Online Appendix A: Supplementary Figures

Figure A1: Hypothesized Paths of Technological Evolution

Figure A1a: Exponential curve representing the Moore's and Kryder's Law

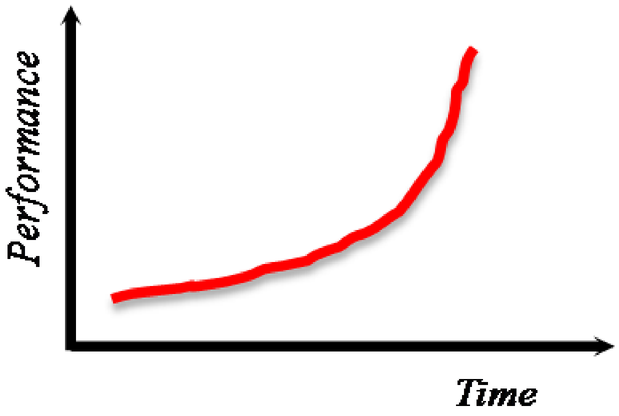


Figure A1b: Sigmoid curve representing the Logistic, Bass and Gompertz Law

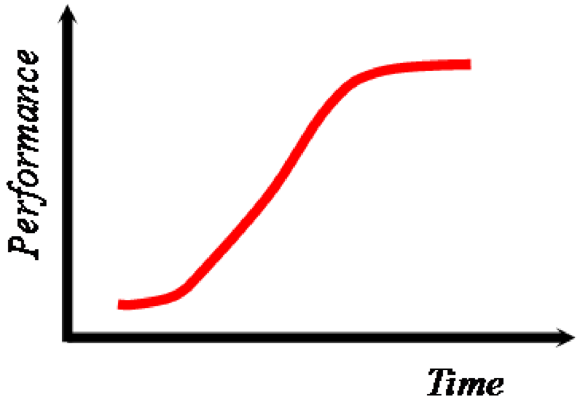


Figure A1c: Step functions representing the Gupta, Tobit II, and SAW models

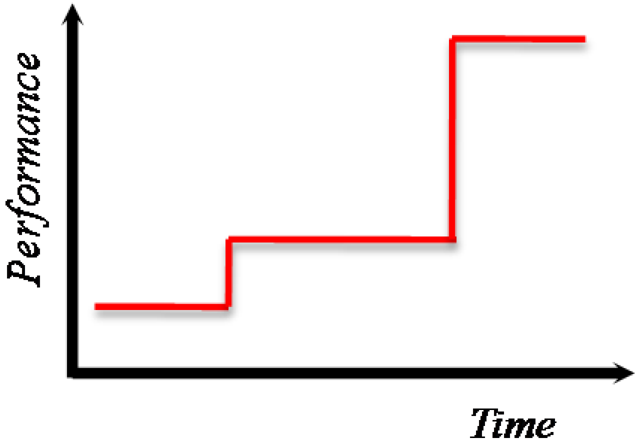


