Online Appendix A:
Procedure

We follow the general rules for data collection for the historical method. We explain eight specific problems we encountered and the rules we used to resolve them. First, we find that the scientific principles on which new technologies are based are often discovered in laboratories, years and sometimes decades before products based on these scientific principles are commercialized. We identify the year of entry of a new technology based on when a new product based on a new technology was first commercialized in the market even if the scientific principle may have been known or described in a technical paper earlier.

Second, we find that many firms make various announcements about new products at various stages of research and development before launching their products. We identify the first firm to introduce a commercialized product in the market even if announcements of prototypes have been reported by other firms before the commercialization. Moreover, we find that in some cases, a number of firms collaborate on the joint research for a new technology and then share licensing rights to products employing that technology. For example, Universal Display Corporation was the forerunner in initial research on OLED display technology. At that time, Universal Display Corporation had Sony, Samsung, DuPont, Motorola, Toyota, Pioneer and the U.S. Army among its strategic partners. We identify the first firm to commercialize the product as the source of new technology in such cases.

Third, the performance of a technology affects the performance of various products based on that technology. To measure the performance of a technology at a point in time, we use the performance of the best-performing product among all products based on that technology. For example, a hard disk has higher storage capacity than a floppy disk even though both are based on the same magnetic storage technology. We searched for records of performance on the primary dimension at the time of entry of each technology in both absolute terms and relative to its rival technologies. We used this precise information to identify whether the technology made an upper or a lower attack at entry. In some cases only cumulative improvements over a period of data are available. We assume a uniform rate of improvement to estimate average annual rate of improvement in such cases.

Fourth, many times a firm loses market share because of competitive pressures within the market not directly related to the threat posed by the new technology. In such cases, if the highest market share passes from one firm to another without a change in the dominant technology on which the products are based, we do not consider it a firm disruption as it is not caused directly by the new technology.

Fifth, we identify the incumbency status of a firm introducing a new technology based on whether the firm was competing in the same market before that introduction. For example, we classify Philips as an incumbent when it introduced MED lamps, because Philips also offered incandescent lamps prior to offering MED lamps. On the other hand, we classify Optware as an entrant because it was founded to develop and commercialize holographic storage. Thus, some new technologies may be introduced by incumbents who may have only a minor market share while others may be introduced by entrants; the new technology may or may not disrupt the dominant technology over time.

Sixth, we collect annual data on firm size for the analyses. However, reliable data on the size of firms is not available for all the firms in the sample at the time disruption occurred especially for firms in the early part of the 20th century and for non-US based firms. Since disruption is already a rare event, we cannot afford to lose these observations because of the lack of data on firm size. Hence, we use a dummy variable to indicate the firm size at entry for the entire analysis. We classify a firm as small if the firm was new and formed only to introduce the new technology.

Seventh, we track whether new technologies were acquired by an incumbent. We did not find any cases where a new technology was acquired before causing firm disruption.

Eighth, for each market, we exclude any early periods for which there is only one technology in a market, because disruption requires at least two technologies.

Online Appendix B:
Operating Principles of Sampled Technologies

Electrical Lighting

1. **Incandescent**: Generates light by heating up thin metallic wires with an electric current
2. **Arc Discharge**: Emits light by arc formed between two electrodes oppositely charged by an electric current in a high-pressure gas chamber
3. **Gas Discharge**: Electrons excited by passing an electric current in a low-pressure gas chamber emit light
4. **LED**: Emits light in n-p transition zone under influence of an electric potential
5. **MED**: Emits light by microwaves from induction coil inside the bulb to excite the gas.

Computer Memory

6. **Punched Cards**: Records data by punching holes into paper or cardboard medium, which can be read electrically or optically
7. **Magnetic**: Records data by passing a frequency modulated (FM) current through the disk drive's magnetic head that magnetizes the particles of the disk's recording surface.
8. **Vacuum Tubes**: Records data by an electronic device that controls the flow of electrons in a vacuum.
9. **Electrostatic Storage tubes**: Records data by writing a grid of dots and dashes (or dim and bright dots) to the CRT, sensed by the collector plate that flips up directly over the CRT (Williams Tube).
10. **Acoustic Delay Lines**: Records data by a simple acoustic delay line, consisting of a delay medium and two transducers.
11. **Semiconductor**: Records data in static memory, where each cell uses a flip-flop made from four or six transistors.
12. **Optical**: Stores data using the laser modulation system, where changes in reflectivity are used to store and retrieve data.
13. **Magneto-optical**: Records data using the magnetic-field modulation system and reads the data with a laser beam.
14. **Holographic storage**: Records data using collinear holography, whereby two lasers, one red and one blue-green, are collimated in a single beam.

Computer Printers

15. **Impact**: Forms an image by striking a hammer against an ink ribbon to print closely spaced dots that form the desired image
16. **Pen Plotter**: Forms graphic images on paper with pens controlled by small electrical motors.
17. **Inkjet:** Forms an image by spraying ionized ink at a sheet of paper through micro-nozzles
18. **Laser**: Forms an image on a photosensitive surface using electrostatic charges, transfers the image on to a paper using toners, and then heats the paper to make the image permanent
19. **Thermal**: Forms an image on paper by heating ink through sublimation or phase change processes.

