Measuring Marketing-Mix Effects in the Video-Game Console Market

Pradeep K. Chintagunta†
Robert Law Professor of Marketing,
Graduate School of Business, University of Chicago,
1101 East 58th Street, Chicago IL 60637.
Email: pradeep.chintagunta@gsb.uchicago.edu
Phone: 773-702-8015
Fax: 773-702-0458

Harikesh S. Nair
Assistant Professor of Marketing,
Graduate School of Business, Stanford University,
518 Memorial Way, Stanford, CA 94305.
Email: harikesh.nair@stanford.edu
Phone: 650-736-4256

R. Sukumar
Director, IPSOS-Insight, Inc., and,
Clinical Professor of Global Business, Thunderbird,
American Graduate School of International Management.
Email: sukumar@pdq.net
Phone: 832-372-8580

September 2006
July 2007

Forthcoming, Journal of Applied Econometrics

† We thank Ester Han, Karen Sperduti and Sima Vasa of the NPD group for generously making available the data used in this research. We also thank P.B. Seetharaman, Ken Singleton, Inseong Song and Naufel Vilcassim for useful comments and suggestions. The first author thanks the Kilts Center for Marketing at the University of Chicago for research support.
Measuring Marketing-Mix Effects in the US 32/64-bit Video-Game Console Market

Abstract

We investigate the short and long run effects of prices and software availability on the category-level diffusion of 32/64 bit video-game consoles in the US. We adopt an estimation framework that allows for a flexible intrinsic growth pattern for the hardware consoles, and uses instrumental variables to control for the potential endogeneity of prices. We find significant long-term effects of prices and software availability on sales. We find that using a non-parametric hazard specification is important: imposing parametric forms results in elasticities that are up to 30% smaller. We use these results to derive implications for marketing policies over the life-cycle of the industry.

Keywords: Video-games, new product diffusion, proportional hazards model, semiparametric estimation, endogeneity.
1. Introduction

It is well known that the effectiveness of marketing mix factors in driving sales of a new product varies over its life-cycle. A large literature in marketing has documented the time-varying impact of prices, advertising, distribution and other marketing mix variables in several product categories (Parsons 1975, Simon 1979, Shoemaker 1986, Lilien and Yoon 1988, Tellis and Fornell 1988, Parker 1994, Parker and Neelamegham 1997, Danaher et al. 2001). As marketing-mix effects change over the life-cycle, the trade-offs to firms of leveraging available marketing instruments for generating sales also changes. An important input to evaluating the tradeoffs are measures of short and long-run effects associated with the variables. Long-run effects may be particularly important if current sales of the product affect its future sales, say by increasing the size of the network (i.e. via word-of-mouth effects) or by inducing supply of complementary products (i.e. via indirect network effects). This is typically the case for most high-technology durable good products. Thus, an issue of importance to firms in these industries is the measurement of the impact of marketing mix variables on current and future sales over time, and the evaluation of the trade-offs involved in investing resources in using one marketing instrument over another. We focus on these issues in this paper in the context of the US market for 32/64-bit video-games consoles.

The 32/64-bit generation of the US video-game industry consisted of the Sony PlayStation, Sega Saturn and Nintendo N64 consoles. The life-cycle of this generation extended roughly between May 1995 and September 2002. Previous literature on the industry (c.f. Coughlan 2001, Shankar and Bayus 2003) identified prices and software (i.e. video-game) availability as the two key drivers of sales of video-game consoles. We investigate the long and short-run impact of these two variables on the category-level sales of the hardware. Our empirical approach is different from previous work in this area because: a) we control for the potential endogeneity of marketing mix elements in the
sales equation, and (b) we are able to accommodate very flexible patterns in the intrinsic growth of the hardware. Both these aspects are found to significantly affect the estimates.

The control for endogeneity is important to obtain a valid measure of the effect of the price and software availability variables. Endogenous prices arise in new product sales models due to classical concerns of simultaneity in demand and supply. Endogeneity arises due to the presence of omitted variables in the sales equation that are correlated with prices. An example would be console advertising expenditures which influence sales and are set jointly with prices by firms, but are unobserved by the researcher. Omitting these generates correlation between prices and the statistical errors in the model, leading to an endogeneity bias. With exclusion restrictions, one can resolve the bias by instrumenting for prices. An alternative is to exploit structure, and to add restrictions from a specified model of pricing conduct to the estimation of the sales equation. To the extent that imposing a potentially misspecified pricing function can bias the obtained demand-side (sales) parameters, instrumental variables methods would be preferred to the latter approach, provided one has access to good instruments for prices.

The second issue we focus on is model flexibility. It is important to note that constraining the intrinsic growth pattern of the hardware (i.e. the hazard) has the potential to influence the nature of the estimates obtained for the price and software availability variables. This point has been made in the household purchase timing literature in the context of marketing mix effects in proportional hazard models (e.g., Jain and Vilcassim 1994). However, discussion of the issue in the new products literature has been limited. In extreme cases, imposing an incorrect parametric form could result in biased estimates of the covariate effects.

Our approach resolves both these issues. We use a flexible estimation framework that enables us to control for endogeneity concerns using instrumental variables, while imposing little structure on the intrinsic hardware growth pattern. Our model for new product sales is a
proportional hazard (PHM) specification, in which the baseline hazard function is estimated semi-parametrically from the data.\textsuperscript{2} This model is based on the PHM specification proposed by Bass, Krishnan and Jain (2000), but modified to allow for added flexibility in the baseline hazard. Prices and software availability affect console sales by multiplicatively shifting the baseline hazard. We present a simple procedure to estimate the parameters of the model from the data. The procedure is based on the attractive property of the model that, conditional on the market potential parameter, the implied hardware sales equation can be linearized via a logarithmic transformation. The estimation procedure involves a nonlinear search for the hardware market potential in an “outer loop,” and estimating the remaining parameters by standard linear regression within an “inner loop.”

In estimation, we also control for the possibility that some potential drivers of sales, including other marketing mix variables, are not observed by the analyst. These unobserved determinants of sales are captured by time-period specific error terms that enter the specification additively with the observed covariates. The nested log-linearization of the sales equation facilitates the application of standard instrumental variables techniques (2SLS) to account for the endogeneity bias caused by the potential correlation of prices with these error terms. An attractive feature of our approach is that the instrumental variables approach can easily be implemented in the context of the PHM model, for any hazard specification, including the semi-parametric case, as well as that derived from the popular Bass (1969) model.

Our empirical results indicate that both prices as well as to the number of software titles significantly affect the sales of hardware consoles. Further, we find significant short and long-term effects of prices and number of software titles on sales, implying that it is important to consider both these effects when deciding which instrument to focus on at various stages of the life-cycle. We

\textsuperscript{2}The relative merits of using the PHM formulation have been discussed extensively elsewhere (Jain 1992, Bass et al. 2000, Danaher et al. 2001, Bass and Srinivasan 2002). The main conclusion appears to be that (a) PHM performs well in terms of fit and predictive ability; (b) researchers need to check the stationarity of covariates before using the PHM specification.
find that prices have a larger effect on sales of consoles in the initial time periods, while the number of software titles in the market has a larger effect in the later time periods.

We find that the baseline hazard for our data, estimated semi-parametrically, has a non-monotonic shape that is not well approximated by standard functional forms. In the context of our data, we find that imposing specific functional form restrictions on the baseline hazard significantly alters the conclusions about the effects of marketing variables; this persists even after accounting for price endogeneity in the parametric models. For example, we find that elasticity estimates are up to 30% lower than our specification, underscoring the need for flexibility in model specification for our application. In contrast to our approach, previous literature that has focused on measuring the impact of prices and other marketing mix variables on new product adoption (c.f. Bass 1980, Dolan and Jeuland 1981, Kamakura and Balsubramanian 1988, Jain and Rao 1990, Bass et al. 1994)\(^3\) has typically proceeded by augmenting the sales function implied by a specific functional form – typically that of the Bass (1969) model – to account for marketing mix covariates.

