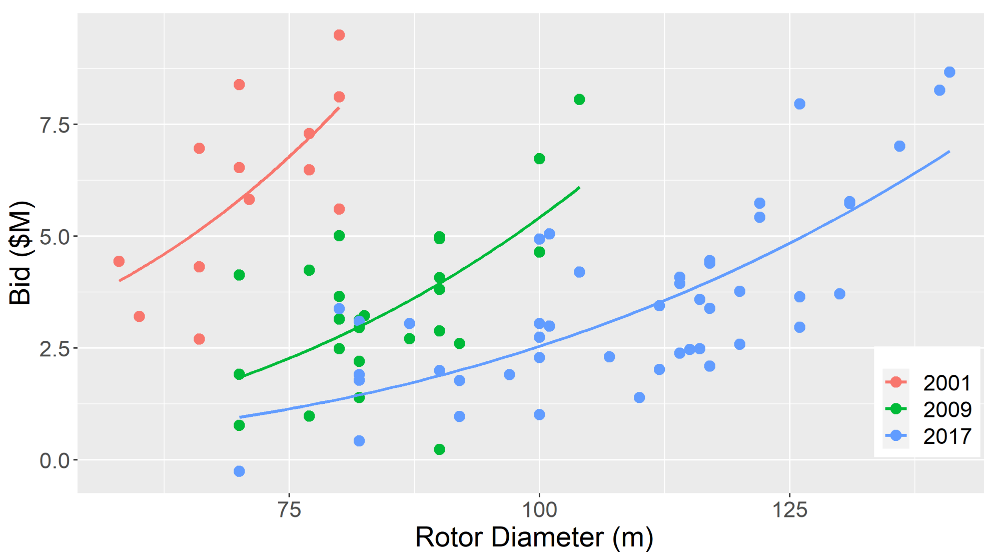
Table 5: Wind Plant Turbine Preferences

	Model 1		Model 2		Model 3		Model 4	
Parameter	Point Est.	Std. Err.	Point Est.	Std. Err.	Point Est.	Std. Err.	Point Est.	Std. Err.
Revenue	0.929	0.179	1.417	0.349	1.290	0.182	1.906	0.379
Revenue x High	-0.583	0.179	-0.835	0.311	-0.770	0.181	-1.107	0.319
Revenue x Country								
AUSTRIA	0.263	0.057	0.449	0.106	0.175	0.057	0.331	0.097
CHINA	-0.122	0.058	-0.179	0.099	-0.367	0.063	-0.513	0.118
DENMARK	0.446	0.071	0.699	0.146	0.425	0.071	0.661	0.138
FRANCE	-0.300	0.029	-0.498	0.085	-0.393	0.030	-0.605	0.093
ITALY	-0.544	0.037	-0.900	0.136	-0.712	0.040	-1.108	0.156
PORTUGAL	0.165	0.062	0.276	0.103	0.097	0.062	0.173	0.098
SPAIN	-0.374	0.042	-0.621	0.104	-0.543	0.044	-0.860	0.128
SWEDEN	-0.125	0.057	-0.262	0.097	-0.103	0.056	-0.200	0.092
UNITED KINGDOM	-0.450	0.035	-0.738	0.115	-0.587	0.037	-0.900	0.130
UNITED STATES	-0.289	0.047	-0.529	0.104	-0.514	0.050	-0.807	0.130
Site/Turbine Class Compatibility								
(Site: 1, IEC: II) x MW	-0.362	0.038	-0.583	0.101	-0.363	0.038	-0.566	0.096
(Site: 1, IEC: III) x MW	-0.772	0.066	-1.293	0.208	-0.782	0.066	-1.259	0.197
(Site: 2, IEC: III) x MW	-0.267	0.032	-0.475	0.085	-0.260	0.032	-0.448	0.080
Home Bias/Manufacturing								
ChinaForeign x MW	-2.378	0.115	-4.052	0.596	-2.425	0.120	-3.957	0.566
Factory x MW	0.334	0.019	0.546	0.080	0.231	0.033	0.360	0.069
Factory x MW x log(N)					0.046	0.014	0.080	0.024
Home x MW	0.655	0.025	0.966	0.132	0.644	0.025	0.952	0.127
Rotor x log(N)					0.009	0.001	0.012	0.002
Heteroskedasticity Over Time								
2001			0.242	0.174			0.131	0.169
2002			0.093	0.171			-0.010	0.168
2003			0.015	0.161			0.047	0.156
2004			-0.095	0.161			-0.156	0.160
2005			-0.456	0.166			-0.414	0.161
2006			-0.494	0.163			-0.454	0.159
2007			-0.468	0.160			-0.448	0.155
2008			-0.644	0.161			-0.620	0.156
2009			-0.434	0.150			-0.426	0.145
2010			-0.394	0.155			-0.344	0.150
2011			-0.406	0.154			-0.366	0.149
2012			-0.571	0.155			-0.576	0.150
2013			-0.523	0.157			-0.467	0.153
2014			-0.457	0.151			-0.413	0.147
2015			-0.481	0.150			-0.467	0.145
2016			-0.600	0.149			-0.555	0.145
2017			-0.539	0.147			-0.492	0.143
2018			-0.614	0.157			-0.579	0.152
2019			-0.652	0.159			-0.648	0.156

The next panel in the table documents the importance of turbine-site compatibility. High wind sites (class 1) appear to strongly dislike turbines designed for low wind sites (classes II and III), and class 2 sites dislike class III turbines. Next we also find considerable heterogeneity in firm preferences across countries. Turbines are large, heavy pieces of equipment, and having a production facility inside the country makes developers significantly more likely to select a firm's turbine. Even conditional on that though, we find that developers have a large willingness to pay for home market turbines, consistent with (Coşar et al., 2015). Finally, in looking at the raw data, it is clear that Chinese developer preferences along these dimensions appear different from the rest of the sample. We rationalize this with the inclusion of a foreign producer dummy for the country, and estimate this effect to be twice as large as the other home bias effects combined.

Figure 7 plots the estimated turbine bids against turbine size for three years in our sample. Looking across these three years, two patterns stand out. First, many turbine sizes are supplied over a long horizon. There are plenty of turbines in the 80m range that are sold in all three years (spanning nearly two decades), and there are also many turbines in the 100m range that are sold in both of the later years we plot. Moreover, across all sizes, the price of a turbine at a given size comes down considerably. Second, although the price of a turbine increases in turbine size, consistent with the discussion in section 2, the price gradient with respect to size flattens out dramatically over time. This change in cost structure is what underlies the remarkable shift towards bigger, therefore more productive, turbines over the past twenty years.

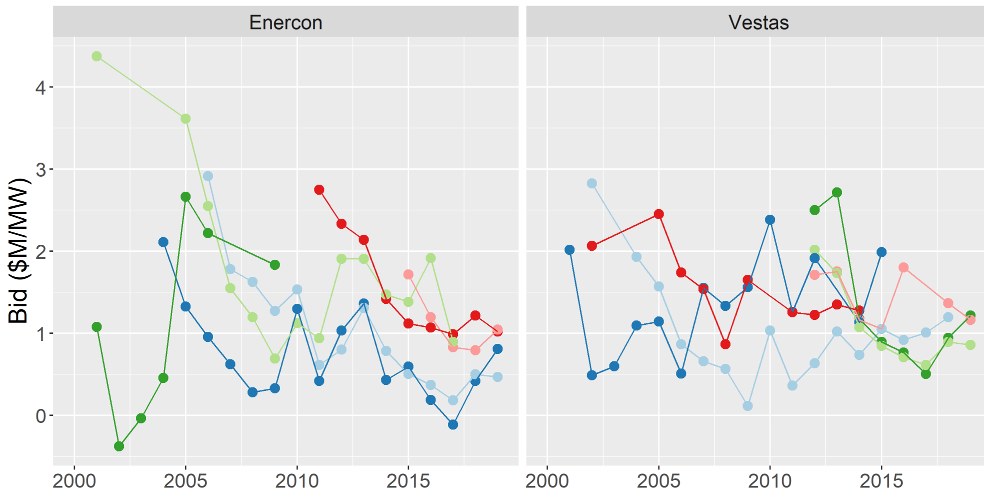
Figure 7: Estimated bids vs rotor size over time



Estimated bids are recovered relative to a reference turbine, which we pick to be the most common Vestas turbine sold that year. To plot levels, we impute the implied cost of this turbine from Vestas financials.

An alternative way to vizualize the estimated bids is to look at the evolution of prices over time within the same device. Figure 8 plots estimated bids for the most popular turbines sold by the two largest European manufacturers, Enercon and Vestas. Although the estimates are inherently a bit noisy, it appears that bids come down on average over time. However, it is important to remember that the direction of prices over time in an oligopolistic market with learning is theoretically ambiguous (Besanko et al., 2010, 2014). On the other hand, some of the *increases* in prices that occur leading up to 2010 are potentially driven by raw materials prices, including steel, which increased nearly 50% between 2005 and 2009.