Figure A2: Experimental Setup

Figure A2a: Sampling of Technologies and Time Periods For Prediction

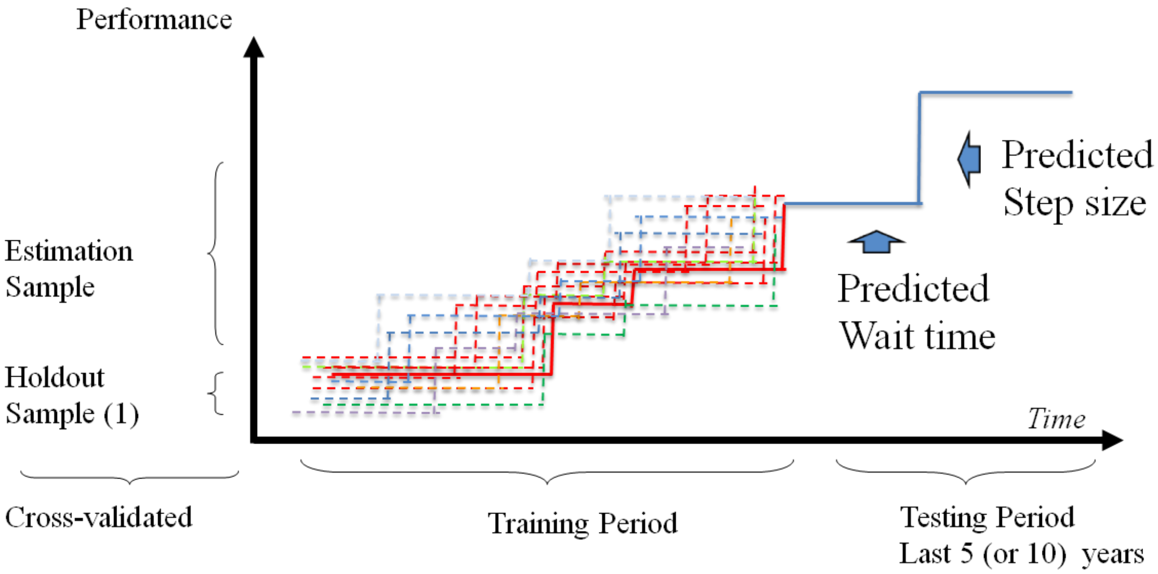


Figure A2b: Unconstrained Fits

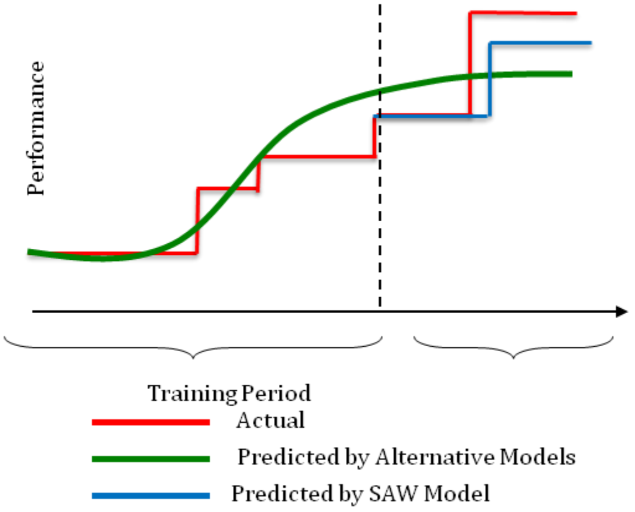


Figure A2c: Constrained Fits

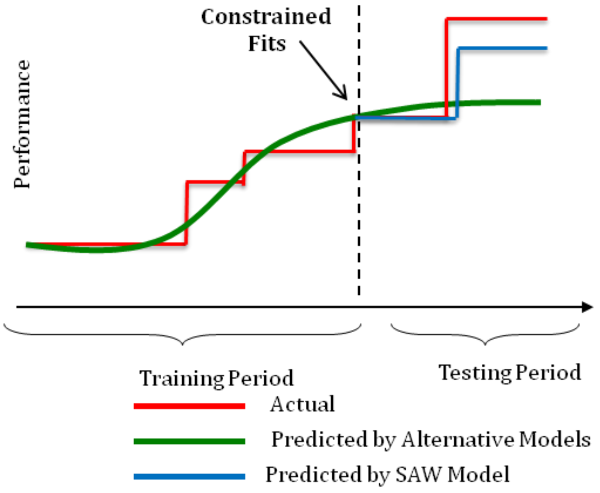


Figure A3: Estimated Parameters (Step Size and Wait Time)

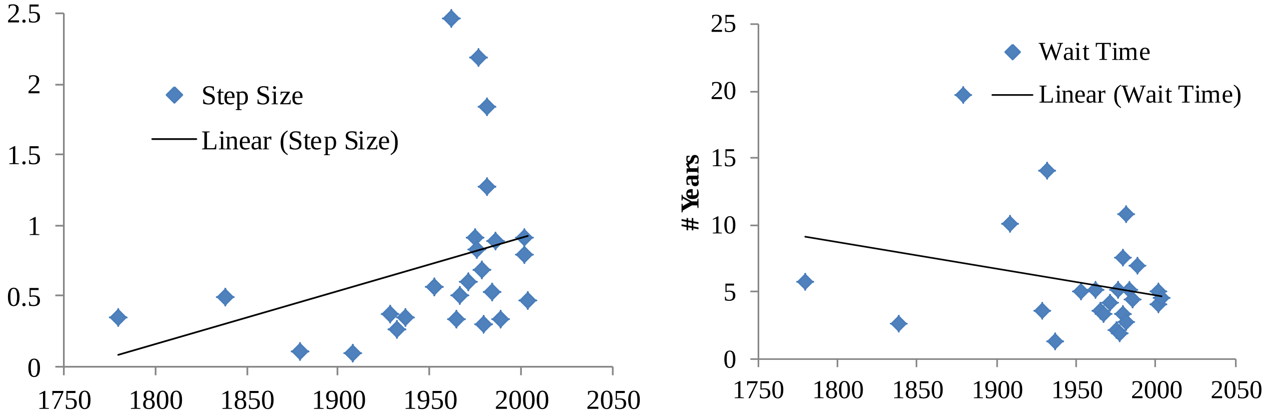
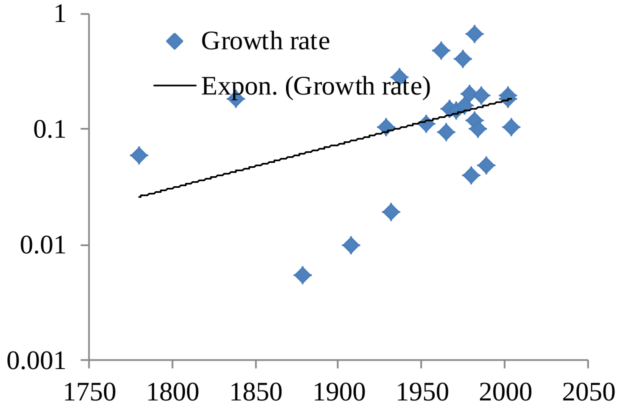