As has been demonstrated by several published studies, the Bass model specification provides a good fit for a wide range of available data. However, there could be situations, such as ours, where imposing functional form restrictions on the shape of the hazard function has the potential to bias the effect of marketing variables, if the data do not in fact support these restrictions. Previous research, cf Bass et al. (1994), has also shown that the Bass model without covariates performs well in terms of model fit when studying the sales of new durable products. Our findings with concur in the sense that most of the variation does arise from the intrinsic sales growth pattern. However, our results also imply that imposing strong functional form restrictions on the shape of the hazard function may further understate the effects of covariates when the underlying intrinsic growth pattern does not, in fact, conform to the parametric specification. Hence, covariates seem to play an

even smaller role when such restrictions are imposed on the data. A pre-requisite for our approach is the availability of a long time-series of data, or data at a higher frequency than typically used in the estimation of macro-level diffusion models. In this sense, we view our approach as complementing the current literature on diffusion models in marketing when such data are available.

The rest of this paper is as follows. The next two sections describe the model specification and estimation procedure. We then describe the data and present the empirical results. We then discuss the substantive implications of the results in terms of long and short-term sales elasticities. The following section presents results from using parametric hazard functions. The last section concludes.

2. Model

We present the basic framework of the model in the first sub-section. We discuss the econometric specification of the model in the subsequent sub-section. Then we discuss the procedure for estimating the parameters of the model.

2.1. Proportional hazard framework

We follow the PHM formulation as in Bass, Jain and Krishnan (2000). The instantaneous probability of adoption at time \( t \) given no adoption prior to \( t \) can be written as:

\[
\lambda(t, Z) = \frac{f(t, Z)}{[1 - F(t, Z)]}
\]  

(1)

where \( Z \) are covariates, \( \lambda(t, Z) \) is the hazard function, \( f(t, Z) \) is the density function, and \( F(t, Z) \) is the distribution function. Integrating (1), we obtain the distribution function:

\[
F(t, Z) = 1 - \exp(-\int_0^t \lambda(u, Z)du)
\]  

(2)
As described previously, our data are at the monthly level. Hence, we need the discrete-time version of
the model. Following Bass et al. (2000), the probability or failure rate in equation (1) can be written in discrete
time as follows:

\[ \lambda(t_{i-1}, t_i) = \frac{F(t_i, Z) - F(t_{i-1}, Z)}{1 - F(t_i, Z)} = 1 - \exp \left( -\int_{t_{i-1}}^{t_i} h(u)e^{Z_iY}du \right) \] (3)

In the above equation, \( \int_{t_{i-1}}^{t_i} h(u)e^{Z_iY}du \) is a positive function of time and covariates such that the
right hand side of equation (3) lies in the interval (0, 1). If the values of the covariates, \( Z \) are constant
during the month, \( Z_{i0} \), then, (3) can be written as:

\[ \lambda(t_{i-1}, t_i) = 1 - \exp(-\exp(\alpha_i + Z_{i}\gamma)) \] (4)

where, \( \exp(\alpha_i) = \int_{t_{i-1}}^{t_i} h(u)du \) (see Seetharaman and Chintagunta 2003, equation 14). In the above
expression, the \( \alpha_i \) term denotes the contribution of the baseline hazard for the time interval \( [t_{i-1}, t_i] \)
and \( Z_{i}\gamma \) denotes the effects of covariates.\(^5\) Hence the \( \alpha_i \) terms capture the “intrinsic” growth in the
sales of the new product over time. Controlling for \( \alpha_i \) helps us disentangle the effect of intrinsic
growth from factors such as falling prices and increased availability of complementary software.
Alternative approaches such as the additive risk formulation (c.f. Seetharaman 2004) can also be
accommodated within (4) in an analogous fashion.

Let \( S_{i} \) denote the sales of the product in period \( t_i \) and \( X_{t,i-1} \) denote the lagged cumulative
sales till time period \( t_{i-1} \). Let \( m \) denote the size of the market (to be estimated from the data). Then,
the sales in the given time period \( (t_{i-1}, t_i] \) can be represented as (see Bass et al. 2000):

\[ S_i = (m - X_{t,i-1})\lambda(t_{i-1}, t_i) = (m - X_{t,i-1})[1 - \exp(-\exp(\alpha_i + Z_{i}\gamma))] \] (5)

\(^4\) The assumption is reasonable as firms typically use these data for decision making purposes.
\(^5\) In principle, the parameters, \( \gamma \) can be time varying. However, we do not have the data that will allow us to estimate
such a model.
This forms the basic framework of the model. We now discuss possible econometric specifications that can be used to take the model presented in (5) to data. Subsequently, we discuss the estimation of the model.

2.2. Econometric specification

The main specification issues for the above model are the choice of the baseline hazard function, $\alpha$, and the control for unobserved drivers of sales. We discuss these issues in turn in the next two subsections.

2.2.1. Choice of baseline hazard

Researchers have two options in specifying a baseline hazard for (5): a) specify $\alpha$ semi-parametrically; this implies that the $\alpha$-s will be estimated as time-period specific fixed effects that do not constrain the baseline hazard to conform to a particular pattern a priori; b) choose a particular functional form for $\alpha$ and impose the restrictions implied by the chosen function in estimation. This is the approach followed by Bass et al. (2000), for example, who impose the hazard corresponding to the Bass (1969) model.

The choice of either option is dictated by the nature of the data and on the goal of the analysis. A semiparametric procedure would be preferred to a parametric one if there are irregular patterns over time that are difficult to capture using a parametric function. For example, suppose at some point in the growth curve of an existing technology, a new technology is introduced which could potentially cannibalize sales of the original technology. Allowing for such a structural change in the market using a parametric specification would necessarily involve explicitly modeling the interactions between the sales of both technologies. However, suppose the primary objective is to measure effects of the marketing activities. To that end, the semi-parametric specification will account for the effects of the new technology implicitly and will ensure that one obtains the
appropriate estimates of the various covariates included in the model. In the context of our videogame data, Sony launched the next generation 128-bit PS2 console even when Sony’s own PS1 (32-bit) console, as well as the Sega Saturn (also 32-bit) and the Nintendo-64 (64-bit) were still on the market and competing with each other. Using a parametric specification might require us to explicitly model the impact of the PS2 launch on the sales of the 32-64-bit generation. With the semi-parametric approach, the $\alpha$ parameters after the launch of the PS2 will reflect the effect of that launch on the sales of the earlier generation machines, which enable us to estimate the effects of prices and number of software titles of that generation. Recall that our primary objective is to obtain appropriate estimates of these effects from the data.

Despite the obvious advantages, a pre-requisite for implementing the semi-parametric approach is the availability of a long time-series of data at a reasonably high frequency (e.g. on a monthly basis). If such data are not available, one might have to impose more structure on the growth pattern by choosing a functional form for the baseline hazard. Further, one cannot make out-of-time-series sales forecasts with the semi-parametric approach, since the $\alpha$ parameters for those time periods are not available. This is not critical for our situation, since our focus is on the measurement of marketing mix effects, rather than on forecasting future sales. However, in many cases, the main goal may be forecasting. One way to address this issue is to first implement the semi-parametric approach to uncover the nature of the underlying distribution describing the sales pattern. One can then choose a parametric form that best represents the pattern uncovered by the $\alpha$ parameters and, subsequently, re-estimate the model parameters using this specification. This latter model can then be used for forecasting purposes. It could also be that a convenient parametric form may not be available a priori. In such cases, one will have to pick the closest available approximation and make sure that the effects of the included covariates are unaffected by the imposition of the functional form. In the estimation section, we discuss how our proposed PHM model can be
estimated both with a semi-parametric specification for the baseline hazard, and for two popular functional forms (viz. that of the Bass 1969 model, and the flexible Expo-power function, c.f. Saha and Hilton, 1997).

We now discuss the control for unobserved factors in the econometric specification and the resulting endogeneity problem that needs to be corrected in estimation.

2.2.2. Control for unobserved factors and implied endogeneity

A practical consideration with the specification in equation (5) is that researchers typically do not have access to all the variables in the vector $Z_{ti}$. Indeed, in most situations, only a subset of the $Z_{ti}$, $W_{ti}$, is observed. Hence, equation (5) can be represented as follows:

$$S_t = (m - X_{t,i-1})[1 - \exp(-\exp(\alpha + W_{ti}\beta + \varepsilon))]$$  \hspace{1cm} (6)

In the above equation, $\varepsilon$ represents the effects of the unobserved drivers of sales in period $t_i$. There are two points to note about the inclusion of the error term in equation (6). First, unlike previous studies such as Bass et al. (2000), Srinivasan and Mason (1986) and others that have used an additive error term to capture a variety of factors including measurement error, the error term enters the right-hand side of equation (6) nonlinearly. Our choice of specification is based on a natural interpretation for the error term as representing factors influencing the sales in time period $t_i$ that are not observed by the researcher. In particular, in any given time period $t_i$, the term $(m - X_{t,i-1})$ is pre-determined. Hence, unobservables are likely to influence the proportion of this pre-determined quantity that gets converted into sales – which is what equation (6) implies. Additionally, the specification in (6) ensures that predicted sales will never be negative – always a possibility with additive error terms. We note however, that there are other approaches to accounting for

---

6 There is an obvious identification issue in separating $\alpha_i$ from $\varepsilon_i$ in (6). In the estimation section, we discuss how this issue is resolved.
unobserved factors influencing the diffusion process – see for example, Boswijk and Franses (2005) for one approach to extending the Bass diffusion model.