Figure 8: Estimated bids for popular turbines at Enercon and Vestas



Estimated bids are recovered relative to a reference turbine, which we pick to be the most common Vestas turbine sold that year. To plot levels, we impute the implied cost of this turbine from Vestas financials.

5 Isolating costs from bids

Although it is optimal for firms to bid their true opportunity costs in the turbine procurement mechanism we assume, the dynamic incentives implied by learning by doing mean that manufacturing costs and opportunity costs are not identical in this setting. As noted by Irwin and Klenow (1994) and Benkard (2004), the fact that sales today may generate cost reductions tomorrow means that rational and forward-looking manufacturers have an incentive to *underprice* recently introduced turbines. Thus, opportunity costs are the sum of true manufacturing costs, which may reflect past production experience, and anticipated learning benefits, or future cost reductions from marginal sales today ("dynamic markdowns"). To learn about the effects of past experience on manufacturing costs, we need to separately control for these dynamic markdowns.

Previous theoretical and applied research on dynamic markdowns explicitly computes them using dynamic programming techniques (Besanko et al., 2014; Benkard, 2004). While this approach is feasible for markets with mature production technologies and a small number of firms and products, it is not helpful in our setting. The wind turbine manufacturing industry has always had several active firms, and in any given year, 30 or more distinct wind turbine models are available for sale. It is not currently computationally possible to compute the equilibria of a dynamic game with this many firms and/or products, especially in a non-stationary setting like this.

Instead of attempting to solve such a complicated dynamic game in order to *exactly* characterize the dynamic markdown, we use a rational expectations method developed in Berry and Pakes (2000). The primary advantage of this approach is that rational expectations techniques provide a way to construct a noisy measure of the dynamic markdown without explicitly solving for the equilibrium of the underlying game firms may be playing.

In this section, we derive the static component of opportunity costs, using the properties of our procurement auction assumptions (Section 5.1). Next, we derive the dynamic markdown term and show how to use the Berry-Pakes method to construct a noisy measure of it (Section 5.2). We then derive the state transition process for this industry, a necessary ingredient to the Berry-Pakes method, in Section 5.3. In Section 5.4 we describe two ways of implementing the rational expectations cash flows needed for the Berry-Pakes method. Finally, we derive our estimating equations (Section 5.5) and discuss our identification strategy (Section 5.6).

5.1 Optimal bidding behavior and static opportunity costs

To relate our observed bids to costs, we will assume that firms choose bids that are dynamically first-order optimal, in a sense we will formalize below. The dynamic component of this problem arises from the fact that past manufacturing experience is a key determinant of today's manufacturing costs, so sales today, which increase manufacturing experience, affect costs tomorrow. To allow for this, we define a state variable x as the cumulative vector of sales experience for <u>all</u> turbines: x_{jt} is the count of all sales of turbine j before period t. This state variable is common to all firms in the industry, which will allow us to account for spillovers in experience across firms, to the extent that they are empirically relevant.[6]

During period t, the industry at state x_t supplies turbines to a set of wind farm projects $i \in W_t$, using the procurement auction mechanism described in the previous section. The aggregate sales from this process are a vector q_t, which, for each $j \in \cup_f K_{ft}$, is defined by:

$$q_{jt} = \sum_{i \in W_t} n_i \mathbb{I}[i \text{ chooses } j \text{ in } t]$$

These sales in turn update the cumulative sales vector for the next period, $x_{t+1} = x_t + q_t$. The probability distribution over q_t, which depends on the set of available projects W_t and the bids firms set b_t, is $dF(q_t \mid W_t, b_t)$, which we derive in Section 5.3.

We assume that firms choose turbine-specific bids at the start of the year, before any sales

[6] Section 6.1 describes the specific functional relationship assumed between manufacturing costs and x.

occur, and submit these fixed bids to each individual procurement auction. In a world with no dynamic incentives, this assumption is innocuous, due the Asker and Cantillon (2008) efficiency result: bids should always equal costs, and so they are common across projects by definition. However, if dynamic pricing incentives are indeed important, then as a firm accumulates experience over the course of a time period, its costs change. Moreover, the learning benefits from selling heterogeneous projects need not be homogenous. For example, bigger projects induce larger changes in experience than smaller projects do. Thus, our assumption that firms set bids once at the start of a period, and submit the same vector of bids to every project's procurement auction is not without loss of generality.

However, we impose this assumption out of necessity. Because we do not directly observe bids, we must infer them from the demand system, in the form of turbine-by-time fixed effects, and there are limits to how aggressibely we can divide the data up in order to generate finer turbine-time fixed effects. Accordingly, in our derivation of optimal bidding that follows, we will assume that firms receive all bidding-relevant information before they set bids. In this sense, our assumption that bids are constant over a time period is similar to a timing assumption in the productivity literature.

Firms choose bids in order to maximize the sum of expected profits from selling to available projects that period (the set $i \in W_t$), and the expected value of a firm-time specific continuation value function of x_{t+1}, where expectations are taken with respect to the distribution $dF(q_t \mid W_t, b_t)$. Firm f's objective is:

$$\max_{b_{ft}} \sum_{i \in W_t} n_i \mathbb{E}\pi_{ift}(b_t, x_t) + \int V_{ft}(x_t + q_t) dF(q_t \mid W_t, b_t)$$

Here, we are making explicit the idea that expected profits depend on the entire vector of bids (b_t). We also allow the expected profit function to depend on x because a firm's present manufacturing costs may depend on its past manufacturing experiences, as well as other firm's manufacturing experiences if spillovers are present.

Optimal bidding for turbine l will satisfy a first-order condition:

$$0 = \underbrace{\sum_{i \in W_t} n_i \frac{\partial}{\partial b_{lt}} \mathbb{E}\pi_{ift}(b_t, x_t)}_{\nabla\pi = \text{marginal static profits}} + \underbrace{\frac{\partial}{\partial b_{lt}} \int V_{ft}(x_t + q_t) dF(q_t \mid W_t, b_t)}_{\nabla V = \text{marginal dynamic benefits}}$$

The first component of this expression shows how marginal pricing changes affect the firm's marginal static profits. In particular, a marginal change in the bid for turbine l induces a change in static profits $(\nabla\pi)_{lt}$:

$$
\begin{aligned}
(\nabla\pi)_{lt} &= \sum_{i\in W_t} n_i \frac{\partial}{\partial b_{lt}} \mathbb{E}\pi_{ift}(b_t, x_t) \\
&= \sum_{i\in W_t} n_i \frac{\partial}{\partial b_{lt}} \left(\log \sum_{j\in\prod_\phi K_{\phi,t}} \exp\left(R_{ijt} - b_{jt}\right) - \log \sum_{k\in\prod_{\phi\neq f} K_{\phi,t}} \exp\left(R_{ikt} - b_{kt}\right) + \sum_{j\in K_{ft}} s_{ijt}(b_{jt} - c_{jt}) \right) \\
&= \sum_{i\in W_t} n_i \left(-s_{ilt} - 0 + \sum_{j\in K_{ft}} \frac{\partial s_{ijt}}{\partial b_{lt}} (b_{jt} - c_{jt}) + s_{ilt} \right) \\
&= \sum_{i\in W_t} n_i \sum_{j\in K_{ft}} \frac{\partial s_{ijt}}{\partial b_{lt}} (b_{jt} - c_{jt})
\end{aligned}
$$

The first equality comes from our assumption that firms set the same turbine bids for all projects in the same time period. The second equality is a result of our procurement acution assumptions: expected (per-turbine) profits on plant i are a "net surplus" term, insensitive to which turbine is actually chosen, plus a share-weighted average bid-cost markup term. The third equality comes from differentiating the expected profit expression with respect to the bid on turbine l. Thanks to the logit errors, the derivative of the net surplus term is simply the negative of the probability that plant i picks turbine l. Similarly, the derivative of the bid-cost markup term is the "standard" two terms common in all discrete choice demand models: the elasticity-weighted sum of changes in margins, plus the gains from infra-marginal buyers. The final expression above includes the first of these standard terms, but not the second, because second price payment rules mean that the firm does not capture infra-marginal benefits from higher prices. As a result, marginal static expected profits are equal to the elasticity weighted average bid-cost markup, summed over all plants in the market.

