Center for
Quantitative
Economics

The Dynamics of Brand Equity: A Hedonic Regression Approach to the Laser Printer Market

Ludwig von Auer† and Mark Trede‡

12/2010

† Department of Economics, University of Trier, Germany

‡ Department of Economics, University of Münster, Germany

wissen.leben
WWU Münster

The Dynamics of Brand Equity: A Hedonic Regression Approach to the Laser Printer Market

Ludwig von Auer[1]
Universität Trier

Mark Trede[2]
Universität Münster

March 2010

[1]Universität Trier – Campus I, Universitätsring 15, 54286 Trier, Germany, Phone: +49-651-201 2716, Fax: +49-0651-201 3968, e-mail: vonauer@uni-trier.de

[2]Corresponding author. Center for Qualitative Economics and Center for Nonlinear Science, Am Stadtgraben 9, 48143 Münster, Germany, Phone: +49-251-83 25006, Fax: +49-251-83 22012, e-mail: mark.trede@uni-muenster.de.

Abstract

The authors develop a dynamic approach to measuring the evolution of comparative brand premium, an important component of brand equity. A comparative brand premium is defined as the pairwise price difference between two products being identical in every respect but brand. The model is based on hedonic regressions and grounded in economic theory. In constrast to existing approaches, the authors explicitly take into account and model the dynamics of the brand premia. By exploiting the premia's intertemporal dependence structure, the Bayesian estimation method produces more accurate estimators of the time paths of the brand premia than other methods. In addition, the authors present a novel yet straightforward way to construct confidence bands that cover the entire time series of brand premia with high probability. The data required for estimation are readily available, cheap, and observable on the market under investigation. The authors apply the dynamic hedonic regression to a large and detailed data set about laser printers gathered on a monthly basis over a four-year period. It transpires that, in general, the estimated brand premia change only gradually from period to period. Nevertheless the method can diagnose sudden downturns of a comparative brand premium. The authors' dynamic hedonic regression approach facilitates the practical evaluation of brand management.

Keywords: brand equity, price premium, hedonic regression, Bayesian estimation, dynamic linear model

1 Introduction

In March 2008 Tata Motors acquired from Ford the two British car manufacturers Jaguar and Land Rover, two of the most famous brands in the car business. Most commentators agreed that a major part of the purchase price of $ 2.3 billion could be rationalized by the brand equity of Jaguar and Land Rover. Providing exact measures for such brand equities, however, is a truly difficult task. It requires both, reliable measurement techniques and sound data. Thanks to the internet, the data base has significantly improved over the years. Data on product prices, product qualities, and sales have become available, providing more reliable information for the measurement of brand equity. However, remarkably few academic studies have explored new statistical measurement techniques capable of fully exploiting these new data sources. The present study develops such a technique and applies it to the market for laser printers.

Regardless of the market considered, brand equity is often defined as 'the set of associations and behavior on the part of a brand's customers, channel members and parent corporation that permits the brand to earn greater volume or greater margins than it could without the brand name' (Leuthesser, 1988). Measures of brand equity are required not only in the context of financial transactions (takeovers, etc.), but also for evaluating past marketing decisions and designing strategies for a successful long term brand management. The 1999 MSI workshop on brand equity set up a list of properties that an ideal measure of brand equity should satisfy, e.g. it should be grounded in economic theory, objective, complete, based on easily available data, intuitive, robust, reliable and stable, yet capable to detect sudden deteriorations of brand equity (Marketing Science Institute, 1999).

It is therefore not surprising that the academic literature has developed a number of sophisticated and also some less sophisticated measures of brand equity. Ailawadi *et al.* (2003), Srinivasan *et al.* (2005), and Shankar *et al.* (2008) give a compact overview about the different categories of brand equity measures. More detailed expositions can be found, for example, in Aaker (1996) and Keller (2008). An obvious measure of brand equity is the purchase price at the time a brand is acquired (e.g., Bahadir *et al.*, 2008). Similarly, the brand equity can be computed from licensing fees and royalties. A different class of measures attempts to assess the loyalties, awareness, and associations that consumers have toward a brand. However, it is difficult to translate these cognitive measures into monetary valuations. Therefore, another class of measures is based on the individual products' market performance. Most of these measures attempt to compute a product's *brand premium*. This is the brand related component of a product's price premium. The latter is the difference between the product's price and its hypothetical price on a fully competitive market, whereas the brand premium is derived from comparing the (actual or hypothetical) price of the branded product to the (actual or hypothetical) price of a no-name product that is perceived as being equivalent to the branded product. The brand premium can be derived by various methods: estimating a structural equilibrium model (Goldfarb *et al.*, 2009), asking consumers, conducting conjoint analysis, or performing hedonic regressions.

Such hedonic regressions have a long and lasting tradition in the literature on brand measurement, e.g. Hjorth-Andersen (1984), Holbrook (1992), Parcell and Schroeder (2007), Roheim *et al.* (2007). Hedonic regressions are attractive for several reasons. They rely on objective market data rather than subjective consumer judgements based on hypothetical situations. Furthermore, they combine the various mechanisms by which the brand name adds value and they express this added value by a single monetary number. The concept of hedonic regressions is grounded in economic theory and it is based on data that are often readily available and cheap to acquire. Once the computational procedure is implemented, it can be conducted on a regular basis producing a steady flow of valuable information on all companies that operate in the market and are covered by the data set. Therefore, market analysts can also greatly benefit from the implementation of such hedonic regressions. Summarizing, our approach satisfies the properties an ideal measure of brand equity should have according to the list set up by the 1999 MSI workshop on brand equity.

Of course, the information generated by hedonic regressions are only as good as the available data base and the applied statistical procedure. Most hedonic regression techniques are static in nature. As a result, they cannot track the evolution of brand premia over time. However, some dynamic techniques also exist: adjacent year regression (Berndt and Rappaport, 2001), continuously changing coefficients (Auer, 2007), the NTP-method (Nelson, Tanguay and Patterson, 1994), linear splines, and semiparametric approaches. A brief discussion of all these techniques can be found in Auer (2007).

In this paper, we introduce a novel dynamic approach to hedonic regressions. In constrast to the existing techniques, our approach has both a more intuitive appeal and a rigorous statistical foundation. The model's parameters are estimable by an implementation of the Bayesian Markov-Chain-Monte-Carlo (MCMC) method, a statistical procedure that is taylor-made for, but has nevertheless not yet been applied to, dynamic hedonic regressions. We further suggest a novel but straightforward technique to create confidence bands that cover the area a curve is likely to be in with high probability. The confidence bands can be more easily interpreted than conventional pointwise confidence bands.

In addition to the methodological contributions we apply the dynamic hedonic regression to answer an empirical question: How do the comparative brand premia of laser printer producers evolve over time? We have gathered a rich panel data set on the market for laser printers in Germany. The data are monthly data covering the years 2003 to 2006.

Our study compiles the evolution of the comparative brand premia of the nine major brands of laser printers. It translates the numerical results into a simple graphical device that conveys to a brand manager the relative performance of the own brand premium relative to each of the competing brands. The use of this graphical device is exemplified by a discussion of the performance of the two brands Hewlett-Packard and Canon.

We organize the rest of the article in the following way: In Section 2, we begin with a presentation of the dynamic hedonic regression model. Section 3 develops the estimation procedure. Computational and theoretical details are relegated to Appendices A and B. In Section 4, we describe the data set and the variables it includes, the specification of the model, the empirical results, and their practical applications. Finally, Section 5 concludes with a summary of our contributions, the study's limitations, and possible future research directions.

2 Dynamic Hedonic Regression Model

In its simplest form, the static hedonic regression model is

$$y_i = \sum_{k=0}^{K} \beta_k q_{ki} + u_i, \quad i = 1, \dots, N \ , \tag{1}$$

where $y_1, \dots, y_N$ are the observable prices of the N products, belonging to some product category, and $q_{1i}, \dots, q_{Ki}$ are K observable quality characteristics (attributes) of the products, $q_{0i} \equiv 1$ being the intercept. The error term u_i has the usual OLS properties. The coefficient β_k can be interpreted as the implicit market price of attribute k. Of course, other specifications than (1) can be utilized leading to other interpretations of the coefficients β_k.

In past research, the hedonic regression approach often has been applied for computing and comparing brand premia. For this purpose, some attributes q_{ki} represent brand attributes, each one being associated with one brand. The brand attribute associated with the producer A, say, takes the value 1 when the product i is manufactured by this producer, and it takes the value 0 otherwise. The corresponding coefficient β_k measures the premium that is paid for this product as compared to the same product when produced by a no-name manufacturer. The brand attributes improve the reliability of the estimations, because they capture hard-to-measure aspects (e.g., "perceived quality") that are important determinants of a product's market price. Without the brand attributes, these determinants would be treated as random disturbances.

