



Munich Personal RePEc Archive

Variable marginal propensities to pirate and the diffusion of computer software

Waters, James

University of Nottingham

10 April 2013

Online at <https://mpra.ub.uni-muenchen.de/46036/>
MPRA Paper No. 46036, posted 12 Apr 2013 15:19 UTC

Variable marginal propensities to pirate and the diffusion of computer software

James Waters *
Nottingham University Business School

Working paper - comments welcome

April 11, 2013

Abstract

In this paper, we empirically investigate the dynamics of the marginal propensity to pirate for computer software. We introduce a state space formulation that allows us to estimate error structures and parameter significance, in contrast to previous work. For data from 1987-92, we find a rising propensity to pirate as the number of existing pirate copies increases, and higher late piracy incidence than implied by static models. We strengthen prior results on the impact of piracy in the spreadsheet market, finding it to be the only significant internal influence on diffusion. However, when we allow for negative error correlation between legal and pirate acquisitions, we contradict earlier work by finding that, in the word processor market, piracy did not contribute to diffusion and only eroded legal sales.

1 Introduction

Social contact has long been implicated in technology diffusion, following Bass (1969). The idea is that existing users of a technology influence non-users to adopt the technology. Similar mechanisms have been proposed for describing markets subject to software piracy, the illegal copying of software. In pirate diffusion literature including Givon *et al.* (1995), Prasad and Mahajan (2003), and Liu *et al.* (2011), influenced non-users may acquire the legal or pirated good. Owners of the pirated good may influence non-users, like legal owners.

The pirate diffusion literature presents a variety of reasons why piracy may be beneficial to legal sellers and consumers. In Givon *et al.* (1995) and Givon *et al.* (1997) it is suggested that pirate acquisitions may accelerate legal diffusion through pirate owners' social contacts with non-users. Prasad and Mahajan (2003) show that legal profits may be increased for

*University of Nottingham Business School. Correspondence: Nottingham University Business School, Jubilee Campus, Nottingham, NG8 1BB, United Kingdom. E-mail: james.waters@nottingham.ac.uk

the same reason, when piracy rates are subject to control by the legal sellers and sales affect price preferences of remaining non-users in a specific way. Liu *et al.* (2011) proposes a similar mechanism, where either piracy or pricing can be selected as routes to obtain optimal diffusion speed prior to mature market sales.

The marginal propensity to pirate is the proportion of pirate acquisitions out of total new acquisitions. Its value and dynamics are critical influences on whether piracy does indeed benefit legal sellers. If the marginal propensity to pirate rises as the market size increases, then legal sellers will capture little of the late market. If they then take measures to avoid piracy, consumer welfare is likely to be affected.

Despite the importance of the marginal propensity to pirate, we are unaware of any pirate diffusion studies that test whether it rises or falls with diffusion. Much pirate diffusion research has been theoretical. The empirical work by Givon *et al.* (1995) and Givon *et al.* (1997) assumes that the marginal propensity to pirate is constant. In Haruvy *et al.* (2004), the ratio of pirate sales to legal sales can fall at a constant exponential rate as the number of users increases. However, the interpretation of the rate in terms of piracy protection and resulting company optimisation function precludes the possibility of an increasing ratio.

There are reasons to believe that increased diffusion could lower or raise the share of piracy in acquisitions. For example, on one hand legal sellers may find it cost-effective to take action against piracy only when it reaches a certain level, so increased piracy could lower the marginal propensity to pirate. On the other hand, widespread piracy may make new piracy less difficult and more socially acceptable, so that the marginal propensity to pirate would rise with higher piracy prevalence.

In this paper, we estimate the level and change in the marginal propensity to pirate for data on spreadsheets and word processors. The statistical significance of the parameters and the models' predictive power is assessed. We compare diffusion when we allow for variable marginal propensity to pirate to diffusion without it.

Our theoretical model is a small modification of that in Givon *et al.* (1995). It introduces an adjustment factor to pirate sales, where the factor is the number of users of pirate goods raised to an estimated coefficient. The adjustment represents the effect of factors promoting or hindering piracy. We also consider an alternative modification where past piracy explicitly adjusts piracy's share of sales.

The stochastic component of the model includes errors in the legal and pirate acquisitions, and allows for their correlation. In order to achieve identification, we restrict the error matrix to depend on a single parameter. However, we consider multiple forms for the error, including positive and negative correlation and heterogeneity.