Display Monitors

20. **Scanning Disk**: Forms an image by first breaking up an image into tiny bits by using a rotating "scanning disk" and converting each slice to a varying electric signal in the

photocell. The signal is then transmitted to a distant receiver and reassembled into a crude picture.

21. **CRT**: Forms an image when electrons, fired from the electron gun, converge to strike a screen coated with phosphors of different colors
22. **LCD**: Forms an image by passing light through molecular structures of liquid crystals
23. **Plasma**: Forms an image by passing a high voltage through a low-pressure electrically neutral highly ionized atmosphere utilizing the polarizing properties of light
24. **OLED**: Forms an image by combining positive and negative excitons (holes emitted by anodes and electrons emitted by cathodes) in a polymer dye through the principle of electroluminescence.
25. **ELD**: Forms an image by switching between a 'reflect' and a 'not-reflect' condition of liquid crystal molecules by applying electric power to transparent electrodes.

Music Recording

26. **Mechanical**: Records sound by producing grooves on a soft rotating cylinder by a stylus and lever assembly.
27. **Electrical**: Records sound by electromagnetic cutting head driven by electronic amplifiers.
28. Magnetic: Digitizes sound in a pattern of 0s and 1s recorded in patterns of magnetic flux on a tape coated with ferric oxide powder by a small, circular electromagnetic head.
29. **Optical**: Sound is digitized in a pattern of 0s and 1s and recorded in a spiral track of microscopic bumps/pits on a plastic surface by a laser.
30. **Semiconductors**: Sound is digitized in a pattern of 0s and 1s and recorded as low and high voltage states in a grid of transistors embedded in a semiconductor chip.

Analgesics

31. **Opioids**: Reduces generation of pain signals by inhibiting the action of Cox enzymes responsible for inflammation.
32. **NSAIDs**: Reduces brain sensitivity to pain by imitating the body's own painkilling chemicals and binding to pain-sensing sites throughout the brain.
33. **Acetaminophen**: Preferentially inhibits pain by disrupting the activation of Cox enzymes.

Data Transfer

34. **Cu/Al**: Transmits data in the form of electrical energy as analog or digital signals.
35. **Fiber Optics**: Transmit data in the form of light pulses through a thin strand of glass using the principles of total internal reflection.
36. **Wireless**: Encodes data in the form of a sine wave and transmits it with radio waves using a transmitter-receiver combination.

Online Appendix C:
Statistical Formulation and Estimation of Hazard Model

This section presents the general formulation and the method of estimation of the hazard model used in this paper. Since the two hazard equations on technology and firm disruption share the same basic structure, we first describe the details in a generic form and then present the joint marginal likelihood function. The log- hazard equation is written in generic form as

$$\ln \mathrm{h(t)} = \beta_0 + \boldsymbol{\beta}_1' \boldsymbol{T}(\mathrm{t}) + \boldsymbol{\beta}_2' X(\mathrm{t}) + \lambda \qquad \text{...(B1)}$$

where T(t) is a vector of piecewise linear spline variables representing duration dependence, X(t) is a vector of exogenous covariates, and λ is a normally distributed residual term representing heterogeneity (unobserved but constant) $\lambda \sim N(0,\sigma^2)$.

The baseline survivor function is then given by:

$$\mathrm{S}_0(t) = \exp\left\{ -\int_{t_q}^{t} e^{\beta_0 + \boldsymbol{\beta}_1' \boldsymbol{T}(\tau)} d\tau \right\} \qquad \text{...(B2)}$$

Where t_q is the time of introduction of the new technology and $t > t_q$. The duration is divided into ranges of time over which the time-varying covariates are constant (we use 1 years periods in the paper). The conditional survivor function then represents the probability that technology *i* will disrupt an old technology after time t, conditional on the sequence of covariates X(t), and on the unobserved residual component λ, and is given by:

$$\mathrm{S}_0(t, X(t), \lambda) = \prod_{s=1}^{S} \left[\frac{\mathrm{S}_0(t^{s+1})}{\mathrm{S}_0(t^{s})} \right]^{\exp\left\{ \boldsymbol{\beta}_2' X(t^s) + \lambda \right\}} \qquad \text{...(B3)}$$

where *S* is the number of subintervals within which covariates **X**(t) are constant, $\mathbf{X}(t^s)$ are the values of the covariates between time t^s and t^{s+1}, and $t^{S+1}=t$ .The conditional likelihood of a completed duration t^* is given by:

$$f(t^*,\ X(t^*),\lambda) = S(t^*,\ X(t^*),\lambda)\ h(t^*,\lambda) \qquad \text{...(B4)}$$