Some markets are characterized by a very small market share of no-name manufacturers. In such cases specification (1) is not optimal, see e.g. the related discussion on identification in Goldfarb *et al.* (2009). It is preferable to include in Equation (1) for each manufacturer its own brand attribute and to drop the intercept and its coefficient (q_{0i} and β_0):

$$y_i = \sum_{k=1}^{K} \beta_k q_{ki} + u_i, \quad i = 1, \ldots, N . \tag{2}$$

With this specification, the coefficient of manufacturer A's brand attribute no longer represents its brand premium over no-name products. Instead, the difference between this coefficient and the respective coefficient of brand B, say, represents brand A's *comparative brand premium* as measured against competitor B. Since the market share of no-name laser printers is very small, the empirical analysis of this market should be based on Equation (2) rather than on Equation (1). Therefore, the following exposition takes Equation (2) as its starting point.

The preferences for the attributes as well as the technologies for producing these attributes change over time. Therefore, the coefficients β_k also change over time. Writing $\boldsymbol{\beta}_t = (\beta_{t1}, \ldots, \beta_{tK})'$, $\mathbf{y}_t = (y_{t1}, \ldots, y_{tN_t})'$, and $\mathbf{u}_t = (u_{t1}, \ldots, u_{tN_t})'$, and collecting all attributes of the products in the matrix $\mathbf{Q}_t$, the hedonic regression model (2) can be expressed in matrix form:

$$\mathbf{y}_t = \mathbf{Q}_t \boldsymbol{\beta}_t + \mathbf{u}_t , \qquad \mathbf{u}_t \sim N(\mathbf{0}, \sigma_t^2 \mathbf{I}_{N_t}) \tag{3}$$

where the matrix $\mathbf{Q}_t$ is of dimension $N_t \times K$. The number of products observed in period t is N_t which may change over time; in contrast, the number of attributes K is constant.

If the data cover more than a single time period, the vector $\boldsymbol{\beta}_t$ could be estimated for each period $t = 1, \ldots, T$ in a simplistic way by running T separate OLS regressions. It is, however, well-known (Arguea and Hsiao, 1993) that this approach often suffers from large standard errors and rather erratic changes in the estimated attribute prices from one period to the next. More sophisticated estimation procedures pool two or more time periods into a single regression but still allow for coefficients β_{tk} that vary over time. The most popular of these procedures is the AYR approach (adjacent year regression approach). However, Auer and Brennan (2007) show that the AYR method produces either biased or at best inefficient estimates. Recently, Auer (2007) suggested the CCC approach (continuously changing coefficients approach). The CCC approach models the evolution of the coefficients by a deterministic polynomial in the time variable. While it might be reasonable to assume that to some extent the coefficients β_{tk} evolve in a deterministic way, there is also a strong unpredictable component in the evolution of these coefficients. Incorporating this random component into the estimation procedure may affect the outcome. We assume that the coefficients follow a random walk process:

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \sim N(\mathbf{0}, \mathbf{W}) \ , \tag{4}$$

for $t = 1, \ldots, T$, where $\mathbf{W}$ is a symmetric, positive definite $(K \times K)$-matrix, and $\mathbf{v}_t$ is a random K-vector. As usual, we assume that the disturbance vectors $\mathbf{u}_t$ and $\mathbf{v}_t$ are independent. The start vector $\boldsymbol{\beta}_0$ is randomly initialized:

$$\boldsymbol{\beta}_0 \sim N(\mathbf{m}, \mathbf{D}) \ . \tag{5}$$

Under the random walk specification (4), shocks to the attribute prices are permanent. In contrast to stationary models with mean reversion, long term changes in, and divergence of, brand premia and other attribute prices are possible. Apart from being symmetric and positive definite, no restrictions are imposed on the covariance matrix $\mathbf{W}$ of the shocks.

3 Estimation

Equations (3), (4) and (5) form a dynamic linear model (West and Harrison, 1997). In a state-space model context, (4) is called transition equation and (3) is called measurement equation. The coefficient vectors $\boldsymbol{\beta}_t$ are called state vectors. If the coefficients $\mathbf{m}, \mathbf{D}, \mathbf{W}$ and $\sigma_1^2, \ldots, \sigma_T^2$ were known, we could utilize the recursive Kalman filter algorithm to estimate the evolution of the unobservable vector $\boldsymbol{\beta}_t$. However, these coefficients are unknown and have to be estimated along with the vectors $\boldsymbol{\beta}_t$. The standard approach to estimation in a state-space model is maximum likelihood (Harvey, 1989). As there is no closed form for the ML estimators, they have to be computed by numerical optimization, which becomes more unstable and error-prone the larger the number of coefficients to be estimated. In our setting, even very parsimonious model specifications have more coefficients than numerical maximum-likelihood estimation methods can reliably cope with.

3.1 Markov Chain Monte Carlo Approach

We therefore propose to estimate all coefficients of interest – $\mathbf{W}$, $\sigma_1^2, \ldots, \sigma_T^2$ and the vectors $\boldsymbol{\beta}_0, \ldots, \boldsymbol{\beta}_T$ – simultaneously by the Markov Chain Monte Carlo (MCMC) method. MCMC is a powerful estimation technique in a Bayesian framework. Both the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$ and the other coefficients $\boldsymbol{\psi} = (\mathbf{W}, \sigma_1^2, \ldots, \sigma_T^2)$ are treated as random vectors. Their prior distribution is assumed to be uninformative for all coefficients. Further, we also assume an uninformative prior distribution for the inital state $\boldsymbol{\beta}_0$ parameterized by $\mathbf{m}$ and $\mathbf{D}$.

What is called for is the joint posterior distribution of the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$ and $\boldsymbol{\psi}$ given the observed data, i.e., the observed prices $\mathbf{y}_t$ and the observed product attributes $\mathbf{Q}_t$ for $t = 1, \ldots, T$. The MCMC method allows to break down the complex joint posterior distribution into two conditional distributions that are easier to handle. An arbitrary number of drawings from the joint distribution can then be generated by Gibbs-sampling (Carter and Kohn, 1994, Frühwirth-Schnatter, 1994, Chib and Greenberg, 1995, Kim and Nelson, 1999). The following two steps have to be iterated until convergence:

1. Given the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$ (and the observed data $\mathbf{y}_t$ and $\mathbf{Q}_t$), generate a realization of the coefficients $\boldsymbol{\psi}$.

2. Given the coefficients $\boldsymbol{\psi}$ (and the observed data $\mathbf{y}_t$ and $\mathbf{Q}_t$), generate a realization of the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$.

The two steps are discussed in detail in Appendix A. Initializing the iteration is straightforward: Suitable starting values for the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$ are calculated by running separate OLS regressions for each period,

$$\hat{\boldsymbol{\beta}}_t = (\mathbf{Q}_t'\mathbf{Q}_t)^{-1}\mathbf{Q}_t'\mathbf{y}_t, \quad t = 1, \ldots, T.$$

The iteration then starts with step 1. Denote the simulated realizations of the state variables of period t in iteration r as $\boldsymbol{\beta}_t^{*r}$.

After a burn-in period, the random drawings of the state variables and $\boldsymbol{\psi}$ are taken from their (unconditional) joint distribution. Note that the Markov chain does not converge to a constant but to its equilibrium distribution. After a burn-in period of R_0 drawings the subsequent R drawings $\boldsymbol{\beta}_1^{*r}, \ldots, \boldsymbol{\beta}_T^{*r}$, $r = R_0+1, \ldots, R_0+R$ are stored and averaged to obtain point estimators of the expectations of the posterior distribution of the state variables, $E(\boldsymbol{\beta}_1|\mathbf{Y}_T), \ldots, E(\boldsymbol{\beta}_T|\mathbf{Y}_T)$, given the observed data $\mathbf{Y}_T = (\mathbf{y}_1, \mathbf{Q}_1, \ldots, \mathbf{y}_T, \mathbf{Q}_T)$. The dependence between subsequent drawings, implied by the Markov property, can be lessened by thinning the process, e.g., keeping only every tenth drawing. The estimator of $E(\boldsymbol{\beta}_t|\mathbf{Y}_T)$ is

$$E\widehat{(\boldsymbol{\beta}_t|\mathbf{Y}_T)} = \frac{1}{R}\sum_{r=1}^{R}\boldsymbol{\beta}_t^{*r}.$$

The evolution of the K components of these vectors are the estimated time paths of the attribute prices. Averaging over the elements of $\boldsymbol{\psi}$ results in point estimates of the other parameters, i.e., the variances $\sigma_1^2, \ldots, \sigma_T^2$ of the measurement equation and the covariance matrix $\mathbf{W}$ of the transition equation.

3.2 Confidence Bands

The point estimators of the time paths do not convey any information about the precision of the estimation. A suitable tool for quantifying and visualizing the precision is a – pointwise or simultaneous – confidence band. Since a large number R of realizations from the joint posterior distribution of all parameters is available, we can easily construct both pointwise and simultaneous confidence bands for the time paths of the attribute prices.

