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

Thomas R. Covert †
Richard L. Sweeney ‡

December 2021

Abstract

We study the extent to which learning and product innovation have reduced costs in the wind industry. As in many settings, a major empirical challenge is the fact that costs are not observed. To overcome this, we leverage a simple but physically realistic model of how observable wind turbine characteristics, like rotor size, relate to power production and manufacturing material needs. We then infer latent costs from wind farm developer and manufacturer preferences. To account for the dynamic aspects of this problem, we introduce a new control that allows us to control for dynamic pricing incentives without estimating or computing a dynamic game. We find that a 10 percent increase in a product's sales causes a 9 - 11 percent reduction in the costs to manufacture it. Sales of other products owned by the same firm generate one tenth the learning benefits, while sales to other firms generate small but nonzero cost reductions. 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 manufacturing costs are generally not observed. In lieu of this, recent empirical literature has made progress by focussing on settings where detailed data on production *inputs* is available.[1] While analyses like this can, in principal, measure the true underlying relationship between manufacturing costs and experience, this kind detailed data is rarely available. 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 when unit level costs or inputs are not available.

This paper presents a new method for measuring learning by doing using only 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. Although our method explicitly acknowledges and controls for these dynamic incentives, it does not require their explicit computation, unlike workhorse dynamic game estimators.

We use this method to estimate the evolution of 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 early growth was fueled by generous subsidies for wind farm project developers, with the explicit goal of instigating future cost declines in the wind turbine manufacturing sector. Existing research has quantified the *static* benefits of these policies, and concluded that they are smaller than their associated costs (Van Benthem et al., 2008; Abrell et al., 2019; Greenstone and Nath, 2020). Thus, whether or not renewable energy policy to date was net beneficial depends on whether or not the wind turbine manufacturing industry exhibits significant learning economies.

We measure the evolution of manufacturing costs using data on the global wind turbine industry which covers the near-universe of wind plants constructed between 2000-2019, the

[1] Benkard (2000) estimates learning and forgetting in the commercial aircraft industry, where he observes the labor inputs into each aircraft produced. Thornton and Thompson (2001) observe the manhours into war ships, and Levitt et al. (2013) observe manhours and output by shift at a major automobile assembly plant.

turbine models those plants chose to install, and engineering estimates of the output each turbine model would generate at each possible plant location. 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. 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 a standard discrete choice model, 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 current period sales. Normally, models like this in the structural industrial organization literature assume that the continuation value satisfies a Bellman equation, 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 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 first order behavior of the dynamic portion of their incentives can be expressed by a rational expectations term. In their main application, this rational expectations term is an explicit function of realized cashflows, or a model thereof. While some of the firms in our setting are publicly traded, and as such do have public accounting data to which we could apply this idea, our methodological innovation is to recognize that the rational expectations term in these Berry and Pakes (2000) expressions is common across all continuous controls a firm sets. Thus, if we have multi-product firms, we can "control for" these dynamic incentives as fixed-effects without explicitly computing or even completely modeling them.

Using this framework, we document considerable learning by doing in wind turbine manufacturing. We find that a 1 percent increase in cumulative sales is associated with an approximately 0.9 percent reduction in turbine costs next period. When we account for potential endogeneity in cumulative sales using an instrumental variable approach in the spirit

of Bartik (1991), we find that this 1.1 percent. We also estimate economically meaningful spillovers across turbines owned by the same firm: every sale of other turbines owned by the same firm is equivalent to a 0.07 to 0.11 increase in sales of the turbine in question. Finally, we look for discrete jumps in the cost function following the introduction a new innovative turbine design.

This paper 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. 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 agenda.

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.

The rest of the paper proceeds as follows. In Section 2, we describe the basics of wind turbine technology and the industry structure. Section 3 discusses the data we use and the filtering criteria we apply to it. Section 4 discusses our approach to recover static costs, and section 6 relates their evolution to firm sales experience. Section 7 discusses policy implications and concludes.