5.2 Dynamic opportunity costs

To characterize the marginal effect of higher prices on future benefits, which we call the dynamic markdown, first note that since we have assumed our continuation values are functions of the present state x_t and the realized aggregate sales vector q_t, turbine price changes only affect the probability distribution of q_t, not the function $V_{ft}(\cdot)$ itself. Thus, our first simplification is:

$$
\frac{\partial}{\partial b_{lt}} \int V_{ft}(x_t + q_t) dF(q_t \mid W_t, \mathbf{b}_t) = \int V_{ft}(x_t + q_t) \frac{\partial}{\partial b_{lt}} dF(q_t \mid W_t, \mathbf{b}_t)
$$

Next, note that for any realizable value of q_t, $dF(q_t \mid W_t, b_t) > 0$, so we can re-arrange terms to get:

$$\begin{aligned}
\int V_{ft}(x_t + q_t) \frac{\partial}{\partial b_{lt}} dF(q_t \mid W_t, b_t) &= \int V_{ft}(x_t + q_t) \frac{\frac{\partial}{\partial b_{lt}} dF(q_t \mid W_t, b_t)}{dF(q_t \mid W_t, b_t)} dF(q_t \mid W_t, b_t) \\
&= \mathbb{E}\left[V_{ft}(x_t + q_t) \frac{\frac{\partial}{\partial b_{lt}} dF(q_t \mid W_t, b_t)}{dF(q_t \mid W_t, b_t)} \mid W_t, b_t \right]
\end{aligned}$$

This re-arrangement makes clear that the dynamic markdown, which is the gradient of expected future benefits with respect to the price of turbine l, can be expressed as the expectation of the *product* of future benefits at a given realization of q and the relative change in the probability that this value of q is realized resulting from the price change. This is the first insight highlighted in Berry and Pakes (2000).

We also employ the second key idea from Berry and Pakes (2000), and assume that firms have rational expectations about this object, conditional on equilibrium bids. We'll assume that:

$$\mathbb{E}\left[V_{ft}(x_t + q_t) \frac{\frac{\partial}{\partial b_{lt}} dF(q_t \mid W_t, b_t)}{dF(q_t \mid W_t, b_t)} \mid W_t, b_t \right] = \underbrace{V_{ft}(x_t + q_t^*)}_{\text{Realized Discounted Cashflows}} \times \underbrace{\frac{\frac{\partial}{\partial b_{lt}} dF(q_t^* \mid W_t, b_t)}{dF(q_t^* \mid W_t, b_t)}}_{\text{Berry-Pakes factor}} + \nu_{lt}$$

where q_t^* is the realized turbine sales vector, $V_{ft}(x_t + q_t^*)$ is the realized discounted cashflows for firm f following time t, and $\mathbb{E}[\nu_{lt} \mid x_t, b_t] = 0$ is a rational expectations error. This assumption is useful because it allows us to decompose the otherwise-infeasible dynamic markdown term into the sum of a product of two feasible expressions, and a rational expectations error. The term $\frac{\frac{\partial}{\partial b_{lt}} dF(q_t^* \mid W_t, b_t)}{dF(q_t^* \mid W_t, b_t)}$ is feasible because it can be computed from knowledge of the realized sales vector q_t^*, the set of projects W_t, and the demand system parameters. We explore two feasible approaches to handling $V_{ft}(x_t + q_t^*)$ in section 5.4.

5.3 State transitions and their sensitivity to bids

To finish our derivation of the firm's pricing FOC, we define $dF(q_t \mid W_t, b_t)$ and compute its gradient with respect to the price of turbine l. There are many allocations of turbines to projects that can generate a given value of q, so the probability that q is realized is the sum of the probabilities of many "feasible" allocations. For example, if there are three projects, A with a demand for 3 turbines, B with a demand for 2 turbines, and C with a demand for 1 turbine, and 2 turbine models, 1 and 2, there are two ways to attain $q_1 = 3$ and $q_2 = 3$. First, turbine 1 could sell to project A, and turbine 2 could sell to projects B and C. The reverse could also happen: 1 sells to B and C while 2 sells to A. This idea generalizes to more than two projects and/or turbines. Recognizing this, let $M(q, W)$ represent the set of allocations of turbines to the sites in W which generate an aggregate sales vector q. An allocation $m \in M(q, W)$ is a vector of turbine choices, so that $m_i = l$ for the chosen turbine l. Let p_m be the unconditional probability of allocation m:

$$p_m = \prod_{i \in W} s_{i,m_i}$$

Then the probability of outcome q with plants W and bids b, denoted by $dF(q \mid W, b)$, is $\sum_{m \in M(q,W)} p_m$. Because $dF(q \mid W, b)$ depends on individual choice probabilities, which, in turn, depend on the vector of turbine bids, the gradient of $dF(q \mid W, b)$ with respect to a given bid is nonzero:

$$\begin{aligned}
\frac{\partial}{\partial b_l} dF(q \mid W, b) &= \frac{\partial}{\partial b_l} \sum_{m \in M(q,W)} \prod_{i \in W} s_{i,m_i} \\
&= \sum_{m \in M(q,W)} \frac{\partial}{\partial b_l} \prod_{i \in W} s_{i,m_i} \\
&= \sum_{m \in M(q,W)} \left(\prod_{i \in W} s_{i,m_i} \right) \sum_{i \in W} \frac{\partial}{\partial b_l} \log s_{i,m_i} \\
&= \sum_{m \in M(q,W)} \left(\prod_{i \in W} s_{i,m_i} \right) \sum_{i \in W} (s_{il} - \mathbb{I}[m_i = l]) \\
&= dF(q \mid W, b) \sum_{i \in W} s_{il} - \sum_{m \in M(q,W)} p_m \sum_{i \in W} \mathbb{I}[m_i = l] \\
&= dF(q \mid W, b) \left(\sum_{i \in W} s_{il} - \frac{\sum_{m \in M(q,W)} p_m \sum_{i \in W} \mathbb{I}[m_i = l]}{dF(q \mid W, b)} \right) \\
&= dF(q \mid W, b) \left(\sum_{i \in W} s_{il} - \mathbb{E}\left[\# \text{ of projects pick } l \mid q, W, b\right] \right)
\end{aligned}$$

A marginal change in bid b_l *proportionally* increases $dF(q \mid W, b)$ by the difference between the unconditional expected number of *projects* which pick turbine l and the same expectation, conditional on the sales vector q. Note that in the special case where $n_i = n$ for all i (homogenous project sizes), in any feasible allocation, the number of projects that pick turbine l is, by definition, equal to q_l. When this happens, $\mathbb{E}\left[\# \text{ of projects pick } l \mid q, W, b\right] = q_l$. For notational brevity, we define $N_l(q, W, B) = \mathbb{E}\left[\# \text{ of projects pick } l \mid q, W, b\right]$.

With this derivation, we can now write the Berry-Pakes factor as:

$$\frac{\frac{\partial}{\partial b_{lt}} dF(q_t^* \mid W_t, b_t)}{dF(q_t^* \mid W_t, b_t)} = \sum_{i \in W_t} s_{ilt} - N_l(q_t, W_t, b_t)$$

5.4 Discounted Cash Flows

The Berry-Pakes approach requires a measure of *realized* discounted future cashflows for each firm and time period, $V_{ft}(x_t + q_t^*)$. In our analyses below, we construct two different types of discounted cash flow measures: those that are implied by our demand and cost models, as suggested in Berry and Pakes (2000), and those derived from public company accounting data, which is available for a subset of our firms and time periods.

In the Berry-Pakes "model" approach, we combine data on realized future turbine choice probabilities and project characteristics with our model for turbine costs. Specifically, we'll compute:

$$V_{ft}^M(x_t + q_t^* \mid \theta) = \sum_{\tau=1}^{T-t} \beta(\tau) \sum_{i \in W_{t+\tau}} n_i \left(R_{i,f,\tau+t} - \sum_{j \in K_{f,t+\tau}} s_{ij} c(E_{f,t+\tau,j}, \theta) \right)$$

where $\beta(t)$ is a discounting factor, $R_{i,f,\tau+t}$ is the expected revenue per turbine for firm f at project i during a procurement auction in period $\tau + t$, and $c(\widetilde{E}_{f,t+\tau,j}, \theta)$ is the cost to firm f of manufacturing a turbine with manufacturing "experience" $E_{f,t+\tau,j}$, which we define in Section 6.1. For all but the final year T, we'll write $\beta(t) = \beta^t$. However, to capture future cash flows that have yet to be realized but which may affect firm's perceptions of the dynamic markdown, we'll annuitize the final year's cashflows and write $\beta(T) = \frac{\beta^T}{1-\beta}$. We assume $\beta = 0.9$.

In this "modeled" approach, future discounted revenues come from our demand system estimates and future discounted costs come from the shares implied by our demand system estimates, realized experience levels, and the functional form and parameters which relate experience to manufacturing costs. Because discounted cash flows are a function of the cost parameters, this future realized cash flow construction may directly affect our estimates of the cost function.