In order to build a pointwise $(1-\alpha)$-confidence band for the time path $\beta_{1k}, \ldots, \beta_{Tk}$ of the k-th component of the attribute price vector, the R simulated drawings $\beta_{tk}^{*1}, \ldots, \beta_{tk}^{*R}$ are ascendingly ordered separately for each time period $t = 1, \ldots, T$. Denote the order statistics as $\beta_{tk}^{*(1)} \leq \ldots \leq \beta_{tk}^{*(R)}$. The pointwise $(1-\alpha)$-confidence band for the k-th attribute price is estimated as

$$\left[\beta_{tk}^{*(\alpha R/2)}, \beta_{tk}^{*((1-\alpha)R/2)}\right], \quad t = 1, \ldots, T.$$

For large R and symmetrically distributed state variables, this confidence band consists of the highest posterior density intervals for each single time period considered separately. However, the pointwise confidence bands do not give an area covering the entire time path of the attribute price simultaneously.

In constrast, a simultaneous $(1-\alpha)$-confidence band shows an area that will cover the entire path with posterior probability $(1-\alpha)$. Obviously, simultaneous confidence bands are in general wider than pointwise bands. How much wider they are, depends on the intertemporal correlation structure of the state variables.

To construct a simultaneous $(1-\alpha)$-confidence band define the i-th upper bound curve of the k-th attribute price as the linear interpolation of the points

$$\beta_{1k}^{*(R-i+1)}, \ldots, \beta_{Tk}^{*(R-i+1)};$$

similarly, the i-th lower bound curve is defined as

$$\beta_{1k}^{*(i)}, \ldots, \beta_{Tk}^{*(i)}.$$

The algorithm for a $(1-\alpha)$-confidence band runs as follows: Start with $i = 1$, i.e. with the first lower and upper bound curves and compute the proportion of simulated time paths lying entirely between the $(i+1)$-th lower and upper bound curves. Increase i step by step until the proportion of covered paths for $(i+1)$ drops below the level $(1-\alpha)$. The simultaneous $(1-\alpha)$-confidence band is given by the i-th lower and upper bound curves. As shown in Appendix B, these confidence bands have an attractive optimality condition: they are the narrowest possible bands if the state variables have a symmetric distribution (e.g. Gaussian).

4 Laser Printer Market

4.1 Data

The data set was compiled from freely available online data generated on the German market for (black and white) laser printers. From January 2003 to December

2006, for each month and each laser printer an average price was computed from the individual prices offered by internet vendors. This generated for each month between 176 and 272 observations. The data cover well above 95 percent of the German market for (black and white) laser printers. In total, the study draws on 10,853 observations. The observed laser printer prices range from € 74.9 to € 8,256.50.

For each laser printer, 25 observable attributes were recorded. The attributes *print speed, processor speed, standard memory size, added memory size, memory size that can still be added, printing resolution, paper capacity of the multi-purpose tray, standard paper capacity of the main paper tray, added paper capacity, optional paper capacity,* and *maintenance cost per page* were measured as continuous variables. Two interaction-variables count the *number of interfaces* of the printers with and without network connectivity. The other twelve attributes were measured as dichotomous variables: *maximum paper size A3, equipped with network connectivity, optional upgrade with network connectivity, printer language PCL5, printer language PCL5 or PCL6, GDI-printer (Graphical Device Interface), equipped with postscript 2, equipped with postscript 3, optional upgrade with postscript, built-in duplex, upgraded with duplex, optional upgrade with duplex.* Nine brands were included in the data set and represented by dichotomous variables (brand attributes). These brands are Brother, Canon, Epson, Hewlett-Packard, Kyocera, Lexmark, Minolta, Oki, and Samsung.

	2003	2004	2005	2006
Brother	11.84	8.93	9.05	10.09
Canon	5.06	3.67	2.38	2.77
Epson	8.03	6.94	5.87	6.77
HP	22.68	25.01	25.05	21.46
Kyocera	18.87	20.90	18.31	18.34
Lexmark	19.50	19.19	21.17	21.98
Minolta	2.43	1.48	2.34	1.84
Oki	6.86	9.21	10.48	8.77
Samsung	4.73	4.67	5.35	7.99
Sum	100.00	100.00	100.00	100.00

Table 1: Proportions of laser printers observed for each brand (in percent).

Table 1 gives the proportions of laser printers observed for each brand. As all printers belong to exactly one of these brands the intercept has to be dropped from the regression to avoid perfect multicollinearity. Alternatively, one of the brands could be treated as reference category. While the latter approach is attractive if the reference group consists of no-name products in the data set, defining one brand as reference category is rather arbitrary in the case of laser printers.

Table 2 presents some annually averaged summary statistics. A more detailed description of the data set used in this study, in conjunction with the data collection process, can be found in Appendix C.

	Means				Standard Deviations			
	2003	2004	2005	2006	2003	2004	2005	2006
# of Observations	2390	2507	2862	3103				
Price	1634	1541	1399	1263	1334	1326	1283	1212
Print Speed	26	29	30	31	11	10	10	10
Processor Speed	221	266	296	330	104	104	128	138
Standard Memory	28	39	47	54	21	28	40	47
Added Memory	10	11	19	22	32	40	61	55
Opt Memory Ext	183	239	256	295	122	147	162	174
Printing Resolution	952	1041	1084	1102	293	269	249	223
Multi-Purp Tray Cap	569	597	627	614	638	634	654	639
Main Tray Cap	140	131	104	96	180	179	119	109
Add Paper Cap	86	92	113	135	352	360	399	430
Opt Paper Cap Ext	1204	1137	1133	1084	1147	1071	1126	1173
Maintenance Cost	0.015	0.013	0.014	0.015	0.007	0.007	0.008	0.008
A3 Paper Size	0.313	0.291	0.269	0.229	0.464	0.454	0.444	0.420
Network Connect	0.422	0.491	0.495	0.546	0.494	0.500	0.500	0.498
Opt Net Connect	0.462	0.409	0.430	0.409	0.499	0.492	0.495	0.492
Interface (Net)	0.241	0.347	0.380	0.453	0.451	0.521	0.550	0.550
Interface (NoNet)	0.464	0.388	0.415	0.377	0.566	0.535	0.549	0.531
PCL 5	0.032	0.035	0.028	0.027	0.175	0.183	0.166	0.162
PCL 5 and 6	0.397	0.374	0.375	0.334	0.489	0.484	0.484	0.472
GDI	0.046	0.042	0.051	0.050	0.210	0.200	0.219	0.218
PostScript 2	0.181	0.075	0.036	0.022	0.385	0.264	0.187	0.146
PostScript 3	0.595	0.761	0.768	0.787	0.491	0.426	0.422	0.409
Opt PostScript	0.116	0.057	0.058	0.042	0.320	0.231	0.234	0.201
Duplex	0.200	0.226	0.226	0.296	0.400	0.418	0.419	0.457
Opt Duplex	0.386	0.398	0.355	0.318	0.487	0.490	0.479	0.466
Added Duplex	0.110	0.117	0.104	0.088	0.314	0.321	0.305	0.283

Table 2: Means and standard deviations of all variables, the monthly observations being averaged over years.

4.2 Model Specification

For the selection of the relevant variables to be included into the estimation procedure, we used the method advocated by Arguea and Hsiao (1993). Computing the condition index in each month for all variables shows that there is hardly any multicollinearity problem. The number of condition indexes above the threshold of 30, suggested by Belsley, Kuh, and Welsh (1980), is almost negligible, only two or three in most months. In any month, roughly 24 out of the 34 variables have condition indexes below 10. Principal component analysis does not suggest, either, that the dimension of the variables should be reduced. We therefore included all attributes as separate variables. The model was specified as given in (3) with log-prices as endogenous variables. Therefore, the estimated coefficients of the attributes ($\hat{\beta}_{tk}$) do not represent attribute prices but approximate the percentage change in a printer's price when the attribute increases by one unit.

The dynamic brand equity model (3) and (4) is estimated by the MCMC methods described above. All prior distributions are chosen as uninformative. The number of MCMC draws is $R = 20,000$ plus a burn-in phase consisting of $R_0 = 5,000$ draws.

4.3 Results

Figure 1 displays (a) the period-by-period OLS estimates of the coefficients of the nine brands along with their pointwise 2-σ-intervals (broken lines), and (b) the MCMC estimated coefficient curves along with their simultaneous 0.95-confidence bands (solid lines). The time series of the coefficients of the technical, non-brand attributes are not reported here but are available in Appendix D.

Evidently, taking into account the intertemporal dependence (4) of the coefficients, results in smoother and much more reliable estimates than the naive period-by-period approach. To assess the substantial gain in precision achieved by the MCMC method, it is important to bear in mind the different meanings of pointwise and simultaneous confidence bands: Even though the proportion of paths inside the pointwise confidence band is, by construction, roughly 95% at each single point of time, the proportion of paths lying entirely inside the band is much smaller: only between about 10% (Lexmark) and 26% (Canon).

In contrast, the simultaneous confidence bands cover 95% of the entire paths. In other words, neglecting the transition equation (4) implies a huge loss of valuable information.