Estimation of the model presented in (6) is complicated owing to the non-linear way in which $\varepsilon_{ti}$ enters the equation and due to the potential for endogeneity bias. This bias can arise if the unobservable term $\varepsilon_{ti}$ is correlated with some of the observable factors in $W_{ti}$. For example, suppose $W_{ti}$ contains the price of the product as a covariate, but advertising expenditures that are known to influence sales are unobserved by the researcher. If pricing and advertising decisions are correlated at the industry level, then $W_{ti}$ will be correlated with $\varepsilon_{ti}$, leading to an endogeneity bias. We control for the endogeneity issue by using cost-side instruments for prices that enable us to infer the true demand-side effects from the data. Intuitively, in estimation, we do not use the price variation that is due to the firms’ response to the changing elasticity of demand each period. Rather, we only use the price variation that is explained by cost-side factors to infer the price coefficient, and thus, the true demand elasticity.\footnote{For other studies that use instrumental variables in the context of duration models, the reader is referred to Bijwaard (2002) and Abbring and van der Berg (2005).}

In the next section, we propose an estimation procedure that accounts for both the nonlinear specification of the error term and controls for potential endogeneity biases using instrumental variables techniques.

3. Estimation

We discuss the estimation procedure in the context of the semi-parametric specification. As discussed before, here we estimate the $\alpha_s$ terms as time-period specific fixed effects that do not constrain the baseline hazard to conform to any particular pattern a priori. The estimation of the model with the Bass and Expo-power hazard specifications is analogous to the semi-parametric case,
and simply involve replacing the $\alpha_i$ terms by the corresponding implied hazards (a description of the estimation of these two models is presented in the appendix).

For estimation, we first modify the sales equation (6) as follows. Note that since we have only a single time series of data available, it is not feasible to estimate a distinct $\alpha$ parameter for each month. Hence, rather than estimate $\alpha_i$ for $t_i = 1, 2, \ldots, T$, we estimate $\alpha_{\tau j}$ for $\tau j = 1, 2, \ldots, \Gamma$ and $\Gamma < T$. Here each interval $\tau j$ encompasses several months or several $t_i$. In the empirical section, we conduct extensive sensitivity analysis to choice of $\Gamma$. Equation (6) is thus modified as:

$$S_m = (m - X_{t,i-1})[1 - \exp(-\exp(\alpha_{\tau j} + W_{\tau j} + \varepsilon_{\tau j}))]$$

(7)

Note that sales, lagged cumulative sales, and covariates are all still measured in the original time units, i.e., months. And hence, the model is estimated using monthly data. However, the “baseline hazard” will be estimated at a lower frequency, bi-monthly, quarterly, semi-annually or annually depending on the ability to precisely estimate the corresponding parameters. Note that introduced in this manner, a component of $\varepsilon_{\tau j}$ also accounts for monthly deviations between $\alpha_{\tau j}$ and $\alpha_{\tau j}$. Further, the identification issue between $\varepsilon_{\tau j}$ and $\alpha_{\tau j}$ is also resolved.

The unknown parameters in our model specification in equation (7) are: \{m, $\beta$, $\alpha_{\tau j}$\}$_{\tau j=1}^\Gamma$. As noted previously, the two issues complicating estimation are (a) the nonlinear way in which $\varepsilon_{\tau j}$ appears in the sales equation, and, (b) the possibility of correlation between $\varepsilon_{\tau j}$ and some or all elements of $W_{\tau j}$. Here we discuss how we address these issues in estimation. We first provide an intuitive description of the estimation procedure and then provide formal details.

First, consider the estimation of $\Theta = \{\beta, \alpha_{\tau j}\}_{\tau = 0}^\Gamma$ conditional on $m$. If $m$ is “known”, then, we can write equation (7) as follows:
\[
\frac{S_a}{(m - X_{t,i-1})} = [1 - \exp(-\exp(\alpha_a + W_a \beta + \epsilon_a))] \tag{8}
\]

The above equation can be linearized using a simple logarithmic transformation to yield:

\[
\ln\left[\ln\left(1 - \frac{S_a}{(m - X_{t,i-1})}\right)\right] = \alpha_a + W_a \beta + \epsilon_a \tag{9}
\]

Note that the left hand side of equation (9) will be a real number as \(S_a < (m - X_{t,i-1})\). Therefore \(K = \ln\left(1 - \frac{S_a}{(m - X_{t,i-1})}\right) < 0\) and hence we compute the natural logarithm of a positive number, \(-K\).

It is clear from equation (9) that the set of parameters, \(\Theta\), can now be estimated using OLS if \(W_a\) and \(\epsilon_a\) are uncorrelated. However, in the event that we are concerned about possible correlation between these two quantities, we need to identify a set of appropriate instruments for the variables in \(W_a\) that are expected to be correlated with \(\epsilon_a\). Two-stage least squares (2SLS) can then be used to estimate the parameters \(\Theta\). Hence, conditional on \(m\), the estimation procedure is straightforward. We now discuss how we estimate the market potential parameter, \(m\).

As we use OLS or 2SLS in the estimation of \(\Theta\), the corresponding objective functions in the estimation, \(\Omega\), are either \(e'e\) or \(e'V(V'V)^{-1}V'e\) respectively, where \(e\) denotes the vector of stacked residuals from equation (9) and \(V\) denotes the matrix of instruments. Note that each value of \(m\) will result in a corresponding value of \(\Omega\). Hence, we simply search the space of \(m\), a scalar (which is also constrained to lie beyond \(X_T\)), that minimizes \(\Omega\).

Formally, the estimation procedure can be described via the following steps.

- **Step 1**: Choose a starting value for the parameter \(m\).
- **Step 2**: Conditional on \(m\), perform the linearization as in equation (9).
- **Step 3**: Use OLS or 2SLS as required to estimate the parameters \(\Theta\), conditional on \(m\).

Compute \(e\), and form the corresponding objective function, \(\Omega = e' e\) or \(e'V(V'V)^{-1}V'e\).
• **Step 4**: Iterate over steps 1 to 3 (i.e., search over the feasible range of the values of m) using any standard nonlinear search routines to obtain the value of m that minimizes $\Omega$.

• **Step 5**: Compute the standard errors for the entire set of estimated parameters, m and $\Theta$, as in Hansen (1982).\(^8\)

In practical terms, steps 1 and 4 constitute the “outer” loop of the estimation procedure, whereas steps 2 and 3 comprise the “inner” loop.\(^9\)

### 3.1. Comparison to Nonlinear Least Squares

An alternative to the above approach is to allow an additive error term into the sales equation, and to estimate the parameters via non-linear least squares (e.g. Lee et al. 2003). That specification, applied to our model or to any PHM-based model, is as follows:

$$ S_{ti} = (m - X_{t,i-1})*[1 - \exp(-\exp(\alpha_t + \beta))] + \varepsilon_{ti} $$

Here $\varepsilon_{ti}$ would be interpreted as measurement error. The parameters of this model can be estimated in a single stage using a standard packaged nonlinear least squares routine. We discuss two problems with this approach. First, it is not straightforward to account for price endogeneity in this case, as linear instrumental variables methods such as 2SLS are no longer applicable in the estimation. Alternatively, one could specify a separate pricing (i.e. supply-side) equation (Bayus 1992, Krishnan et al. 1999), and estimate both this and the sales equation jointly using nonlinear least squares.

However, if good instruments for prices are available, we would prefer IV methods to this approach, so as to avoid biasing the demand-side (sales) estimates by imposing a potentially misspecified supply-side model. A second problem arises if there exists serial correlation in the error term. Serial correlation implies that $\varepsilon_{ti}$ is correlated with $\varepsilon_{t,i-1}$. And therefore, $X_{t,i-1}$ also correlated with $\varepsilon_{t,i-1}$ via $S_{t,i-1}$.

