Technological Forecasting & Social Change xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

How well do experience curves predict technological progress? A method for making distributional forecasts

François Lafond[a,b,c,*], Aimee Gotway Bailey[d], Jan David Bakker[e], Dylan Rebois[b], Rubina Zadourian[f,a,g], Patrick McSharry[b,h,i,j], J. Doyne Farmer[a,f,k,l]

[a] Institute for New Economic Thinking, Oxford Martin School, UK
[b] Smith School for Enterprise and the Environment, University of Oxford, UK
[c] London Institute for Mathematical Sciences, UK
[d] U.S. Department of Energy, USA
[e] Department of Economics, University of Oxford, UK
[f] Mathematical Institute, University of Oxford, UK
[g] Max Planck Institute for the Physics of Complex Systems, Germany
[h] Carnegie Mellon University Africa, Rwanda
[i] African Center of Excellence in Data Science, University of Rwanda, Rwanda
[j] Oxford Man Institute of Quantitative Finance, University of Oxford, UK
[k] Computer Science Department, University of Oxford, UK
[l] Santa Fe Institute, USA

ARTICLE INFO

JEL classification:
C53
O30
Q47

Keywords:
Forecasting
Technological progress
Experience curves

ABSTRACT

Experience curves are widely used to predict the cost benefits of increasing the deployment of a technology. But how good are such forecasts? Can one predict their accuracy a priori? In this paper we answer these questions by developing a method to make distributional forecasts for experience curves. We test our method using a dataset with proxies for cost and experience for 51 products and technologies and show that it works reasonably well. The framework that we develop helps clarify why the experience curve method often gives similar results to simply assuming that costs decrease exponentially. To illustrate our method we make a distributional forecast for prices of solar photovoltaic modules.

1. Introduction

Since Wright's (1936) study of airplanes, it has been observed that for many products and technologies the unit cost of production tends to decrease by a constant factor every time cumulative production doubles (Thompson, 2012). This relationship, also called the experience or learning curve, has been studied in many domains.[1] It is often argued that it can be useful for forecasting and planning the deployment of a particular technology (Ayres, 1969; Sahal, 1979; Martino, 1993). However in practice experience curves are typically used to make point forecasts, neglecting prediction uncertainty. Our central result in this paper is a method for making *distributional forecasts*, that explicitly take prediction uncertainty into account. We use historical data to test this and demonstrate that the method works reasonably well.

Forecasts with experience curves are usually made by regressing historical costs on cumulative production. In this paper we recast the experience curve as a time series model expressed in first-differences: the change in costs is determined by the change in experience. We derive a formula for how the accuracy of prediction varies as a function of the time horizon for the forecast, the number of data points the forecast is based on, and the volatility of the time series. We are thus able to make distributional rather than point forecasts. Our approach builds on earlier work by Farmer and Lafond (2016) that showed how to do this for univariate forecasting based on a generalization of Moore's law (the autocorrelated geometric random walk with drift). Here we apply our new method based on experience curves to solar photovoltaic modules (PV) and compare to the univariate model.

Other than Farmer and Lafond (2016), the two closest papers to our contribution here are Alberth (2008) and Nagy et al. (2013). Both papers tested the forecast accuracy of the experience curve model, and performed

* Corresponding author.
E-mail addresses: francois.lafond@inet.ox.ac.uk (F. Lafond), doyne.farmer@inet.ox.ac.uk (J.D. Farmer).

[1] See Anzanello and Fogliatto (2011), Dutton and Thomas (1984), Yelle (1979) and for energy technologies Candelise et al. (2013), Isoard and Soria (2001), Junginger et al. (2010), Kahouli-Brahmi (2009), Neij (1997), Nemet (2006).

https://doi.org/10.1016/j.techfore.2017.11.001
Received 17 March 2017; Received in revised form 28 October 2017; Accepted 1 November 2017

comparisons with the time trend model. Alberth (2008) performed forecast evaluation by keeping some of the available data for comparing forecasts with actual realized values.[2] Here, we build on the methodology developed by Nagy et al. (2013) and Farmer and Lafond (2016), which consists of performing systematic hindcasting. That is, we use an estimation window of a constant (small) size and perform as many forecasts as possible. As in Alberth (2008) and Nagy et al. (2013), we use several datasets and we pool forecast errors to construct a distribution of forecast errors. We think that out-of-sample forecasts are indeed good tests for models that aim at predicting technological progress. However, when a forecast error is observed, it is generally not clear whether it is "large" or "small", from a statistical point of view. And it is not clear that it makes sense to aggregate forecast errors from technologies that are more or less volatile and have high or low learning rates.

A distinctive feature of our work is that we actually calculate the expected forecast errors. As in Farmer and Lafond (2016), we derive an approximate formula for the theoretical variance of the forecast errors, so that forecast errors from different technologies can be normalized, and thus aggregated in a theoretically grounded way. As a result, we can check whether our empirical forecast errors are in line with the model. We show how in our model forecast errors depend on future random shocks, but also parameter uncertainty, as is only seldomly acknowledged in the literature (for exceptions, see Vigil and Sarper, 1994 and Van Sark, 2008).

Alberth (2008) and Nagy et al. (2013) compared the forecasts from the experience curve, which we call Wright's law, with those from a simple univariate time series model of exponential progress, which we call Moore's law. While Alberth (2008) found that the experience curve model was vastly superior to an exogenous time trend, our results (and method and dataset) are closer to the findings of Nagy et al. (2013): univariate and experience curves models tend to perform similarly, due to the fact that for many products cumulative experience grows exponentially. When this is the case, we cannot expect experience curves to perform much better than an exogenous time trend unless cumulative experience is very volatile, as we explain in detail here.

We should emphasize that this comparison is difficult because the forecasts are conditioned on different variables: Moore's law is conditioned on time, while Wright's law is conditioned on experience. Which of these is more useful depends on the context. As we demonstrate, Moore's law is more convenient and just as good for business as usual forecasts for a given time in the future. However, providing there is a causal relationship from experience to cost, Wright's law makes it possible to forecast for policy purposes (Way et al., 2017).

Finally, we depart from Alberth (2008), Nagy et al. (2013) and most of the literature by using a different statistical model. As we explain in the next section, we have chosen to estimate a model in which the variables are first-differenced, instead of kept in levels as is usually done. From a theoretical point of view, we believe that it is reasonable to think that the stationary relationship is between the increase of experience and technological progress, instead of between a level of experience and a level of technology. In addition, we will also introduce a moving average noise, as in Farmer and Lafond (2016). This is meant to capture some of the complex autocorrelation patterns present in the data in a parsimonious way, and increase theoretical forecast errors so that they match the empirical forecast errors.

Our focus is on understanding the forecast errors from a simple experience curve model[3]. The experience curve, like any model, is only an approximation. Its simplicity is both a virtue and a detriment. The virtue is that the model is so simple that its parameters can usually be estimated well enough to have predictive value based on the short data sets that are typically available[4]. The detriment is that such a simple model neglects many effects that are likely to be important. A large literature starting with Arrow (1962) has convincingly argued that learning-by-doing occurs during the production (or investment) process, leading to decreasing unit costs. But innovation is a complex process relying on a variety of interacting factors such as economies of scale, input prices, R&D and patents, knowledge depreciation effects, or other effects captured by exogenous time trends.[5] For instance, Candelise et al. (2013) argue that there is a lot of variation around the experience curve trend in solar PV, due to a number of unmodelled mechanisms linked to industrial dynamics and international trade, and Sinclair et al. (2000) argued that the relationship between costs and experience is due to experience driving expectations of future production and thus incentives to invest in R&D. Besides, some have argued that simple exponential time trends are more reliable than experience curves. For instance Funk and Magee (2015) noted that significant technological improvements can take place even though production experience did not really accumulate, and Magee et al. (2016) found that in domains where experience (measured as annual patent output) did not grow exponentially, costs still had an exponentially decreasing pattern, breaking down the experience curve. Finally, another important aspect that we do not address is reverse causality (Kahouli-Brahmi, 2009; Nordhaus, 2014; Witajewski-Baltvilks et al., 2015): if demand is elastic, a decrease in price should lead to an increase in production. Here we have intentionally focused on the simplest case in order to develop the method.

2. Empirical framework

2.1. The basic model

Experience curves postulate that unit costs decrease by a constant factor for every doubling of cumulative production[6]. This implies a linear relationship between the log of the cost, which we denote y, and the log of cumulative production which we denote x:

$$y_t = y_0 + \omega x_t. \tag{1}$$

This relationship has also often been called "the learning curve" or the experience curve. We will often call it "Wright's law" in reference to Wright's original study, and to express our agnostic view regarding the causal mechanism. Generally, experience curves are estimated as

$$y_t = y_0 + \omega x_t + \iota_t, \tag{2}$$

where ι_t is i.i.d. noise. However, it has sometimes been noticed that residuals may be autocorrelated. For instance Womer and Patterson (1983) noticed that autocorrelation "seems to be an important problem" and Lieberman (1984) "corrected" for autocorrelation using the Cochrane-Orcutt procedure.[7] Bailey et al. (2012) proposed to estimate Eq. (1) in first difference

$$y_t - y_{t-1} = \omega(x_t - x_{t-1}) + \eta_t, \tag{3}$$

where η_t are i.i.d errors $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$. In Eq. (3), noise accumulates so that in the long run the variables in level can deviate significantly from the deterministic relationship. To see this, note that (assuming $x_0 = \log(1) = 0$) Eq. (3) can be rewritten as

[2] Alberth (2008) produced forecasts for a number (1,2, … 6) of doublings of cumulative production. Here instead we use time series methods so it is more natural to compute everything in terms of calendar forecast horizon.

[3] We limit ourselves to showing that the forecast errors are compatible with our model being correct, and we do not try to show that they could be compatible with the experience curve model being spurious.

[4] For short data sets such as most of those used here, fitting more than one parameter often results in degradation in out-of-sample performance (Nagy et al., 2013).

[5] For examples of papers discussing these effects within the experience curves framework, see Argote et al. (1990), Berndt (1991), Isoard and Soria (2001), Papineau (2006), Söderholm and Sundqvist (2007), Jamasb (2007), Kahouli-Brahmi (2009), Bettencourt et al. (2013), Benson and Magee (2015) and Nordhaus (2014).

[6] For other parametric models relating experience to costs see Goldberg and Touw (2003) and Anzanello and Fogliatto (2011).

[7] See also McDonald (1987), Hall and Howell (1985), and Goldberg and Touw (2003) for further discussion of the effect of autocorrelation on different estimation techniques.

$$y_t = y_0 + \omega x_t + \sum_{i=1}^{t} \eta_i,$$

which is the same as Eq. (2) except that the noise is accumulated across the entire time series. In contrast, Eq. (2) implies that even in the long run the two variables should be close to their deterministic relationship.

If y and x are I(1)[8], Eq. (2) defines a cointegrated relationship. We have not tested for cointegration rigorously, mostly because unit root and cointegration tests have uncertain properties in small samples, and our time series are typically short and they are all of different length. Nevertheless, we have run some analyses suggesting that the difference model may be more appropriate. First of all in about half the cases we found that model (2) resulted in a Durbin-Watson statistic lower than the R^2, indicating a risk of spurious regression and suggesting that first-differencing may be appropriate[9]. Second, the variance of the residuals of the level model was generally higher, so that the tests proposed in Harvey (1980) generally favored the first-difference model. Third, we ran cointegration tests in the form of Augmented Dickey-Fuller tests on the residuals of the regression (2), again generally suggesting a lack of cointegration. While a lengthy study using different tests and paying attention to differing sample sizes would shed more light on this issue, in this paper we will use Eq. (3) (with autocorrelated noise). The simplicity of this specification is also motivated by the fact that we want to have the same model for all technologies, we want to be able to calculate the variance of the forecast errors, and we want to estimate parameters with very short estimation windows so as to obtain as many forecast errors as possible.