In the Berry-Pakes "accounting" approach, we instead rely on public financial reporting which is available for a subset of the industry. Vestas, Nordex and Gamesa, as well as some of the larger Chinese manufacturers, are publicly traded companies that are either wind turbine "pure plays" or have detailed wind segment reporting. These firms directly report the revenues earned and costs incurred selling wind turbines, and we can use this information to directly construct discounted cash flows, again annuitizing the final year to capture future cash flows which have yet to be realized:

$$V^A_{ft}(x_t + q^*_t) = \sum_{\tau=1}^{T-t} \beta(\tau) \left(R^{\text{Accounting}}_{f\tau} - C^{\text{Accounting}}_{f\tau} \right)$$

5.5 Estimating Equations

To recap, the firm's first order condition for dynamically optimal bidding is:

$$0 = \sum_{i \in W_t} n_i \sum_{j \in K_{ft}} \frac{\partial s_{ijt}}{\partial b_{lt}} (b_{jt} - c_{jt}) + V_{ft}(x_t + q^*_t) \left(\sum_{i \in W_t} s_{ilt} - N_l(q_t, W_t, b_t) \right) + \nu_{lt} \tag{3}$$

We can collect these terms to describe a practical estimating equation. Let b_{ft} and c_{ft} represent vectors of the bids and marginal cost functions for firm f's turbines in period t. Let the matrix of project-size weighted demand elasticities among firm f's turbines be $\Delta_{ft} = \sum_{i \in W_t} n_i \sum_{j \in K_{ft}} \nabla s_{ift}$, where s_{ift} is the vector of firm f's choice probabilities at project i. Similarly, let $N_{ft}(q_t, W_t, b_t)$ be the vector of the expected number of projects that pick f's turbine's in market realization q_t for projects W_t with the full vector of bids b_t. Finally, let $s_{ft} = \sum_{i \in W_t} s_{if}$. Then our FOC in vector form is:

$$0 = \Delta_{ft}(b_{ft} - c_{ft}) + V_{ft}(x_t + q^*_t) \left(s_{ft} - N_{ft}(q_t, W_t, b_t) \right) + \nu_{ft}$$

We can re-arrange this to express the vector of firm f's bids in terms of its costs, its dynamic incentives, and a rational expectations shock in period t:

$$b_{ft} = c_{ft} - V_{ft}(x_t + q^*_t)(\Delta_{ft})^{-1} \left(s_{ft} - N_{ft}(q_t, W_t, b + t) \right) - (\Delta_{ft})^{-1} \nu_{ft} \tag{4}$$

To specify this for a single turbine j, let ξ_{jt} be the j-th entry of the vector $(\Delta_{ft})^{-1} \left(s_{ft} - \widehat{N}_{ft}(q_t, W_t, b + t) \right)$, and let the j-th entry of $(\Delta_{ft})^{-1} \nu_{ft}$ be $\widetilde{\nu}_{jt}$. Then our estimating equation becomes:

$$b_{jt} = c_{jt} - V_{ft}(x_t + q^*_t) \xi_{jt} - \widetilde{\nu}_{jt}$$

To the extent that our discounted cash flow calculations do not capture all of the firm's dynamic bidding incentives, bids and $V_{ft}(x_t + q^*_t)\xi_{jt}$ will not necessary covary one-to-one. This could happen because our discount factor is wrong (we are assuming a nominal discount rate of 10% per year), because we do not account for future fixed cost expenditures, or because the price of turbine j may only affect a portion of future discounted cash flows. To allow for this, we'll estimate a coefficient μ on the dynamic markdown terms, so that our estimating equation becomes:

$$b_{jt} = c_{jt} + \mu V_{ft}(x_t + q_t^*)\xi_{jt} - \widetilde{\nu}_{jt}$$

If our model were perfectly specified, we would expect to find $\mu = -1$.

Finally, recall that we do not actually observe or estimate the leve of bid b_{jt}, as our discrete choice model does not have an outside option. Instead, our discrete choice model estimates deliver $\widehat{b}_{jt} = b_{jt} - b_{0t}$ for a pre-specified base turbine 0. Accounting for this, our estimating equation becomes:

$$\widehat{b}_{jt} = c_{jt} + \mu V_{ft}(x_t + q_t^*)\xi_{jt} - b_{0t} - \widetilde{\nu}_{jt} \tag{5}$$

We capture the effects of the base turbine bid b_{0t} using year fixed effects.

5.6 Identification

We estimate the above model using control-function nonlinear least squares techniques and generalized method of moments estimators. In order for either approach to deliver consistent estimates of the cost function, we must make assumptions about how the unobservable terms are correlated with observables or other variables we may use as instruments. We envision two sources of identification challenges in this setting.

First, any specification that includes a version of our rational expectations dynamic markdown term will depend on q_t^* and other variables which are not known by firms when they set bids. As a result, these variables cannot serve as valid instruments, as they will be mechanically correlated with the rational expectations shock. If we had a fully-specified data generating process for the set of projects that come to the market in each year, the power prices they face, and evolution of the market for wind turbine materials, then functions of the current value of those state variables would be valid instruments. This is the approach in Berry and Pakes (2000) and other rational expectations settings. However, we have left these details unspecified, focusing only on various components of world/firm/turbine experience as state variables in the firms' dynamic problems. Thus, we need other variables which correlate with $V_{ft}(x_t + q_t^*)\left(\sum_{i\in W_t} s_{ilt} - N_l(q_t, W_t, b_t)\right)$ but which are uncorrelated with the rational expectations shock.

We obtain these variables by recognizing that the endogenous term we have constructed, $V_{ft}(x_t+q_t^*)\left(\sum_{i\in W_t} s_{ilt} - N_l(q_t, W_t, b_t)\right)$, is a noisy measure of the *ideal* term we'd like to construct: $\mathbb{E}\left[\frac{\partial}{\partial b_{lt}} V_{ft}(x_t + q_t) \mid W_t, b_t\right]$. In fact, the relationship between the observed quantity and the ideal quantity in this setting exactly satisfies the standard classical measurement error assumption, that the difference between the true and observed quantity is independent of the truth. This means that we can construct instruments using tools from the classical measurement error literature.

We follow the results from Lewbel (1997) and create instrumental variables from (centered) higher order moments of bids, Berry-Pakes factors, and components of the discounted cash flow terms. Lewbel (1997) shows that these variables are excludable (uncorrelated with ν) and relevant (correlated with $V_{ft}(x_t + q_t^*) \left(\sum_{i \in W_t} s_{ilt} - N_l(q_t, W_t, b_t)\right)$ whenever $\mathbb{E}\left[\frac{\partial}{\partial b_{lt}} V_{ft}(x_t + q_t) \mid W_t, b_t\right]$ has nonzero skew.

Second, although our exposition of the above model envisions the unobservable term in our estimating equation as a purely expectational shock, it is possible that there are other unobservable determinants of bids which may be correlated with observable determinants of costs, like firm- or turbine-specific experience. This would be the case if, for example, firms had serially correlated productivity shocks. If this were true, high productivity firms in previous years would have lots of experience and also lower than average costs in the present year. To deal with this potential source of endogeneity, we construct predictions of turbine sales that are driven exclusively by home-market bias and changes to country level demands. We then use these predictions to construct predictions of the experience firms accumulate from these forces alone, which we use as instrumental variables for our observed experience measures.[7] The validity of this "Bartik" style instrument relies on the assumption that year-to-year changes in total demand for wind turbines across countries are uncorrelated with unobservable determinants of manufacturer costs.