---

\(^8\) Letting $G' = \left[\partial\varepsilon_{ti}/\partial\theta\right]_{\theta=\theta^{*}}$, the gradient of the moments evaluated at the final parameter values, and $W^*$ denote $E[G^* G^*']$, the asymptotic variance of $\theta = \{m,\Theta\}$ is given by $(G^* G^*)^{-1}W^*G^* (G^* G^*)^{-1}$.

\(^9\) The nonlinear search with the nested regression is analogous to the procedure proposed in the recent literature on estimating aggregate discrete choice models of demand (c.f. Berry, Levinsohn and Pakes 1995.)
In this situation, the error term in the sales equation is correlated with one of the right-hand-side variables, rendering the estimates obtained biased and inconsistent. To resolve this issue one might want to instrument for the lagged cumulative sales variable, but once again, linear instrumental variables methods will not be appropriate in this case. On the other hand, our approach does not suffer from this problem even if there is serial correlation in the error term given the manner in which our estimates are obtained. Indeed, such correlation in our case will only influence the estimated standard errors which can be corrected in a straightforward fashion via Newey-West (1987) or Conley (1999) approaches.

4. Data

Our data are for one generation of the videogames industry. In particular, our dataset contains monthly data on the number of units sold of all 32/64-bit consoles that include Sony PS1, Sega Saturn (both 32-bit machines) and Nintendo N64 (64-bit machine). The data are at the category level (i.e., not for each firm’s console) and span the range from May 1995 (when Sega Saturn was introduced) till September 2002. During that last month, the 3 consoles combined sold a total of 129,163 units with a cumulative sales of over 45 million units. Sony launched its PS1 console in September 1995 and Nintendo launched the N64 in September 1996. Note that while Sony launched its PS2 game console in November 2000 and Microsoft and Nintendo the following year, our model specification will be able to account for this as the estimated $\alpha$ parameters in the period after the launch of those consoles will reflect those actions. This is one of the advantages of the proposed semiparametric specification. Moreover, although the category sales combine sales from the three platforms, competition was intense across these platforms. This provides some justification for treating the three platforms as a single “category” (see Shankar and Bayus 2003).
Besides the category unit sales of consoles, our data also include information on two key covariates. These are the average console price in each month as well as the total number of software titles available across all 3 platforms that had positive sales in each month. Price is a key variable in this category. The number of software titles is another key variable that drives console sales in the video-game industry according to most industry observers (see Coughlan 2001) as well as academics (Shankar and Bayus 2003). Thus, an issue of interest in this industry is the relative size of the effect of the number of software titles on hardware sales as compared to that of prices and other “intrinsic” product growth effects.

In Figure 1 we plot the unit sales over the time horizon for our videogames data. We see that while the cumulative unit sales in Figure 1(b) is typical of model diffusion patterns, the unit sales in Figure 1(a) exhibit a lot of variation. Interestingly, not all the “peaks” in that figure can be attributed to holiday/seasonal effects. Indeed, in our data, it appears that we can associate some of the peaks with either price reductions or launch of new software titles. So we would, in this product category, expect marketing activities to be a significant driver of sales. Additionally, the absence of advertising and distribution related variables in the data suggests that including the unobservable term in our model specification may be appropriate. The empirical results will reveal whether endogeneity is also an issue with these data.

We also examined the price pattern for videogame consoles and find that it follows the typical pattern of most technology/durable goods. Prices start high and decline steadily over time from a high of about $400 to about $60 at the end of the time horizon. The first big drop in prices is around period 4/5. This corresponds to the entry of PS1 into the market. The subsequent drop is just prior to the entry of N64. A regression of prices on exogenous variables also indicates that prices are correlated with various factor costs of product as well as season dummy variables. We will use some of these exogenous variables as instruments in the analysis (discussed later).
We also find that there is considerable volatility in the number of titles in each month in the earlier stages of the adoption of the 32/64-bit consoles. This reflects to some extent the fact that many games introduced early on were eventual failures, and had no sales in the months following their introduction. Following Bass and Srinivasan (2002), we also check the price and software series for stationarity. We find that these series do have the property, alleviating concerns that one might have about “spurious regression” effects for these data.

As noted previously, one of the concerns in the estimation is the potential endogeneity of prices. We use an instrumental variables procedure to overcome the potential biases arising from the endogeneity concerns. We use factor prices that determine the costs of manufacturing consoles as instruments for prices. The main components of that generation videogame consoles are semiconductors, random-access-memory (RAM) and other electronic components. Further, these consoles were manufactured in several different Asian countries (Japan, Malaysia, Taiwan, etc.). However, the producer price indices (PPIs) from those countries are not readily available to us. Instead, we assume that the PPIs for the console components in their countries of manufacture are correlated with the corresponding PPIs in the U.S. and use these PPIs as instruments for videogame prices. In our empirical analysis we used current and 6 period lagged PPIs as the instruments and found that the only input price that did not correlate well with videogame console prices was that for RAM. Accordingly, we used semiconductors and other electronic components (current and 6

---

10 The motivation for factor prices as instruments for prices implicitly assumes that macro-level shocks to demand for consoles do not result in large shifts in aggregate factor demand. This assumption is valid if factor markets are perfectly competitive. In other contexts, this assumption is predicated on the interpretation of demand shocks as either unobserved (to the econometrician) product characteristics that shift demand, or changing average tastes of consumers for products that impact on demand over time or seasonality. In the case of technology products like consoles, product attributes are not time-varying, and do not explain time-series variation in demand. The effect of the holiday season is picked up by a seasonality dummy. Hence, we are left with a changing tastes interpretation. Controlling for game availability and seasonality, we do not expect changing tastes are large enough to produce big enough shifts in aggregate console demand so as the affect factor prices. Industry trade-press also do not report large macro shocks to demand in this industry during the time period of our data.
period lags) as our instruments. The $R^2$ of the regression of prices just on the instruments was 0.78 indicating that these instruments do a good job in explaining the variation in prices over time.

Our analysis here assumes that number of software titles is exogenous. One justification for this assumption is the institutional feature that decisions regarding making a title available in a particular time period are made well in advance, reducing the possibility of these being correlated with unobserved factors in the hardware sales equation in the current period. In a related analysis of the video-game industry, Clements and Ohashi (2006) instrument for software availability using the average age of software titles provided to a console, and find little differences between OLS and 2SLS estimates of software availability effects (see Columns H1 and H2, Table 3, Clements and Ohashi 2006.)

5. Empirical Results

We first discuss the parameter estimates and then provide implications for the intrinsic probabilities of adoption or failure rates obtained from the estimation after controlling for the effects of the covariates. Subsequently, we discuss the substantive implications of the effects of the covariates and provide short- and long-term elasticities for the price and number of titles variables. In carrying out the estimation, we tried out various intervals at which we estimated the parameters of the “baseline” hazard, i.e. $\alpha_{\tau_j}$ for $\tau_j = 1, 2, \ldots, \Gamma$. We estimated these parameters on a bi-monthly (44 $\alpha_{\tau_j}$ parameters), quarterly (30 $\alpha_{\tau_j}$ parameters), semi-annual (15 $\alpha_{\tau_j}$ parameters) and annual (8 $\alpha_{\tau_j}$ parameters) basis. The substantive nature of the results was consistent across the different time intervals. Here, we provide results only from the semi-annual specification. Having 6 data points to estimate each $\alpha_{\tau_j}$ parameter seemed reasonable and hence we picked this specification. Results from the other models are available from the authors.
5.1. Parameter estimates

Table 1 provides the parameter estimates and the standard errors for two specifications – one in which we account for price endogeneity via instruments and the other in which prices are assumed to be exogenous. In this latter specification however, we constrained the market size parameter (m) to be the same as that under the specification in which prices are endogenous so as to be able to compare the estimates across the two specifications. The actual estimated market size in the exogenous prices case was slightly smaller than that reported in Table 1. As we have 89 monthly observations of data and as the baseline parameters were estimated for six-monthly periods, we have 15 parameters with the first parameter accounting for only 5 months.