We use data on legal software sales taken from Givon *et al.* (1995), which is also used in Haruvy *et al.* (2004) and (for calibration) in Liu *et al.* (2011). As with these prior authors, we have no piracy data. Givon *et al.* (1995) estimates their model by non-linear

least squares while omitting joint error specifications between the legal and pirate data, and so do not report parameter standard errors. Standard errors are also missing from non-linear least squares estimates in Givon *et al.* (1997), simulated annealing estimates in Haruvy *et al.* (2004), and calibrated model solutions in Liu *et al.* (2011).

We estimate our model by formulating it in state space form, with pirate acquisitions as an unobserved state variable. We calculate one-step ahead predictions using the Kalman filter, and maximise the resulting likelihood to give parameter estimates. We allow for cross-sectional relations between legal and pirate diffusion, and present parameter standard deviations unlike prior work. We use the continuous time, discrete observation extended Kalman filter to avoid time interval bias, in common with Xie *et al.* (1997). However, whereas they use the filter projection as a Bayesian updating procedure for parameter estimates and sales simultaneously, we leave the parameters outside the state variable, and so make them available for classical estimation. As extensive piracy represents a hidden phenomenon of uncertain impact, limiting the impact of prior beliefs is a prudent approach to analysis and permits classical inference.

Our central estimates show that the share of pirate acquisitions out of current acquisitions rose with past piracy. The expanded specification offer gains in fit and assumption plausibility that were robust across different deterministic and stochastic specifications, but frequently lacked parameter certainty. Predictive performance is mixed. We find dynamic estimates of piracy that are higher than static estimates at long time scales.

Givon *et al.* (1995) find that past piracy is an important internal influence on spreadsheet diffusion. We strengthen their result, finding that piracy was the only economically and statistically significant internal influence on spreadsheet diffusion, with no role for past legal sales. Our finding is consistent across all specifications. Givon *et al.* (1995) also find that piracy influenced word processor diffusion. In many specifications we obtain similar results. However, when we allow for negative correlation between legal and pirate errors, piracy is a negligible influence on diffusion and only serves to displace legal sales. The negative correlation stochastic specification outperforms models with no correlation, and is our preferred specification. We interpret the difference between the results on piracy's effect as being due to stochastic correlations being incorrectly ascribed to deterministic links when no correlations are allowed.

In section 2 we present our model and in section 3 we look at the data and empirical method. Results are in section 4 and section 5 concludes.

2 Model

In this section we present our model of diffusion with increasing marginal propensity to pirate. The model is a small deterministic variation on the one described in Givon *et al.* (1995), and a larger stochastic variation. It describes the joint evolution of a technology's

acquisition by legal and pirate means. The deterministic component is similar to a bivariate Bass model in that either source or an external advertiser can inform a non-user about the technology, who then adopts. The stochastic specification allows for correlation in the adoptions by either route.

There is a population of agents of constant size m who are able to buy a computer. The adoption process for computers follows a standard univariate Bass model. At any time, some of the population will have acquired a computer, and the rest will not yet have bought one and remain potential buyers. Initially, there are no owners. The potential users are subject to external advertising, so that a fixed proportion p of them are contacted by advertisers and then buy the computer in any time period. There is also a word-of-mouth effect by which an additional share of potential adopters adopts in the period, where the share is proportional to the number of previous adopters with constant of proportionality equal to q/m . The diffusion pattern for computers thus follows the differential equation

$$dN_t/dt = (p + q\frac{N_t}{m})(m - N_t) \quad (1)$$

An agent can acquire a computer software product only if they own a computer. Of the computer owning population, a number Z_t of these potential software users will have acquired the software and the remainder totalling $N_t - Z_t$ will not yet have acquired it. They can acquire it only once. Initially there are no computer software users. The software can be produced as a legal or pirate copy. The number of legal owners is X_t and the number of pirate owners is Y_t , so $X_t + Y_t = Z_t$.

Non-users are subject to external influence so that they acquire legal copies at an instantaneous rate of a . They are also subject to internal influences from current legal and pirate owners. Legal owners influence them to acquire either legal or pirate copies at a rate that is linear in the number of legal owners, b_1X_t/N_t . Pirate owners influence them to acquire software by either route at a rate linear in the number of pirate owners, b_2Y_t/N_t . For non-users who are internally influenced to acquire the software, a share α acquires the legal good, while the remaining $1 - \alpha$ intend to acquire a pirate copy. These number of these motivated non-users who adopt the pirate copy is then either magnified or diminished by the number of existing pirate copies. For example, it may be magnified if current pirates make piracy more acceptable, or diminished if increased piracy leads to anti-piracy measures being taken. The magnification or diminution is represented by a multiplier applied to the number of pirate adopters, $\max(Y_{t-1}, 1)^\epsilon$, depending on whether ϵ is greater or less than zero. Givon *et al.* (1995) constrain $\epsilon = 0$.