6 Results

6.1 Cost parameterization

Our goal is to measure the extent to which manufacturing costs correlate with various notions of manufacturing experience, at the turbine, firm, and industry-level. To do this, we need to impose more structure on the cost term c_{jt} in equation 5. We assume turbine manufacturing costs factor into two terms: *resources* and *experience*. By "resources," we mean that bigger turbines require more material inputs to produce, conditional on firm experience. As we discussed in section 2, material inputs for manufacturing a wind turbine grow approximately cubically in rotor size r. However, firms with more experience may be able to economize on unobservable inputs, like labor, wasted materials, etc. We assume that this learning reduces

[7] Specifically, we compute choice probabilities for each turbine-project pair using a demand system that includes turbine-project revenues and country-by-firm dummies. These choice probabilities reflect changes in each country's overall demand and the share of demand in a country that each turbine would typically expect to receive, but do not depend on year-to-year changes in turbine pricing.

firms' cost per unit turbine volume, such that the marginal cost of manufacturing turbine j at time t is

$$c_{jt} = \omega_{jt} r_j^3 = \omega_0 \left(\widetilde{E}_{jt}\right)^{\alpha} r_j^3 \tag{6}$$

where ω_0 reflects the baseline cost of volume in the industry. The actual cost faced by turbine j at time t declines with "effective" experience $\widetilde{E}_{jt}$. Following Irwin and Klenow (1994), we parameterize this as

$$\widetilde{E}_{jt} = \beta_J E_{jt} + (E_{ft} - E_{jt}) + \beta_W (E_t - E_{f(j),t}) \tag{7}$$

Effective experience is a linear combination of a firm's experience manufacturing a specific turbine (E_{jt}), the firm's experience manufacturing *other* turbines in its portfolio ($E_{f(j),t} - E_{jt}$), and the experience that all other firms in the industry have thus far accumulated ($E_t - E_{f(j),t}$).[8] We measure experience as the total cumulate *volume* that firm $f(j)$ has previously shipped, so these experience terms can be constructed from the state vector x_t.[9] The parameter β_J thus measures the extent to which experience manufacturing turbine j is more or less useful than the firm's experience producing other turbines. Similarly, the parameter β_W measures the extent to which other firms' aggregate experience is as useful as a firm's own experience making other turbines. That is, a unit of turbine-specific experience is β_J times as valuable as firm experience, and a unit of industry experience is β_W times as valuable as firm experience. Finally, the parameter α represents learning economies. Holding everything else equal, a 1% increase in effective experience increases costs by α%. When $\alpha < 0$, the learning-by-doing literature often reports $1-2^{\alpha}$, the so-called "Spence coefficient" which measures the proportional effect of doubling of effective experience on costs.

Our fully parameterized estimating equation is now :

$$\widehat{b}_{jt} = \omega_0 \left(\beta_J E_{jt} + (E_{ft} - E_{jt}) + \beta_W (E_t - E_{f(j),t})\right)^{\alpha} r_j^3 + \mu V_{ft}(x_t + q_t^*)\xi_{jt} - b_{0t} - \widetilde{\nu}_{jt} \tag{8}$$

6.2 Static results

We first estimate equation 8 under the assumption that pricing choices are purely static, ignoring the dynamic learning benefit term, $V_{ft}\xi_{jt}$. Table 6 presents these results. We restrict the estimating sample to include turbines sold by one of the top nine firms globally (Table 1), with at least two sales in the demand estimation sample in year t, and estimate

[8] Formally, firm experience $E_{f(j),t} = \sum_{l \in K_{f(j),t}} E_{lt}$, and industry experience $E_t = \sum_f E_{ft}$.

[9] This state vector also admits construction of experience in terms of megawatts, turbine shipments, etc.

its parameters using nonlinear least squares.[10] To handle the endogeneity of the experience variables, we use the control function approach of Newey et al. (1999).[11]

In column 1, we assume that the cost of volume from equation 7 only depends on firm experience. The estimated "initial" cost per unit volume, when effective experience is normalized to one, is 11.2.[12] To put this number in perspective, at this cost, a 90 meter turbine made by an industry with this initial level of experience would cost about \$1 million. However, costs are negatively correlated with experience, to the point that a 1% increase in effective experience decreases marginal costs by about 0.2% (α in equation 8). In column 2 we allow a turbine's effective experience to include the experience of other firms. The estimates in the row labelled "IndustryExpc," which represents β_W in equation 7, imply that a one unit increase in global experience at other firms generates 12 percent of the learning benefits as a unit increase on own-firm experience. In column 3, we allow effective experience to evolve differently across turbines within the same firm. The row labelled "TurbineExpc," which represents β_J in equation 7, implies that a one unit increase in experience for turbine j provides more than one hundred times the learning benefits generated by the firm's other turbines. In this model, where turbine j's experience has been separated out, global experience is now worth 22 percent of firm experience selling other turbines.

[10] The tenth largest firm, Suzlon, sells nearly all of it's turbines in India. India was excluded from the demand estimation sample because we could not reliably measure the output price each developer faced.

[11] To construct these control functions, we first regress our observed experience covariates onto polynomial functions of the full vector of the instruments described above, as well as all the exogenous terms in equation 8. We then include a polynomial function of these first stage residuals in the nonlinear least squares procedure.

[12] To improve numerical stability and readability, we divide all the experience terms in equation 7 by the global experience level E_t in the year 2000. Given this, ω_0 can be interpreted as the cost state when effective experience is equal to global experience at the start of the sample.

Table 6: Static Learning Regressions

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Base Cost (ω_0)	10.2	26.0	41.7	11.3	40.5	65.1
	(1.1)	(3.5)	(5.9)	(1.6)	(11.0)	(12.4)
Learning Exponent (α)	-0.20	-0.44	-0.50	-0.21	-0.41	-0.47
	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)	(0.03)
IndustryExpc (β_W)		0.10	0.18		0.47	0.67
		(0.03)	(0.06)		(0.31)	(0.22)
Turbine Expc (β_J)			103.8			267.0
			(31.7)			(70.2)
Observations	983	983	983	983	983	983
Adjusted Pseudo R^2	0.12	0.12	0.15	0.21	0.23	0.26
FE	Year	Year	Year	Year, Firm	Year, Firm	Year, Firm
Spence Coef.	0.13	0.26	0.29	0.13	0.25	0.28

All models estimated with nonlinear least squares. To account for endogeneity in turbine and firm experience, all models include control function using the approach of Newey et al. (1999). Robust standard errors presented in parentheses.

Columns 4 through 6 repeat these models but include manufacturer (OEM) fixed effects, to allow for other time invariant unobservable cost difference across firms that might be correlated with sales. Comparing these estimates to the first set, the estimated initial cost levels (ω_0) are larger, and learning coefficients (α) are smaller. However, the confidence intervals in columns 4-6 contain the point estimates from columns 1-3, so statistically these estimates are similar. Looking within manufacturer, the effects of both global experience and turbine experience are larger than they are in the cross section, although the estimates are noisier. Across all six specifications, the Spence coefficients ranges from 14% to 29%.

6.3 Dynamic results

A potential concern with all of the models in the previous section is the possibility that the dynamic markdowns we've ignored contain valuable information about learning economies and/or spillovers. We address this in Table 7, which implements the strategies to control for dynamic markdowns developed in section 5.4. Column 1 repeats the model with both world and turbine experience (column 3) from table 6, but uses GMM-IV rather than NLS with a control function. These estimates suggest greater learning economies and smaller spillovers than their counterparts in Table 6, though they are statistically similar. Column

2 includes firm-time fixed effects, in an attempt to "simply" control for unobserved dynamic markdowns. In addition to subsuming b_{0t}, the bid on the base turbine, this approach would also control for any dynamic bidding incentives that were common across turbines sold by the same firm at the same point in time. Though it does not exactly map to the incentives derived above, it is considerably easier to construct than our other dynamic corrections. Compared to model 1, model 2 shows somewhat smaller learning economies, and that turbine specific experience and world spillovers are relatively more important, but the estimates are quite noisy.

Table 7: Dynamic Markdown Controls

	(1)	(2)	(3)	(4)	(5)
Base Cost (ω_0)	57.44	73.81	94.49	66.84	64.58
	(13.39)	(23.60)	(25.11)	(20.59)	(19.05)
Learning Exponent (α)	-0.53	-0.48	-0.57	-0.67	-0.67
	(0.05)	(0.04)	(0.04)	(0.09)	(0.09)
Industry Expc (β_W)	0.07	0.21	0.22	0.05	0.05
	(0.04)	(0.16)	(0.11)	(0.03)	(0.03)
Turbine Expc (β_J)	74.56	362.06	161.76	44.20	35.76
	(35.91)	(216.46)	(76.67)	(26.67)	(21.57)
$\mu \times 10^4$			-0.94		-1.43
			(0.22)		(0.46)
N	983	983	983	505	505
FE	Year	Firm x Year	Year	Year	Year
Dynamics	None	None	BP (Model)	None	BP (Accounting)
Sample	All	All	All	Accounting	Accounting

ω_0 is the "initial" cost of a turbine per unit of materials. α is the Irwin & Klenow exponent. μ is the coefficient on the dynamic markdown term.

Column 3 implements the model-based Berry-Pakes control strategy. Compared to the first two columns, the estimated initial cost state was larger. Relatedly, the implied Spence coefficient increases to 0.32 from 0.3 in column 1. World experience and turbine experience are twice as important, relative to firm experience, in this model compared to model 1. The estimated coefficient on the dynamic markdown term, μ is statistically significant, and has the correct sign, but is economically quite small. Though we can easily reject a hypothesis test that $\mu = -1$, the fact that turbine-specific experience and industry-spillovers are both larger in a model that accounts for dynamic markdowns is consistent with the Benkard (2004) argument that prices should respond less to experience accumulation that costs do.