We will compare our forecasts using Wright's law with those of a univariate time series model which we call Moore's law

$$y_t - y_{t-1} = \mu + n_t. \tag{4}$$

This is a random walk with drift. The forecast errors have been analyzed for i.i.d. normal n_t and for $n_t = \nu_t + \theta\nu_{t-1}$ with i.i.d. normal ν_t (keeping the simplest forecasting rule) in Farmer and Lafond (2016). As we will note throughout the paper, if cumulative production grows at a constant logarithmic rate of growth, i.e. $x_{t+1} - x_t = r$ for all t, Moore's and Wright's laws are observationally equivalent in the sense that Eq. (3) becomes Eq. (4) with $\mu = \omega r$. This equivalence has already been noted by Sahal (1979) and Ferioli and Van der Zwaan (2009) for the deterministic case. Nagy et al. (2013), using a dataset very close to ours, showed that using trend stationary models to estimate the three parameters independently (Eq. (2) and regressions of the (log) costs and experience levels on a time trend), the identity $\hat{\mu} = \hat{\omega}\hat{r}$ holds very well for most technologies. Here we will replicate this result using difference stationary models.

2.2. Hindcasting and surrogate data procedures

To evaluate the predictive ability of the models, we follow closely Farmer and Lafond (2016) by using hindcasting to compute as many forecast errors as possible and using a surrogate data procedure to test their statistical compatibility with our models. Pretending to be in the past, we make pseudo forecasts of values that we are able to observe and compute the errors of our forecasts. More precisely, our procedure is as follows. We consider all periods for which we have $(m + 1)$ years of observations (i.e. m year-to-year growth rates) to estimate the parameters, and at least one year ahead to make a forecast (unless otherwise noted we choose $m = 5$). For each of these periods, we estimate the parameters and make all the forecasts for which we can compute forecast errors. Because of our focus on testing the method and comparing with univariate forecasts, throughout the paper we assume that cumulative production is known in advance. Having obtained a set of forecast errors, we compute a number of indicators, such as the distribution of the forecast errors or the mean squared forecast error, and compare the empirical values to what we expect given the size and structure of our dataset.

To know what we expect to find, we use an analytical approach as well as a surrogate data procedure. The analytical approach simply consists of deriving an approximation of the distribution of forecast errors. However, the hindcasting procedure generates forecast errors which, for a single technology, are not independent[10]. However, in this paper we have many short time series so that the problem is somewhat limited (see Appendix A, Figs. 15–16). Nevertheless, we deal with it by using a surrogate data procedure: we simulate many datasets similar to ours and perform the same analysis, thereby determining the sampling distribution of any statistics of interest.

2.3. Parameter estimation

To simplify notation a bit, let $Y_t = y_t - y_{t-1}$ and $X_t = x_t - x_{t-1}$ be the changes of y and x in period t. We estimate Wright's exponent from Eq. (3) by running an OLS regression through the origin. Assuming that we have data for times $i = 1...(m + 1)$, minimizing the squared errors gives[11]

$$\hat{\omega} = \frac{\sum_{i=2}^{m+1} X_i Y_i}{\sum_{i=2}^{m+1} X_i^2}. \tag{5}$$

Substituting $\omega X_i + \eta_i$ for Y_i, we have

$$\hat{\omega} = \omega + \frac{\sum_{i=2}^{m+1} X_i \eta_i}{\sum_{i=2}^{m+1} X_i^2}. \tag{6}$$

The variance of the noise σ_η^2 is estimated as the regression standard error

$$\hat{\sigma}_\eta^2 = \frac{1}{m-1} \sum_{i=2}^{m+1} (Y_i - \hat{\omega} X_i)^2. \tag{7}$$

For comparison, parameter estimation in the univariate model Eq. (4) as done in Farmer and Lafond (2016) yields the sample mean $\hat{\mu}$ and variance $\hat{K}^2$ of $Y_t \sim \mathcal{N}(\mu, K^2)$.

2.4. Forecast errors

Let us first recall that for the univariate model Eq. (4), the variance of the forecast errors is given by (Sampson, 1991; Clements and Hendry, 2001; Farmer and Lafond, 2016)

$$E[\mathscr{E}_{M,\tau}^2] = K^2\left(\tau + \frac{\tau^2}{m}\right), \tag{8}$$

where τ is the forecast horizon and the subscript $_M$ indicates forecast errors obtained using "Moore"'s model. It shows that in the simplest model, the expected squared forecast error grows due to future noise accumulating (τ) and to estimation error (τ^2/m). These terms will reappear later so we will use a shorthand

$$A \equiv \tau + \frac{\tau^2}{m}, \tag{9}$$

We now compute the variance of the forecast errors for Wright's model. If we are at time $t = m + 1$ and look τ steps ahead into the future,

[8] A variable is I(1) or integrated of order one if its first difference $y_{t+1} - y_t$ is stationary.

[9] Note, however, that since we do not include an intercept in the difference model and since the volatility of experience is low, first differencing is not a good solution to the spurious regression problem.

[10] For a review of forecasting ability tests and a discussion of how the estimation scheme affects the forecast errors, see West (2006) and Clark and McCracken (2013).

[11] Throughout the paper, we will use the hat symbol for estimated parameters when the estimation is made using only the $m + 1$ years of data on which the forecasts are based. When we provide full sample estimates we use the tilde symbol.

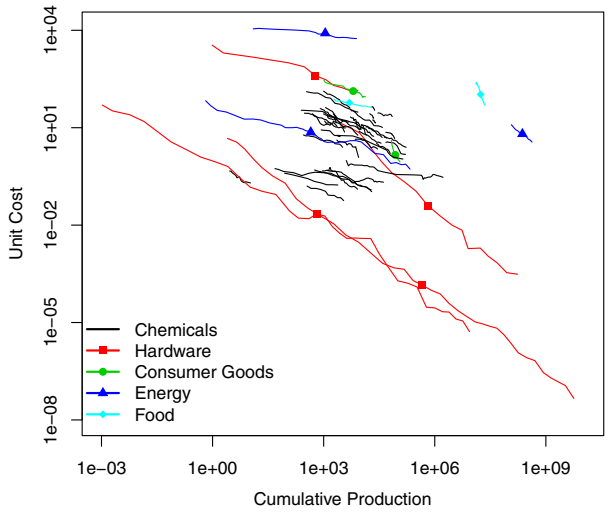

Fig. 1. Scatter plot of unit costs against cumulative production.

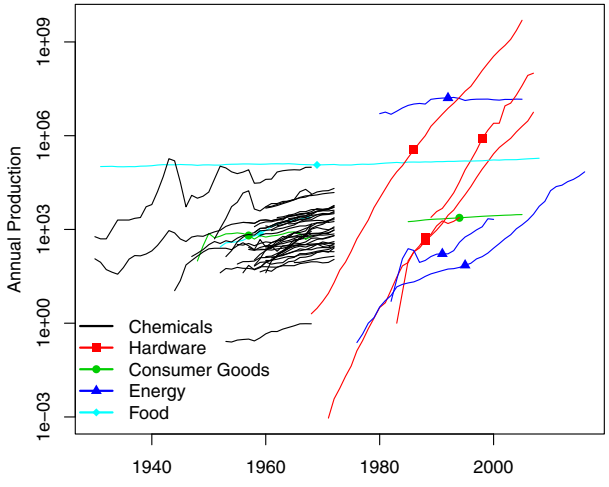

Fig. 2. Production time series.

we know that

$$y_{t+\tau} = y_t + \omega(x_{t+\tau} - x_t) + \sum_{i=t+1}^{t+\tau} \eta_i. \tag{10}$$

To make the forecasts we assume that the future values of x are known, i.e. we are forecasting costs conditional on a given growth of future experience. This is a common practice in the literature (Meese and Rogoff, 1983; Alberth, 2008). More formally, the point forecast at horizon τ is

$$\hat{y}_{t+\tau} = y_t + \hat{\omega}(x_{t+\tau} - x_t). \tag{11}$$

The forecast error is the difference between Eqs. (10) and (11), that is

$$\mathscr{E}_\tau \equiv y_{t+\tau} - \hat{y}_{t+\tau} = (\omega - \hat{\omega}) \sum_{i=t+1}^{t+\tau} X_i + \sum_{i=t+1}^{t+\tau} \eta_i. \tag{12}$$

We can derive the expected squared error. Since the X_is are known constants, using $\hat{\omega}$ from Eq. (6) and the notation $m+1=t$, we find

$$E[\mathscr{E}_\tau^2] = \sigma_\eta^2 \left(\tau + \frac{(\sum_{i=t+1}^{t+\tau} X_i)^2}{\sum_{i=2}^{t} X_i^2} \right). \tag{13}$$

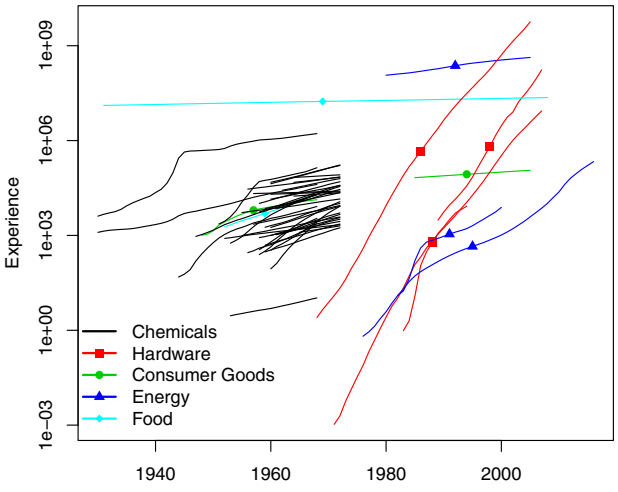

Fig. 3. Experience time series.

2.5. Comparison of Wright's law and Moore's law

Sahal (1979) was the first to point out that in the deterministic limit the combination of exponentially increasing cumulative production and exponentially decreasing costs gives Wright's law. Here we generalize this result in the presence of noise and show how variability in the production process affects this relationship.

Under the assumption that experience growth rates are constant ($X_i = r$) and letting $m = t-1$, Eq. (13) gives the result that the variance of Wright's law forecast errors are precisely the same as the variance of Moore's law forecast errors given in Eq. (8), with $\widehat{K} = \hat{\sigma}_\eta$. To see how the fluctuations in the growth rate of experience impact forecast errors we can rewrite Eq. (13) as

$$E[\mathscr{E}_\tau^2] = \sigma_\eta^2 \left(\tau + \frac{\tau^2}{m} \frac{\hat{r}_{(f)}^2}{\hat{\sigma}_{x,(p)}^2 + \hat{r}_{(p)}^2} \right), \tag{14}$$

where $\hat{\sigma}_{x,(p)}^2$ refers to the estimated variance of past experience growth rates, $\hat{r}_{(p)}$ to the estimated mean of past experience growth rates, and $\hat{r}_{(f)}$ to the estimated mean of future experience growth rates.[12]

This makes it clear that the higher the volatility of experience (σ_x^2), the lower the forecast errors. This comes from a simple, well-known fact of regression analysis: high variance of the regressor makes the estimates of the slope more precise. Here the high standard errors in the estimation of ω (due to low σ_x^2) decrease the part of the forecast error variance due to parameter estimation, which is associated with the term τ^2/m.

This result shows that, assuming Wright's law is correct, for Wright's law forecasts to work well (and in particular to outperform Moore's law), it is better to have cumulative production growth rates that fluctuate a great deal. Unfortunately for our data this is typically not the case. Instead, empirically cumulative production follows a fairly smooth exponential trend. To explain this finding we calculated the stochastic properties of cumulative production assuming that production is a geometric random walk with drift g and volatility σ_q. In Appendix B, using a saddle point approximation for the long time limit we find that $E[X] \equiv r \approx g$ and

$$\mathrm{Var}[X] \equiv \sigma_x^2 \approx \sigma_q^2 \tanh(g/2), \tag{15}$$

where tanh is the hyperbolic tangent function. We have tested this remarkably simple relationship using artificially generated data and we find that it works reasonably well.