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:

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

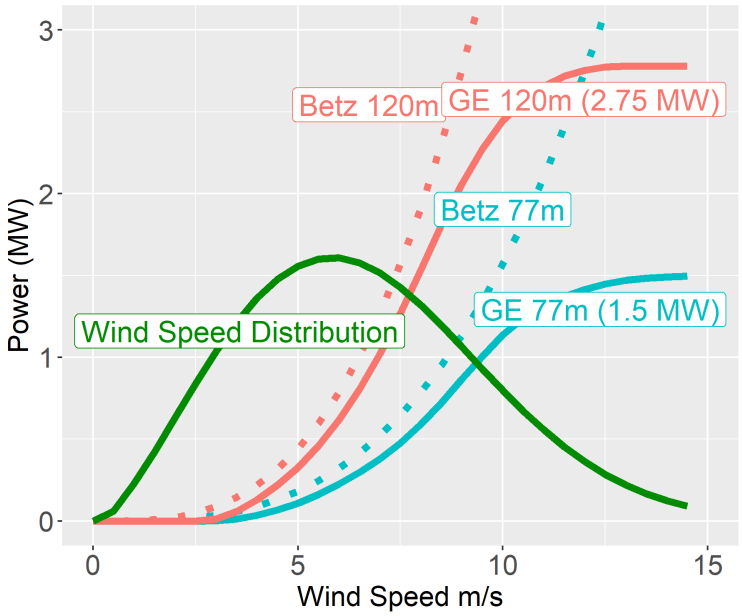
where C_p is the "power coefficient", or the ratio of the power flowing through the device that is captured, and d is the density of the air the turbine is exposed to. Betz (1926) demonstrated that the theoretical limit on C_p is $C_p^{\max} = \frac{16}{27} \approx 0.593$.

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

Figure 1: Example Power Curves

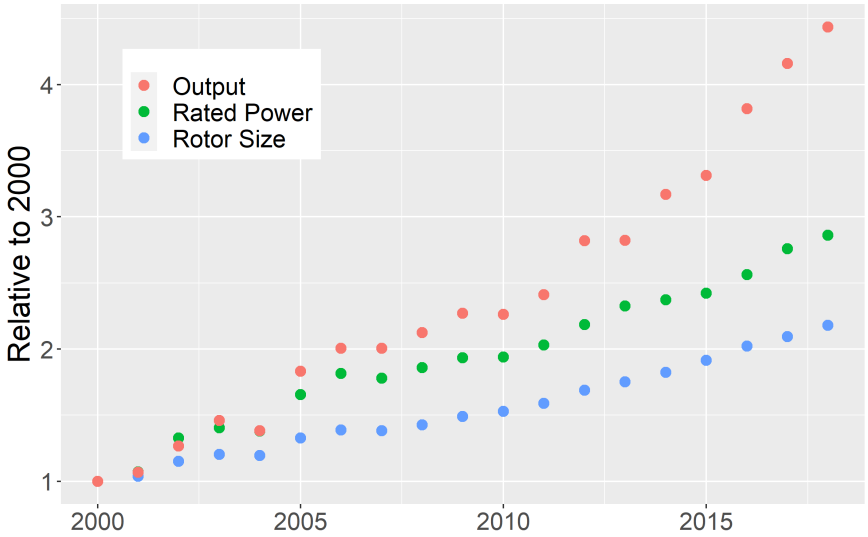


nearly fourfold (Figure 2).

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 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.

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.

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

Figure 3: Turbine Mass vs Size (NREL)