Considering the MCMC estimates, most of the estimated coefficients show a rather monotonic evolution. This suggests that the brand premia (expressed here not in monetary units but in percent) do not change dramatically from year to year but move only slowly, if at all. For those experts primarily interested in

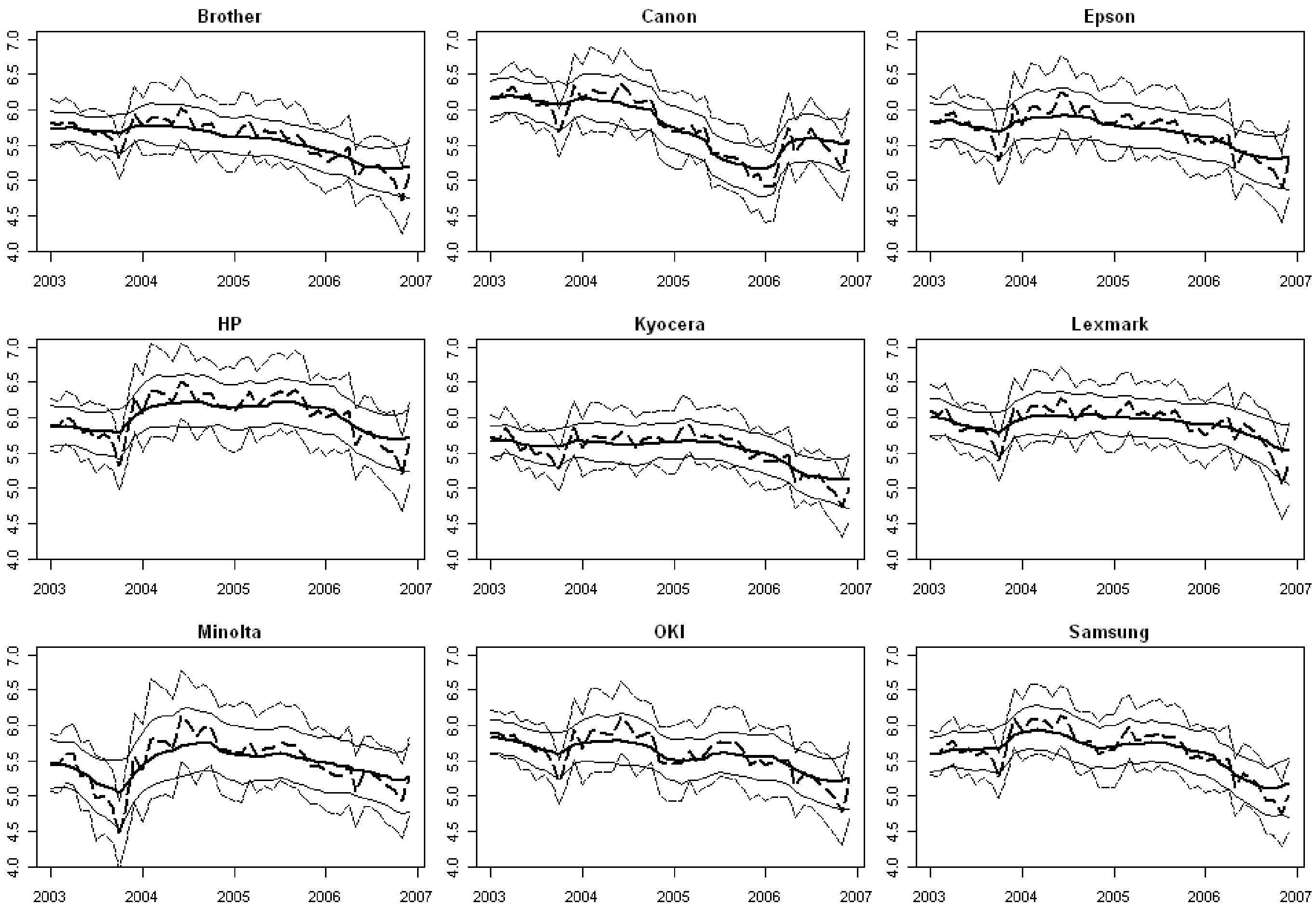


Figure 1: Time series of the period-by-period OLS estimates with pointwise 2-σ-intervals (broken lines) and the MCMC estimates with 0.95-confidence bands (solid lines) of the brand coefficients

the theoretical side of this study, Figure 1 presents the complete findings on the evolution of brand premia in a detailed though compact form.

Table 3 shows the estimated correlation coefficients of the brand attribute price shocks. Evidently, they are all highly positively correlated, indicating that there is a strong common factor affecting all brands at the same time.

	Bro	Can	Eps	HP	Kyo	Lex	Min	Oki	Sam
Brother	1.000	0.409	0.832	0.759	0.813	0.735	0.661	0.738	0.775
Canon		1.000	0.402	0.193	0.201	0.362	0.357	0.320	0.210
Epson			1.000	0.830	0.738	0.741	0.783	0.820	0.761
HP				1.000	0.721	0.778	0.808	0.791	0.819
Kyocera					1.000	0.741	0.560	0.609	0.722
Lexmark						1.000	0.720	0.681	0.671
Minolta							1.000	0.754	0.602
Oki								1.000	0.701

Table 3: Coefficients of correlation of the brand attribute price shocks

4.4 Practical Application

For brand managers the visualization of Figure 1 is not optimal. The success, or otherwise, of brand management cannot be evaluated by looking at a single brand in Figure 1 in isolation. A brand manager of Canon, for example, would like to have a clear graphical indication of how well the Canon brand premium fares in comparison to its competitors. This information is provided in the upper part of Figure 2. It depicts the evolution of the comparative brand premia of all brands in comparison to the Canon premium. The diagram shows that throughout 2003 Canon commanded the largest brand premium. However, after a long period of slow but steady decline, during the early summer of 2005 Canon experienced a sudden fall in its brand premium with a strong recovery during the first half of 2006.

A closer look at the data set and at external sources reveals that it was Canon's own pricing policy that triggered the dramatic decline of its brand premium. Having a market share of well below 2% of the laser printers sold in Germany, Canon decided in June 2005 to slash the price of its high-end laser printer Canon LBP 2000. The strategy proved successful. The market share of Canon laser printers more than doubled. By April 2006, Canon had replaced these printers by their

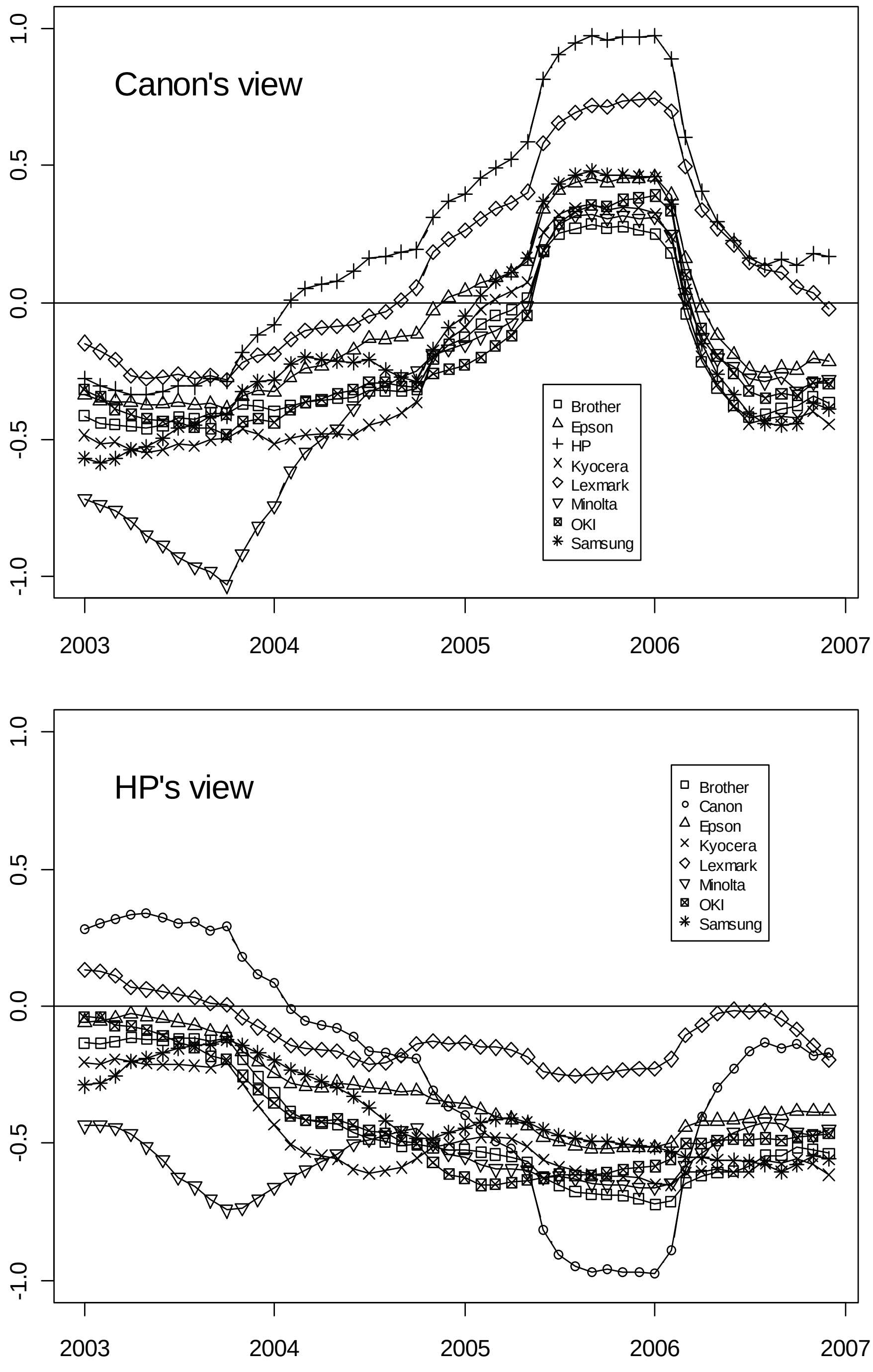


Figure 2: Comparative Brand Premia of Canon and HP.

much higher priced Canon LBP 3460. As a result, Canon's brand premium recovered strongly and returned almost back to its 2003 level. Its market share remained above 3 percent.