Figure A3a: Size of Steps **Figure A3b: Wait Times**

Figure A3c: Growth Rate: Log (Ratio of Steps and Waits)



Online Appendix B:

Prediction and Fitting of Comparison Models

In their standard forms, Moore's law, Kryder's law and the logistic, Bass, and Gompterz model do not directly incorporate covariates into their predictions. In order to provide a comparison with SAW, which does allow for the inclusion of covariates, we fit two modified versions of these methods. In the first implementation we used a non-linear mixed effects model which fit the standard functional forms of each method but modeled the various parameters as random effects coming from a Gaussian distribution. The parameters for the Gaussian distribution were estimated using all technologies simultaneously and hence built strength in a similar fashion to SAW. Our second implementation of these methods also modeled the parameters using a random effects formulation but in addition incorporated the covariates as a multiplicative adjustment to the original prediction using (11). Next, we discuss both the *Mixed Effects Model* and the *Mixed Effects Model with Covariate Effects* for each method. By comparison the Gupta and Tobit II models involve covariates so must be fit to all curves simultaneously to estimate the population level covariate coefficients.

Extensions of Moore's Law and Kryder's Law (Exponential Model)

Mixed Effects Model

Moore's Law and Kryder's Law each state that the rate of change in the performance of a technology is exponential with given constants. Thus, if we model performance of the technology as an exponential function of time, the coefficient of time represents the constant rate at which the technology improves. To test the applicability of Moore's Law and Kryder's Law, we model the relationship between time and technology performance using the following exponential function:

$$P_{ij} = \tau_{i1} e^{\tau_{i2} t_{ij}} e^{\epsilon_{ij}} \qquad (14)$$

where, P_{ij} is performance for technology i at time t_{ij}, and τ_{i1} and τ_{i2} are modeled as random variables coming from Gaussian distributions.

We estimate the coefficients for the exponential model associated with Moore's Law and Kryder's Law using linear mixed effects software applied to the log-transformed data. Kryder's Law assumes $\tau_2 = \frac{12}{13}\log 2$ while Moore's Law assumes $\tau_2 = \frac{12}{18}\log 2$. After fitting the mixed effects model to each technology we use the fitted parameters to form predictions and compare the estimate for $\tau_{i2}$with that predicted by Kryder's and Moore's Laws.

Mixed Effects Model with Covariate Effects

Let $X_{ij1}, X_{ij2}, \ldots, X_{ijq}$ represent q different covariates measured at time t_{ij}. We incorporate these covariates into the exponential model in a multiplicative fashion using the following formulation,

$$P_{ij} = \tau_{i1} e^{\tau_{i2} t_{ij}} \times \exp\left(\beta_0 + \sum_{k=1}^{q} \beta_k X_{ijk}\right) e^{\epsilon_{ij}}, \quad \tau_{i1} \sim N(\mu_1, \sigma_1^2), \;\; \tau_{i2} \sim N(\mu_2, \sigma_2^2) \tag{15}$$

Note that the parameter τ_{i1} are modeled as random effects because they are specific to a particular technology but the β coefficeints are treated as fixed effects because they are common to all technologies.