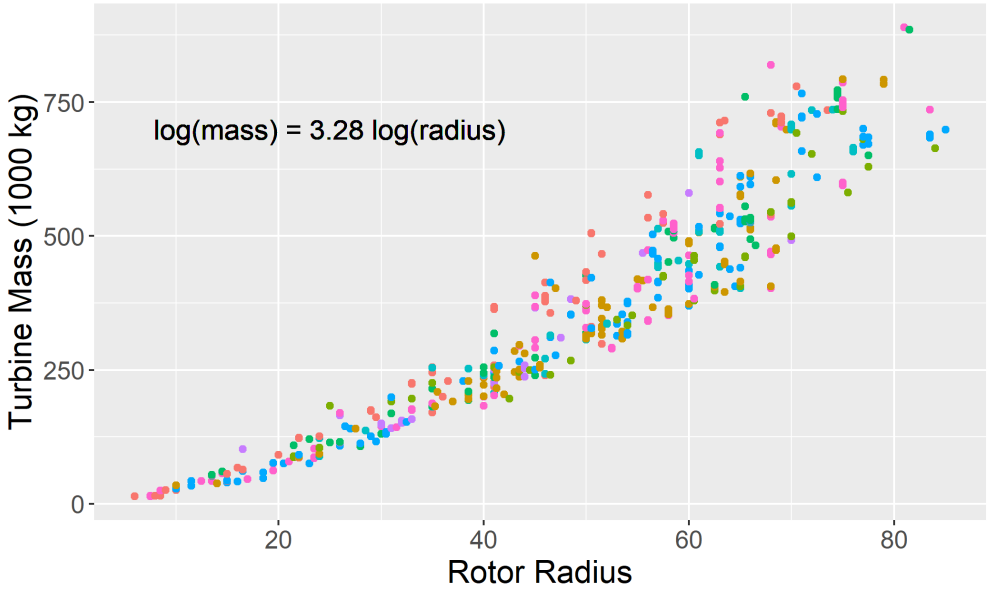


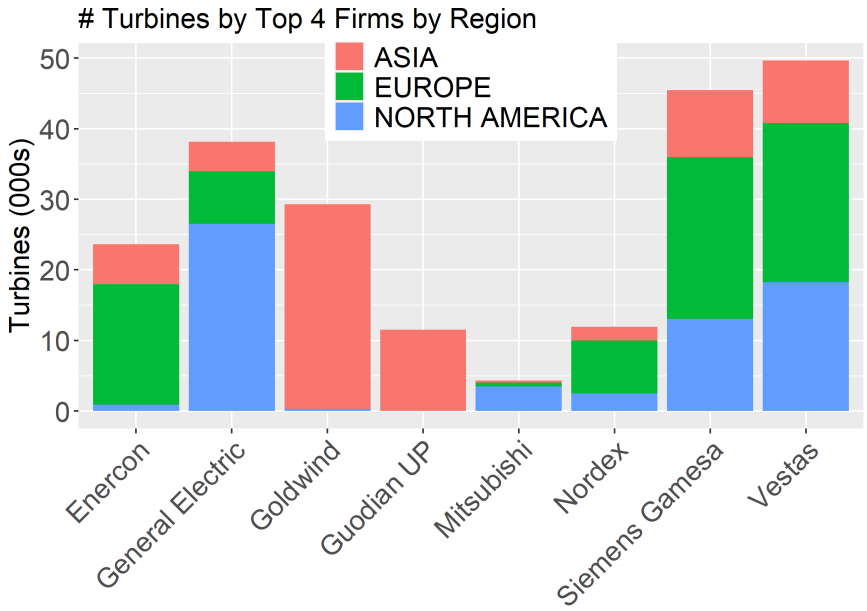
Table 1: Sales by Firm

Manufacturer	Turbines	Capacity (MW)	Wind Farms
Vestas	47,324	85,726	3,585
Siemens Gamesa	44,611	72,549	2,126
General Electric	37,054	64,142	1,471
Goldwind	26,500	42,143	702
Enercon	22,093	37,968	3,233
Nordex	11,084	23,484	1,121
Guodian	9,625	15,750	292
Sinovel	9,098	14,432	225
Suzlon	8,370	14,537	992
Dongfang	7,268	11,962	199
Other	58,847	95,055	3,517

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 4: Sales by Region



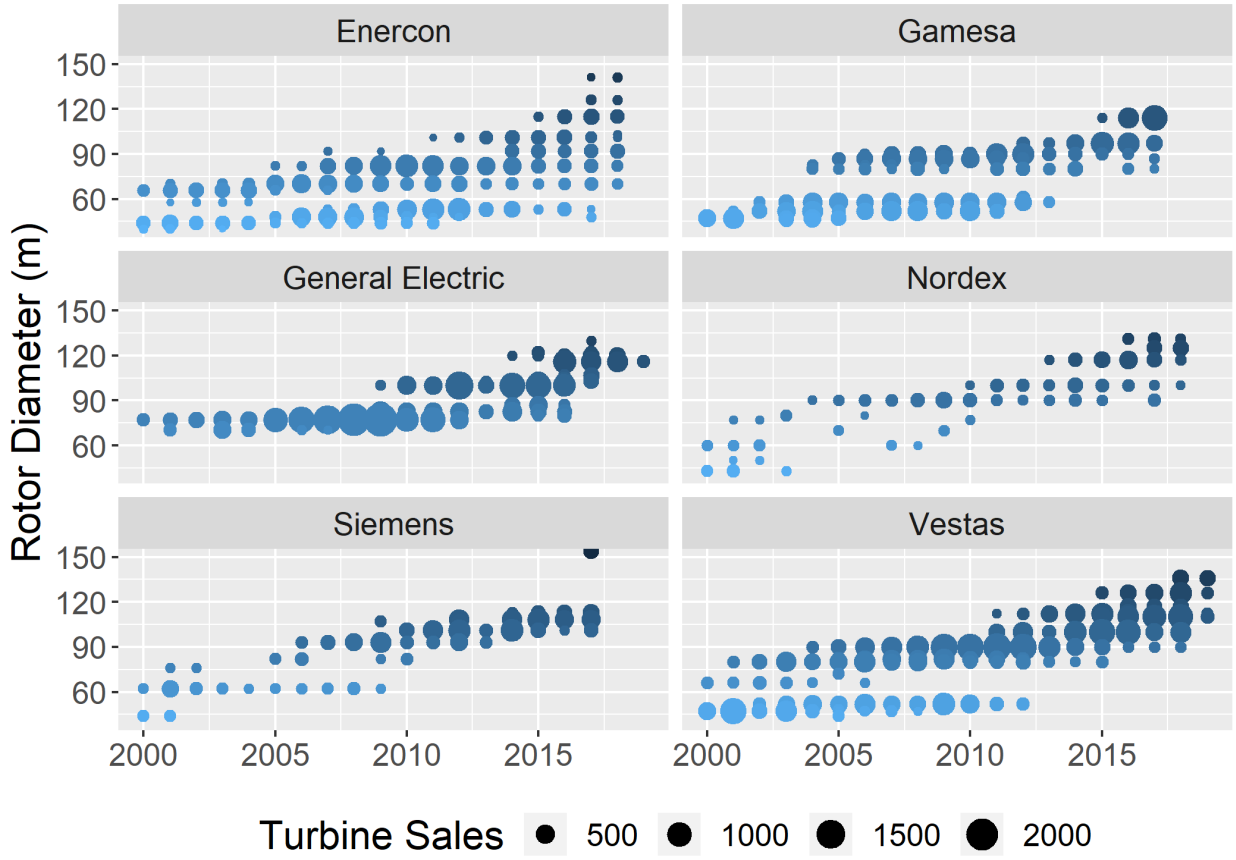
5 also shows that when new, larger turbines are introduced, they do not immediately take over all of the market share from existing smaller turbines, despite being more productive. Instead, these larger turbines gradually gain market share over time, which provides suggestive evidence that the relative 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 5: Sales by Turbine Size, Top Western Firms