These results show that cumulative production grows at the same

[12] The past refers to data at times (1,..., t) and the future to times ($t+1$,..., $t+\tau$).

Table 1
Parameter estimates.

		Cost		Production		Cumul. prod.		Wright's law		
	T	$\widetilde{\mu}$	$\widetilde{K}$	$\widetilde{g}$	$\widetilde{\sigma}_q$	$\widetilde{r}$	$\widetilde{\sigma}_x$	$\widetilde{\omega}$	$\widetilde{\sigma_\eta}$	$\widetilde{\rho}$
Automotive	21	−0.076	0.047	0.026	0.011	0.027	0.000	−2.832	0.048	1.000
Milk	78	−0.020	0.023	0.008	0.020	0.007	0.000	−2.591	0.023	0.019
Isopropyl alcohol	9	−0.039	0.024	0.022	0.074	0.023	0.002	−1.677	0.024	−0.274
Neoprene rubber	13	−0.021	0.020	0.015	0.055	0.015	0.001	−1.447	0.020	0.882
Phthalic anhydride	18	−0.076	0.152	0.061	0.094	0.058	0.006	−1.198	0.155	0.321
Titanium dioxide	9	−0.037	0.049	0.031	0.029	0.031	0.001	−1.194	0.049	−0.400
Sodium	16	−0.013	0.023	0.013	0.081	0.012	0.001	−1.179	0.023	0.407
Pentaerythritol	21	−0.050	0.066	0.050	0.107	0.050	0.006	−0.954	0.068	0.343
Methanol	16	−0.082	0.143	0.097	0.081	0.091	0.004	−0.924	0.142	0.289
Hard disk drive	19	−0.593	0.314	0.590	0.307	0.608	0.128	−0.911	0.364	−0.569
Geothermal electricity	26	−0.049	0.022	0.043	0.116	0.052	0.013	−0.910	0.023	0.175
Phenol	14	−0.078	0.090	0.088	0.055	0.089	0.004	−0.853	0.092	−1.000
Transistor	38	−0.498	0.240	0.585	0.157	0.582	0.122	−0.849	0.226	−0.143
Formaldehyde	11	−0.070	0.061	0.086	0.078	0.085	0.005	−0.793	0.063	0.489
Ethanolamine	18	−0.059	0.042	0.076	0.076	0.080	0.005	−0.748	0.041	0.355
Caprolactam	11	−0.103	0.075	0.136	0.071	0.142	0.009	−0.746	0.071	0.328
Ammonia	13	−0.070	0.099	0.096	0.049	0.102	0.007	−0.740	0.095	1.000
Acrylic fiber	13	−0.100	0.057	0.127	0.126	0.137	0.020	−0.726	0.056	−0.141
Ethylene glycol	13	−0.062	0.059	0.089	0.107	0.083	0.006	−0.711	0.062	−0.428
DRAM	37	−0.446	0.383	0.626	0.253	0.634	0.185	−0.680	0.380	0.116
Benzene	16	−0.056	0.083	0.087	0.114	0.087	0.012	−0.621	0.085	−0.092
Aniline	12	−0.072	0.095	0.110	0.099	0.113	0.008	−0.620	0.097	−1.000
Vinyl acetate	13	−0.082	0.061	0.131	0.080	0.129	0.010	−0.617	0.065	0.341
Vinyl chloride	11	−0.083	0.050	0.136	0.085	0.137	0.008	−0.613	0.049	−0.247
Polyethylene LD	15	−0.085	0.076	0.135	0.075	0.139	0.009	−0.611	0.075	0.910
Acrylonitrile	14	−0.084	0.108	0.121	0.178	0.134	0.025	−0.605	0.109	1.000
Styrene	15	−0.068	0.047	0.112	0.089	0.113	0.008	−0.585	0.050	0.759
Maleic anhydride	14	−0.069	0.114	0.116	0.143	0.119	0.013	−0.551	0.116	0.641
Ethylene	13	−0.060	0.057	0.114	0.054	0.114	0.005	−0.526	0.057	−0.290
Urea	12	−0.062	0.094	0.121	0.073	0.127	0.011	−0.502	0.093	0.003
Sorbitol	8	−0.032	0.046	0.067	0.025	0.067	0.002	−0.473	0.046	−1.000
Polyester fiber	13	−0.121	0.100	0.261	0.132	0.267	0.034	−0.466	0.094	−0.294
Bisphenol A	14	−0.059	0.048	0.136	0.136	0.135	0.012	−0.437	0.048	−0.056
Paraxylene	11	−0.103	0.097	0.259	0.326	0.228	0.054	−0.417	0.104	−1.000
Polyvinyl chloride	22	−0.064	0.057	0.137	0.136	0.144	0.024	−0.411	0.062	0.319
Low density polyethylene	16	−0.103	0.064	0.213	0.164	0.237	0.069	−0.400	0.071	0.473
Sodium chlorate	15	−0.033	0.039	0.076	0.077	0.084	0.006	−0.397	0.039	0.875
Titanium sponge	18	−0.099	0.099	0.196	0.518	0.241	0.196	−0.382	0.075	0.609
Photovoltaics	41	−0.121	0.153	0.315	0.202	0.318	0.133	−0.380	0.145	−0.019
Monochrome television	21	−0.060	0.072	0.093	0.365	0.130	0.093	−0.368	0.074	−0.444
Cyclohexane	17	−0.055	0.052	0.134	0.214	0.152	0.034	−0.317	0.057	0.375
Polyethylene HD	15	−0.090	0.075	0.250	0.166	0.275	0.074	−0.307	0.079	0.249
Carbon black	9	−0.013	0.016	0.046	0.051	0.046	0.002	−0.277	0.016	−1.000
Laser diode	12	−0.293	0.202	0.708	0.823	0.824	0.633	−0.270	0.227	0.156
Aluminum	17	−0.015	0.044	0.056	0.075	0.056	0.004	−0.264	0.044	0.761
Polypropylene	9	−0.105	0.069	0.383	0.207	0.414	0.079	−0.261	0.059	0.110
Beer	17	−0.036	0.042	0.137	0.091	0.146	0.016	−0.235	0.043	−1.000
Primary aluminum	39	−0.022	0.080	0.088	0.256	0.092	0.040	−0.206	0.080	0.443
Polystyrene	25	−0.061	0.086	0.205	0.361	0.214	0.149	−0.163	0.097	0.074
Primary magnesium	39	−0.031	0.089	0.135	0.634	0.158	0.211	−0.131	0.088	−0.037
Wind turbine	19	−0.038	0.047	0.336	0.570	0.357	0.337	−0.071	0.050	0.750

rate as production. More importantly, since $0 < \tanh(g/2) < 1$ (and here we assume $g > 0$), the volatility of cumulative production is lower than the volatility of production. This is not surprising: it is well-known that integration acts as a low pass filter, in this case making cumulative production smoother than production. Thus if production follows a geometric random walk with drift, experience is a smoothed version, making it hard to distinguish from an exogenous exponential trend. When this happens Wright's law and Moore's law yield similar predictions. This theoretical result is relevant to our case, as can be seen in the time series of production and experience plotted in Fig. 2 and Fig. 3, and the low volatility of experience compared to production reported in Table 1 below.

2.6. Autocorrelation

We now turn to an extension of the basic model. As we will see, the data shows some evidence of autocorrelation. Following Farmer and Lafond (2016), we augment the model to allow for first order moving average autocorrelation. For the autocorrelated Moore's law model ("Integrated Moving Average of order 1")

$$y_t - y_{t-1} = \mu + v_t + \theta v_{t-1},$$

Farmer and Lafond (2016) obtained a formula for the forecast error variance when the forecasts are performed assuming no autocorrelation

$$E[\mathscr{E}^2_{M,\tau}] = \sigma_v^2\left[-2\theta + \left(1 + \frac{2(m-1)\theta}{m} + \theta^2\right)A\right], \tag{16}$$

where $A = \tau + \frac{\tau^2}{m}$ (Eq. (9)). Here we extend this result to the autocorrelated Wright's law model

$$y_t - y_{t-1} = \omega(x_t - x_{t-1}) + u_t + \rho u_{t-1}, \tag{17}$$

where $u_t \sim \mathcal{N}(0, \sigma_u^2)$. We treat ρ as a known parameter. Moreover, we will assume that it is the same for all technologies and we will estimate it as the average of the $\widetilde{\rho}_j$ estimated on each technology separately (as

described in the next section). This is a delicate assumption, but it is motivated by the fact that many of our time series are too short to estimate a specific ρ_j reliably, and assuming a universal, known value of ρ allows us to keep analytical tractability.

The forecasts are made exactly as before, but the forecast error now is

$$\mathscr{E}_\tau = \sum_{j=2}^{m+1} H_j [v_j + \rho v_{j-1}] + \sum_{T=t+1}^{t+\tau} [v_T + \rho v_{T-1}], \tag{18}$$

where the H_j are defined as

$$H_j = -\frac{\sum_{i=t+1}^{t+\tau} X_i}{\sum_{i=2}^{t} X_i^2} X_j. \tag{19}$$

The forecast error can be decomposed as a sum of independent normally distributed variables, from which the variance can be computed as

$$E[\mathscr{E}_\tau^2] = \sigma_u^2 \left(\rho^2 H_2^2 + \sum_{j=2}^{m} (H_j + \rho H_{j+1})^2 + (\rho + H_{m+1})^2 + (\tau - 1)(1+\rho)^2 + 1 \right). \tag{20}$$

When generating real forecasts (Section 4), we will take the rate of growth of future cumulative production as constant. If we also assume that the growth rates of past cumulative production were constant, we have $X_i = r$ and thus $H_i = -\tau/m$ for all i. As expected from Sahal's identity, simplifying Eq. (20) under this assumption gives Eq. (16) where θ is substituted by ρ and σ_v is substituted by σ_u,

$$E[\mathscr{E}_\tau^2] = \sigma_u^2 \left[-2\rho + \left(1 + \frac{2(m-1)\rho}{m} + \rho^2 \right) A \right]. \tag{21}$$

In practice we compute $\hat{\sigma}_\eta$ using Eq. (7), so that σ_u may be estimated as

$$\hat{\sigma}_u = \sqrt{\hat{\sigma}_\eta^2 / (1+\rho^2)},$$

suggesting the normalized error $\mathscr{E}_\tau^2 / \hat{\sigma}_\eta^2$. To gain more intuition on Eq. (21), and propose a simple, easy to use formula, note that for $\tau \gg 1$ and $m \gg 1$ it can be approximated as

$$E\left[\left(\frac{\mathscr{E}_\tau}{\sigma_\eta} \right)^2 \right] \approx \frac{(1+\rho)^2}{1+\rho^2} \left(\tau + \frac{\tau^2}{m} \right). \tag{22}$$

For all models (Moore and Wright with and without autocorrelated noise), having determined the variance of the forecast errors we can normalize them so that they follow a Standard Normal distribution

$$\frac{\mathscr{E}_\tau}{\sqrt{E[\mathscr{E}_\tau^2]}} \sim \mathcal{N}(0,1) \tag{23}$$

In what follows we will replace σ_η^2 by its estimated value, so that when $\mathscr{E}$ and $\hat{\sigma}_\eta$ are independent the reference distribution is Student. However, for small sample size m the Student distribution is only a rough approximation, as shown in Appendix A, where it is also shown that the theory works well when the variance is known, or when the variance is estimated but m is large enough.