Extensions of Logistic Model

Mixed Effects Model

The generalized form of the logistic curve is a relatively flexible model that can also capture an S-shape of the path of technological evolution. This model has been widely used in prior literature (e.g. Young 1993; Meade and Islam 1998).

$$P_{ij} = \frac{\tau_{i3}}{1 + \tau_{i4} \times \exp(-\tau_{i5} t)} e^{\epsilon_{ij}} \tag{16}$$

where, P_{ij} is the performance of technology i at time t_{ij}, and τ_3, τ_4 and τ_5 are modeled as random variables coming from Gaussian distributions.

We fit equation (16) using non-linear mixed effects software applied to the log-transformed data.

Mixed Effects Model with Covariate Effects

We incorporate the covariates into the logistic model using a multiplicative formulation, thus

$$P_{ij} = \frac{\tau_{i3}}{1+\tau_{i4}\times\exp(-\tau_{i5}t_{ij})} \times \exp\left(\beta_0 + \sum_{k=1}^{q} \beta_k X_{ijk}\right) e^{\epsilon_{ij}} \tag{17}$$

Note that the parameters $\tau_{31}, \tau_{i4}, \tau_{i5}$ are modeled as random effects because they are specific to a particular technology but the β coefficients are treated as fixed effects because they are common to all technologies.

Extensions of Bass Model

Mixed Effects Model

The Bass model (Bass 1969) is a special case of the Gamma/shifted-Gompertz distribution that can capture an S-shape plus a variety of other shapes that approximate the S-curve depending on the values of the parameters. We use the operational form of the Bass model used previously for modeling technology evolution (Young and Ord 1989; Young 1993):

$$p_{ij} = \left(\tau_{i6} + \tau_{i7}P_{ij} + \tau_{i8}P_{ij}^2\right)e^{\epsilon_{ij}} \tag{18}$$

where, P_{ij} is the performance of technology i at time t_{ij},

p_{ij} is marginal performance improvement at time t_{ij},

τ_{i6}, τ_{i7} and τ_{i8} are modeled as random variables coming from a Gaussian distribution. We fit equation (18) using non-linear mixed effects software applied to the log-transformed data.

Mixed Effects Model with Covariate Effects

We incorporate the covariates into the Bass model using a multiplicative formulation, thus:

$$p_{ij} = (\tau_{i6} + \tau_{i7} P_{ij} + \tau_{i8} P_{ij}^2) \times \exp\left(\beta_0 + \sum_{k=1}^{q} \beta_k X_{ijk}\right) e^{\epsilon_{ij}} \tag{19}$$

Note that the parameters $\tau_{i6}, \tau_{i7}, \tau_{i8}$ are modeled as random effects because they are specific to a particular technology but the β coefficients are treated as fixed effects because they are common to all technologies.

Extensions of Gompertz' Model

Mixed Effects Model

The Gompertz' Model used to estimate Gompertz' Law takes the functional form:

$$P'_{ij} = r P_{ij} \log \frac{K}{P_{ij}} \tag{20}$$

Where, P_{ij} is the performance of technology i at time t_{ij},

P'_{ij} is the corresponding derivative of the performance,

r is the intrinsic growth rate,

K is the final technology level.

Notice that this equation gives slow growth when P_{ij} is either low or close to K, and rapid growth in between. The solution to this differential equation is the following double exponential function:

$$P_{ij} = \tau_{i9} \exp(-\tau_{i10} e^{-\tau_{i11} t_{ij}})\, e^{\epsilon_t} \tag{21}$$

where τ_9, τ_{10} and τ_{11} are modeled as random variables coming from a Gaussian distribution. We fit equation (21) using non-linear mixed effects software applied to the log-transformed data.

Mixed Effects Model with Covariate Effects

We incorporate the covariates into the Gompertz model using a multiplicative formulation, thus:

$$P_{ij} = \tau_{i9}\exp(-\tau_{i10}e^{-\tau_{i11}t_{ij}}) \times \exp\left(\beta_0 + \sum_{k=1}^{q}\beta_k X_{ijk}\right)e^{\epsilon_t} \tag{22}$$

Note that the parameters $\tau_{i9}, \tau_{i10}, \tau_{i11}$ are modeled as random effects because they are specific to a particular technology but the β coefficients are treated as fixed effects because they are common to all technologies.