The lower part of Figure 2 depicts the brand premium evolution as seen from the perspective of Hewlett-Packard. In 2003 not only Canon but also Lexmark commanded higher brand premia than Hewlett-Packard. However, over the years, all competitors' brand premia have fallen relative to the Hewlett-Packard premium such that by 2004 all relative price premia had become negative and have remained negative throughout all subsequent years. The pronounced Canon bump in 2005 is clearly visible again.

5 Conclusions

This paper develops a dynamic approach to measuring brand premia based on hedonic regressions. The model, consisting of a measurement equation and a transition equation, is formulated in state-space form. In contrast to a naive period-by-period OLS hedonic regression, the model incorporates valuable additional information by exploiting the intertemporal dependence of brand premia (and other attribute prices) as given in the transition equation. As a result, the estimators of brand premia are smoother and much more precise. We present a simple way to construct confidence bands that cover the entire time series of the brand premia with high probability.

Estimation of the model by maximum likelihood is not feasible due to the rather large number of parameters. Instead, we propose a Markov Chain Monte Carlo approach that is both numerically robust and easy to implement. Being a Bayesian method, it requires (informative or uninformative) priors about the parameters and the state variables. Their joint posterior distribution is obtained by Gibbs sampling.

The dynamic hedonic regression is applied to a large and detailed data set about black-and-white laser printers on the German market. It turns out that, in general, the brand premium changes only gradually from period to period. An exception is Canon, experiencing a drastic, but temporary, brand premium decline in 2005.

In order to make the main implications of the model estimation more transparent for brand managers, we define a graphic measure of comparative brand premia showing at a glance the relative strength of each competitor.

There are some limitations of our study that could be addressed by future research. First, no sales data are utilized. While brand premia, as calculated with our method, are an important and valuable source of information for brand management, the volume aspect needs to be addressed in order to weigh the brand premia. Disregarding sales data has the advantage that all data required for our procedure are in general easily accessible. Second, brand equity is usually not only incorporated in a single product category but rather a range of different products from a potentially large number of categories. Our method only concentrates on a single product category. Effects such as brand spill-overs are not part of our framework. Third, hedonic regression does not work properly if the market under consideration is inefficient (Hjorth-Andersen, 1984). As competition is strong on the laser printer market, this is likely to be of minor importance in our empirical implementation. Further, the method may still be applied in inefficient markets and might yield valuable insights into the market power of the different producers.

The extension of the static hedonic regression approach to a dynamic setting may also be fruitful in other areas than brand management, for instance, the evaluation and forecasting of house price dynamics.

References

[1] Aaker, David A. (1996), *Building Strong Brands*, Free Press, New York.

[2] Ailawadi, Kusum L., Donald R. Lehmann, and Scott A. Neslin (2003), "Revenue Premium as an Outcome Measure of Brand Equity," *Journal of Marketing*, 67, 1-17.

[3] Arguea, Nestor M. and Cheng Hsiao (1993), "Econometric Issues of Estimating Hedonic Price Functions," *Journal of Econometrics*, 56, 241-267.

[4] Auer, Ludwig v. (2007), "Hedonic Price Measurement: The CCC Approach," *Empirical Economics*, 33, 289-311.

[5] Auer, Ludwig v. and John E. Brennan (2007), "Bias and Inefficiency in Quality-Adjusted Hedonic Regression Analysis," *Applied Economics*, 39, 95-107.

[6] Bahadir, S. Cem, Sundar G. Bharadwaj, and Rajendra K. Srivastava (2008), "Financial Value of Brands in Mergers and Acquisitions: Is Value in the Eye of the Beholder?," *Journal of Marketing*, 72, 49-64.

[7] Belsley, David A., Edwin Kuh, and Roy E. Welsch (1980), *Regression Diagnostics*, Wiley, New York.

[8] Berndt, E.R. and N.J. Rappaport (2001), "Price and Quality of Desktop and Mobile Personal Computers: a Quarter-Century Historical Overview," *American Economic Review*, 91, 268-273.

[9] Carter, C. K. and R. Kohn (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541-553.

[10] Chib, Siddhartha (1993), "Bayes Regression with Autoregressive Errors," *Journal of Econometrics*, 58, 275-294.

[11] Chib, Siddhartha and Edward Greenberg (1995), "Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-Varying Parameter Models," *Journal of Econometrics*, 68, 339-360.

[12] Frühwirth-Schnatter, Sylvia (1994), "Data Augmentation and Dynamic Linear Models," *Journal of Time Series Analysis*, 15, 183-202.

[13] Goldfarb, Avi, Qiang Lu, and Sridhar Moorthy (2009), "Measuring Brand Value in an Equilibrium Framework," *Marketing Science*, 28, 69-86.

[14] Harvey, Andrew C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

[15] Hjorth-Andersen, Christian (1984), "The Concept of Quality and the Efficiency of Markets for Consumer Products," *Journal of Consumer Research*, 11, 708-718.

[16] Holbrook, Morris B. (1992), "Product Quality, Attributes, and Brand Name as Determinants of Price: The Case of Consumer Electronics," *Marketing Letters*, 3, 71-83.

[17] Keller, Kevin L. (2008), *Strategic Brand Management*, 3rd ed., Pearson Education, Upper Saddle River (New Jersey).

[18] Kim, Chang-Jin and Charles R. Nelson (1999), *State-Space Models with Regime Switching*, MIT Press, Cambridge.

[19] Leuthesser, Lance (1988), "Defining, Measuring and Managing Brand Equity," Marketing Science Institute, Report no 88-104, Cambridge MA.

[20] Marketing Science Institute (1999), Value of the Brand, workshop at Marketing Science Institute Conference on Marketing Metrics, Washington DC, October 7-8.

[21] Nelson, R.A., T.L. Tanguay, and C.D. Patterson (1994), "A Quality-Adjusted Price Index for Personal Computers," *Journal of Business and Economic Statistics*, 12, 23–31.

[22] Parcell, Joseph L. and T. C. Schroeder (2007), "Hedonic Retail Beef and Pork Product Prices," *Journal of Agricultural and Applied Economics*, 39, 29-46.

[23] Roheim, Cathy A., Lacey Gardiner, and Frank Asche (2007), "Value of Brands and Other Attributes: Hedonic Analysis of Retail Frozen Fish in the UK," *Marine Resource Economics*, 22, 239-253.

[24] Shankar, Venkatesh, Pablo Azar, and Matthew Fuller (2008), "BRAN*EQT: A Multicategory Brand Equity Model and Its Application at Allstate," *Marketing Science*, 27, 567-584.

[25] Srinivasan, V., Chan .S. Park, and Dae R. Chang (2005), "An Approach to the Measurement, Analysis, and Prediction of Brand Equity and Its Sources," *Management Science*, 51, 1433-1448.

[26] West, Mike and Jeff Harrison (1997), *Bayesian Forecasting and Dynamic Models*, Springer, New York.

A Gibbs sampling

The Gibbs-sampling consists of the following two iteration steps.

Step 1: Consider the measurement equation (3) and the transition equation (4) in turn. Conditional on the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$, the only unknown coefficient in the measurement equation (3) for period t is σ_t^2. A standard noninformative prior distribution for the error variance is given by the density $p\left(\sigma_t^2|\boldsymbol{\beta}_t\right) \propto 1/\sigma_t^2$. The posterior distribution of σ_t^2 is inverse gamma with pdf

$$p\left(\sigma_t^2|\boldsymbol{\beta}_t, \mathbf{Q}_t, \mathbf{y}_t\right) \propto \left(\sigma_t^2\right)^{-A-1} \exp\left(-\frac{B}{\sigma_t^2}\right)$$

with parameters $A = N_t/2$ and $B = \left(\mathbf{y}_t - \mathbf{Q}_t\boldsymbol{\beta}_t\right)'\left(\mathbf{y}_t - \mathbf{Q}_t\boldsymbol{\beta}_t\right)/2$. A realization from this distribution can be generated by standard methods. Note that, given the state variables, (a) the measurement equations for each period are no longer linked, and (b) the transition equation is no longer linked to the measurement equation. Hence, the coefficients $\sigma_1^2, \ldots, \sigma_T^2$ of the measurement equation can be generated independently.