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.

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	18,680	0.84
Onshore	18,483	0.83

Source: BNEF Project Database

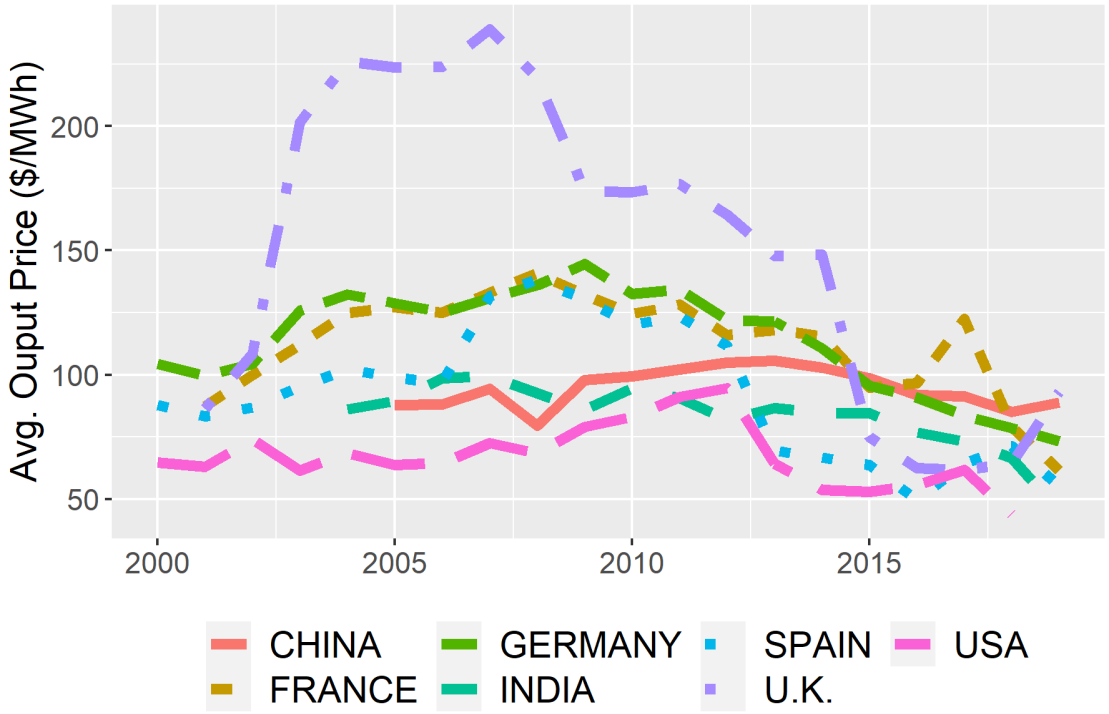
Table 3: Sample Summary

Variable	mean	sd	max	median	min
Capacity (MW)	27.36	40.00	600.00	12.30	1.00
# Turbines	16.14	23.83	460.00	7.00	1.00
Rated Power (MW)	1.91	0.79	9.50	2.00	0.05
Wind speed (m/s)	6.65	1.00	15.86	6.55	2.36
Rotor Radius (M)	42.05	11.08	83.50	41.00	7.50
	644.86	379.56	4,419.43	593.36	5.80

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 procurred in a confidential, developer specific process. Once completed, the terms agreed upon for specific deals are rarely disclosed. Thus, before estimating manufacturing costs using the inversion of a demand system, we must recover the prices they set. To do this, we specify and estimate a model of turbine procurement that allows us to infer what prices manufacturers bid their turbines at each year.