Constrained Parameters

Several of the parameters in the above mentioned models are constrained to by positive. To operationalize this constraint we parameterized the corresponding coefficients as $\tau = \exp(\tau^*)$ where τ^* was modeled as coming from a Gaussian distribution. This formulation ensured that τ would always be positive.

Extensions of Gupta Model

The interpurchase time model uses an Erlang-2 distribution to model the time until a purchase, or in our case wait time until a jump in technology, with the Erlang parameter modeled as a function of a set of explanatory variables. Specifically,

$$f_{ij}(t) = \alpha_{ij}^2 t \exp(-\alpha_{ij}t) \tag{23}$$

$$\alpha_{ij} = \exp(-\delta_0 - \delta_1 X_{ij1} - \cdots - \delta_p X_{ijp}) \tag{24}$$

where $f_{ij}(t)$ is the probability density, at time t_{ij}, of the wait until the next jump for technology i, α_{ij} is the Erlang-2 scale parameter and X_{ijk} is the kth covariate at t_{ij}. The time until purchase, or technology jump, is predicted using the Erlang-2 mean, $\frac{2}{\alpha_{ij}}$.

Gupta (1988) models the purchase quantity using a logistic distribution because his sales data is categorical in terms of size e.g. 16 oz, 32 oz, etc. However, our jump size data is continuous so we follow Gupta (1988)'s recommendation to model such data using a standard linear regression, with explanatory variables $Y_{ij1}, Y_{ij2}, \dots, Y_{ijq}$. Consistent with our other comparison models, we use the log transformed jump size as the dependent variable.

The interpurchase time and purchase quantity models are fit separately. The interpurchase time model is fit using an iterative reweighted least squares algorithm that maximizes the likelihood function associated with the Erlang-2 distribution. The purchase quantity model is fit using a standard linear regression least squares procedure with the log transformed jumps as the dependent variable and the values of the covariates at the associated time points as the independent variables. Combining the interpurchase time model, which predicts time until the next jump, with the purchase quantity model, which predicts jump size, we can use the explanatory variables to estimate the remaining evolution for any given technology.

Tobit II Model

For the i^{th} technology, at time t_{ij} , Tobit II models the probability of a step (P_1^*) as a function of explanatory variables $X_{ij1}, X_{ij2}, \ldots, X_{ijp}$, and the size of the step (P_2^*) as a function of explanatory variables $Y_{ij1}, Y_{ij2}, \ldots, Y_{ijq}$ thus:

$$P_{1ij}^* = \frac{e^{\delta_0+\delta_1 X_{ij1}+\cdots+\delta_p X_{ijp}}}{1+e^{\delta_0+\delta_1 X_{ij1}+\cdots+\delta_p X_{ijp}}} + U_{ij} \tag{25}$$

(26) $$P_{2ij}^{*} = \chi_0 + \chi_1 Y_{ij1} + \cdots + \chi_q Y_{ijq} + V_{ij}$$

where $$(U, V) \sim N\left(0, \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}\right)$$

We observe a step conditional on the probability of a jump exceeding a cutoff value, for example $P_{1ij} = 1(P_{1ij}^{*} > 0.1)$ always, but observe the size of step P_{2ij} only when the step occurs i.e., $P_{1ij} = 1$. We use standard software to estimate the joint outcome – probability and size of step – in the Tobit II model (Tellis 1988).

Online Appendix C: Fitting of SAW

We provide details on fitting of SAW:

Fitting of SAW

Fitting SAW requires estimating a number of parameters. We use a maximum likelihood approach. Suppose we have observed n_i steps of technology *i* at times $t_{i1},\ldots,t_{in_i}$. Let $J_{i.} = (J_{i1}, \ldots, J_{in_i})$ represent the series of observed steps and $T_{i.} = (T_{i1}, \ldots, T_{in_i})$ be the times between these steps. In addition we assume covariates X_{ijk} and Y_{ijk} have been observed at times t_{ij}.