Before discussing the specific model parameters, one question of interest is: what is the proportion of total variation in \(\ln[-\ln(1-\frac{(S\alpha_i)(m-X_{t,i+1})}{(m-X_{t,i+1}))}]\) explained by the 3 components – (a) the baseline hazard (\(\alpha\)); (b) covariates (\(\beta\)); and (c) the unobserved error term (\(\epsilon\)). In the case of our videogame data we find these percentages to be 74%, 23% and 3% respectively. This implies the following. First, the baseline or the intrinsic adoption pattern accounts for most of the total variation in the dependent variable (conditional on market size). Second, covariates do account for a sizable portion of the variation and are jointly important to understand the nature of the adoption process in the videogames market. Third, once the baseline and the covariates are incorporated, most the total variation is accounted for, as the contribution of the unobservables term is very small in percentage terms. These findings are quite consistent with the Bass model literature (see for example, Bass et al. 1994 or Bass et al. 2000) which notes that while covariates are significant, it is the basic adoption pattern that is critical to explaining the observed pattern in the data and that the adoption model typically fits the data extremely well. The finding that the unobservable effects account for a small fraction of the total variation also indicates that our assumption that the baseline
parameters remain fixed over 6-month intervals is reasonable. Having established the relative importance of these parameters, we explore next the estimates of these parameters.

Table 1 indicates that the parameters characterizing the baseline hazard are, with a few exceptions, declining over time. This pattern is consistent across the two specifications, although as we would expect a priori, the standard errors in the instrumented case are much larger than those in the exogenous prices case. To understand the pattern implied by these parameters better, in Figure 2 we translate these parameters into the actual baseline hazard \((1-\exp(-\exp(\alpha_t)))\). Since there is a large drop-off in the baseline hazard after about two years, we plot it in two time horizons – months 1 to 23 (Figure 2(a)) and months 24 to 89 (Figure 2(b)). For completeness, we plot the density corresponding to the baseline hazard in Figure 2(c). Figure 2 indicates that till around month 40, the baseline hazard represents a very non-monotonic pattern. Indeed, the baseline hazard decreases, increases and decreases once again. Subsequent to month 40, the pattern is close to a monotonically declining path. The non-monotonic shape of the hazard function in Figure 2 implies that it would be very difficult for a parametric form to accurately reflect this pattern. Figure 2 also indicates that not accounting for endogeneity does have some effect on the baseline hazard in the first two time intervals. However, given the estimated standard errors, the differences are not statistically significant. We return to this difference subsequently, in the context of the estimated elasticities.

Overall, the conclusions from Figure 2 are (a) that the baseline hazard has an irregular pattern that cannot easily be represented via a parametric specification; and (b) not accounting for the endogeneity of prices appears only to have a small impact on the estimated baseline hazard parameters.

Returning to the remaining parameters in Table 1, we see that the effects of all 3 included variables – price, number of software titles (i.e. number of videogames) as well as the “low” season January through March – all have statistically significant effects on the adoption of videogame consoles. We find that accounting for price endogeneity strengthens the price effect somewhat
(again, elasticity implications are investigated below). We also note that the effect of the software titles variable is largely unchanged when prices are instrumented for. We now turn to a discussion of the elasticities computed from the model and their implications.

5.2. Elasticities

It is important to note that we can compute two types of elasticities – “short term” or contemporaneous elasticities and “long term” elasticities. By short term, we mean the percentage change in the current month’s sales if there is a 1% change in price or software titles in this month. By long term elasticities we mean the change in subsequent months’ sales due to a 1% change in price or software titles in this month. This long run effect comes about because sales in period t+1 depends upon the cumulative sales till period t. Suppose there is a price cut in period t. This will increase the cumulative sales till period t. Hence the potential market available in period t+1 is smaller as \( m - X_t \) is smaller if \( X_t \) is larger. If there is no price change in period t+1 then the sales level in that period would be smaller than without the price cut in period t due to the smaller potential market available in that time period.

The contemporaneous “short term” elasticities are presented in Table 2 and the “long term” elasticities are presented in Table 3. We discuss these in turn in the next two subsections.

5.2.1. Short-term elasticities

The mean monthly price elasticity (i.e. averaged over the 89 months) for the videogame consoles estimated using our data is -3.680. This indicates considerable price sensitivity for these consoles. While the magnitude of the elasticity is quite comparable to those reported in Shankar and Bayus (2003), those authors were looking at brand-level sales of a previous generation (16-bit) console. We discuss how these elasticities vary over time periods. Rather than report the elasticities for each month, we report the averages of these monthly elasticities over the 15 time periods for which we estimated the baseline hazard parameters. Table 2(a) contains these short term price elasticities.
Further, we report the elasticities for both endogenous and exogenous prices cases. Additionally, we provide the unit sales effect in addition to the elasticity figure to better understand the substantive content of the elasticities. The price elasticities imply a declining pattern over time. This pattern is consistent with previous findings in the literature on durable products (e.g., Parker and Neelamegham 1997). Nevertheless, it is important to answer why the elasticities decline over time. Note that in the initial time periods, the prices are higher and so is the potential market available for adoption. And the sales levels are low. Hence a price reduction at this point can draw from a larger pool of customers and since the sales levels are relatively low, the number of customers that can be attracted in percentage terms is high. Consequently, the initial elasticities are high. As adoption progresses, the potential market declines and the sales level rises. So in percentage terms, the attractiveness of a price cut is smaller thereby explaining the smaller elasticities. We also computed the price elasticities in each time period at the average price level across all time periods. In this case, the elasticities would largely reflect temporal differences in the intrinsic growth pattern. We find that these elasticities increase over time from -3.36 to -3.73, reflecting the notion that consumers who adopt later have intrinsically higher price elasticities than those who adopt earlier.

The next observation from Table 2(a) is that there is a statistically significant difference in the elasticities for the first 4 6-month periods between the endogenous and exogenous prices specifications. In other words, there is a difference for the first 2 years of data. This could partly be explained by greater correlation between monthly prices and monthly unobserved factors such as advertising in the first two years of the data compared to the subsequent months. We also computed the exogenous price elasticities at the average price level across time periods. This resulted in elasticities anywhere between 10 and 15% smaller than those when prices are treated as being endogenous.
The differences in elasticities are also reflected in differences in predicted unit sales levels in columns 4 and 5 of Table 2(a). Looking at the unit sales effects also highlights the extent of price sensitivity in the market. Specifically, in months 6-11, a 1% price change (of the order of about $3) in any of the 6 months would have led to an increase in approximately 24,000 units of videogame consoles. The average sales level in those months was about 425,000 units. And the average console price was $264. Cutting price by $3 would have reduced the revenues on the 425,000 units by about $1.28MM. And the increased revenues from the incremental units would be approximately $6.26MM. While this looks attractive from a revenue perspective, note that when thinking about margins, while the loss on existing units is still $1.28MM, only a fraction of the revenues accrues as margin to the firm. Hence $6.26MM will have to be adjusted downwards from revenues terms to margin terms to make the appropriate comparison. Nevertheless, such a computation highlights the tradeoffs involved when one is considering a price change in this market.

Table 2(b) highlights the contrast in effects between prices and number of software titles. In particular, we find that the effect of the number of software titles increases over time. In the first few periods, the effect is very small. However, halfway through the time period after month 47, the elasticity exceeds 1 in magnitude and gets as high as 2.07 towards the end of the 89-month period. Consistent with our findings from the parameters estimates in Table 1, we see almost no difference in elasticities between when prices are assumed to be endogenous versus exogenous.