Thus, we have the following model

$$\begin{aligned}
dX_t &= \left(\left[a + \alpha \frac{b_1 X_t + b_2 Y_t}{N_t} \right] (N_t - X_t - Y_t) \right) dt + dw_1 \\
dY_t &= \left(\left[(1 - \alpha) \max(Y_t, 1)^\epsilon \frac{b_1 X_t + b_2 Y_t}{N_t} \right] (N_t - X_t - Y_t) \right) dt + dw_2
\end{aligned} \tag{2}$$

where $\mathbf{dw} = (dw_1, dw_2) \sim N(0, \mathbf{Q}dt)$ is a normal error term with covariance matrix $\mathbf{Q}dt$. The explicit introduction of general errors and allowance for their covariance is a novelty over the specification in Givon *et al.* (1995), or the sampling errors in Haruvy *et al.* (2004). We consider their structure in the estimation section.

3 Estimation

In this section, we describe our empirical approach, presenting the data and estimation method.

3.1 Data

The data we use is from Givon *et al.* (1995). It consists of legal sales of personal computers using a DOS operating system, of spreadsheets, and of word processors in the UK, and is reported monthly from January 1987 to August 1992 inclusive. As with Givon *et al.* (1995), we assume that DOS personal computers were introduced in October 1981 and the two software products were introduced in October 1982. These assumptions are used to determine initial values for cumulative sales in January 1987.

3.2 Estimation by maximum likelihood

We now present the estimation method for our model and its restriction to the Givon *et al.* (1995) pirate model. It is maximum likelihood estimation with tracking of the likelihood function through an extended Kalman filter. The approach generates estimates of the joint error structure in legal and pirate acquisitions, and parameter standard errors.

3.2.1 The extended Kalman filter with continuous state and discrete observations

The extended Kalman filter with continuous state and discrete observations operates on state space models of the form

$$d\boldsymbol{\xi}_{t+1}/dt = f(\boldsymbol{\xi}_t, \mathbf{u}_t) + \mathbf{v}_{t+1} \quad (3)$$

$$\mathbf{z}_t = h(\boldsymbol{\xi}_t) + \mathbf{w}_t \quad (4)$$

where \mathbf{z}_t is a vector of variables observed at time t and $\boldsymbol{\xi}_t$ is a state vector of possibly unobserved variables. \mathbf{u}_t is a vector of exogenous variables, while f and h are differentiable functions. The error vectors \mathbf{v}_{t+1} and \mathbf{w}_{t+1} are mutually uncorrelated white noise, with contemporaneous error variances given by $E(\mathbf{v}_t \mathbf{v}_t^T) = \mathbf{Q}_t$ and $E(\mathbf{w}_t \mathbf{w}_t^T) = \mathbf{R}$.

The filter generates repeated linear forecasts based on past data given at discrete intervals, with forecasts generated recursively through a linearised approximation to the continuous generating system. It proceeds by two steps at each period in a time series, alternating between forecasting based on past data and projection based on current data. It tracks the forecasted state variable $\boldsymbol{\xi}_t$ given data available at time $t-1$ (when the forecast is denoted $\boldsymbol{\xi}_{t|t-1}$) and the projected state variable given data available at time t (the projection is denoted $\boldsymbol{\xi}_{t|t}$). The forecast mean squared errors are also tracked. They are denoted $\mathbf{P}_{t|t-1} = E((\boldsymbol{\xi}_t - \boldsymbol{\xi}_{t|t-1})(\boldsymbol{\xi}_t - \boldsymbol{\xi}_{t|t-1})^T)$ and $\mathbf{P}_{t|t} = E((\boldsymbol{\xi}_t - \boldsymbol{\xi}_{t|t})(\boldsymbol{\xi}_t - \boldsymbol{\xi}_{t|t})^T)$.

In detail, the filter stages are as follows.

Initialisation

Estimates are made of the state vector and its mean squared error matrix in the absence of any information at time zero, that is, of $\boldsymbol{\xi}_{0|0}$ and $\mathbf{P}_{0|0}$.