Conditional on λ_i and ω the distribution of $T_{i\cdot}$ is

$$f(T_{i.}|\lambda_i, \omega) = \frac{1}{\Gamma(\mathrm{K})^{n_i}} \left(\prod_{j=1}^{n_i} \omega_{i(j-1)}^{-K} T_{ij}^{K-1} \right) \lambda_i^{-Kn_i} \exp\left(-\lambda_i^{-1} \sum_{j=1}^{n_i} \omega_{i(j-1)}^{-1} T_{ij} \right)$$

Similarly, the distribution of λ_i^{-1} conditional on κ and ω is

$$f(\lambda_i^{-1}|\kappa, \theta) = \frac{\lambda_i^{-(\kappa-1)}}{\Gamma(\kappa)\theta^\kappa} \exp\left(\frac{-\lambda_i^{-1}}{\theta}\right)$$

Hence, the distribution of $T_{i\cdot}$ conditional on ω, κ, K and θ is

(27)

$$f(T_{i.}|\kappa, \theta, \omega, K) =$$

$$\int \frac{1}{\Gamma(\mathrm{K})^{n_i}} \left(\prod_{j=1}^{n_i} \omega_{i(j-1)}^{-K} T_{ij}^{K-1} \right) \lambda_i^{-Kn_i} \exp\left(-\lambda_i^{-1} \sum_{j=1}^{n_i} \omega_{i(j-1)}^{-1} T_{ij} \right) \frac{\lambda_i^{-(\kappa-1)}}{\Gamma(\kappa)\theta^\kappa} \exp\left(\frac{-\lambda_i^{-1}}{\theta}\right) d\lambda_i^{-1}$$

$$= \frac{\frac{\Gamma(Kn_i + \kappa)}{\Gamma(\mathrm{K})^{\mathrm{n_i}}\Gamma(\kappa)} \left(\prod_{j=1}^{n_i} \theta^K T_{ij}^{K-1} \omega_{i(j-1)}^{-1} \right)}{\left(1 + \sum_{j=1}^{n_i} \theta \omega_{i(j-1)}^{-1} T_{ij} \right)^{\kappa + Kn_i}}$$

We can use Equation (27) to write down the log likelihood function,

$$(28)\qquad l_T(\kappa,\theta,\beta,K)=\sum_{i=1}^{N}\log f(T_{i.}|\kappa,\theta,\omega,K)=\sum_{i=1}^{N}\log\left(\frac{\Gamma(Kn_i+\kappa)}{\Gamma(K)^{n_i}\Gamma(\kappa)}\right)-\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left((1-K)\log(T_{ij})+K(sJ_{i(j-1)}+\beta_0+\sum_{k=1}^{p}\beta_k X_{i(j-1)k})\right)-\sum_{i=1}^{N}(\kappa+Kn_i)log\Big(1+\sum_{j=1}^{n_i}T_{ij}\exp\big(-(sJ_{i(j-1)}+\beta_0+\sum_{k=1}^{p}\beta_k X_{i(j-1)k})\big)\Big)$$

where $\theta=\exp(-\beta_0)$. Equation (28) is a convex function provided the n_i's are large enough. Hence standard optimization techniques can maximize Equation (28) in terms of κ and the β's.

An analogous argument shows that the log likelihood function for the Step sub-model is

$$(29)\qquad l_J(\rho,\eta,\alpha,M)=\sum_{i=1}^{N}\log f(J_{i.}|\rho,\eta,\alpha,M)=\sum_{i=1}^{N}\log\left(\frac{\Gamma(Mn_i+\rho)}{\Gamma(M)^{n_i}\Gamma(\rho)}\right)-\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left((1-M)\log(J_{ij})+M(rT_{ij}+\alpha_0+\sum_{k=1}^{q}\alpha_k Y_{ijk})\right)-\sum_{i=1}^{N}(\rho+Mn_i)log\Big(1+\sum_{j=1}^{n_i}J_{ij}\exp\big(-(rT_{ij}+\alpha_0+\sum_{k=1}^{q}\alpha_k Y_{ijk})\big)\Big)$$

where $\eta=\exp(-\alpha_0)$, which can similarly be optimized by standard techniques. We calculate maximum likelihood estimates for all the parameters and produce future predictions using Equations (8) and (9). Note that the joint log likelihood of both T and J is equal to,

$$l_{T,J}(\kappa,\theta,\beta,K,\rho,\eta,\alpha,M)=\sum_{i=1}^{N}\log f(T_{i.},J_{i.}|\kappa,\theta,\omega,K,\rho,\eta,\alpha,M)$$

$$=\ l_T(\kappa,\theta,\beta,K)+l_J(\rho,\eta,\alpha,M).$$

Hence, the joint likelihood is separable into the sum of the Wait and Step likelihoods.

Comparison of one step and two step approaches to estimation

As a direct consequence the two step approach of maximizing the Wait and Step likelihoods individually is mathematically identical to the one step approach of maximizing the joint SAW likelihood. This is an advantage of the SAW model because it significantly reduces the fitting procedure's complexity.