2.7. Comparing Moore and Wright at different forecast horizons

One objective of the paper is to compare Moore's law and Wright's law forecasts. To normalize Moore's law forecasts, Farmer and Lafond (2016) used the estimate of the variance of the delta log cost time series $\widehat{K}^2$, as suggested by Eq. (8), i.e.

$$\epsilon_M = \mathscr{E}_M / \widehat{K}, \tag{24}$$

To compare the two models, we propose that Wright's law forecast errors can be normalized by the very same value

$$\epsilon_W = \mathscr{E}_W / \widehat{K}, \tag{25}$$

Using this normalization, we can plot the normalized mean squared errors from Moore's and Wright's models at each forecast horizon. These are directly comparable, because the raw errors are divided by the same value, and these are meaningful because Moore's normalization ensures that the errors from different technologies are comparable and can reasonably be aggregated. In the context of comparing Moore and Wright, when pooling the errors of different forecast horizons we also use the normalization from Moore's model (neglecting autocorrelation for simplicity), $A \equiv \tau + \tau^2/m$ (see Farmer and Lafond (2016) and Eqs. (8) and (9)).

3. Empirical results

3.1. The data

We mostly use data from the performance curve database[13] created at the Santa Fe Institute by Bela Nagy and collaborators from personal communications and from Colpier and Cornland (2002), Goldemberg et al. (2004), Lieberman (1984), Lipman and Sperling (1999), McDonald and Schrattenholzer (2001), Moore (2006), Neij et al. (2003), Nemet (2006), Zhao (1999) and Schilling and Esmundo (2009). We augmented the dataset with data on solar photovoltaics modules taken from public releases of the consulting firms Strategies Unlimited, Navigant and SPV Market Research, which give the average selling price of solar PV modules, and that we corrected for inflation using the US GDP deflator.

This database gives a proxy for unit costs[14] for a number of technologies over variable periods of time. In principle we would prefer to have data on unit costs, but often these are unavailable and the data is about prices[15]. Since the database is built from experience curves found in the literature, rather than from a representative sample of products/technologies, there are limits to the external validity of our study but unfortunately we do not know of a database that contains suitably normalized unit costs for all products.

We have selected technologies to minimize correlations between the different time series, by removing technologies that are too similar (e.g. other data on solar photovoltaics). We have also refrained from including very long time series that would represent a disproportionate share of our forecast errors, make the problem of autocorrelation of forecast errors very pronounced, and prevent us from generating many random datasets for reasons of limited computing power. Starting with the set of 53 technologies with a significant improvement rate from Farmer and Lafond (2016), we removed DNA sequencing for which no production data was available, and Electric Range which had a zero production growth rate so that we cannot apply the correction to cumulative production described below. We are left with 51 technologies belonging to different sectors (chemical, energy, hardware, consumer durables, and food), although chemicals from the historical reference (Boston Consulting Group, 1972) represent a large part of the dataset. For some technologies the number of years is slightly different from Farmer and Lafond (2016) because we had to remove observations for

[13] The data can be accessed at pcdb.santafe.edu.

[14] In a few cases (milk and automotive), a measure of performance is used instead of costs, and automotive's experience is computed based on distance driven. The main results would not be severely affected by the exclusion of these two time series. Also, unit cost is generally computed from total cost and production of a year or batch, not from actual observation of every unit cost. Different methods may give slightly different results (Gallant, 1968; Womer and Patterson, 1983; Goldberg and Touw, 2003), but our dataset is too heterogenous to attempt any correction. Obviously, changes in unit costs do not come from technological progress only, and it is difficult to account for changes in quality, but unit costs are nevertheless a widely used and defensible proxy.

[15] This implies a bias whenever prices and costs do not have the same growth rate, as is typically the case when pricing strategies are dynamic and account for learning effects (for instance predatory pricing). For a review of the industrial organization literature on this topic, see Thompson (2010).

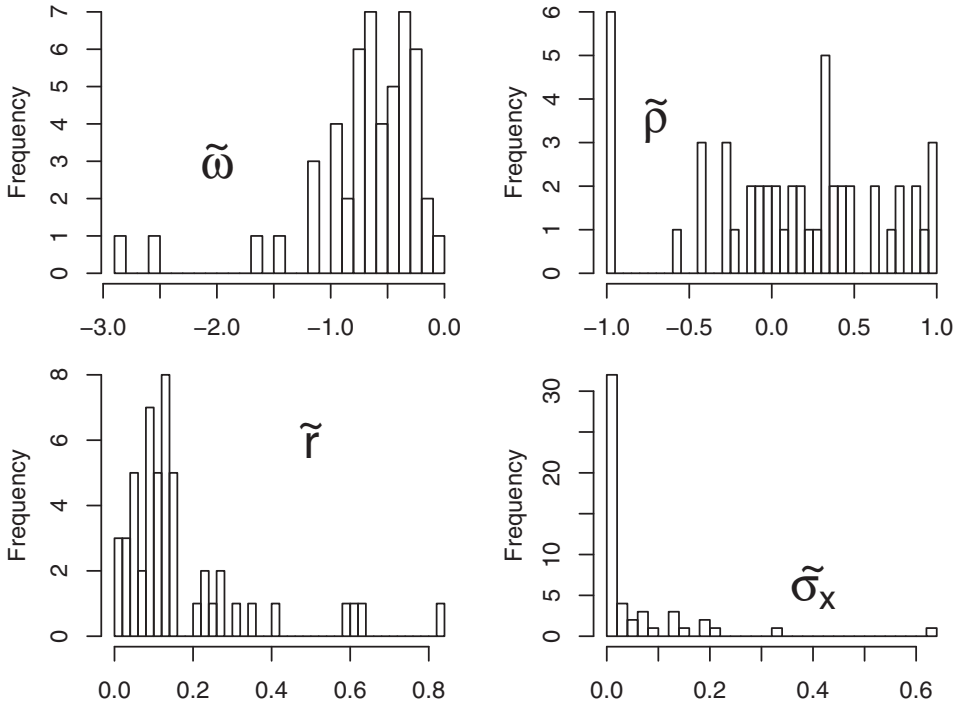

Fig. 4. Histograms of estimated parameters.

which data on production was not available.

Fig. 1 shows the experience curves, while Figs. 2 and 3 show production and experience time series, suggesting at least visually that experience time series are "smoothed" versions of production time series.

3.2. Estimating cumulative production

A potentially serious problem in experience curve studies is that one generally does not observe the complete history of the technology, so that simply summing up observed production misses the experience previously accumulated. There is no perfect solution to this problem. For each technology, we infer the initial cumulative production using a procedure common in the "R&D capital" literature (Hall and Mairesse, 1995), although not often used in experience curve studies (for an exception see Nordhaus (2014)). It assumes that production grew as $Q_{t+1} = Q_t(1 + g_d)$ and experience accumulates as $Z_{t+1} = Z_t + Q_t$, so that it can be shown that $Z_{t_0} = Q_{t_0}/g_d$. We estimate the discrete annual growth rate as $\hat{g}_d = \exp(\log(Q_T/Q_{t_0})/(T-1)) - 1$, where Q_{t_0} is production during the first available year, and T is the number of available years. We then construct the experience variable as $Z_{t_0} = Q_{t_0}/\hat{g}_d$ for the first year, and $Z_{t+1} = Z_t + Q_t$ afterwards.

Note that this formulation implies that experience at time t does not include production of time t, so the change in experience from time t to $t+1$ does not include how much is produced during year $t+1$. In this sense we assume that the growth of experience affects technological progress with a certain time lag. We have experimented with slightly different ways of constructing the experience time series, and the aggregated results do not change much, due to the high persistence of production growth rates.

A more important consequence of this correction is that products with a small production growth rate will have a very important correction for initial cumulative production. In turn, this large correction

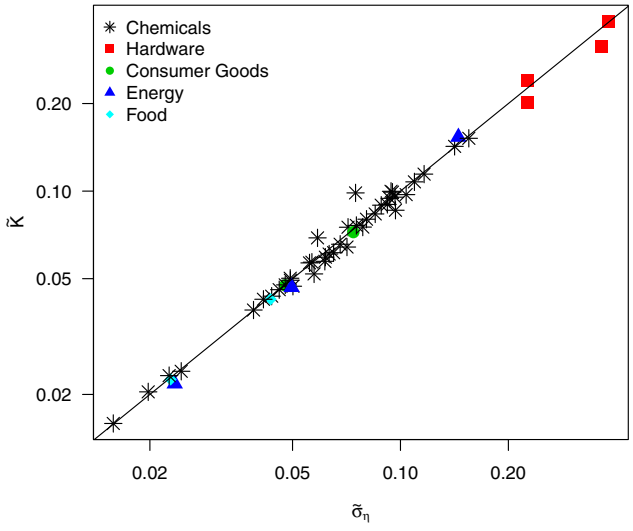

Fig. 5. Comparison of residuals standard deviation from Moore's and Wright's models.

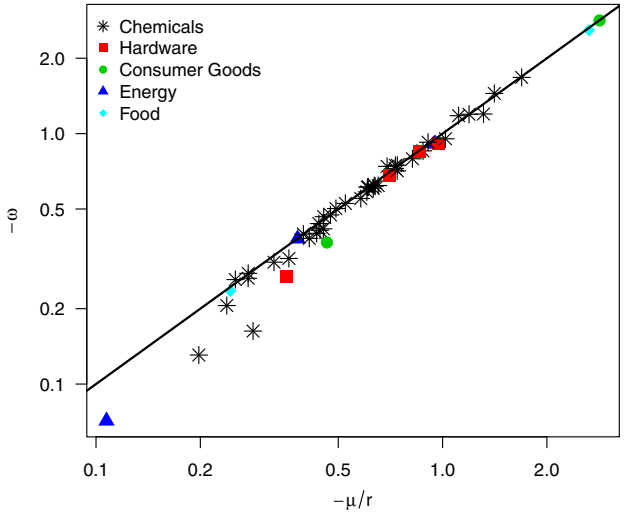

Fig. 6. Illustration of Sahal's identity.

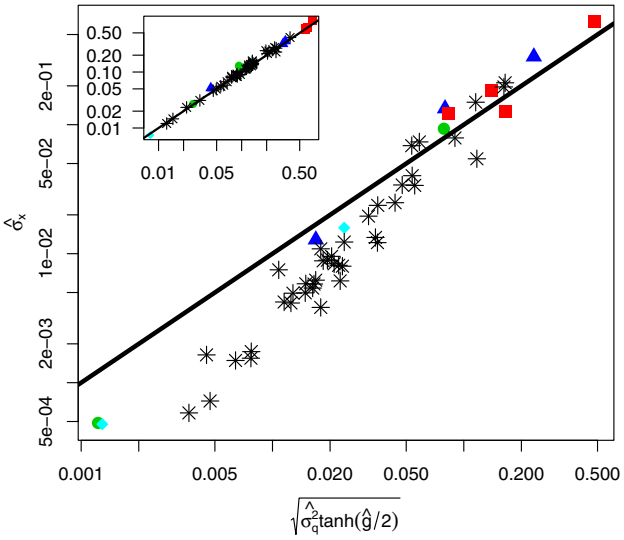

Fig. 7. Test of Eq. (15) relating the volatility of cumulative production to the drift and volatility of production. The inset shows the drift of cumulative production $\hat{r}$ against the drift of production $\hat{g}$.

of initial cumulative production leads to significantly lower values for the annual growth rates of cumulative production. As a result, the experience exponent $\hat{\omega}$ becomes larger than if there was no correction. This explains why products like milk, which have a very low rate of growth of production, have such a large experience exponent. Depending on the product, this correction may be small or large, and it may be meaningful or not. Here we have decided to use this correction for all products.

3.3. Descriptive statistics and Sahal's identity

Table 1 summarizes our dataset, showing in particular the parameters estimated using the full sample. Fig. 4 complements the table by showing histograms for the distribution of the most important parameters. Note that we denote estimated parameters using a tilde because we use the full sample (when using the small rolling window, we used the hat notation). To estimate the $\widetilde{\rho_j}$, we have used a maximum likelihood estimation of Eq. (17) (letting $\widetilde{\omega}_{MLE}$ differ from $\widetilde{\omega}$).