The survivor and density functions are combined into the conditional likelihood notation covering both the censored and uncensored cases

$$L(\lambda)=\begin{cases} S(t^{**},\ X(t^{**}),\lambda) & \text{if censored (at t=t**)} \\ f(t^*,\ X(t^*),\lambda) & \text{if uncensored} \end{cases} \qquad \text{...(B5)}$$

When only one hazard process is considered (e.g. when estimating the uncorrelated hazards), the "marginal" likelihood function is obtained by integrating over the range of the heterogeneity component, λ, and is given by

$$L = \int_{\lambda} \frac{1}{\sigma_{\lambda}} \phi(\frac{\lambda}{\sigma_{\lambda}}) L(\lambda)\ d\lambda \qquad \text{...(B6)}$$

When we account for correlation between the two hazards, we assume that all correlation is captured by the heterogeneity components. Given this, the joint conditional likelihood of the set of observed outcomes is the product of the conditional probabilities of the individual outcomes. The full marginal likelihood is obtained by integrating over the range of the heterogeneity components, λ_T and λ_F, and is given by,

$$L = \int_{\lambda_T} \int_{\lambda_F} \frac{\phi(\frac{\lambda_T}{\sigma_T}, \frac{\lambda_F}{\sigma_F} \mid \rho_{\sigma_T \sigma_F})}{\sigma_{\lambda_T} \sigma_{\lambda_F}} \prod_{i}^{N} (L^T(\lambda_T)(L^F(\lambda_F) d\lambda_T d\lambda_F \qquad \text{...(B7)}$$

Where both integrals range from -∞ to +∞. Estimation is based on maximization of the marginal likelihood (for details see Panis and Lillard 1993).

Online Appendix D:

Step-by-Step Instructions of Out-of-Sample Prediction

Stage 1: Prepare data and estimate the model iteratively in aML using batch mode in DOS

1. Prepare datasets for estimation using SAS.
   a. For prediction at entry, prepare *n* datasets by excluding one target technology at a time, where n is the number of technologies in the sample.
   b. For one year updated prediction, prepare $\sum_{j=1}^{n} m_j$ datasets where j^{th} technology has m_j years after introduction. For each dataset, include data on target technology only till year *t*. Note that this method generates an estimate of the predicted time to disruption for each year the target technology is in the market.
2. Export datasets in ASCII (raw, text) format. The order in which variables are written out is important. Refer documentation on aML for details (www.applied-ml.com).
3. Create and run a control file using any text editor to convert ASCII data into aML-suitable binary format using raw2aml program within aML. This file specifies the input data and variable names and needs an .r2a extension. Note that a unique .r2a file is required for each dataset. This step creates a data file for aML with extension ".dat" and completes the data preparation stage.
4. Create a model specification file using any text editor to estimate the model. The model specification file consists of three parts and needs an .aml extension. The first part specifies the name of the data set. The second part defines "building blocks" of models and specifies the model(s). The third part initializes parameter values for the initial round of optimization.
5. Estimate the model(s) in aML
   a. aML parses the model specification file, reads the data, estimates the model equation(s), and writes the results both to the screen and an output file (by default with extension ".out"). Check results for convergence and any highlighted errors.

b. Note that you can use the batch mode in DOS to run aML to estimate the model(s) for all the datasets created in Step 1.

Stage 2: Use estimates from aML to make predictions in Excel

1. Extract estimates from .out files using any text editor. We use SAS's text parsing capabilities to extract the results into a database since the number of files is large.
2. Multiply these estimates with the values of the independent variables in Excel to get the predicted occurrence (or not) of disruption for each iteration. Compare the value with a cutoff point to predict a disruption.
3. Compare prediction with actual event and classify predictive accuracy as follows:

		Predicted Event	
		Yes	No
Actual Event	Yes	True Positive	False Negative
	No	False Positive	True Negative

4. Calculate metrics for predictive accuracy
   a. **Predictive Statistics**: Count total number of each of four outcomes in Step 3. Calculate Specificity and Sensitivity as per Equations 6 and 7 in the manuscript (see Table 5).
   b. **Graphical Comparison**: Using predictions from Step 1b above, plot the number of disruptions predicted by the model versus the actual number of disruptions for each technology as follows:
      i. For each technology, identify whether the model predicts disruption in the actual year of disruption (for observed events) or in the year of censoring (for censored cases).
      ii. Plot the number of disruptions, both predicted and actual, for each year from Year 1 till the last event in the sample (see Figure 5a and 5b).
   c. **Error in the Prediction**: We calculate this error in two ways (see Table 5).

i. Compute the mean absolute error in predicting disruption (1) or not (0), across all technologies.

ii. Compute the mean absolute error in predicted time to disruption, across all years:

True Positive:	|Actual Year – Predicted Year|
True Negative:	0
False Negative:	|Actual Year – Year of Prediction|
False Positive:	|Year at Censor – Predicted Year|

Compute standard error as $SE = \sqrt{\frac{\sum(|Y - Y'|)^2}{(N-1)}}$ where $(|Y - Y'|)$ is the error in prediction as defined above and N is the number of predictions.