Forecasting

Given $\boldsymbol{\xi}_{t|t}$ and $\mathbf{P}_{t|t}$ for any t , we integrate the equations

$$d\boldsymbol{\xi}_t/dt = f(\boldsymbol{\xi}_t, \mathbf{u}_t) \quad (5)$$

$$d\mathbf{P}_t/dt = \mathbf{F}_t \mathbf{P}_t + \mathbf{P}_t \mathbf{F}_t^T + \mathbf{Q}_t \quad (6)$$

where $\mathbf{F}_t = df/d\boldsymbol{\xi}^T$ is the Jacobian of f evaluated at $(\boldsymbol{\xi}_t, \mathbf{u}_t)$. The integration is performed from t to $t+1$, with the initial $\boldsymbol{\xi}_t = \boldsymbol{\xi}_{t|t}$ and $\mathbf{P}_t = \mathbf{P}_{t|t}$ in the first and second equations respectively. We set $\boldsymbol{\xi}_{t+1|t}$ and $\mathbf{P}_{t+1|t}$ to be the integrated values at the corresponding end points. The forecasted value for the observation equation and its MSE are then

$$\mathbf{z}_{t+1|t} = \mathbf{H} \boldsymbol{\xi}_{t+1|t} \quad (7)$$

$$MSE(\mathbf{z}_{t+1|t}) = \mathbf{H}_{t+1} \mathbf{P}_{t+1|t} \mathbf{H}_{t+1}^T + \mathbf{R}_{t+1} \quad (8)$$

where $\mathbf{H}_{t+1} = dh/d\xi^T$ is the Jacobian of h evaluated at $\xi_{t+1|t}$.

Updating

Given $\xi_{t+1|t}$ and $\mathbf{P}_{t+1|t}$, we update the forecasts with data z_{t+1} at time $t + 1$ using the formulae

$$\xi_{t+1|t+1} = \xi_{t+1|t} + \mathbf{K}_{t+1}(z_{t+1} - h(\xi_{t+1|t})) \quad (9)$$

$$\mathbf{P}_{t+1|t+1} = (\mathbf{I} - \mathbf{K}_{t+1}\mathbf{H}_{t+1})\mathbf{P}_{t+1|t} \quad (10)$$

where \mathbf{I} is the identity matrix with dimension equal to the number of state variables, and

$$\mathbf{K}_{t+1} = \mathbf{P}_{t+1|t}\mathbf{H}_{t+1}^T(\mathbf{H}_{t+1}\mathbf{P}_{t+1|t}\mathbf{H}_{t+1}^T + \mathbf{R}_{t+1})^{-1} \quad (11)$$

3.2.2 State space representation of the pirate diffusion model

We may represent our extended pirate diffusion model in the state space models with the following definitions:

$$\xi_t = \begin{pmatrix} X_t \\ Y_t \end{pmatrix} \quad (12)$$

$$f(\xi_t, \mathbf{u}_t) = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \quad (13)$$

$$\mathbf{u}_t = 0 \quad (14)$$

$$\mathbf{F} = \begin{pmatrix} F_{1,1} & F_{1,2} \\ F_{2,1} & F_{2,2} \end{pmatrix} \quad (15)$$

$$\mathbf{v}_t = \begin{pmatrix} \epsilon_X \\ \epsilon_Y \end{pmatrix} \quad (16)$$

$$z_t = (X_t) \quad (17)$$

$$h(\xi_t) = \xi_t \quad (18)$$

$$\mathbf{H} = (1 \ 0) \quad (19)$$

$$\mathbf{w}_t = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (20)$$

where the components of the vector $f(\xi_t, \mathbf{u}_t)$ are given by

$$f_1 = \left[a + \alpha \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \quad (21)$$

$$f_2 = \left[(1 - \alpha) \max(Y_{t-1}, 1)^\epsilon \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \quad (22)$$

and the components of the matrix \mathbf{F} are given by the following expressions:

$$\mathbf{F}_{1,1} = \alpha b_1 - a - 2\alpha \frac{b_1}{N_t} X_{t-1} - \alpha \frac{b_1 + b_2}{N_t} Y_{t-1} \quad (23)$$

$$\mathbf{F}_{1,2} = \alpha b_2 - a - \alpha \frac{b_1 + b_2}{N_t} X_{t-1} - 2\alpha \frac{b_2}{N_t} Y_{t-1} \quad (24)$$

$$\begin{aligned} \mathbf{F}_{2,1} = & \left[(1 - \alpha) \max(Y_{t-1}, 1)^\epsilon \frac{b_1}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \\ & - \left[(1 - \alpha) \max(Y_{t-1}, 1)^\epsilon \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] \end{aligned} \quad (25)$$

$$\begin{aligned} \mathbf{F}_{2,2} = & \left[(1 - \alpha) \epsilon \max(Y_{t-1}, 1)^{\epsilon-1} \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] \\ & \times (N_t - X_{t-1} - Y_{t-1}) \\ & + \left[(1 - \alpha) \max(Y_{t-1}, 1)^\epsilon \frac{b_2}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \\ & - \left[(1 - \alpha) \max(Y_{t-1}, 1)^\epsilon \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] \end{aligned} \quad (26)$$

The stochastic component of our model includes all sources of error. In contrast, the formulations in Schmittlein and Mahajan (1982) and Basu *et al.* (1995) allocate all error to differences from multinomial sampling of adoption timing. Thus, while both we and these authors use maximum likelihood estimation, we are not susceptible to the type of criticism levelled in Srinivasan and Mason (1986) that our error specification leads to underestimation of errors.