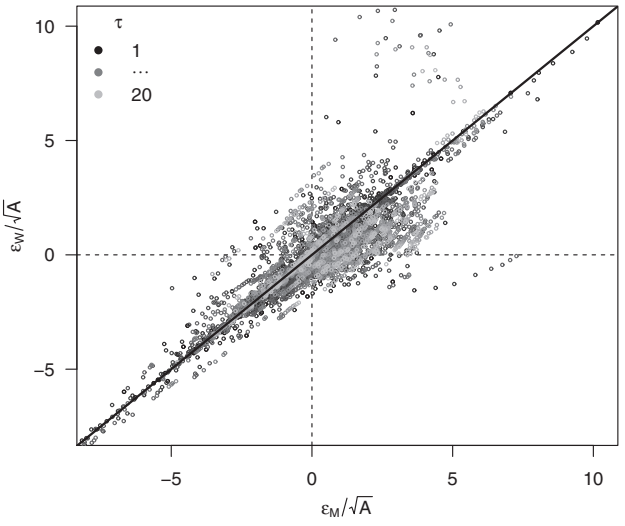

Fig. 8. Scatter plot of Moore-normalized forecast errors ϵ_M and ϵ_W (forecasts are made using $m = 5$). This shows that in the vast majority of cases Wright's and Moore's law forecast errors have the same sign and a similar magnitude, but not always.

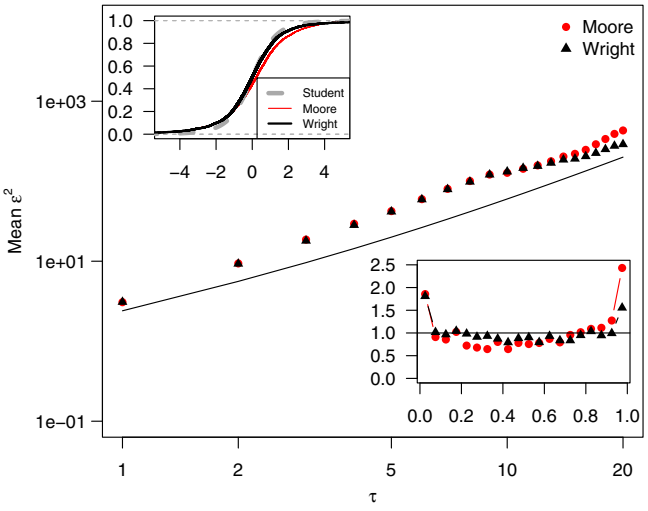

Fig. 9. Comparison of Moore-normalized forecast errors from Moore's and Wright's models. The main chart shows the mean squared forecast errors at different forecast horizons. The insets show the distribution of the normalized forecast errors as an empirical cumulative distribution function against the Student distribution (top left) and as a probability integral transform against a uniform distribution (bottom right).

Fig. 5 compares the variance of the noise estimated from Wright's and Moore's models. These key quantities express how much of the change in (log) cost is left unexplained by each model; they also enter as direct factor in the expected mean squared forecast error formulas. The lower the value, the better the fit and the more reliable the forecasts. The figure shows that for each technology the two models give similar values; see Table 1.

Next, we show Sahal's identity as in Nagy et al. (2013). Sahal's observation is that if cumulative production and costs both have exponential trends r and μ, respectively, then costs and production have a power law (constant elasticity) relationship parametrized by $\omega = \mu/r$. One way to check the validity of this relationship is to measure μ, r and ω independently and plot ω against μ/r. Fig. 6 shows the results and confirms the relevance of Sahal's identity.

To explain why Sahal's identity works so well and Moore's and Wright's laws have similar explanatory power, in Section 2.4 we have shown that in theory if production grows exponentially, cumulative production grows exponentially with an even lower volatility. Fig. 7 shows how this theoretical result applies to our dataset. Knowing the drift and volatility of the (log) production time series, we are able to predict the drift and volatility of the (log) cumulative production time series fairly well.

3.4. Comparing Moore's and Wright's law forecast errors

Moore's law forecasts are based only on the cost time series, whereas Wright's law forecasts use information about future experience to predict future costs. Thus, we expect that in principle Wright's forecasts should be better. We now compare Wright's and Moore's models in a number of ways. The first way is simply to show a scatter plot of the forecast errors from the two models. Fig. 8 shows this scatter plot for the errors normalized by $\widehat{K}\sqrt{A}$ (i.e. "Moore-normalized", see Eqs. (8)–(9) and (24) and (25)), with the identity line as a point of comparison. It is clear that they are highly correlated. When Moore's law over (under)predicts, it is likely that Wright's law over(under)predicts as well, and when Moore's law leads to a large error, it is likely that Wright's law leads to a large error as well.

In Fig. 9, the main plot shows the mean squared Moore-normalized forecast error ϵ_M^2 and ϵ_W^2 (see Section 2.7), where the average is taken over all available forecast errors of a given forecast horizon (note that some technologies are more represented than others). The solid

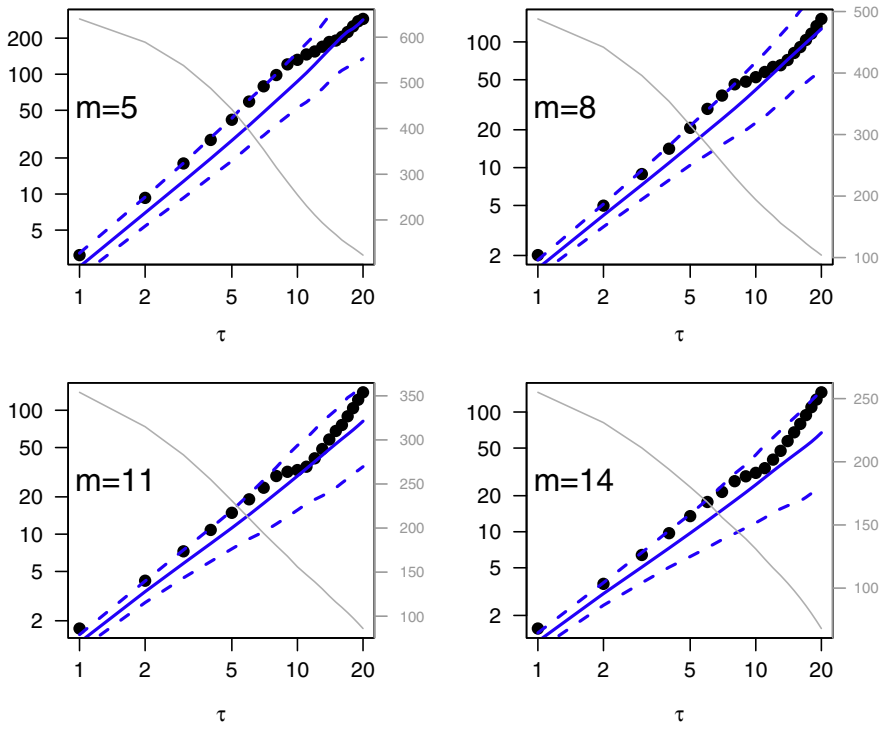

Fig. 10. Mean squared Moore-normalized forecast errors of the Wright's law model (Mean ϵ_W^2) versus forecast horizon. The 95% intervals (dashed lines) and the mean (solid line) are computed using simulations as described in the text. The grey line, associated with the right axis, shows the number of forecast errors used to make an average.

diagonal line is the benchmark for the Moore model without autocorrelation, i.e. the line $y = \frac{m-1}{m-3}A$ (Farmer and Lafond, 2016). Wright's model appears slightly better at the longest horizons, however there are not many forecasts at these horizons so we do not put too much emphasis on this finding. The two insets show the distribution of the rescaled Moore-normalized errors $\epsilon/\sqrt{A}$, either as a cumulative distribution function (top left) or using the probability integral transform[16] (bottom right). All three visualizations confirm that Wright's model only slightly outperform Moore's model.

3.5. Wright's law forecast errors

In this section, we analyze in detail the forecast errors from Wright's model. We will use the proper normalization derived in Section 2.4, but since it does not allow us to look at horizon specific errors we first look again at the horizon specific Moore-normalized mean squared forecast errors. Fig. 10 shows the results for different values of m (for $m = 5$ the empirical errors are the same as in Fig. 9). The confidence intervals are created using the surrogate data procedure described in Section 2.2, in which we simulate many random datasets using the autocorrelated Wright's law model (Eq. (17)) and the parameters of Table 1 forcing $\rho_j = \rho^* = 0.19$ (see below). We then apply the same hindcasting, error normalization and averaging procedure to the surrogate data that we did for the empirical data, and show with blue lines the mean and 95% confidence intervals. This suggests that the empirical data is compatible with the model (17) in terms of Moore-normalized forecast errors at different horizons.

We now analyze the forecast errors from Wright's model normalized using the (approximate) theory of Section 2.4. Again we use the hindcasting procedure and unless otherwise noted, we use an estimation window of $m = 5$ points (i.e. 6 years) and a maximum forecasting horizon $\tau_{\max} = 20$. To normalize the errors, we need to choose a value of ρ. This is a difficult problem, because for simplicity we have assumed that ρ is the same for all technologies, but in reality it probably is not. We have experimented with different methods of choosing ρ based on modelling the forecast errors, for instance by looking for the value of ρ which makes the distribution of normalized errors closest to a Student distribution. While these methods may suggest that $\rho \approx 0.4$, they generally give different values of ρ for different values of m and $\tau_{\max}$ (which indicates that some non-stationarity/misspecification is present). Moreover, since the theoretical forecast errors do not exactly follow a Student distribution (see Appendix A) this estimator is biased. For simplicity, we use the average value of $\tilde{\rho}$ in our dataset, after removing the 9 values of $\tilde{\rho}$ whose absolute value was greater than 0.99 (which may indicate a misspecified model). Throughout the paper, we will thus use $\rho^* = 0.19$.

In Fig. 11, we show the empirical cumulative distribution function of the normalized errors (for ρ^* and $\rho = 0$) and compare it to the Student prediction. In Fig. 12 we show the probability integral transform of the normalized errors (assuming a Student distribution, and using $\rho = \rho^*$). In addition, Fig. 12 shows the confidence intervals obtained by the surrogate data method, using data simulated under the assumption $\rho = \rho^*$. Again, the results confirm that the empirical forecast errors are compatible with Wright's law, Eq. (17).

4. Application to solar photovoltaic modules

In this section we apply our method to solar photovoltaic modules. Technological progress in solar photovoltaics (PV) is a very prominent example of the use of experience curves.[17] Of course, the limitations of using experience curve models are valid in the context of solar PV modules; we refer the reader to the recent studies by Zheng and Kammen (2014) for a discussion of economies of scale, innovation output and policies; by de La Tour et al. (2013) for a discussion of the effects of input prices; and by Hutchby (2014) for a detailed study of levelized costs (module costs represent only part of the cost of

[16] The Probability Integral Transform is a transformation that allows to compare data against a theoretical distribution by transforming it and comparing it against the Uniform distribution. See for example Diebold et al. (1998), who used it to construct a test for evaluating density forecasts.

[17] Alberth (2008), Candelise et al. (2013), Isoard and Soria (2001), Junginger et al. (2010), Kahouli-Brahmi (2009), Neij (1997), Nemet (2006), Papineau (2006), Schaeffer et al. (2004), Swanson (2006), Van der Zwaan and Rabl (2004).