Once the state variables are given as data, the only remaining unknown parameter belonging to the transition equation (4) is the covariance matrix $\mathbf{W}$. Using a noninformative inverse Wishart prior distribution, the posterior distribution is inverse Wishart again and standard methods for generating realizations are available.

Step 2: The distribution of the state variables $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T$ conditional on the model's other coefficients $\boldsymbol{\psi}$ and the prior about $\boldsymbol{\beta}_0$ is jointly normal. Their expectation vector and covariance matrix can be computed recursively. Denote the data known up to and including period t as $\mathbf{Y}_t = (\mathbf{y}_1, \mathbf{Q}_1, \ldots, \mathbf{y}_t, \mathbf{Q}_t)$. The joint posterior distribution $p(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T)$ can be factorized as (Kim and Nelson, 1999, chap. 8)

$$\begin{aligned}
p(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T) &= p\left(\boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T\right) \times p(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{T-1}|\boldsymbol{\beta}_T, \boldsymbol{\psi}, \mathbf{Y}_T) \\
&= p\left(\boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T\right) \times p\left(\boldsymbol{\beta}_{T-1}|\boldsymbol{\beta}_T, \boldsymbol{\psi}, \mathbf{Y}_T\right) \\
&\quad \times p(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{T-2}|\boldsymbol{\beta}_{T-1}, \boldsymbol{\beta}_T, \boldsymbol{\psi}, \mathbf{Y}_T) \\
&= \ldots \\
&= p\left(\boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T\right) \prod_{t=1}^{T-1} p\left(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t+1}, \ldots, \boldsymbol{\beta}_T, \boldsymbol{\psi}, \mathbf{Y}_T\right) \\
&= p\left(\boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T\right) \prod_{t=1}^{T-1} p\left(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t+1}, \boldsymbol{\psi}, \mathbf{Y}_T\right) \\
&= p\left(\boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T\right) \prod_{t=1}^{T-1} p\left(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t+1}, \boldsymbol{\psi}, \mathbf{Y}_t\right).
\end{aligned}$$

Obviously, the recursion starts in the last period by generating a realization from $p(\boldsymbol{\beta}_T|\boldsymbol{\psi}, \mathbf{Y}_T)$ and works backward to the first period, always conditioning on the realizations of the subsequent periods. The posterior distribution of $\boldsymbol{\beta}_T$ is Gaussian

$$\boldsymbol{\beta}_T \sim N\left(\boldsymbol{\beta}_{T|T}, \boldsymbol{\Sigma}_{T|T}\right). \tag{6}$$

Its expectation vector $\boldsymbol{\beta}_{T|T}$ and covariance matrix $\boldsymbol{\Sigma}_{T|T}$ can be found by the Kalman filter (Harvey, 1989, chap. 3.2): Let $\boldsymbol{\beta}_{t-1|t-1}$ denote the optimal estimator of $\boldsymbol{\beta}_{t-1}$ using $\mathbf{Y}_{t-1}$, i.e., all observable information from period 1 up to period $t-1$. Denote the covariance matrix of the estimation error as

$$\boldsymbol{\Sigma}_{t-1|t-1} = E\left[\left(\boldsymbol{\beta}_{t-1} - \boldsymbol{\beta}_{t-1|t-1}\right)\left(\boldsymbol{\beta}_{t-1} - \boldsymbol{\beta}_{t-1|t-1}\right)'\right].$$

Using only the observations up to period $t-1$ and equation (4), the optimal estimator of $\boldsymbol{\beta}_t$ is simply $\boldsymbol{\beta}_{t|t-1} = \boldsymbol{\beta}_{t-1|t-1}$, and the covariance matrix of the estimation error is

$$\boldsymbol{\Sigma}_{t|t-1} = \boldsymbol{\Sigma}_{t-1|t-1} + \mathbf{W}.$$

Once the observation of period t is available, the estimator and the covariance matrix of the errors are updated (see Harvey, 1989, p. 106),

$$\begin{aligned}\boldsymbol{\beta}_{t|t} &= \boldsymbol{\beta}_{t|t-1} + \boldsymbol{\Sigma}_{t|t-1}\mathbf{Q}_t'\mathbf{F}_t^{-1}\left(\mathbf{y}_t - \mathbf{Q}_t\boldsymbol{\beta}_{t|t-1}\right)\\ \boldsymbol{\Sigma}_{t|t} &= \boldsymbol{\Sigma}_{t|t-1} - \boldsymbol{\Sigma}_{t|t-1}\mathbf{Q}_t'\mathbf{F}_t^{-1}\mathbf{Q}_t\boldsymbol{\Sigma}_{t|t-1}\end{aligned}$$

where

$$\mathbf{F}_t = \mathbf{Q}_t\boldsymbol{\Sigma}_{t|t-1}\mathbf{Q}_t' + \sigma_t^2\mathbf{I}.$$

The recursion is anchored in the first period utilizing the initial distribution (5) of $\boldsymbol{\beta}_0$ which can be made uninformative by setting the variances in $\mathbf{D}$ to some large value. Hence, the Kalman filter is a forward-recursive procedure working from $\boldsymbol{\beta}_{1|1}, \boldsymbol{\Sigma}_{1|1}$ until $\boldsymbol{\beta}_{T|T}, \boldsymbol{\Sigma}_{T|T}$. Using the full information estimators $\boldsymbol{\beta}_{T|T}, \boldsymbol{\Sigma}_{T|T}$ a realization of $\boldsymbol{\beta}_T$ is readily generated from (6).

Due to the multivariate normality, the distribution of $\boldsymbol{\beta}_t$, conditioned on the next period's $\boldsymbol{\beta}_{t+1}$ and the data $\mathbf{Y}_t$, is also multivariate normal (see Kim and Nelson, 1999, chap. 8),

$$\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t+1}, \mathbf{Y}_t \sim N\left(\boldsymbol{\beta}_{t|t,\beta_{t+1}}, \Sigma_{t|t,\beta_{t+1}}\right) \tag{7}$$

with

$$\begin{aligned}\boldsymbol{\beta}_{t|t,\beta_{t+1}} &= E\left(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t+1}, \mathbf{Y}_t\right)\\ &= \boldsymbol{\beta}_{t|t} + \boldsymbol{\Sigma}_{t|t}\left(\boldsymbol{\Sigma}_{t|t} + \mathbf{W}\right)^{-1}\left(\boldsymbol{\beta}_{t+1} - \boldsymbol{\beta}_{t|t}\right) \qquad (8)\end{aligned}$$

and

$$\begin{aligned}\mathbf{\Sigma}_{t|t,\beta_{t+1}} &= Cov\left(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t+1},\mathbf{Y}_t\right)\\ &= \mathbf{\Sigma}_{t|t}-\mathbf{\Sigma}_{t|t}\left(\mathbf{\Sigma}_{t|t}+\mathbf{W}\right)^{-1}\mathbf{\Sigma}_{t|t}.\end{aligned}\tag{9}$$

Hence, starting from $\boldsymbol{\beta}_T$ we can work backwards down to $\boldsymbol{\beta}_1$. To summarize, the state variables are generated by (a) running the Kalman filter as a forward recursion in order to generate $\boldsymbol{\beta}_T$, and (b) using (7), (8) and (9) as a backward recursion to generate the remaining state variables $\boldsymbol{\beta}_{T-1},\ldots,\boldsymbol{\beta}_1$. Frühwirth-Schnatter (1994) termed this procedure 'forward-filtering-backward-sampling'.