We now discuss how the pattern of elasticities are driven by the nature of the time-series of prices and software-availability, as well as the functional form of the model. From the sales equation (7), note that the marginal effect, and the elasticity of sales with respect to prices are,
The model implies that the marginal effects of marketing mix variables on sales is a function of the levels of all variables in that period, as well as the size of the potential market remaining in that period. In our empirical application, we find that these marginal effects (i.e. the term marginal) are declining over time. This is intuitive since in later periods, the pool of customers from which price cuts or game availability can draw sales is lesser. The elasticities are a function of marginal effects and proportional to the level of variables. The equation above implies that the % change in marginal effects over time is the same for all included variables. Hence, the differences in the pattern of elasticities between prices and software availability are driven mainly by the pattern of evolution these variables over time.\(^{11}\)

One of the interesting implications of the results in Table 2 is its use as a potential input into deciding which marketing instrument to focus on at different points in time if one was interested only in the short term sales impact of these instruments. Table 2 indicates that in the initial time periods since price has a bigger effect on sales (in both elasticity and units terms), it would be the appropriate instrument to use to influence sales. However, as one gets to the later time periods, the number of software titles variable have about as big an effect in terms of unit sales as does the price variable. Indeed, by the months 84-89, a 1% change in the number of software titles results in a unit sales effect of 1,828 as compared to a change of 1,511 due to a 1% change in price. However, it

\(^{11}\) Note that the model can be extended to allow the % change in marginal effects to differ by each marketing mix variable. One approach would be to allow for different price (or software availability) effects over time by interacting prices with, for instance, year-fixed effects. We could not get reliable estimates under this approach given the short time series of the data, and the fact that we have already included month-specific fixed effects in the sales equation. Another approach to flexibility is to allow for higher order terms for prices/software in the sales equation. For these data, we did not find these terms to be significant. We thank an anonymous reviewer for alerting us to this point.
must be noted that the change in sales in unit terms is quite small compared to the price effect in earlier time periods (e.g. a change of 51,000 units in the months 36-41). These results indicate that the optimal marketing policy for firms in this industry is to set low margins initially (i.e. penetration pricing) and to then increase them over time.\footnote{Note that the pattern of falling consoles prices observed in the data does not rule out the possibility that this is exactly what firms are doing. The trade-press reports significant economies of scale in the production of hardware, implying that console production costs are falling over time. Thus, if costs fall faster than prices, then firms could optimally be setting margins low initially and increasing them over time, as implied by these estimates.} Further, the results imply that firms should focus more on market software provision in the later stages of the life cycle. Using a different model and estimation approach, Clements and Ohashi (2006) find similar patterns of effects for the video-game industry.

The important points to take away from our discussion of Table 2 are: (a) the 32-64 bit videogame console market is characterized by consumers who are sensitive to prices as well as to the number of software titles; (b) the model formulation allows us to capture the time-varying effects of these variables; and (c) characterizing the time path of the effects of price and number of software titles enables us to decide the instrument most suitable to influence sales in any given time period.

5.2.2. Long-term elasticities

In Table 3, we provide long run effects for the price and titles variables expressed in terms of changes in unit sales. Specifically, we report the following numbers. Suppose the price of videogame consoles is increased by 1% in month 1. Then, the “current month” column indicates the change in unit sales in that month, i.e. in month 1. The “average of future months” column indicates the change in unit sales in each of the subsequent months (i.e. 2 through 89) averaged over the appropriate number of months (88 in this example). To obtain the total effect over all subsequent months, one can multiply the number in the “average of future months” column by the appropriate number of months (in this illustration that would be 55.57*88). Table 3 indicates that the long term effects of changes in the variables of interest are not negligible. A price increase of 1% in month 21
(sales = 3.3MM units, price = $158) results in decline in sales of 127,000 units. The short term net revenue loss for the industry is $15MM. However, since there is a sales increase of 1520 units in each of the subsequent 68 time periods, the undiscounted total revenue increase in subsequent periods is $11.7MM (this number would be smaller with discounting). So the net effect of the 1% price increase for the firms is about $3.3MM – much smaller than indicated by the short term calculations. In the same way, we can see that the net benefit from a price cut in a given period is much smaller than that implied by the short run effect alone.

Table 3 also show how the short and long term effects of the two variables – prices and number of software titles vary at different points in time. Pursuant to our earlier discussion regarding the choice of appropriate instrument to generate sales, we note that we need to look not just at the tradeoff between the effects of the variables in the short run, but also the net effect after factoring in the long term effects of the two variables. As an illustration, consider the net effect of a price increase of 1% in period 61 and an increase in number of titles by 1% in the same period. Looking at the just the short run effect, we see that there is net loss of 3679-2405 = 1274 units. However, once we factor in the long term effects, the net loss is only about 686 units.

To summarize, the key points of interest from Table 3 are: (a) there are significant long-term effects of prices and number of software titles on sales; and (b) when deciding which instrument to focus on in order to drive sales, one must consider not just the short run tradeoffs between the two instruments but also the long run effects.

6. Parametric functional forms and the forecasting issue

From our discussion of the “baseline” hazard and from Figure 2, it is clear that the parametric forms that imply a monotonic baseline hazard model may not be the most appropriate specification for these data, if the focus is on obtaining the appropriate estimates for the effects of price and number of software titles. However, as noted previously, in many circumstances, we are primarily interested
in forecasting. Here we ask the question: how close to the estimated price and software title effects (from the semiparametric model) can we get using a flexible parametric specification? Further, what are the implications of imposing specific parametric structures on the intrinsic adoption of this generation of videogame consoles? We explore these issues in the context of our data using the Bass and Expo-power hazard specifications, and also using with a higher-order polynomial function of time that approximates the hazard. All models are estimated using the linearization and nested-regression procedure discussed in section 3. Analogous to the semi-parametric specification, we control for price endogeneity in the estimation of all three models.

6.1. Bass model hazard

We first estimated the PHM model imposing the hazard function implied by the Bass model.\(^\text{13}\) Note that the Bass model cannot intrinsically capture the sales fluctuations evidenced in Figure 1 beyond what is implied by the variation in the explanatory variables. Hence, we added month dummies as covariates to capture any seasonal effects in the data. Additionally, we had to include a dummy variable for “holidays” which was an ad-hoc variable created largely to explain sales peaks not accounted for by the monthly dummy variables. So instead of estimating 19 parameters as in the proposed model (Table 1), we estimated 18 parameters in this case (viz. \(p, q, m\), 12 monthly dummies, 1 holiday dummy, 1 price effect, 1 software effect). In Table 4 we present the estimates for the key parameters – \(m, p, q\) and the coefficients for price and number of software titles for the exogenous as well as the endogenous prices scenarios. Once again we constrained the estimates of \(p, q\) and \(m\) across the two specifications in order to ensure comparability of the estimates of the covariates.

Table 4 indicates that we obtain reasonable values for the model parameters estimated. In particular the values of \(p\) and \(q\) lie between 0 and 1. The estimate of \(p\) does appear to be a bit high

\(^{13}\) We also estimated the parameters using the standard additive error term formulation. However, in this case, there is no easy way of accounting for price endogeneity due to the nonlinear model structure.
relative to the estimate for $q$. The estimated market size parameter is quite close to the cumulative sales at the end of the last period (about 45 million units). This property of the model has been discussed elsewhere (e.g. Van den Bulte and Lilien 1997, Venkatesan et al. 2003). Moreover, it is smaller in magnitude compared to that estimated for the proposed model (although since the standard errors are large, the difference is not statistically significant). Further, we find that the effects of marketing activities are understated when we impose the parametric form of the Bass model on the data. Indeed the corresponding elasticities are over 30% smaller than those under the proposed specification. Previous research (c.f. Bass et al. 1994) has shown that the Bass model without covariates performs well in terms of model fit. Our findings with the proposed model concur in the sense that the majority of the variation does come from the intrinsic growth pattern. What our results also seem to imply is that imposing a specific parametric form may further understate the effects of covariates, when it is imposed on data where the underlying intrinsic growth pattern does not, in fact, conform to the parametric model. So covariates play an even smaller role when such a model is imposed on the data.

We also compared the in-sample model fit under the two specifications. We find that the proposed model has an in-sample MAPD (mean absolute percentage deviation) of 0.496 compared to a value of 0.824 for the Bass model. Figure 3 plots the fit of the proposed model and Bass model (with monthly and holiday dummy variables). We see that in general, the proposed model fits the data better except in the 2 peak periods in months 56 and 68 due to the presence of the additional dummy variables. Hence for these data, it does appear that the added flexibility provided by the proposed model helps in fitting the empirical data better.

6.2. Expo-power hazard

In their review article, Seetharaman and Chintagunta (2003) find that one of the most flexible parametric specifications for the baseline hazard is the expo-power distribution (Saha and Hilton
1997). This specification allows for the hazard to be flat, increasing, decreasing, U-shaped, or inverted U-shaped. Further, it requires the estimation of 3 parameters. With the market size parameter as the fourth, this specification will have one parameter more than the Bass model.

We fit the expo-power hazard model to the videogames data accounting for price endogeneity in the same way as before.\textsuperscript{14} We find that the estimated elasticities from this model are about 10-15\% smaller than those from the proposed specification. So while an improvement relative to the Bass-PHM, the specification still results in slightly lower elasticities. Further, the MAPD for this specification is 0.719 which is higher than that from the proposed model. Therefore, we obtain some benefits from the added flexibility of the expo-power specification relative to the Bass-PHM, although these benefits are not very big especially on the model fit dimension. Nevertheless, if forecasting is the primary objective, we may need to use such a parametric form in the estimation.