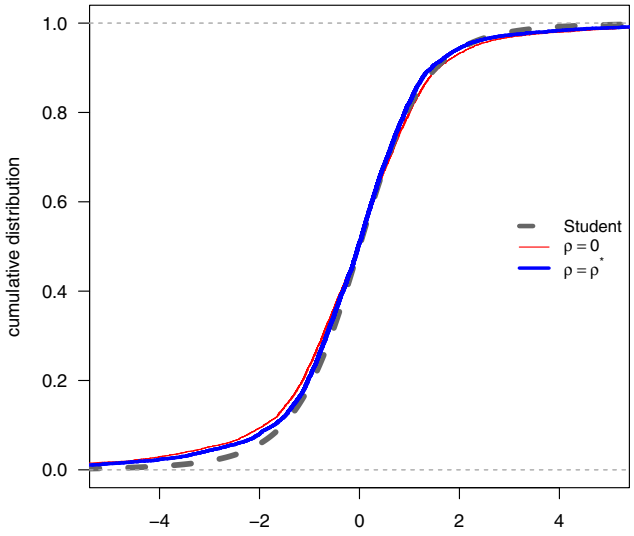

Fig. 11. Cumulative distribution of the normalized forecast errors, $m = 5$, $\tau_{max} = 20$, and the associated $\rho^* = 0.19$.

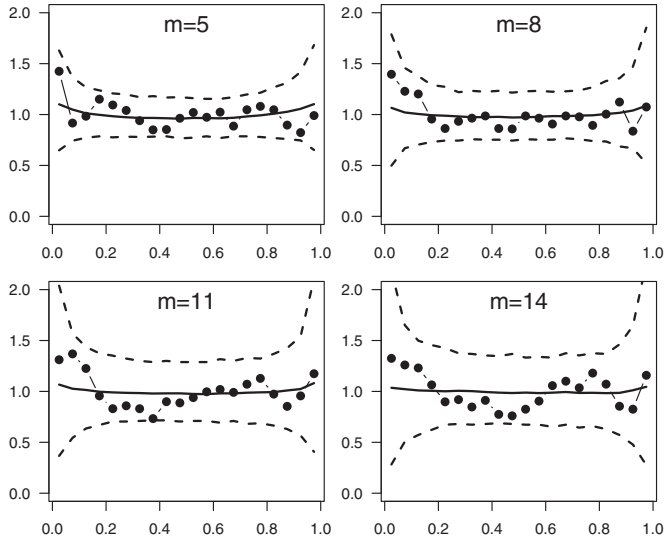

Fig. 12. Probability Integral Transform of the normalized forecast errors, for different values of m.

producing solar electricity), and to Farmer and Makhijani (2010) for a prediction of levelized solar photovoltaic costs made in 2010.

Historically (Wright, 1936; Alchian, 1963), the estimation of learning curves suggested that costs would drop by 20% for every doubling of cumulative production, although these early examples are almost surely symptoms of a sample bias. This corresponds to an estimated elasticity of about $\omega = 0.33$. As it turns out, estimations for PV have been relatively close to this number. Here we have a progress ratio of $2^{-0.38} = 23\%$. A recent study (de La Tour et al., 2013) contains a more complete review of previous experience curve studies for PV, and finds an average progress ratio of 20.2%. There are some differences across studies, mostly due to data sources, geographic and temporal dimension, and choice of proxy for experience. Our estimate here differs also because we use a different estimation method (first-differencing), and because we correct for initial cumulative production (most other studies either use available data on initial experience, or do not make a correction; it is quite small here, less than a MW).

In order to compare a forecast based on Wright's law, which gives cost as a function of cumulative production, with a forecast based on

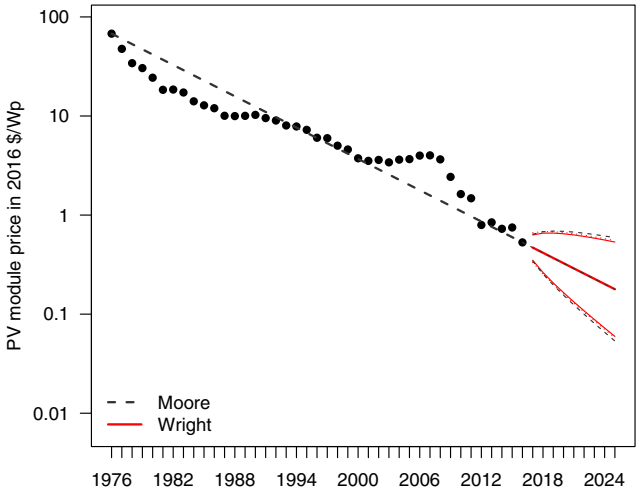

Fig. 13. Comparison of Moore's and Wright's law distributional forecasts (95% prediction intervals).

Moore's law, which gives cost as a function of time, we need to make an assumption about the production at a given point in time. Here we provide a distributional forecast for the evolution of PV costs under the assumption that cumulative production will keep growing at exactly the same rate as in the past, but without any variance, and we assume that we know this in advance. One should bear in mind that this assumption is purely to provide a point of comparison. The two models have a different purpose: Moore's law gives an unconditional forecast at a given point in time, and Wright's law a forecast conditioned on a given growth rate of cumulative production. While it is perfectly possible to use values from a logistic diffusion model or from expert forecasts, here we illustrate our method based on the exponential assumption to emphasise again the similarity of Moore's and Wright's laws.

Our distributional forecast is

$$y_{T+\tau} \sim \mathcal{N}(\hat{y}_{T+\tau}, V(\hat{y}_{T+\tau})) \tag{26}$$

Following the discussion of Sections 2.4 and 2.6, the point forecast is

$$\hat{y}_{T+\tau} = y_T + \widetilde{\omega}(x_{T+\tau} - x_T)$$

and the variance is

$$V(y_{T+\tau}) = \frac{\hat{\sigma}_\eta^2}{1+\rho^{*2}}\left(\rho^{*2}H_2^2 + \sum_{j=2}^{T-1}(H_j + \rho^{*2}H_{j+1})^2 + (\rho^* + H_T)^2 + (\tau - 1)(1+\rho^*)^2 + 1\right). \tag{27}$$

where

$$H_j = -\frac{\sum_{i=T+1}^{T+\tau} X_i}{\sum_{i=2}^{T} X_i^2} X_j,$$

and recalling that $X_t \equiv y_t - y_{t-1}$.

We now assume that the growth of cumulative production is exactly $\tilde{r}$ in the future, so the point forecast simplifies to

$$\hat{y}_{T+\tau} = y_T + \widetilde{\omega}\tilde{r}\tau \tag{28}$$

Regarding the variance, Eq. (27) is cumbersome but it can be greatly simplified. First, we assume that past growth rates of experience were constant, that is $X_i = \tilde{r}$ for $i = 1 \ldots T$ (i.e. $\hat{\sigma}_x^2 \approx 0$), leading to the equivalent of Eq. (21). As long as production grows exponentially this approximation is likely to be defensible (see Eq. (15)). Although from Table 1 we see that solar PV is not a particularly favourable example, we find that this assumption does not affect the variance of forecast errors significantly, at least for short or medium run horizons.

Since Eq. (21) is still a bit complicated, we can further assume that $\tau \gg 1$ and $T \gg 1$, so that we arrive at the equivalent of Eq. (22), that is

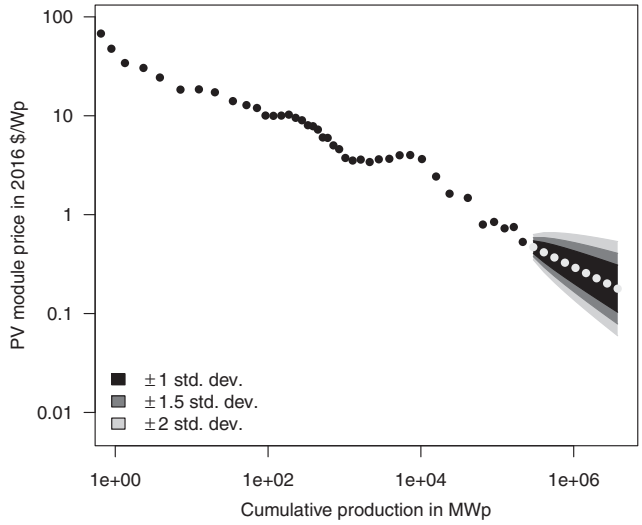

Fig. 14. Distributional forecast for the price of PV modules up to 2025, using Eqs. (26), (27) and (28).

$$V(y_{T+\tau}) \approx \hat{\sigma}_\eta^2 \frac{(1+\rho^*)^2}{1+\rho^{*2}}\left(\tau + \frac{\tau^2}{T-1}\right). \tag{29}$$

This equation is very simple and we will see that it gives results extremely close to Eq. (27), so that it can be used in applications.

Our point of comparison is the distributional forecast of Farmer and Lafond (2016) based on Moore's law with autocorrelated noise, and estimating $\theta^* = 0.23$ in the same way as ρ^* (average across all technologies of the measured MA(1) coefficient, removing the $|\widetilde{\theta}_j| \approx 1$). All other parameters are taken from Table 1. Fig. 13 shows the forecast for the mean log cost and its 95% prediction interval for the two models. The point forecasts of the two models are almost exactly the same because $\widehat{\omega}\widehat{r} = -0.1209 \approx \widehat{\mu} = -0.1213$. Moreover, Wright's law prediction intervals are slightly smaller because $\widehat{\sigma}_\eta = 0.145 < \widehat{K} = 0.153$. Overall, the forecasts are very similar as shown in Fig. 13. Fig. 13 does also show the prediction intervals from Eq. (29), in red dotted lines, but they are so close to those calculated using Eq. (27) that the difference can barely be seen.

In Fig. 14, we show the Wright's law-based distributional forecast, but against cumulative production. We show the forecast intervals corresponding to 1, 1.5 and 2 standard deviations (corresponding approximately to 68, 87 and 95% confidence intervals, respectively). The figure also makes clear the large scale deployment assumed by the forecast, with cumulative PV production (log) growth rate of 32% per year. Again, we note as a caveat that exponential diffusion leads to fairly high numbers as compared to expert opinions (Bosetti et al., 2012) and the academic (Gan and Li, 2015) and professional (Masson et al., 2013; International Energy Agency, 2014) literature, which generally assumes that PV deployment will slow down for a number of reasons such as intermittency and energy storage issues. But other studies (Zheng and Kammen, 2014; Jean et al., 2015) do take more optimistic assumptions as working hypothesis, and it is outside the scope of this paper to model diffusion explicitly.

5. Conclusions

We presented a method to test the accuracy and validity of experience curve forecasts. It leads to a simple method for producing distributional forecasts at different forecast horizons. We compared the experience curve forecasts with those from a univariate time series model (Moore's law of exponential progress), and found that they are fairly similar. This is due to the fact that production tends to grow exponentially[18], so that cumulative production tends to grow exponentially with low fluctuations, mimicking an exogenous exponential time trend. We applied the method to solar photovoltaic modules, showing that if the exponential trend in diffusion continues, they are likely to become very inexpensive in the near future.

There are a number of limitations and caveats that are worth reiterating here: our time series are examples from the literature so that the dataset is likely to have a strong sample bias, which limits the external validity of the results. Also, many time series are quite short, measure technical performance imperfectly, and we had to estimate initial experience in a way that is largely untested. Clearly, the experience curve model also omits important factors such as R&D. Finally, we make predictions conditional on future experience, which is not the same as doing prediction solely based on time. In settings where production is a decision variable, e.g. Way et al. (2017), and where Wright's Law is not spurious, forecasts conditional on experience are the most useful. However, it remains true that to make an unconditional forecast for a point in time in the future, using Wright's law also requires an additional assumption about the speed of technology diffusion. Thus in a situation of business as usual where experience grows exponentially, using Moore's law is simpler and almost as accurate.