B Confidence bands

We now show that the confidence bands described in section 3.2 are consistent estimators of the narrowest possible bands for symmetric distributions. Let $f_{\boldsymbol{\beta}}(x_1,\ldots,x_T)=f_{\beta_1,\ldots,\beta_T}(x_1,\ldots,x_T)$ denote the joint density of the state variables. A $(1-\alpha)$-confidence band consists of upper and lower boundaries g_t^{up} and g_t^{low} for $t=1,\ldots,T$. A desirable optimality property is to minimize the area covered by the band, i.e. to minimize

$$\sum_{t=1}^{T}\left(g_t^{up}-g_t^{low}\right)$$

with respect to $g_1^{low},g_1^{up},\ldots,g_T^{low},g_T^{up}$, subject to

$$\int_{g_1^{low}}^{g_1^{up}}\cdots\int_{g_T^{low}}^{g_T^{up}} f_{\boldsymbol{\beta}}(x_1,\ldots,x_T)dx_T\ldots dx_1=1-\alpha,$$

i.e. the simultaneous coverage probability must equal $1-\alpha$. Using a standard Lagrangian approach we derive the first order conditions,

$$\int\limits_{\mathbf{G}_{(-t)}} f_{\boldsymbol{\beta}}(x_1,\ldots,g_t^{low},\ldots,x_T)d\mathbf{x}_{(-t)}=\int\limits_{\mathbf{G}_{(-s)}} f_{\boldsymbol{\beta}}(x_1,\ldots,g_s^{up},\ldots,x_T)d\mathbf{x}_{(-s)}\tag{10}$$

for $s,t=1,\ldots,T$, where the integration takes place over

$$\mathbf{G}_{(-t)}=\left[g_1^{low},g_1^{up}\right]\times\ldots\times\left[g_{t-1}^{low},g_{t-1}^{up}\right]\times\left[g_{t+1}^{low},g_{t+1}^{up}\right]\times\ldots\times\left[g_T^{low},g_T^{up}\right].$$

If $f_{\boldsymbol{\beta}}$ is symmetric (around the expectation) in each component, (10) implies that

$$\int\limits_{-\infty}^{g_t^{low}}\int\limits_{\mathbf{G}_{(-t)}} f_{\boldsymbol{\beta}}(x_1,\ldots,z,\ldots,x_T)d\mathbf{x}_{(-t)}dz=\int\limits_{g_t^{up}}^{\infty}\int\limits_{\mathbf{G}_{(-t)}} f_{\boldsymbol{\beta}}(x_1,\ldots,z,\ldots,x_T)d\mathbf{x}_{(-t)}dz,$$

i.e. the probability of leaving the confidence interval $[g_t^{low}, g_t^{up}]$ at time t downward must equal the probability of leaving it upward. The algorithm given in section 3.2 is based on an arbitrarily large number of MCMC realizations; hence it yields consistent estimators of the optimal $g_1^{low}, g_1^{up}, \ldots, g_T^{low}, g_T^{up}$. The confidence bands calculated by this algorithm are central in the sense that the probability of leaving them downward at least once equals the probability of leaving it upward at least once.

C Data Collection and Description

The data set of this study was compiled from freely available online data generated on the German market for (black and white) laser printers. From January 2003 to December 2006, for each month and each laser printer an average price was computed from the individual prices offered by internet vendors. Furthermore, for each laser printer, 25 observable attributes were recorded. Some of the laser printers were offered to the customers in extended versions. For example, the customers could order the printer with an enlarged paper capacity. Such an extension has been treated as a separate observation. It deviates from the basic printer merely with respect to its added paper capacity and its price.

In total, the study draws on 10,853 observations. Due to entry and exit, the data set represents an unbalanced panel data set. For each of the 48 months between 176 and 272 observations exist. The data cover well above 95 percent of the German market for (black and white) laser printers.

The dependent variable was

- LogPrice: The logarithm of the laser printer's average price during the month considered (measured in Euro).

The following list describes the regression's explanatory variables. The plus or minus sign in squared brackets indicates the expected sign of the associated coefficient.

- Print Speed: [+] The speed of printing a standard page (measured as pages per minute).
- Processor Speed: [+] The speed of the printer's processor (measured in MHz).
- Standard Memory: [+] The printer's standard RAM (measured in MB).

Some printers allow for later additions to its standard RAM.

- Added Memory: [+] RAM that the customer has added to the printer's standard memory size (measured in MB).

- Optional Memory Extension: [+] This variable is obtained by subtracting the sum of the printer's "Standard Memory" and "Added Memory" from its maximum RAM capacity.

Many printers do not allow for extensions of their RAM. For these printers, the variables "Added Memory" and "Optional Memory Extension" take the value 0.

- Printing Resolution: [+] The printing resolution of the laser printer (measured in dots per inch).

Most printers have a multi-purpose paper tray and a main paper tray. Some printers allow for later additions to its standard paper capacity. In our study, four variables describe the paper capacity properties of a printer.

- Paper Capacity of Multi-Purpose Tray: [+] The actual capacity of the printer's multi-purpose paper tray (measured in pages).

- Paper Capacity of Main Tray: [+] The actual capacity of the printer's main paper tray (measured in pages).

The sum of the variables "Paper Capacity of Multi-Purpose Tray" and "Paper Capacity of Main Tray" gives the laser printer's actual paper capacity.

- Added Paper Capacity: [-] This variable measures the part of the printer's actual paper capacity that is not part of the printer's standard equipment but is due to a customer-ordered extension (measured in pages).

- Optional Paper Capacity Extension: [+] The paper capacity that can still be added to the printer's actual paper capacity (measured in pages). It is identical to the printer's maximum paper capacity minus the sum of the variables "Paper Capacity of Multi-Purpose Tray" and "Paper Capacity of Main Tray".

- Maintenance Cost: [-] The cost that is caused by the printer's maintenance. For this purpose the price and life expectancy (measured in pages) of a printer's toner cartridge and drum unit have been recorded and translated into a price per page (measured in Cent).

- Paper Size A3: [+] A dummy variable that takes the value 1 if the laser printer is able to print A3-size pages. Otherwise it takes the value 0.

- Network Connectivity: [+] A dummy variable that takes the value 1 if the laser printer is equipped with network connectivity. Otherwise it takes the value 0.

- Optional Network Connectivity: [+] A dummy variable that takes the value 1 if the laser printer is not equipped with but can later be upgraded with network connectivity. Otherwise it takes the value 0.

- Interfaces (Network): [+] An interaction-variable that counts the number of additional interfaces if the printer is equipped with network connectivity ("Network Connectivity" = 1). If the printer is not equipped with network connectivity or the total number of interfaces is not larger than 1, the variable takes the value 0.

- Interfaces (No Network): [+] An interaction variable that counts the number of additional interfaces if the printer is not equipped with network connectivity ("Network Connectivity" = 0). If the printer is equipped with network connectivity or the total number of interfaces is not larger than 1, the variable takes the value 0.

- PCL 5: [+] A dummy variable that takes the value 1 if the laser printer works with the protocol PCL 5 but not with the protocol PCL 6. Otherwise it takes the value 0.

- PCL 5 and 6: [+] A dummy variable that takes the value 1 if the laser printer works with both the protocol PCL 5 and the protocol PCL 6. Otherwise it takes the value 0.

- GDI: [-] A dummy variable that takes the value 1 if the laser printer works neither with protocol PCL 5 nor with protocol PCL 6, but outsources the picture processing to the computer. Otherwise it takes the value 0.

- PostScript 2: [+] A dummy variable that takes the value 1 if the laser printer is equipped with PostScript II. Otherwise it takes the value 0.

- PostScript 3: [+] A dummy variable that takes the value 1 if the laser printer is equipped with PostScript III. Otherwise it takes the value 0.

- Optional PostScript 3: [+] A dummy variable that takes the value 1 if the laser printer is not equipped with Postscript but can be upgraded to PostScript III (upgrades to PostScript II are never offered). Otherwise it takes the value 0.

- Duplex: [+] A dummy variable that takes the value 1 if the laser printer is equipped with duplex printing. Otherwise it takes the value 0.
- Added Duplex: [+] A dummy variable that takes the value 1 if, due to a customer-ordered extension, the laser printer is equipped with duplex even though duplex is not part of the printer's standard equipment. Otherwise it takes the value 0.
- Optional Duplex: [+] A dummy variable that takes the value 1 if the laser printer is not equipped with but can be upgraded to duplex printing. Otherwise it takes the value 0.

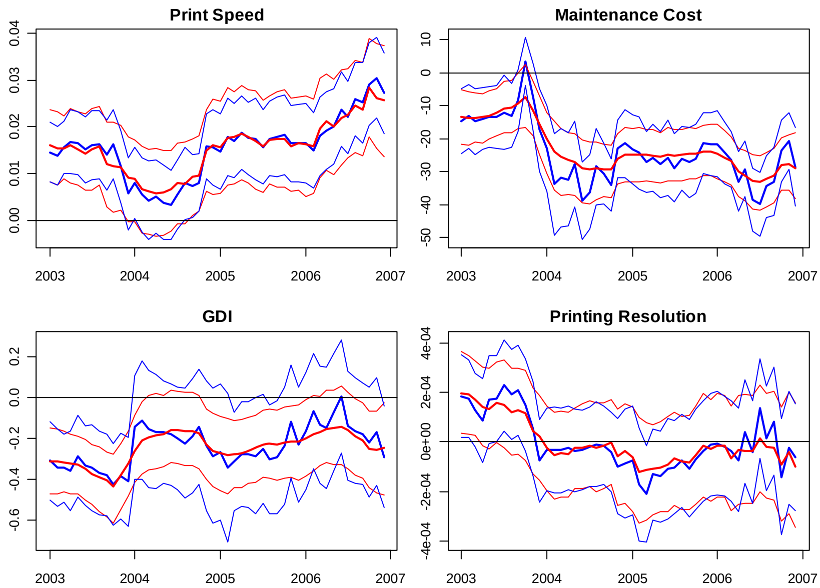
Furthermore, nine brand dummies are included in the regression.

- Brother: A dummy variable that takes the value 1 if the laser printer is produced by Brother. Otherwise it takes the value 0.