In columns 4 and 5 we implement the accounting based dynamic control term. A panel of reliable financials was only available for five firms (Vestas, Gamesa, Nordex, Goldwind and Guodian) whose turbines collectively cover about half of our overall sample. To have a baseline, no-dynamics comparison, we first estimate the static model (column 1) on this smaller sample. The learning coefficient in this sample is larger, and the world and turbine experience terms are relatively smaller, compared to the model in column 1. In column 5, we include the accounting-based dynamic control. As in column 3, the multiplier on that term, μ is precisely estimated, and has the corect sign, but is economically small. Although the estimates are quite noisy given the small sample, the point estimates are very similar across the two models.

6.4 Alternative Learning Models

In table 8 we explore the extent to which our results are robust to alternative assumptions about the nature of learning by doing in this industry, including different measures of experience and different functional forms relating experience and costs. For comparison, column 1 repeats our primary specification from table 6.

In the models above, we measure experience as the cumulative sum of turbine *volume* sold as of time t, motivated by the fact that the key materials cost variable we use is also volume. In columns 2 and 3 of Table 8, we instead use cumulative megawatts and cumulative turbines sold as the measure of experience. Both of these alternative measures suggest larger learning economies and spillovers, as well as a larger role for own-turbine experience.

Table 8: Learning Function Alternatives

Model:	(1)	(2)	(3)	(4)	(5)
Base Cost (ω_0)	44.5	42.9	58.1	32.7	20.5
	(6.8)	(8.4)	(21.4)	(4.8)	(3.0)
Turbine Expc (β_J)	122.1	144.9	471.4	122.5	108.4
	(37.8)	(48.3)	(245.5)	(48.7)	(143.3)
IndustryExpc (β_W)	0.22	0.27	0.62	0.25	0.26
	(0.07)	(0.10)	(0.30)	(0.11)	(0.26)
Learning Exponent (α)	-0.50	-0.59	-0.77		
	(0.03)	(0.04)	(0.05)		
Terminal Cost (ω_T)				0.34	2.0
				(0.93)	(0.88)
Expc Measure	Volume	MW	Turbines	Volume	Volume
Learning Model	Unbounded	Unbounded	Unbounded	Accumulation	Replacement
Fixed-effects					
Year	Yes	Yes	Yes	Yes	Yes
Observations	983	983	983	983	983
Adjusted Pseudo R^2	0.15	0.15	0.15	0.15	0.15

All models estimated with nonlinear least squares. To account for endogeneity in turbine and firm experience, all models include control function using the approach of Newey et al. (1999). Robust standard errors presented in parentheses.

So far, we have used the same functional form relating costs to experience as most other work in the learning by doing literature (Wright, 1936; Thornton and Thompson, 2001; Benkard, 2000). However, this model is *unbounded* in the sense that costs approach zero as firms accumulate infinite experience. As noted by Thompson (2007), there are other tractable functional forms for LBD that admit more realistic long-term predictions about cost, while still allowing for an additive structure governing effective experience. In columns 4 and 5, we implement two *bounded* learning models suggested in Thompson (2007). In these models, ω_T represents estimated *terminal* costs, after all learning opportunities have been exhausted, so these estimates indicate that experience can eventually reduce costs by one to two orders of magnitude.[13] Moreover, the initial cost parameter ω_0 and the weights on turbine-specific experience and industry spillovers are similar to those in column 1. Taken together, these results indicate that the learning economies and spillover magnitudes we measure are robust to a variety of alternative ways of measuring experience and modelling

[13] In bounded learning models, there is no learning exponent α. Instead, there is a "step-size" parameter, representing the fraction of the gap between inintial and terminal costs that each unit of effective experience delivers which we calibrate to 0.05.

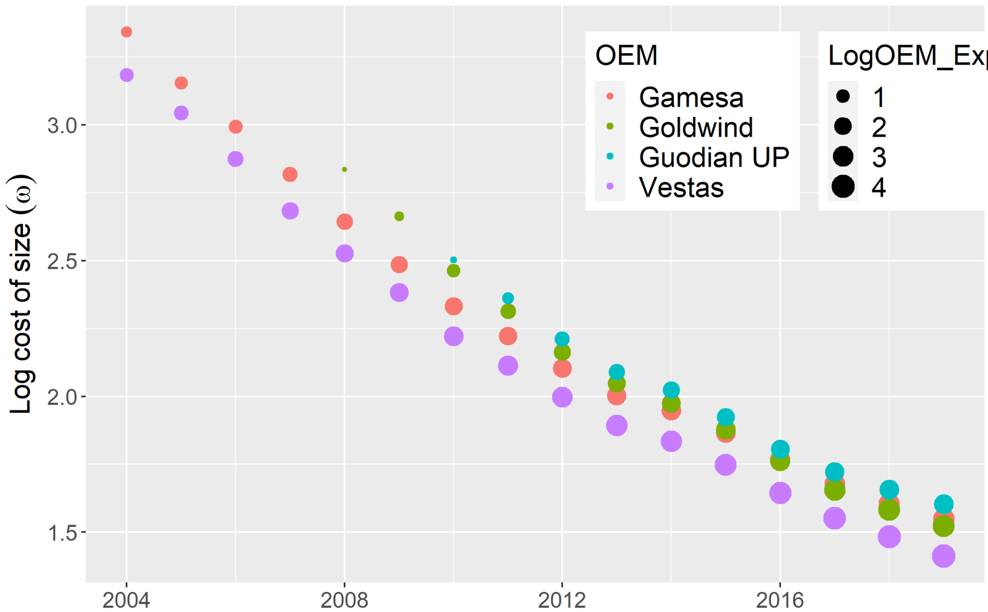
costs.

7 Discussion

We consistently find that experience today generates large cost reductions tomorrow, with a doubling of effective experience reducing costs by up to 29%. At the same time, while spillovers do exist, the effect of a turbine's own experience on its costs is two orders of magnitude larger than the effect experience coming from the manufacturers other turbines. Taken together, these two facts present a puzzle: if learning effects are large, and within-turbine experience generates 100 times the experience benefits of other experience within the firm, why do manufactures *ever* introduce new turbines? Note that turbine lifespans are fairly short. For turbines introduced between 2000 and 2015, the median duration on the market was six years, and most sales occur in three or fewer years.

One reason why firms may introduce new turbines, even when existing turbines have learning-driven cost advantages, is that future gains to learning for mature turbines are eventually small, due to the decreasing returns we estimate. While the *level* of a mature turbine's cost can be quite low after a few years of sales, the marginal gains to future sales are also quite small, while new turbines have an entire learning curve ahead of them. Moreover, the initial cost of a new turbine is declining over time due to spillovers within and across firms. Thus, firms may introduce new turbines in spite of temporary cost disadvantages, as a way of investing in new learning opportunities. Our cost function estimates support this idea. Figure 9 presents the inital cost of size, by year, for four large firms. On average, initial costs are falling by more than half a log point every four years. As the cost of size declines, the value proposition of introducing a new, larger turbine, with valuable future learning opportunities, becomes more attractive, even as costs of existing smaller turbines remain relatively small.

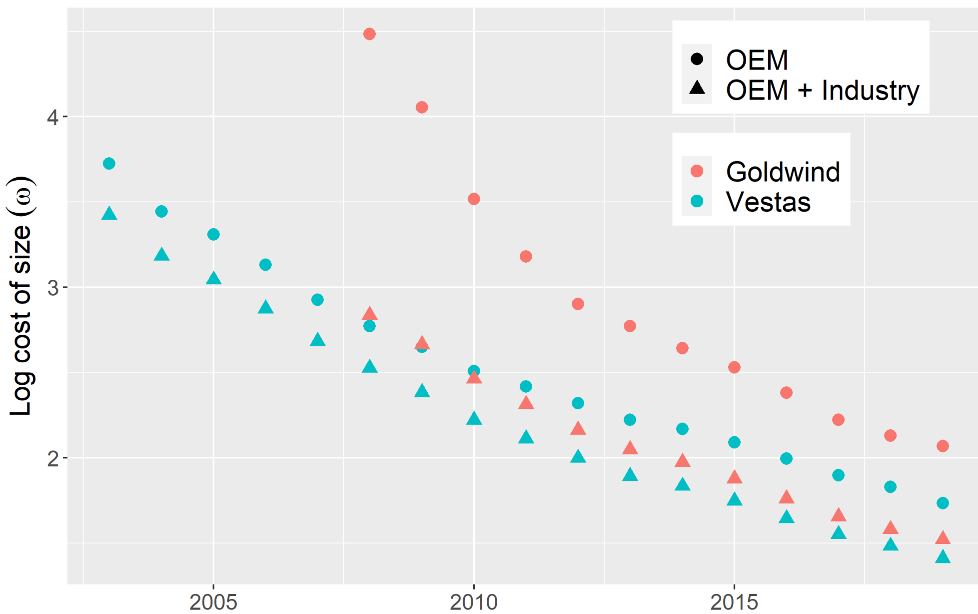
Figure 9: Example Estimated New Turbine Cost of Size



This figure presents the (log) estimated cost of size (ω_{ft}) for a newly introduced turbine ($E_{jt} = 0$), using the results from column 3 of table 6. Point sizes are proportional to cumulative manufacturer ("OEM") experience at the start of the year.