The method we introduce here is closely analogous to that introduced in Farmer and Lafond (2016). Although Moore's law and Wright's law tend to make forecasts of similar quality, it is important to emphasize that when it comes to policy, the difference is potentially very important. While the correlation between costs and cumulative production is well-established, we should stress that the causal relationship is not. But to the extent that Wright's law implies that cumulative production causally influences cost, costs can be driven down by boosting cumulative production. In this case one no longer expects the two methods to make similar predictions, and the method we have introduced here plays a useful role in making it possible to think about not just what the median effect would be, but rather the likelihood of effects of different magnitudes.

Acknowledgements

This project was primarily supported by the European Commission project FP7-ICT-2013-611272 (GROWTHCOM) and by Partners for a New Economy. We also gratefully acknowledge the support from the European Commission project H2020-730427 (COP21 RIPPLES) and the Institute for New Economic Thinking. We are grateful to Giorgio Triulzi and two anonymous referees for their comments on an earlier draft.

Appendix A. Comparison of analytical results to simulations

To check whether the analytical theory is reasonable we use the following setting. We simulate 200 technologies for 50 periods. A single time series of cumulative production is generated by assuming that production follows a geometric random walk with drift $g = 0.1$ and volatility $\sigma_q = 0.1$ (no correction for previous production is made). Cost is generated assuming Wright's law with $\sigma_\eta = 0.1$, $\omega = -0.3$ and $\rho = 0.6$.

Forecast errors are computed by the hindcasting methodology, and normalized using either the true ρ or $\rho = 0$. The results are presented in Fig. 15 for $m = 5{,}40$ and for estimated or true variance ($\hat{\sigma}_v = \hat{\sigma}_\eta/\sqrt{1+\rho^2}$ or $\sigma_v = \sigma_\eta/\sqrt{1+\rho^2}$). In all cases, using the proper normalization factor $\rho = \rho^*$ makes the distribution very close to the predicted distribution (Normal or Student). When $m = 5$ and the variance is estimated, we observe a slight

[18] Recall that we selected technologies with a strictly positive growth rate of production.

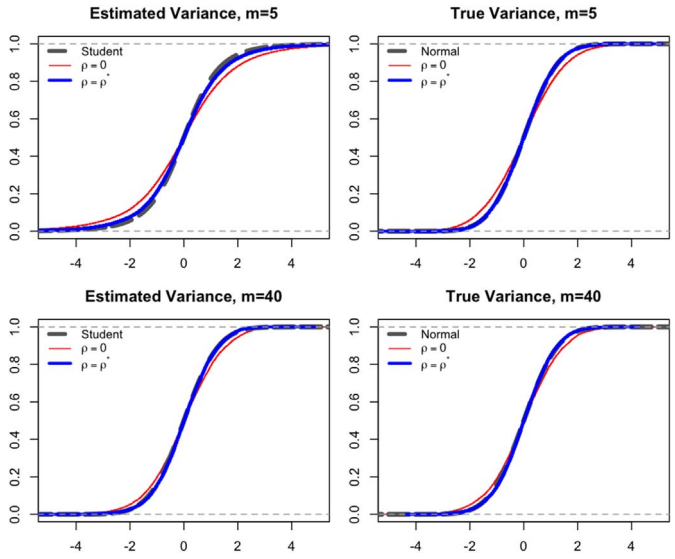

Fig. 15. Test of the theory for forecast errors. Top: $m = 5$; Bottom: $m = 40$. Left: estimated variance; Right: True variance.

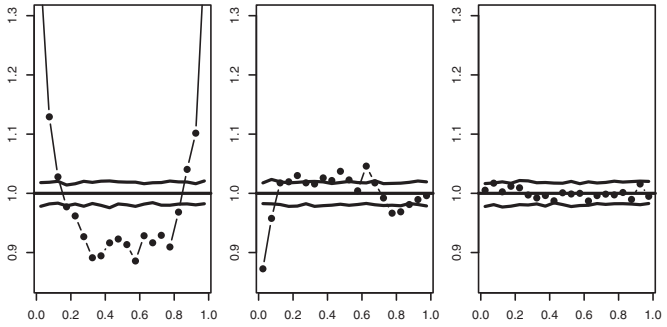

Fig. 16. Test of the theory for forecast errors. In the 3 cases $m = 5$. On the left, the variance is estimated. In the centre, errors are normalized using the true variance. On the right, we also used the true variance but the errors are i.i.d.

departure from the theory as in Farmer and Lafond (2016), which seems to be lower for large m or when the true σ_v is known.

To see the deviation from the theory more clearly, we repeat the exercise but this time we apply the probability integral transform to the resulting normalized forecast errors. We use the same parameters, and another realization of the (unique) production time series and of the (200) cost time series. As a point of comparison, we also apply the probability integral transform to randomly generated values from the reference distribution (Student when variance is estimated, Normal when the known variance is used), so that confidence intervals can be plotted. This allows us to see more clearly the departure from the Student distribution when the variance is estimated and m is small (left panel). When the true variance is used (center panel), there is still some departure but it is much smaller. Finally, for the latest panel (right), instead of generating 200 series of 50 periods, we generated 198000 time series of 7 periods, so that we have the same number of forecast errors but they do not suffer from being correlated due to the moving estimation window (only one forecast error per time series is computed). In this case we find that normalized forecast errors and independently drawn normal values are similar.

Overall these simulation results confirm under what conditions our theoretical results are useful (namely, m large enough, or knowing the true variance). For this reason, we have used the surrogate data procedure when testing the errors with small m and estimated variance, and we have used the normal approximation when forecasting solar costs based on almost 40 years of data.

Appendix B. Derivation of the properties of cumulative production

Here we give an approximation for the volatility of the log of cumulative production, assuming that production follows a geometric random walk with drift g and volatility σ_q. We use saddle point methods to compute the expected value of the log of cumulative production $E[\log Z]$, its variance $\mathrm{Var}(\log Z)$ and eventually our quantity of interest $\sigma_x^2 \equiv \mathrm{Var}(X) \equiv \mathrm{Var}(\Delta \log Z)$. The essence of the saddle point method is to approximate the integral by taking into account only that portion of the range of the integration where the integrand assumes large values. More specifically in our calculation we find the maxima of the integrands and approximate fluctuations around these points keeping quadratic and neglecting higher order terms. Assuming the initial condition $Z(0) = 1$, we can write the cumulative production at time t as $Z_t = 1 + \sum_{i=1}^{t} e^{gi+\sum_j^i a_j}$, where $a_1,\ldots,a_t$ are normally distributed i.i.d. random variables with mean zero and variance σ_q^2, describing the noise in the production process. $E[\log Z]$ is defined by the (multiple) integral over a_i

$$E(\log Z) = \int_{-\infty}^{\infty} \log Z \prod_{i=1}^{t} \frac{da_i}{\sqrt{2\pi\sigma_q^2}} \exp\left[-\frac{a_i^2}{2\sigma_q^2}\right] = \int_{-\infty}^{\infty} e^{S(\{a_i\})} \prod_{i=1}^{t} \frac{da_i}{\sqrt{2\pi\sigma_q^2}} \tag{30}$$

with $S(\{a_i\}) = \log(\log Z) - \sum_{i=1}^{t} \frac{a_i^2}{2\sigma_q^2}$, which we will calculate by the saddle point method assuming $\sigma_q^2 \ll 1$.

The saddle point is defined by the system of equations $\partial_i S(\{a_i^*\}) = 0$, $\partial_i = \partial/\partial_{a_i}$, $i = 1 \cdots t$ for which we can write

$$S(\{a_i\}) = S(\{a_i^*\}) + \sum_{ij} (a_i - a_i^*)(a_j - a_j^*) G_{ij} + \mathcal{O}(\{(a_i - a_i^*)^3\}) \tag{31}$$

where a_i^* is the solution of the saddle point equations and $G_{ij} = \frac{1}{2}\partial_i \partial_j S(\{a_i\})|_{a_i = a_i^*}$. In the saddle point approximation we restrict ourselves to quadratic terms in the expansion (31) which makes the integral (30) Gaussian. Then we obtain

$$E(\log Z) = (\det G)^{-1/2} \frac{e^{S(\{a_i^*\})}}{(2\sigma_q^2)^{t/2}} \tag{32}$$

The saddle point equation leads to $\partial_n \left(\log(\log Z) - \frac{a_n^2}{2\sigma_q^2} \right) = 0$, which can be written as

$$a_i = \sigma_q^2 \frac{\partial_i Z}{Z \log Z} = a_i^* + \mathcal{O}(\sigma_q^4)$$

$$a_i^* = \sigma_q^2 \frac{\partial_i Z}{Z \log Z}\bigg|_{a_i = 0} \tag{33}$$

Substituting this a_i^* into the $e^{S(\{a_i^*\})}$ term in Eq. (32) we obtain after some algebra

$$e^{S(\{a_i^*\})} = \left(\log Z + \frac{\sigma_q^2}{2} \frac{\sum_{i=1}^{t} (\partial_i Z)^2}{Z^2 \log Z} \right)\bigg|_{a_i = 0} + O(\sigma_q^4) \tag{34}$$

The calculation of G_{ij} as a second derivative gives

$$G_{i,j} = \frac{1}{2\sigma_q^2} \left(\delta_{i,j} + \sigma_q^2 \left(\frac{(1 + \log Z)\partial_i Z \partial_j Z}{Z^2 \log^2 Z} - \frac{\partial_i \partial_j Z}{Z \log Z} \right)\bigg|_{a_i = 0} \right) + O(\sigma_q^2), \tag{35}$$

which leads to

$$(2\sigma_q^2)^{-t/2} (\det G)^{-1/2} = 1 + \frac{\sigma_q^2}{2} \left(\frac{\sum_{i=1}^{t} \partial_i^2 Z}{Z \log Z} - \frac{\sum_{i=1}^{t} (\partial_i Z)^2 (1 + \log Z)}{Z^2 \log^2 Z} \right)\bigg|_{a_i = 0} + O(\sigma_q^4). \tag{36}$$

Here we used the formula $\det G = \exp(tr \log G)$ and an easy expansion of $\log G$ over σ_q^2. Now putting formulas (34) and (A.1) into (32) we obtain

$$E(\log Z) = \log Z|_{a_i = 0} + \frac{\sigma_q^2}{2} \sum_{i=1}^{t} \left(\frac{\partial_i^2 Z}{Z} - \frac{(\partial_i Z)^2}{Z^2} \right)\bigg|_{a_i = 0} + O(\sigma_q^4) \tag{37}$$

The calculation of Z and its derivatives at $a_i = 0$ is straightforward. If $g > 0$, for large t it gives the very simple formula

$$E(\log Z(t))|_{t \to \infty} = g(t + 1) - \log(e^g - 1) + \frac{\sigma_q^2}{4 \sinh(g)} + O(\sigma_q^4) \tag{38}$$

With the same procedure as for Eqs. (30)–(38) we calculate the expectation value of $E(\log {}^2 Z)$, $E(\log Z(t) \log Z(t+1)) - E(\log Z(t)) E(\log Z(t+1))$ which leads to similar formulas as Eqs. (A.1)–(37), but with different coefficients. The result for $g > 0$ and $t \to \infty$ reads

$$\mathrm{Var}(\log Z(t)) = E(\log^2 Z) - E(\log Z)^2 = \sigma_q^2 \left(\frac{2e^g + 1}{1 - e^{2g}} + t \right) + O(\sigma_q^4) \tag{39}$$

and

$$\mathrm{Var}(\Delta \log Z) = E((\log Z(t+1) - \log Z(t))^2) - (E(\log Z(t+1) - \log Z(t)))^2 = \sigma_q^2 \tanh\left(\frac{g}{2}\right) + O(\sigma_q^4). \tag{40}$$

References

Alberth, S., 2008. Forecasting technology costs via the experience curve: myth or magic? Technol. Forecast. Soc. Chang. 75 (7), 952–983.

Alchian, A., 1963. Reliability of progress curves in airframe production. Econometrica 31 (4), 679–693.