Figure 6: Average Wind Output Price by Country



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.

These revenues 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

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

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, 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. We make the assumption that manufacturing costs are constant within a given year, so that if two projects i and i' are built in the same year, $b_{ijt} = b_{i'jt}$. 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

[2] This means we are ruling out situations where a developer is choosing between 3 1 MW turbines or 2 1.5 MW turbines, and plans to keep output fixed.

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. 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 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 margins 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:

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

The variable R_{ijt} represents all wind revenues, while the variable R^H_{ijt} represents the subset of those revenues which accrue at high wind speeds.[3] 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 output. The vector of parameters $\{\alpha_c\}$ represent the country-specific marginal utility for the subset of those revenues which are produced at especially high wind 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 wind speed distribution sufficiently 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.

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

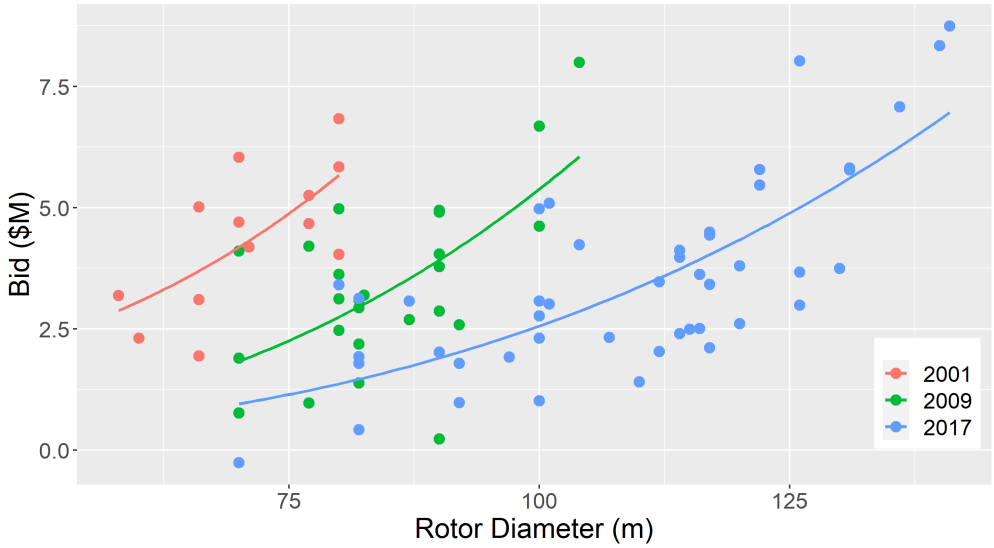
[3]We define "high" wind speeds as those greater than 7.5 m/s. We allow developers to have distinct preferences over wind production revenues that accrue at high wind speeds for several reasons. First, episodes of high wind speeds often coincide with electricity grid transmission constraints, meaning 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.

choice, we estimate the b's, the α's and the vector β^W using maximum likelihood.

4.2 Bid estimation results

Here we provide some graphical evidence on the estimated turbine bids, which are taken as data in the second stage of our estimation routine. 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 (spaning 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