6.3. Polynomial in time hazard

We also explored using a polynomial function of time to approximate the semi-parametrically estimated baseline hazard. A fifth-order polynomial was found to provide the best fit. We find that the shape of the hazard implied by this model is quite similar to that from the more parsimonious expo-power specification. The polynomial fits the data reasonably well in the later time periods of the data, especially after month 25. However, it did not capture the highly variable pattern prior to month 25 (see figure 2). Overall, in the absence of a suitable parametric form, this approach can be used for forecasting purposes.

7. Conclusions

We investigate the short and long run effects of prices and software availability on the sales of 32/64 bit video-game consoles in the US. To measure these effects empirically, we adopt an estimation

\textsuperscript{14} An alternative flexible specification is the “semi-non parametric” approach of Gallant and Nychka (1987). This specification yielded a baseline hazard pattern similar to that of the expo-power hazard.
framework that controls for the potential endogeneity of marketing mix elements, and allows for a flexible intrinsic growth pattern for the hardware. Flexibility is important since an inappropriate specification for the intrinsic growth rate of the product can bias the effect of marketing variables. We use a proportional hazards formulation for the sales pattern over time, and allow for a flexible baseline hazard via a semiparametric approach. We show how to estimate the parameters of the model while controlling for potential endogeneity issues using instrumental variables.

Our main empirical findings using videogame console sales, prices and software titles data from the 32-64 bit generation are as follows. Both prices and software availability have significant effects on hardware growth. Prices have a bigger effect on sales of consoles in the initial time periods, while the number of software titles in the market has a bigger effect in the later time periods. This supporting a policy for console manufacturers of pricing aggressively initially, followed by an increased focus in generating games in the market in the mature stages of the life-cycle. Additionally, there are significant long-term effects of prices and number of software titles on sales. Characterizing the time path of the effects of price and number of software titles in terms of their long and short-term effects enables one to decide which instrument to focus on to drive sales over time.

We find that the estimated baseline hazard has an irregular pattern that cannot easily be represented via a parametric specification. Our findings indicate that imposing a specific parametric form on the intrinsic growth pattern results in elasticities that are up to 30% smaller than those from the proposed model. Not accounting for the endogeneity of prices appears only to have a small impact on the estimated baseline hazard parameters. However, accounting for the endogeneity impacts the estimated price elasticities for the earlier time periods in the data.

Overall, our study makes methodological as well as substantive contributions. On the methodological side, we provide an approach for relaxing the parametric form for the sales pattern
while accounting for endogeneity in firms’ decision-making variables. On the substantive side we are able to estimate the short- and long-run price and software elasticities for a specific generation of the videogame console market. Further, we able to obtain implications for which marketing instrument needs to be emphasized at various points of the life cycle. Accordingly, this could have implications for marketers of future console generations or of other technology products where price and software are the key drivers of sales.

There are several directions for future research that can be explored. With brand level data, one can study category sales and brand shares and the role of endogeneity at the two levels. Previous research has found the effects of endogeneity to be important at the brand share level (Besanko et al. 1998) and that might also be the case here. The diffusion literature has drawn a distinction between entering covariates as levels or as changes. It might be of interest to think of ways in which the endogeneity issue can be addressed when changes are used in place of levels as we do in this paper. Another potential limitation of our analysis is that the model parameters could be heterogeneous across sub-populations in the market. Given the short time-series of our data, we did not obtain reliable estimates of model specifications in which response parameters were allowed to be heterogeneous in the population. Nevertheless, with richer datasets, this would be an approach worth investigating. We hope that our approach encourages more research on evaluating the impact of specific functional forms for the intrinsic growth rate and the role of endogeneity on the estimated effects of marketing activities in new product models.

References


Appendix – Model Estimation with Bass & Expo-Power Hazards

Bass model specification

The sales equation (5) for the Bass model is (see Bass et al. 2000):

\[
S_n = (m - X_{t,i-1}) \left[ 1 - \exp\left\{ - \exp(W \beta + \epsilon_i) * \ln \left( \frac{b + \exp(at_i)}{b + \exp(at_{i-1})} \right) \right\} \right]
\]

where \( a = p + q \), and \( b = q / p \), and,

\[
g(t_i, Z) = \exp(W \beta + \epsilon_i) * \ln \left( \frac{b + \exp(at_i)}{b + \exp(at_{i-1})} \right) > 0
\]

where, \( p \) and \( q \) are the usual Bass model parameters. Note that for comparability with the proposed model, we have introduced the error term into the Bass model in the same “structural” way as before. The main advantage of this approach is that, conditional on \( m \), \( p \) and \( q \) we can once again linearize the above expression. This form of the model is given as:

\[
\ln \left[ \frac{\ln \left\{ 1 - S_n / (m - X_{t,i-1}) \right\}}{\ln \left\{ (b + \exp(at_i)) / (b + \exp(at_{i-1})) \right\}} \right] = W \beta + \epsilon_i
\]

With linearization, we are also able to account for price endogeneity within the context of the Bass model in a manner similar to the proposed model. The parameters \( m \), \( p \) and \( q \) are then obtained by searching over the space of these parameters that minimized the quadratic form \( \epsilon'V(V'V)^{-1}V'\epsilon \)

where as before, \( \epsilon \) denotes the vector of stacked residuals and \( V \) is the matrix of instruments.

Expo-power specification

The expo-power distribution (Saha and Hilton 1997) is a three parameter distribution that allows for the hazard to be flat, increasing, decreasing, U-shaped, or inverted U-shaped. The corresponding formulae for this function are:

\[
g(t, Z) = \exp(W \beta + \epsilon_i) * \left\{ \gamma \left( \exp(\theta t_i^\alpha) - \exp(\theta t_{i-1}^\alpha) \right) \right\}, \gamma, \alpha > 0
\]

And, the estimation equation is:

\[
\ln \left[ \frac{\ln \left\{ 1 - S_n / (m - X_{t,i-1}) \right\}}{(\exp(\theta t_i^\alpha) - \exp(\theta t_{i-1}^\alpha)) \gamma / \theta} \right] = W \beta + \epsilon_i
\]
Table 1: Parameter estimates and their standard errors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Endogenous prices</th>
<th>Exogenous prices</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
<td>Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Market Size*1e-7 (m)</td>
<td>5.3553*</td>
<td>3.1311</td>
<td>-2.0858</td>
<td>1.4308</td>
</tr>
<tr>
<td>α1</td>
<td>-1.3213</td>
<td>1.2016</td>
<td>-1.8399</td>
<td>0.9818</td>
</tr>
<tr>
<td>α2</td>
<td>-1.3887</td>
<td>1.2512</td>
<td>-1.8399</td>
<td>0.9818</td>
</tr>
<tr>
<td>α3</td>
<td>-5.2719</td>
<td>0.9340</td>
<td>-5.4829</td>
<td>0.6795</td>
</tr>
<tr>
<td>α4</td>
<td>-2.0718</td>
<td>0.9087</td>
<td>-2.2488</td>
<td>0.6282</td>
</tr>
<tr>
<td>α5</td>
<td>-3.3956</td>
<td>0.9477</td>
<td>-5.3333</td>
<td>0.4869</td>
</tr>
<tr>
<td>α6</td>
<td>-4.4119</td>
<td>0.9149</td>
<td>-4.4469</td>
<td>0.5557</td>
</tr>
<tr>
<td>α7</td>
<td>-3.0536</td>
<td>1.0885</td>
<td>-3.0266</td>
<td>0.4548</td>
</tr>
<tr>
<td>α8</td>
<td>-3.4651</td>
<td>1.3091</td>
<td>-3.4609</td>
<td>0.4151</td>
</tr>
<tr>
<td>α9</td>
<td>-5.2968</td>
<td>1.5099</td>
<td>-5.2017</td>
<td>0.4609</td>
</tr>
<tr>
<td>α10</td>
<td>-4.5300</td>
<td>1.8108</td>
<td>-4.4056</td>
<td>0.5135</td>
</tr>
<tr>
<td>α11</td>
<td>-5.8429</td>
<td>2.3019</td>
<td>-5.6775</td>
<td>0.5753</td>
</tr>
<tr>
<td>α12</td>
<td>-4.9767</td>
<td>2.8350</td>
<td>-4.8316</td>
<td>0.6098</td>
</tr>
<tr>
<td>α13</td>
<td>-6.7523</td>
<td>3.2900</td>
<td>-6.5387</td>
<td>0.6717</td>
</tr>
<tr>
<td>α14</td>
<td>-6.1706</td>
<td>3.4991</td>
<td>-6.0262</td>
<td>0.6030</td>
</tr>
<tr>
<td>α15</td>
<td>-7.5204</td>
<td>3.8433</td>
<td>-7.2320</td>
<td>0.7201</td>
</tr>
<tr>
<td>Price</td>
<td>-0.0257</td>
<td>0.0061</td>
<td>-0.0226</td>
<td>0.0053</td>
</tr>
<tr>
<td>Titles</td>
<td>0.0016</td>
<td>0.0005</td>
<td>0.0015</td>
<td>0.0007</td>
</tr>
<tr>
<td>Season (Jan-Mar)</td>
<td>-1.2258</td>
<td>0.3027</td>
<td>-1.1283</td>
<td>0.3418</td>
</tr>
</tbody>
</table>