Given that spillovers are relatively small, why are new turbine costs declining so much? In figure 9, effective experience only depends on firm experience and global experience. And while these are worth approximately one percent of turbine experience, they are applied of a much larger set of turbines. So, while individual turbine spillovers are small, across all sales, they create a meaningful reduction over time. Figure 10 demonstrates this for Vestas, the largest manufacturer at the start of the sample, and Goldwind, the largest Chinese manufacture, which not in the market until the mid 2000s. Despite having no experience when it enters in 2008, it's costs were much closer to Vesta's initially due to spillovers. By 2015, Goldwind's cost of size with world spillovers was already equivalent to Vestas' without it.

Figure 10: New Turbine Cost of Size with and without Spillovers



This figure presents the (log) estimated cost of size (ω_{ft}) for a newly introduced turbine ($E_{jt} = 0$), using the results from column 3 of table 6. In the series labeled "OEM", spillovers from other firms' experience are turned off ($\beta_W = 0$). In the models labeled "OEM + World", the estimated of β_W is taken from column 3 of table 6.

The other reason why firms introduce new turbines is that they have to in order to stay competitive. As noted in Stein (1997), firms who hope to move down the learning curve for a newly introduced product are often disappointed when their rivals quickly respond with a better product. At that point, the now-lagging firm faces a tradeoff: continue to produce its existing, and now inferior product, at ever lower costs, or abandon those learning opportunities and introduce its own newer product. Some combination of these two forces is presumably what drives the ever increasing sizes observed in Figure 5.

8 Conclusion

We estimate the extent to which learning by doing and spillovers have reduced costs in the wind turbine industry. Because neither costs nor inputs are recorded in public data, we infer latent manufacturing costs from a structural model of turbine procurement which we estimate using the universe of wind turbine models procured by the near-universe of wind plants built in the last twenty years. To distinguish between manufacturing costs and the dynamic benefits inherent in any environment where learning by doing is present, we leverage insights from Berry and Pakes (2000) which allow us to control for dynamic pricing incentives without estimating or computing a dynamic game. We find that a doubling of manufacturing experience reduces manufacturing costs by 14 to 29 percent. Only 1 to 2 percent of experience spills over to other turbine models produced by the same firm, and spillovers to turbines produced by other firms are on the order 0.1 to 0.6 percent. Though relatively small, we show that, in aggregate, spillovers have generated significant cost reductions over time. These results are consistent with policymaker motivation for generously subsidizing the industry.

References

Abrell, Jan, Mirjam Kosch, and Sebastian Rausch, "Carbon abatement with renewables: Evaluating wind and solar subsidies in Germany and Spain," *Journal of Public Economics*, 2019, *169*, 172–202.

Acemoglu, Daron, David Hemous, Lint Barrage, Philippe Aghion et al., "Climate change, directed innovation, and energy transition: The long-run consequences of the shale gas revolution," in "2019 Meeting Papers" number 1302 Society for Economic Dynamics 2019.

Aldy, Joseph E, Todd D Gerarden, and Richard L Sweeney, "Investment versus Output Subsidies: Implications of Alternative Incentives for Wind Energy," Technical Report 24378, National Bureau of Economic Research April 2019.

Arrow, Kenneth J, "The Economic Implications of Learning by Doing," *The Review of economic studies*, 1962, *29* (3), 155–173.

Asker, John and Estelle Cantillon, "Properties of scoring auctions," *The RAND Journal of Economics*, 2008, *39* (1), 69–85.

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin, "Estimating dynamic models of imperfect competition," *Econometrica*, 2007, *75* (5), 1331–1370.

Benkard, C. Lanier, "Learning and Forgetting: The Dynamics of Aircraft Production," *American Economic Review*, September 2000, *90* (4), 1034–1054.

Benkard, C Lanier, "A dynamic analysis of the market for wide-bodied commercial aircraft," *The Review of Economic Studies*, 2004, *71* (3), 581–611.

Benthem, Arthur Van, Kenneth Gillingham, and James Sweeney, "Learning-by-doing and the optimal solar policy in California," *The Energy Journal*, 2008, *29* (3).

Berry, Steve and Ariel Pakes, "Estimation from the Optimality Conditions for Dynamic Controls," December 2000.

Besanko, David, Ulrich Doraszelski, and Yaroslav Kryukov, "The economics of predation: What drives pricing when there is learning-by-doing?," *American Economic Review*, 2014, *104* (3), 868–97.

— , — , — , **and Mark Satterthwaite**, "Learning-by-doing, organizational forgetting, and industry dynamics," *Econometrica*, 2010, *78* (2), 453–508.

Betz, Albert, *Wind-Energie und ihre Ausnutzung Durch Windmühlen*, Vandenhoeck, 1926.

Coşar, A Kerem, Paul L E Grieco, and Felix Tintelnot, "Borders, Geography, and Oligopoly: Evidence from the Wind Turbine Industry," *The Review of Economics and Statistics*, July 2015, *97* (3), 623–637.

Dasgupta, Partha and Joseph Stiglitz, "Learning-by-Doing, Market Structure and Industrial and Trade Policies," *Oxford Economic Papers*, 1988, *40* (2), 246–268.

Greenstone, Michael and Ishan Nath, "Do renewable portfolio standards deliver cost-effective carbon abatement?," *Becker-Friedman Institute Working Paper*, 2020.

Irwin, Douglas A and Peter J Klenow, "Learning-by-Doing Spillovers in the Semiconductor Industry," *The Journal of Political Rconomy*, 1994, *102* (6), 1200–1227.

Jaffe, Adam B, Richard G Newell, and Robert N Stavins, "A tale of two market failures: Technology and environmental policy," *Ecological economics*, 2005, *54* (2-3), 164–174.

Knittel, Christopher R, "Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector," *The American Economic Review*, December 2011, *101* (7), 3368–3399.

Levitt, Steven D, John A List, and Chad Syverson, "Toward an understanding of learning by doing: Evidence from an automobile assembly plant," *Journal of Political Economy*, 2013, *121* (4), 643–681.

Lewbel, Arthur, "Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D," *Econometrica: journal of the econometric society*, 1997, pp. 1201–1213.

Newell, Richard G, Adam B Jaffe, and Robert N Stavins, "The induced innovation hypothesis and energy-saving technological change," *The Quarterly Journal of Economics*, 1999, *114* (3), 941–975.

Newey, Whitney K, James L Powell, and Francis Vella, "Nonparametric estimation of triangular simultaneous equations models," *Econometrica*, 1999, *67* (3), 565–603.

Rust, John, "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, 1987, pp. 999–1033.

Stein, Jeremy C, "Waves of Creative Destruction: Firm-Specific Learning-by-Doing and the Dynamics of Innovation," *The Review of Economic Studies*, April 1997, *64* (2), 265–288.

Thompson, Peter, "How Much Did the Liberty Shipbuilders Forget?," *Management Science*, 2007, *53* (6), 908–918.

Thornton, Rebecca Achee and Peter Thompson, "Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding," *American Economic Review*, 2001, *91* (5), 1350–1368.

Wright, Theodore P, "Factors affecting the cost of airplanes," *Journal of the aeronautical sciences*, 1936, *3* (4), 122–128.