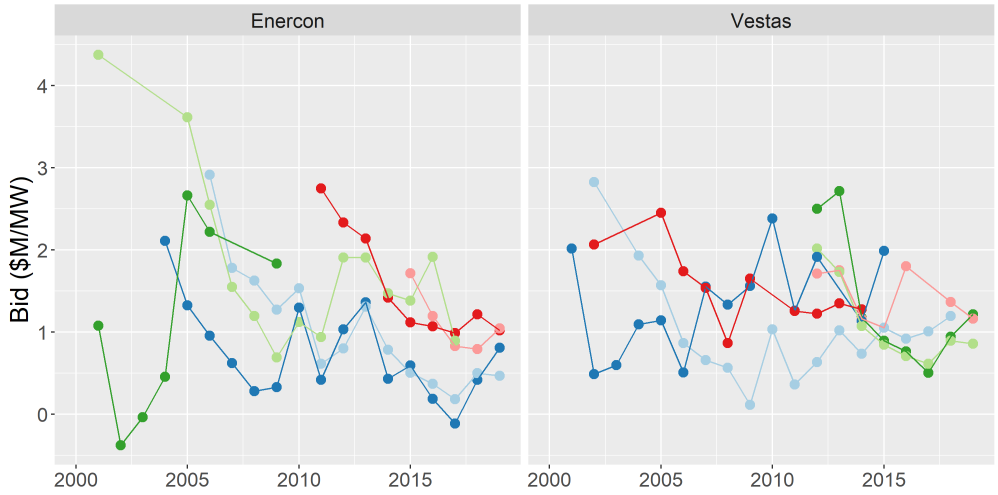


Estimated bids are recovered relative to a reference turbine, which we pick to be the most common Vestas turbine sold that year. To plot levels, we impute the implied cost of this turbine from Vestas financials.

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

5.1 A feasible control for markdowns

Although our assumptions about the turbine procurement mechanism mean that bids are equivalent to opportunity costs, 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. To learn about the effects of past experience on manufacturing costs, we need a method for separating these two sources of opportunity cost. Our approach is to write down a model of dynamically optimizing turbine manufacturer behavior, use its first-order optimal pricing behavior to characterize what these dynamic learning benefits must be, and show how we can control for their presence in estimation without explicitly computing them.

We assume that costs may reflect past manufacturing experience, either directly experienced by the firm, or through other firm's experiences in a spillover mechanism. To allow for this, we define a state variable x as the cumulative vector of sales experience for all turbines: x_{jt} is the count of all past sales of turbine j turbine before period t. This state variable is common to all firms in the industry.

During period t, a set of wind farm projects $i \in W_t$ solicit turbine bids from the market, 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, $x_{t+1} = x_t + q_t$. The probability distribution over q_t, which depends on the set of available projects W_t and the bids firms set b_t, is $dF(q_t \mid W_t, b_t)$.

Firms choose turbine bids 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}. 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 have made 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, and as we'll describe below, has a relatively standard representation, thanks to our discrete choice setup. However, the second component is less standard, so we explore it first.

To characterize the marginal effect of higher prices on future benefits, 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 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 we borrow from Berry and Pakes (2000).

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

$$\mathbb{E}\left[V_{ft}(x_t + q_t) \frac{\frac{\partial}{\partial b_{lt}} dF(q_t \mid W_t, b_t)}{dF(q_t \mid W_t, b_t)} \mid W_t, b_t \right] = V_{ft}(x_t + q_t^*) \left(\frac{\frac{\partial}{\partial b_{lt}} dF(q_t^* \mid W_t, b_t)}{dF(q_t^* \mid W_t, b_t)} \right) + \nu_{lt}$$

where q_t^* is the realized turbine sales vector and $\mathbb{E}\left[\nu_{lt} \mid x_t, b_t\right] = 0$. Aside from the rational expectations error, the one unobserved term in this expression, $V_{ft}(x_t + q_t^*)$, does not depend on l, the turbine in question. This means that if firms are setting bids for more than one turbine, the first order conditions for these bids will have a *common* unobservable term, one that we can control for as the firm-by-time slope on the observable term $\left(\frac{\frac{\partial}{\partial b_{jt}} dF(q_t^* \mid W_t, b_t)}{dF(q_t^* \mid W_t, b_t)} \right)$. Thus, our estimation strategy requires the presence of multi-product firms.

To finish our derivation of the dynamic portion 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 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 finalize the expression for the dynamic portion of the

firm's FOC:

$$\frac{\partial}{\partial b_{lt}} \int V_{ft}(x_t + q_t) dF(q_t \mid W_t, b_t) = V_{ft}(x_t + q_t^*) \left(\sum_{i \in W_t} s_{ilt} - N_l(q_t, W_t, b_t) \right) + \nu_{lt}$$

As noted above, both components of the expression in the parentheses can be computed from the demand system estimates. Thus, as long as a firm offers two or more turbines in a given year, we can estimate the dynamic portion of the firm's FOC as a firm-by-time specific slope on an observable quantity.

To derive the static portion of the firm's FOC, note that 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 third equality shows that the derivative of the logit inclusive value with respect to the bid of a given turbine is simply the negative of the probability that the plant picks that turbine. It also shows that the inclusive value, excluding turbines from firm f, is insensitive to the bids for f's turbines. Finally, it includes the two "standard" terms in first order conditions for discrete choice demand models: the elasticity-weighted sum of changes in margins, plus the gains from infra-marginal buyers. This final expression above includes the first of theses term, but not the second, because second price payment rules mean that the firm does not capture infra-marginal benefits from higher prices.