Anzanello, M.J., Fogliatto, F.S., 2011. Learning curve models and applications: literature review and research directions. Int. J. Ind. Ergon. 41 (5), 573–583.

Argote, L., Beckman, S.L., Epple, D., 1990. The persistence and transfer of learning in industrial settings. Manag. Sci. 36 (2), 140–154.

Arrow, K.J., 1962. The economic implications of learning by doing. Rev. Econ. Stud. 155–173.

Ayres, R.U., 1969. Technological Forecasting and Long-range Planning. McGraw-Hill Book Company.

Bailey, A., Bui, Q.M., Farmer, J.D., Margolis, R., Ramesh, R., 2012. Forecasting

technological innovation. In: ARCS Workshops (ARCS), 2012, pp. 1–6.
Benson, C.L., Magee, C.L., 2015. Quantitative determination of technological improvement from patent data. PloS One 10 (4) e0121635.
Berndt, E.R., 1991. The Practice of Econometrics: Classic and Contemporary. Addison-Wesley Reading, MA.
Bettencourt, L.M., Trancik, J.E., Kaur, J., 2013. Determinants of the pace of global innovation in energy technologies. PloS One 8 (10), e67864.
Bosetti, V., Catenacci, M., Fiorese, G., Verdolini, E., 2012. The future prospect of PV and CSP solar technologies: an expert elicitation survey. Energy Policy 49, 308–317.
Boston Consulting Group, 1972. Perspectives on Experience, 3 edn. The Boston Consulting Group, Inc., One Boston Place, Boston, Massachusetts 02106.
Candelise, C., Winskel, M., Gross, R.J., 2013. The dynamics of solar PV costs and prices as a challenge for technology forecasting. Renew. Sust. Energ. Rev. 26, 96–107.
Clark, T., McCracken, M., 2013. Advances in forecast evaluation. In: Handbook of Economic Forecasting. 2. pp. 1107–1201.
Clements, M.P., Hendry, D.F., 2001. Forecasting with difference-stationary and trend-stationary models. Econ. J. 4 (1), 1–19.
Colpier, U.C., Cornland, D., 2002. The economics of the combined cycle gas turbine, an experience curve analysis. Energy Policy 30 (4), 309–316.
de La Tour, A., Glachant, M., Ménière, Y., 2013. Predicting the costs of photovoltaic solar modules in 2020 using experience curve models. Energy 62, 341–348.
Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. Int. Econ. Rev. 863–883.
Dutton, J.M., Thomas, A., 1984. Treating progress functions as a managerial opportunity. Acad. Manage. Rev. 9 (2), 235–247.
Farmer, J.D., Lafond, F., 2016. How predictable is technological progress? Res. Policy 45 (3), 647–665.
Farmer, J., Makhijani, A., 2010. A U.S. nuclear future: not wanted, not needed. Nature 467, 391–393.
Ferioli, F., Van der Zwaan, B., 2009. Learning in times of change: a dynamic explanation for technological progress. Environ. Sci. Technol. 43 (11), 4002–4008.
Funk, J.L., Magee, C.L., 2015. Rapid improvements with no commercial production: how do the improvements occur? Res. Policy 44 (3), 777–788.
Gallant, A., 1968. A note on the measurement of cost/quantity relationships in the aircraft industry. J. Am. Stat. Assoc. 63 (324), 1247–1252.
Gan, P.Y., Li, Z., 2015. Quantitative study on long term global solar photovoltaic market. Renew. Sust. Energ. Rev. 46, 88–99.
Goldberg, M.S., Touw, A., 2003. Statistical Methods for Learning Curves and Cost Analysis. Institute for Operations Research and the Management Sciences (INFORMS).
Goldemberg, J., Coelho, S.T., Nastari, P.M., Lucon, O., 2004. Ethanol learning curve, the Brazilian experience. Biomass Bioenergy 26 (3), 301–304.
Hall, B.H., Mairesse, J., 1995. Exploring the relationship between R&D and productivity in French manufacturing firms. J. Econometr. 65 (1), 263–293.
Hall, G., Howell, S., 1985. The experience curve from the economist's perspective. Strateg. Manag. J. 6 (3), 197–212.
Harvey, A.C., 1980. On comparing regression models in levels and first differences. Int. Econ. Rev. 707–720.
Hutchby, J.A., 2014. A “Moore's law”-like approach to roadmapping photovoltaic technologies. Renew. Sust. Energ. Rev. 29, 883–890.
International Energy Agency, 2014. Technology roadmap: solar photovoltaic energy (2014 ed.). In: Technical report, OECD/IEA, Paris.
Isoard, S., Soria, A., 2001. Technical change dynamics: evidence from the emerging renewable energy technologies. Energy Econ. 23 (6), 619–636.
Jamasb, T., 2007. Technical change theory and learning curves: patterns of progress in electricity generation technologies. Energy J. 28 (3), 51–71.
Jean, J., Brown, P.R., Jaffe, R.L., Buonassisi, T., Bulović, V., 2015. Pathways for solar photovoltaics. Energy Environ. Sci. 8 (4), 1200–1219.
Junginger, M., van Sark, W., Faaij, A., 2010. Technological Learning in the Energy Sector: Lessons for Policy, Industry and Science. Edward Elgar Publishing.
Kahouli-Brahmi, S., 2009. Testing for the presence of some features of increasing returns to adoption factors in energy system dynamics: an analysis via the learning curve approach. Ecol. Econ. 68 (4), 1195–1212.
Lieberman, M.B., 1984. The learning curve and pricing in the chemical processing industries. RAND J. Econ. 15 (2), 213–228.
Lipman, T.E., Sperling, D., 1999. Forecasting the costs of automotive pem fuel cell systems: using bounded manufacturing progress functions. In: Paper presented at IEA International Workshop, Stuttgart, Germany, 10–11 May..
Magee, C.L., Basnet, S., Funk, J.L., Benson, C.L., 2016. Quantitative empirical trends in technical performance. Tech. Forcasting Soc. Chang. 104, 237–246.
Martino, J.P., 1993. Technological Forecasting for Decision Making. McGraw-Hill, Inc.
Masson, G., Latour, M., Rekinger, M., Theologitis, I.-T., Papoutsi, M., 2013. Global Market Outlook for Photovoltaics 2013–2017. European Photovoltaic Industry Association, pp. 12–32.
McDonald, A., Schrattenholzer, L., 2001. Learning rates for energy technologies. Energy Policy 29 (4), 255–261.
McDonald, J., 1987. A new model for learning curves, DARM. J. Bus. Econ. Stat. 5 (3), 329–335.
Meese, R.A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: do they fit out of sample? J. Int. Econ. 14 (1), 3–24.
Moore, G.E., 2006. Behind the Ubiquitous Microchip.
Nagy, B., Farmer, J.D., Bui, Q.M., Trancik, J.E., 2013. Statistical basis for predicting technological progress. PloS One 8 (2) e52669.
Neij, L., 1997. Use of experience curves to analyse the prospects for diffusion and adoption of renewable energy technology. Energy Policy 25 (13), 1099–1107.
Neij, L., Andersen, P.D., Durstewitz, M., Helby, P., Hoppe-Kilpper, M., Morthorst, P., 2003. Experience Curves: A Tool for Energy Policy Assessment. Environmental and Energy Systems Studies, Univ.
Nemet, G.F., 2006. Beyond the learning curve: factors influencing cost reductions in photovoltaics. Energy Policy 34 (17), 3218–3232.
Nordhaus, W.D., 2014. The perils of the learning model for modeling endogenous technological change. Energy Journal 35 (1), 1–13.
Papineau, M., 2006. An economic perspective on experience curves and dynamic economies in renewable energy technologies. Energy Policy 34 (4), 422–432.
Sahal, D., 1979. A theory of progress functions. AIIE Trans. 11 (1), 23–29.
Sampson, M., 1991. The effect of parameter uncertainty on forecast variances and confidence intervals for unit root and trend stationary time-series models. J. Appl. Economet. 6 (1), 67–76.
Schaeffer, G.J., Seebregts, A., Beurskens, L., Moor, H. d., Alsema, E., Sark, W. v., Durstewicz, M., Perrin, M., Boulanger, P., Laukamp, H., et al., 2004. Learning from the sun; analysis of the use of experience curves for energy policy purposes: the case of photovoltaic power. Final report of the photex project. In: Report ECN DEGO: ECN-C-04-035, ECN Renewable Energy in the Built Environment.
Schilling, M.A., Esmundo, M., 2009. Technology s-curves in renewable energy alternatives: analysis and implications for industry and government. Energy Policy 37 (5), 1767–1781.
Sinclair, G., Klepper, S., Cohen, W., 2000. What's experience got to do with it? Sources of cost reduction in a large specialty chemicals producer. Manag. Sci. 46 (1), 28–45.
Söderholm, P., Sundqvist, T., 2007. Empirical challenges in the use of learning curves for assessing the economic prospects of renewable energy technologies. Renew. Energy 32 (15), 2559–2578.
Swanson, R.M., 2006. A vision for crystalline silicon photovoltaics. Prog. Photovolt. Res. Appl. 14 (5), 443–453.
Thompson, P., 2010. Learning by doing. In: Handbook of the Economics of Innovation. 1. pp. 429–476.
Thompson, P., 2012. The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing. J. Econ. Perspect. 26 (3), 203–224.
Van der Zwaan, B., Rabl, A., 2004. The learning potential of photovoltaics: implications for energy policy. Energy Policy 32 (13), 1545–1554.
Van Sark, W., 2008. Introducing errors in progress ratios determined from experience curves. Technol. Forecast. Soc. Chang. 75 (3), 405–415.
Vigil, D.P., Sarper, H., 1994. Estimating the effects of parameter variability on learning curve model predictions. Int. J. Prod. Econ. 34 (2), 187–200.
Way, R., Lafond, F., Farmer, J.D., Lillo, F., Panchenko, V., 2017. Wright Meets Markowitz: How Standard Portfolio Theory Changes when Assets are Technologies Following Experience Curves. arxiv 1705.03423.
West, K.D., 2006. Forecast evaluation. In: Handbook of Economic Forecasting. 1. pp. 99–134.
Witajewski-Baltvilks, J., Verdolini, E., Tavoni, M., 2015. Bending the learning curve. Energy Econ. 52, S86–S99.
Womer, N.K., Patterson, J.W., 1983. Estimation and testing of learning curves. J. Bus. Econ. Stat. 1 (4), 265–272.
Wright, T.P., 1936. Factors affecting the cost of airplanes. J. Aeronaut. Sci. 3 (4), 122–128.
Yelle, L.E., 1979. The learning curve: historical review and comprehensive survey. Dec. Sci. 10 (2), 302–328.
Zhao, J., 1999. The Diffusion and Costs of Natural Gas Infrastructures.
Zheng, C., Kammen, D.M., 2014. An innovation-focused roadmap for a sustainable global photovoltaic industry. Energy Policy 67, 159–169.

Dr. François Lafond is a researcher at the Institute for New Economic Thinking at the Oxford Martin School and at the Smith School for Enterprise and the Environment.

Dr. Aimee Bailey is a Principal Energy Analyst at EDF Innovation Lab, the Silicon Valley R&D outpost of EDF.

Mr. Jan Bakker is a doctoral student at the University of Oxford.

Mr. Dylan Rebois works as a management consultant.

Ms. Rubina Zadourian is a doctoral student at Max Planck Institute.

Patrick McSharry is a Professor at Carnegie Mellon University Africa, Director of the African Center of Excellence in Data Science, University of Rwanda and an Associate Member of the Oxford Man Institute of Quantitative Finance and the Oxford Internet Institute, Oxford University.

J. Doyne Farmer is Director of the Complexity Economics program at the Institute for New Economic Thinking at the Oxford Martin School, Professor in the Mathematical Institute at the University of Oxford, and an External Professor at the Santa Fe Institute.