* Estimates for the “exogenous prices” model are obtained by constraining the market size to be the same as for the “endogenous prices” case. This ensures comparability of the other estimates.
Table 2a: Contemporaneous elasticities & unit sales effects of a 1% increase in price

<table>
<thead>
<tr>
<th>Average per month</th>
<th>Elasticities</th>
<th>Unit sales effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endogenous prices</td>
<td>Exogenous prices</td>
</tr>
<tr>
<td>Month 1-5</td>
<td>-9.2894*</td>
<td>-8.2148*</td>
</tr>
<tr>
<td>Month 6-11</td>
<td>-6.5338*</td>
<td>-5.7686*</td>
</tr>
<tr>
<td>Month 12-17</td>
<td>-5.0567*</td>
<td>-4.4595*</td>
</tr>
<tr>
<td>Month 18-23</td>
<td>-4.4284*</td>
<td>-3.9039*</td>
</tr>
<tr>
<td>Month 24-29</td>
<td>-3.7491</td>
<td>-3.3035</td>
</tr>
<tr>
<td>Month 30-35</td>
<td>-3.3537</td>
<td>-2.9545</td>
</tr>
<tr>
<td>Month 36-41</td>
<td>-3.2307</td>
<td>-2.8458</td>
</tr>
<tr>
<td>Month 42-47</td>
<td>-3.2084</td>
<td>-2.8261</td>
</tr>
<tr>
<td>Month 48-53</td>
<td>-3.1164</td>
<td>-2.7450</td>
</tr>
<tr>
<td>Month 54-59</td>
<td>-2.5145</td>
<td>-2.2139</td>
</tr>
<tr>
<td>Month 60-65</td>
<td>-2.6432</td>
<td>-2.3275</td>
</tr>
<tr>
<td>Month 66-71</td>
<td>-2.4716</td>
<td>-2.1761</td>
</tr>
<tr>
<td>Month 72-77</td>
<td>-2.3344</td>
<td>-2.0552</td>
</tr>
<tr>
<td>Month 78-83</td>
<td>-2.5043</td>
<td>-2.2051</td>
</tr>
<tr>
<td>Month 84-89</td>
<td>-1.6952</td>
<td>-1.4919</td>
</tr>
</tbody>
</table>

* Difference under endogenous and exogenous prices significantly different from zero at the 5% level

Table 2b: Contemporaneous elasticities & unit sales effects of a 1% increase in number of software titles

<table>
<thead>
<tr>
<th>Average per month</th>
<th>Elasticities</th>
<th>Unit sales effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endogenous prices</td>
<td>Exogenous prices</td>
</tr>
<tr>
<td>Month 1-5</td>
<td>0.0169</td>
<td>0.0161</td>
</tr>
<tr>
<td>Month 6-11</td>
<td>0.1022</td>
<td>0.0979</td>
</tr>
<tr>
<td>Month 12-17</td>
<td>0.0513</td>
<td>0.0491</td>
</tr>
<tr>
<td>Month 18-23</td>
<td>0.3006</td>
<td>0.2878</td>
</tr>
<tr>
<td>Month 24-29</td>
<td>0.3541</td>
<td>0.3390</td>
</tr>
<tr>
<td>Month 30-35</td>
<td>0.3058</td>
<td>0.2927</td>
</tr>
<tr>
<td>Month 36-41</td>
<td>0.4581</td>
<td>0.4385</td>
</tr>
<tr>
<td>Month 42-47</td>
<td>0.9005</td>
<td>0.8619</td>
</tr>
<tr>
<td>Month 48-53</td>
<td>1.5682</td>
<td>1.5009</td>
</tr>
<tr>
<td>Month 54-59</td>
<td>1.7236</td>
<td>1.6496</td>
</tr>
<tr>
<td>Month 60-65</td>
<td>1.8369</td>
<td>1.7580</td>
</tr>
<tr>
<td>Month 66-71</td>
<td>2.0581</td>
<td>1.9696</td>
</tr>
<tr>
<td>Month 72-77</td>
<td>2.0484</td>
<td>1.9603</td>
</tr>
<tr>
<td>Month 78-83</td>
<td>2.0654</td>
<td>1.9766</td>
</tr>
<tr>
<td>Month 84-89</td>
<td>1.9356</td>
<td>1.8524</td>
</tr>
</tbody>
</table>
Table 3: Current month and average of future months’ change in sales for a 1% increase in current month’s price and in current month’s number of software titles

<table>
<thead>
<tr>
<th>Month at which the price is changed</th>
<th>Price Current month</th>
<th>Price Average of future months</th>
<th>Number of software titles Current month</th>
<th>Number of software titles Average of future months</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-5768.6</td>
<td>55.5771</td>
<td>5.54286</td>
<td>-0.0534</td>
</tr>
<tr>
<td>11</td>
<td>-1342.2</td>
<td>14.4492</td>
<td>24.051</td>
<td>-0.2589</td>
</tr>
<tr>
<td>21</td>
<td>-127219</td>
<td>1519.57</td>
<td>20996.1</td>
<td>-250.79</td>
</tr>
<tr>
<td>31</td>
<td>-5762.3</td>
<td>77.4551</td>
<td>383.343</td>
<td>-5.1528</td>
</tr>
<tr>
<td>41</td>
<td>-45103</td>
<td>653.191</td>
<td>7553.5</td>
<td>-109.39</td>
</tr>
<tr>
<td>51</td>
<td>-7857.3</td>
<td>128.003</td>
<td>3749.65</td>
<td>-61.085</td>
</tr>
<tr>
<td>61</td>
<td>-3679.0</td>
<td>57.9968</td>
<td>2404.55</td>
<td>-37.906</td>
</tr>
<tr>
<td>71</td>
<td>-6207.2</td>
<td>71.3778</td>
<td>5619.76</td>
<td>-64.623</td>
</tr>
<tr>
<td>81</td>
<td>-1282.3</td>
<td>12.7016</td>
<td>1000.89</td>
<td>-9.9141</td>
</tr>
</tbody>
</table>

Table 4: Bass model estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Endogenous prices</th>
<th>Exogenous prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Market Size*1e-7 (m)</td>
<td>4.6000*</td>
<td>2.2349</td>
</tr>
<tr>
<td>p</td>
<td>0.3004*</td>
<td>0.1786</td>
</tr>
<tr>
<td>q</td>
<td>0.4167*</td>
<td>0.2387</td>
</tr>
<tr>
<td>Price</td>
<td>-0.0165</td>
<td>0.0039</td>
</tr>
<tr>
<td>Titles</td>
<td>0.0009</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

* Estimates for the “exogenous prices” model are obtained by constraining the market size (m), p and q to be the same as for the “endogenous prices” case. This ensures comparability of the estimates of the covariates – price and number of titles.
Figure 1: Monthly unit sales and cumulative unit sales over time

1(a) Unit Sales

1(b) Cumulative Unit Sales
Figure 2(a): “Baseline” hazard for months 1-23

![Graph showing baseline hazard for months 1-23 with endogenous and exogenous price lines.]

Figure 2(b): “Baseline” hazard for months 24-89

![Graph showing baseline hazard for months 24-89 with endogenous and exogenous price lines.]

Figure 2(c) Density corresponding to the baseline hazard

Figure 3: Model Fit

Sales*1e-7

Month

Data

Semiparametric

Bass Model with Monthly & Holiday dummies