Appendix A Computational details

A.0.1 Computation of N_l

How do we compute $N_l(q, W, b)$, the expected number of projects out of market W that pick turbine l, when the bids are the vector b and the realized turbine sales vector is q? As indicated above, in the case of homogenously sized projects, this object is simply the vector q. However, there is meaningful heterogeneity in size across projects which we must account for. Recall that the formal definition of this object is:

$$\begin{aligned} N_l(W_t, q_t, b_t) &= \mathbb{E}\left[\# \text{ of projects pick } l \mid q_t, b_t, W_t\right] \\ &= \frac{\sum_{m \in M(q_t, W_t)} \left(\prod_{i \in W_t} s_{i,m_i}\right) \sum_{i \in W_t} \mathbb{I}[m_i = l]}{\sum_{m \in M(q_t, W_t)} \prod_{i \in W_t} s_{i,m_i}} \end{aligned}$$

Exact computation of this requires a full enumeration of the set $M(q_t, W_t)$, which can be defined as the set of integer solutions to an under-determined system of linear equations with integer coefficients. Let μ_{ij} for $i \in W_t$ and $j \in \cup_f K_{ft}$ be the entries of a matrix representing an allocation of turbines to plants. When plant i receives turbine j, $\mu_{ij} = 1$, and zero otherwise. In addition to the requirement that all entries of μ_{ij} are binary, feasible values of μ_{ij} satisfy two constraints. First, each plant chooses exactly one turbine, so for all i:

$$\sum_j \mu_{ij} = 1$$

Second, each turbine must sell in the quantities we ultimately observe in the aggregate sales vector q, so for all j:

$$\sum_i n_i \mu_{ij} = q_j$$

If the number of wind farms and turbines were sufficiently small, it would be possible to exhaustively enumerate all feasible solutions μ_{ij}, either using an integer linear programming solver, or specialized software for finding the vertices of the polyhedron defined by these equations.[14] However, with hundreds of projects per year and as many as 50 turbines in some years, complete enumeration is computationally impossible. Moreover, many feasible allocations may have vanishingly small probabilities of occurring, and as such may not contribute much to the exact value of $N_l(q, W, b)$.

In light of this, we approximate this object by using an integer linear programming solver in the "solution pool" mode.[15] We ask the solver to find the $L = 200$ "best" feasible solutions, where solution quality is the log-likelihood of the observed allocation, or $\sum_i \sum_j \mu_{ij} \log s_{ij}$.[16] Let $\widehat{M}(q, W)$ be our approximate set of solutions for realized sales q under market W. Then

[14] See, for example, `http://cgm.cs.mcgill.ca/~avis/C/lrs.html`

[15] Both CPLEX and Gurobi offer this option. We have used Gurobi here.

[16] This problem is NP-hard, and so instead of allowing the solver to run indefinitely, we collect the 200 best solutions available after 15 minutes of solution time per market.

we approximate $N_l(q_t, W_t, b_t)$ with:

$$\widehat{N}_l(q_t, W_t, b_t) = \frac{\sum_{m\in\widehat{M}(q_t,W_t)} \left(\prod_{i\in W_t} s_{i,m_i}\right) \sum_{i\in W_t} \mathbb{I}[m_i = l]}{\sum_{m\in\widehat{M}(q_t,W_t)} \prod_{i\in W_i} s_{i,m_i}}$$

for each l in $\cup_f K_{ft}$.

Appendix B Additional Tables and Figures

Table A.1: Demand Estimation Observations by Country, Total Capacity

Year	AU-DK-PT	CHINA	DEU	FRA	ITA	SPAIN	SWE	U.K.	U.S.A.	Excluded	Share
2000	120	42	1183	39	167	721	14	30	40	367	0.87
2001	51	0	2003	23	144	837	16	78	1425	650	0.88
2002	227	0	2204	43	53	931	20	79	643	628	0.87
2003	286	101	2092	56	493	1350	61	99	1562	1097	0.85
2004	480	46	1499	120	364	2649	42	110	376	1246	0.82
2005	768	359	1601	418	286	1591	15	372	2145	1782	0.81
2006	836	775	1808	790	474	2077	94	607	2534	3758	0.73
2007	551	2245	1287	843	798	2285	162	322	5278	3217	0.81
2008	659	4451	923	1234	767	2378	151	488	8275	4231	0.82
2009	762	10045	1696	1220	1227	2410	288	741	9596	6128	0.82
2010	632	13660	992	1374	1127	1053	516	560	4234	7060	0.77
2011	421	15846	1342	986	855	976	508	491	6397	9988	0.74
2012	408	10175	1354	700	775	824	645	1008	12712	8987	0.76
2013	722	10087	1603	649	393	331	578	1288	604	7457	0.69
2014	681	12074	3493	1031	135	0	950	820	4957	14183	0.63
2015	546	18731	3372	1084	193	14	723	550	8232	14415	0.70
2016	708	13785	4279	1378	320	13	491	748	8757	12761	0.70
2017	401	9817	4860	1782	250	47	160	2079	6008	12265	0.67
2018	502	7828	2491	1233	398	242	497	688	6785	11661	0.64
2019	191	6664	915	1320	490	1892	728	495	6148	13743	0.58
2020	193	17193	1296	1166	93	1568	1851	90	0	15586	0.60

B.1 Additional Results: Static Estimation

Table A.2: Static Learning Regressions - Year Heterogeneity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Base Cost (ω_0)	15.9	39.7	60.1	18.2	57.0	88.1
	(1.7)	(4.9)	(7.9)	(2.5)	(12.2)	(14.5)
Learning Exponent (α)	-0.21	-0.45	-0.49	-0.23	-0.42	-0.46
	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)	(0.03)
IndustryExpc (β_W)		0.10	0.16		0.34	0.54
		(0.03)	(0.05)		(0.18)	(0.17)
Turbine Expc (β_J)			105.4			247.1
			(35.1)			(69.4)
Fixed-effects						
Year	Yes	Yes	Yes	Yes	Yes	Yes
OEM				Yes	Yes	Yes
Observations	983	983	983	983	983	983
Adjusted Pseudo R^2	0.11	0.11	0.13	0.19	0.20	0.23

All models estimated with nonlinear least squares. To account for endogeneity in turbine and firm experience, all models include control function using the approach of Newey et al. (1999). Robust standard errors presented in parentheses.

Table A.3: Static Learning Regressions - Size Heterogeneity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Base Cost (ω_0)	15.6	41.4	57.6	18.3	60.7	77.3
	(1.2)	(4.5)	(7.1)	(1.8)	(12.8)	(17.3)
Learning Exponent (α)	-0.15	-0.39	-0.44	-0.19	-0.37	-0.42
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
IndustryExpc (β_W)		0.15	0.24		0.57	0.63
		(0.03)	(0.07)		(0.33)	(0.34)
Turbine Expc (β_J)			63.8			119.2
			(17.8)			(61.3)
Fixed-effects						
Year	Yes	Yes	Yes	Yes	Yes	Yes
OEM				Yes	Yes	Yes
Observations	983	983	983	983	983	983
Adjusted Pseudo R^2	0.15	0.17	0.20	0.25	0.27	0.30

All models estimated with nonlinear least squares. To account for endogeneity in turbine and firm experience, all models include control function using the approach of Newey et al. (1999). Robust standard errors presented in parentheses.

B.2 Additional Results: Markdown Correction

Table A.4: Dynamic Markdown Controls - Year Heterogeneity

	(1)	(2)	(3)	(4)	(5)
Base Cost (ω_0)	87.24	115.45	134.63	99.33	93.11
	(21.18)	(40.25)	(37.31)	(31.54)	(27.50)
Learning Exponent (α)	-0.54	-0.48	-0.56	-0.67	-0.68
	(0.06)	(0.05)	(0.04)	(0.10)	(0.10)
Industry Expc (β_W)	0.06	0.19	0.19	0.04	0.03
	(0.04)	(0.16)	(0.11)	(0.03)	(0.02)
Turbine Expc (β_J)	82.50	490.69	198.14	45.41	31.43
	(42.58)	(320.97)	(98.16)	(29.24)	(20.10)
$\mu \times 10^4$			-1.37		-3.08
			(0.39)		(0.77)
N	983	983	983	505	505
FE	Year	Firm x Year	Year	Year	Year
Dynamics	None	None	BP (Model)	None	BP (Accounting)
Sample	All	All	All	Accounting	Accounting

ω_0 is the "initial" cost of a turbine per unit of materials. α is the Irwin & Klenow exponent. μ is the coefficient on the dynamic markdown term.

Table A.5: Dynamic Markdown Controls - Size Heterogeneity

	(1)	(2)	(3)	(4)	(5)
Base Cost (ω_0)	80.12	107.34	135.55	103.75	101.56
	(12.00)	(20.32)	(24.21)	(22.21)	(21.16)
Learning Exponent (α)	-0.41	-0.44	-0.47	-0.56	-0.56
	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)
Industry Expc (β_W)	0.12	0.26	0.35	0.10	0.09
	(0.04)	(0.13)	(0.13)	(0.04)	(0.04)
Turbine Expc (β_J)	26.56	89.78	87.09	30.67	26.66
	(8.98)	(37.87)	(30.28)	(12.80)	(10.98)
$\mu \times 10^4$			-1.64		-1.50
			(0.33)		(0.59)
N	983	983	983	505	505
FE	Year	Firm x Year	Year	Year	Year
Dynamics	None	None	BP (Model)	None	BP (Accounting)
Sample	All	All	All	Accounting	Accounting

ω_0 is the "initial" cost of a turbine per unit of materials. α is the Irwin & Klenow exponent. μ is the coefficient on the dynamic markdown term.