Putting the pieces together, we can write the firm's FOC for the bid for turbine l at time t as:

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

We can collect these terms to describe a feasible estimation equation. Let b_{ft} and c_{ft} represent vectors of the bids and marginal costs 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 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 firm f's bids in terms of its costs, its dynamic incentives, and a rational expectations shock in period t:

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

At this point, everything in the above expression is either observed, possibly after demand estimation (bids, demand elasticities, choice probabilities, and the N term), a set of parameters (costs and the V_{ft}'s) or a rational expectations shock.

5.2 Cost function

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. Consistent with this, 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 discussed in section 2, material inputs for manufacturing a wind turbine grow approximately cubically in rotor size r.

Firms with more experience may be able to economize on unobservable inputs, like labor, wasted materials, etc. We assume that this learning reduces firms' cost per unit turbine volume, such that the marginal cost of manufacturing turbine j at time t is

$$c_{jt} = \delta\left(\widetilde{E}_{jt}\right)^{\alpha} r_j^3$$

where "effective" experience $\widetilde{E}_{jt}$ is given by

$$\widetilde{E}_{jt} = E_{j,t} + \beta(E_{ft} - E_{jt}) + \gamma(E_t - E_{f(j),t})$$

and is analogous to the similar term in Irwin and Klenow (1994).

Effective experience is a linear combination of a firm's experience manufacturing a specific turbine (E_{jt}), the firm's additional experience manufacturing other turbines ($E_{f(j),t} - E_{jt}$),

and the additional experience that all other firms in the industry have thus far accumulated $(E_t - E_{f(j),t})$.[4] We measure experience as the number of *turbines* that have previously shipped, so these experience terms can be constructed from the state vector x_t. The parameter β thus measures the extent to which the firm's experience manufacturing other turbines is as useful in manufacturing turbine j as turbine j specific experience. Similarly, the parameter γ measures the extent to which other firm's aggregate effective experience is as useful as j-specific experience. That is, a unit of own-firm experience is β times as valuable as turbine-specific experience, and a unit of industry experience is γ times as valuable as turbine-specific experience. Finally, the parameter α represents learning economies. Holding everything else equal, a 1% increase in effective experience should increase 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.

5.3 Estimation

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 estimation equation should be:

$$b_{jt} = \delta\left(E_{j,t} + \beta(E_{f(j),t} - E_{j,t}) + \gamma(E_t - E_{f(j),t})\right)^{\alpha} r_j^3 - V_{ft}(x_t + q_t^*)\xi_{jt} - \widetilde{\nu}_{jt}$$

However, we do not actually observe or estimate the 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. Thus, a feasible version of the estimating equation is:

$$\widehat{b}_{jt} = w_{jt}\delta\left(E_{j,t} + \beta(E_{f(j),t} - E_{j,t}) + \gamma(E_t - E_{f(j),t})\right)^{\alpha} - V_{ft}(x_t + q_t^*)\xi_{jt} - b_{0t} - \widetilde{\nu}_{jt} \qquad (5)$$

We estimate this using non-linear least squares, with year fixed effects (to capture b_{0t}) and firm-by-time slopes on ξ_{jt} (to capture $V_{ft}(x_t + q_t^*)$).

6 Results

Table 4 presents the results from estimating equation 5. In the first column, we begin with an estimate that excludes our dynamic learning benefit term, $V_{ft}\xi_{jt}$. This specification captures what a "standard" inversion of demand estimates would reveal the relationship between costs and experience to be, since it does not include any dynamic controls. Under this model, we find that costs are negatively correlated with experience, to the point that a 1% increase in

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

effective experience decreases marginal costs by about 0.92% (α in equation 5). The estimate of β is positive, and precisely estimated, suggesting that there are within-firm spillovers of experience. An additional sale of a firm's turbines, other than turbine j, reduces costs by 8% as much as an additional sale of turbine j itself. Though the point estimate on γ is positive, it is not precisely estimated at conventional levels.