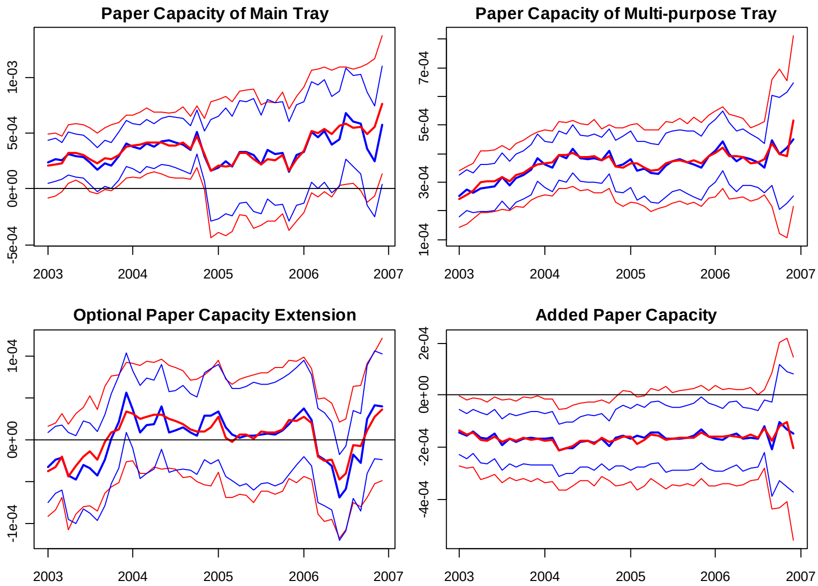
The brand dummies of the other manufactures (Canon, Epson, Hewlett-Packard, Kyocera, Lexmark, Minolta, OKI, Samsung).

D Further Estimation Results

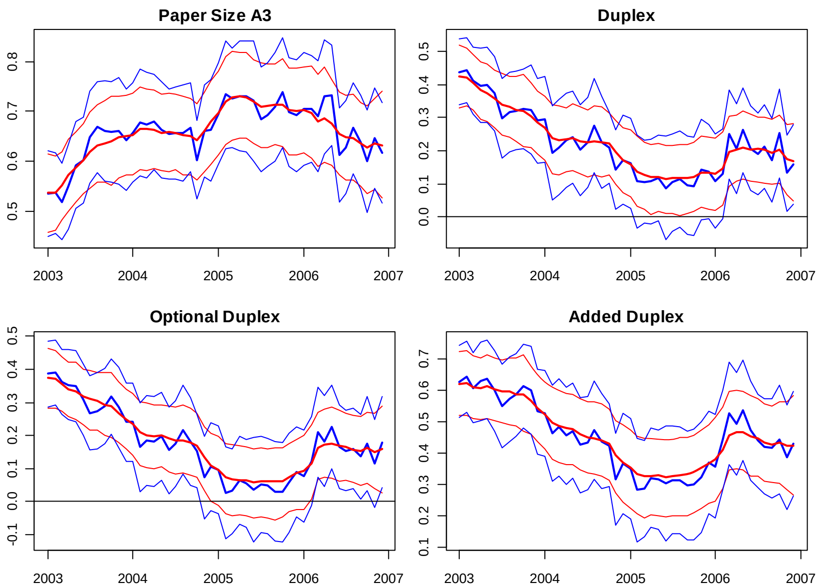
The evolution of the nine brands' coefficients was shown in the main body of the paper. Here, the evolution of the other coefficients is presented. The first graph shows the coefficients associated with the variables "Print Speed", "Maintenance Cost", "GDI", and "Printing Resolution".



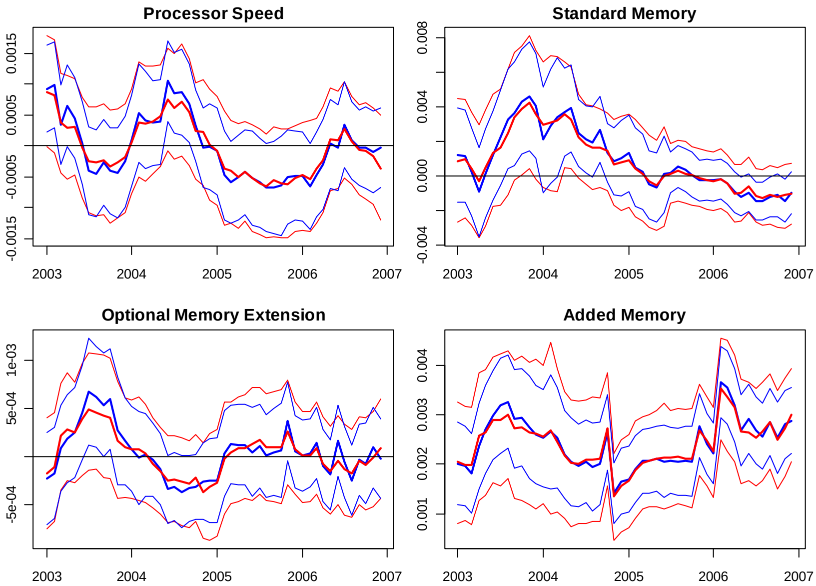
The second graph shows the coefficients associated with the variables "Paper Capacity of Main Tray", "Paper Capacity of Multi-Purpose Tray", "Added Paper Capacity", and "Optional Paper Capacity Extension".



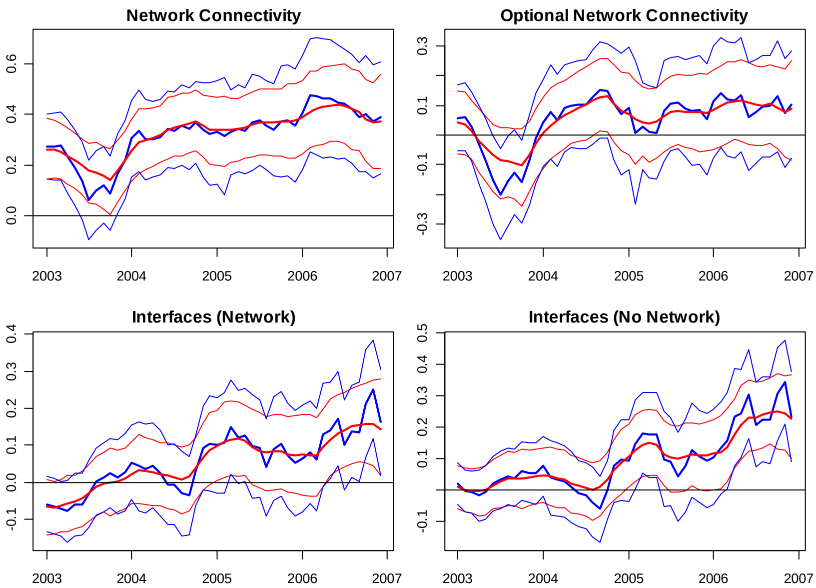
The third graph shows the coefficients associated with the variables "Paper Size A3", "Duplex", "Optional Duplex", and "Added Duplex".



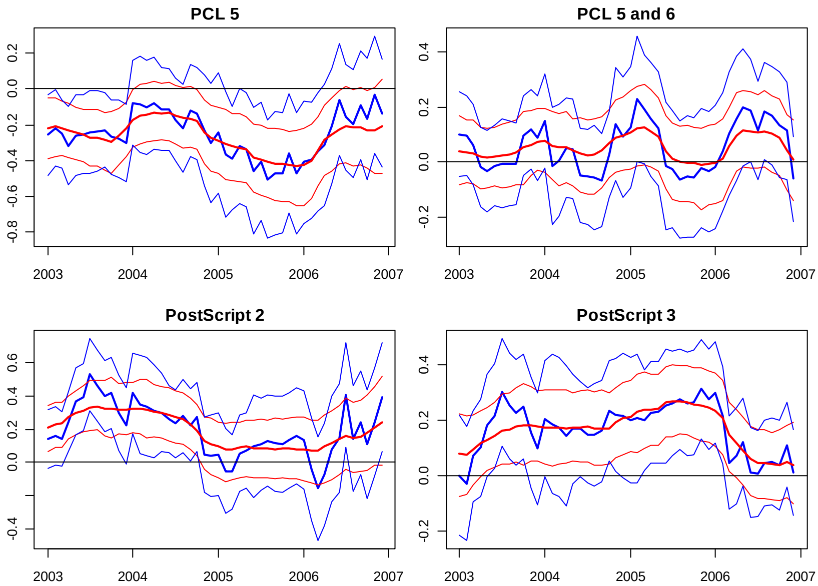
The fourth graph shows the coefficients associated with the variables "Processor Speed", "Standard Memory", "Optional Memory Extension", and "Added Memory".



The fifth graph shows the coefficients associated with the variables "Network Connectivity", "Optional Network Connectivity", "Interfaces (Network)", and "Interfaces (No Network)".



The sixth graph shows the coefficients associated with the variables "PCL 5", "PCL 5 and 6", "PostScript 2", and "PostScript 3".



The seventh graph shows the coefficient associated with the variable "Optional PostScript 3".