Table 4: Learning Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
beta	0.08	0.06	0.03	0.11	0.07	0.10	0.05
	(0.04)	(0.03)	(0.04)	(0.05)	(0.03)	(0.06)	(0.03)
gamma	0.04	0.04	0.03	0.05	0.02	0.07	0.01
	(0.03)	(0.02)	(0.02)	(0.05)	(0.02)	(0.06)	(0.02)
alpha	-0.92	-0.86	-0.64	-0.94	-0.48	-1.10	-0.37
	(0.19)	(0.18)	(0.16)	(0.38)	(0.11)	(0.53)	(0.13)
MW	0.23	0.93	1.00	1.03	0.97	1.03	0.98
	(0.18)	(0.29)	(0.33)	(0.40)	(0.43)	(0.40)	(0.42)
Innovations					-3.41		-3.83
					(2.23)		(2.51)
BP Control				X	X	X	X
IV Control						X	X
Observations	655	655	655	655	655	655	655
Year fixed effects	✓	✓		✓	✓	✓	✓
Firm fixed effects		✓		✓	✓	✓	✓
Year×Firm fixed effects			✓				

Standard errors clustered by year.

In the next specification, we include firm-fixed effects as well, to account for the possibility that some firms may have persistently lower or higher costs than others for reasons unrelated to experience. These results are qualitatively similar to those in column 1. In column 3, we replace the individual firm and time fixed effects with a combined firm-by-time fixed effect. In addition to subsuming b_{0t}, the bid on the base turbine, it 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 much simpler to construct than our dynamic correction term. Compared to model 2, model 3 suggests that both α and within-firm spillovers are smaller.

In column 4, we revisit the firm panel regression setup from model 2, and include our formal dynamic correction term (referred to in the table as the "BP control"). Comparing columns 3 and 4 shows that a firm-by-time fixed effect does not provide qualitatively similar

estimates. Compared with model 2, the learning parameter α is slightly higher but statistically similar. However the estimated within firm-spillovers are nearly twice as large. In sum, these results confirm that experience is associated with lower costs (as α is generally large in magnitude and negative), and that there are meaningful within-firm experience spillovers (as β is positive and precisely estimated).

In column 5 we consider the fact that, when firms develop new, innovative turbines, they may generate knowledge which spills over to other products. We define the number of innovations present at the time a firm releases turbine j as the cumulative number of previous turbine size increases observed at that firm in the data (seen for select firms in Figure 5). Though the estimates on this term are noisy, they are large and negative, consistent with innovation also reducing manufacturing costs in this industry. The inclusion of this control reduces the estimated firm and industry spillover terms, suggesting that part of what the previous estimates were capturing might have been more about innovation than learning. Compared to column 4 the estimated learning parameter is nearly half as large. This translates into a decline in Spence's learning coeffcient from 0.48 to 0.28.

The specifications we've considered so far explicitly assume that cumulative experience at the start of time t is uncorrelated with other unmeasured determinants of turbine bids in time t. However, it is easy to imagine how this could be violated. For example, if a firm experiences a favorable cost shock in time t, it will sell more turbines. If that cost shock persists, it will also sell more in time $t+1$, which our esitmator would infer was *caused* by the increased experience gained in time t.

To deal with these potential sources of endogeneity, we leverage strong home market bias and large panel variation in the aggregate level of demand to construct "predicted" turbine sales each year and use these to construct "predicted" experience, which we can use as instrumental variables for our experience measures. 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.

Columns 6 and 7 supplement the specifications in columns 4 and 5 with a control function meant to deal with the endogeneity of experience, using the approach of Newey et al. (1999). To construct these control functions, we first regress our observed experience covariates onto the full vector of the instruments described above, as well as all of the exogenous terms in equation 5. We then include the residuals of these first stage regressions as covariates in the nonlinear least squares procedure. While the spillover parameters in models 4 and 6 similar, α is economical larger in the model with the control function (although the confidence intervals are wide). Turning to the model which includes rotor innovations, α is significantly

lower after the inclusion of the IV control.

7 Conclusion

We estimate the extent to which learning by doing and product innovation have reduced costs in the wind industry. To overcome the common empirical barrier that cost and inputs are not observed, we infer latent manufacturing costs from a structural model of turbine procurement. To separate the bids in this model from dynamic benefits inherent in any environment where learning by doing is present, we introduce a new correction term which allows researchers to control for dynamic pricing incentives without estimating or computing a dynamic game. We estimate this using rich data from the universe of commercial wind turbines pruchases worldwide between 2000 and 2019, and find evidence of significant learning both within and across turbine models. This is 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.

Bartik, Timothy J, "Who Benefits from State and Local Economic Development Policies?," *Books from Upjohn Press*, 1991.

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.

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.

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 vectoris 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 (**I think??**) 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.[5] 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.[6] We ask the solver to find the $L = 200$ "best" feasible solutions, where solution quality is the log-likelihood of the observed allocation, or $\sum_i \sum_j \mu_{ij} \log s_{ij}$.[7] Let $\widehat{M}(q, W)$ be our approximate set of solutions for realized sales q under market W. Then

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

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

[7] 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 $\cup_f K_{ft}$.