Our general variance-covariance matrix specification allows for contemporaneous correlation only and is given by

$$\mathbf{Q}_t = \begin{pmatrix} q_{1,1} & q_{1,2} \\ q_{1,2} & q_{2,2} \end{pmatrix} \quad (27)$$

$$\mathbf{R}_t = \mathbf{0} \quad (28)$$

for scalars $q_{1,1}$, $q_{1,2}$, $q_{2,1}$, and $q_{2,2}$.

We have assumed that the errors occur in the state equation rather than the observation equation, so that errors are persistent over time. This approach is consistent with the accumulating errors used in estimation methods including OLS (Bass, 1969), NLS (Srinivasan and Mason, 1986; Jain and Rao, 1990), and MLE (Schmittlein and Mahajan, 1982; Basu *et al.*, 1995) specifications of the deterministic-stochastic Bass model.

The restriction on the \mathbf{R} matrix reduces the number of parameters in our model. State space models are typically underidentified in maximum likelihood estimation (Hamilton, 1994, pp.387-8). To achieve identification, we further restrict the parameters in the \mathbf{Q} matrix. Our initial specification sets $q_{1,1} = q_{2,2} = \sigma^2$ for some constant σ^2 and $q_{1,2} = q_{2,1} = 0$. Later, we consider alternative specifications for the fixed parameters.

The initial state vector $\boldsymbol{\xi}_{0|0}$ in January 1987 is generated by iterating on the system in equation 2 from October 1982 using the parameters estimated in Givon *et al.* (1995). The initial mean squared error matrix $\mathbf{P}_{0|0}$ is assumed to be the zero matrix, so the starting state vector is known with certainty.

For the forecasting stage, we integrate equations 5 and 6 numerically over ten iterations. We also require estimates of N_t . From equation 1, it follows a Bass model. Givon *et al.* (1995) make the same assumption and fit the equation by non-linear least squares. We retain their estimated parameters of $p = 0.00037$, $q = 0.0316$, and $m = 15,386,100$.

3.2.3 Maximum likelihood estimation

The one step ahead forecasts for the observation $z_{t+1|t}$ and its mean squared error $MSE(z_{t+1|t})$ are described by equations 5 and 6. Given a distribution f_{Z_t} of the next observation dependent on these two parameters, we may construct the sample log likelihood as

$$\sum_{t=0}^{T-1} \log f_{Z_{t+1}}(z_{t+1}) \quad (29)$$

where the distribution $f_{Z_{t+1}}$ is conditioned on $z_{t+1|t}$ and $MSE(z_{t+1|t})$. Under the assumption of normal distributions for $\boldsymbol{\xi}_{0|0}$, v_t , and w_t , the log likelihood is (Hamilton, 1994, p.385)

$$\begin{aligned}
\log f_{z_{t+1}}(z_{t+1}) = & (2\pi)^{-n/2} |\mathbf{H}_{t+1} \mathbf{P}_{t+1|t} \mathbf{H}_{t+1}^T + \mathbf{R}_{t+1}|^{-1/2} \\
& \times \exp\{-(1/2)(z_{t+1} - H_t \boldsymbol{\xi}_{t+1|t})^T (\mathbf{H}_{t+1} \mathbf{P}_{t+1|t} \mathbf{H}_{t+1}^T + \mathbf{R}_{t+1})^{-1} \\
& \times (z_{t+1} - H_t \boldsymbol{\xi}_{t+1|t})\}
\end{aligned} \tag{30}$$

where $|\mathbf{M}|$ is the determinant of \mathbf{M} and n is the dimension of w_t .

We maximise the log likelihood function in equation 30 numerically. The maximisation has to give estimates in feasible parameter regions, with positive q variance parameters, a , b_1 , and b_2 contact parameters that are positive and bounded by unity, and the same for the α share parameter. We further constrain the ϵ parameter to lie between -0.2 and 0.2 , the α parameter to be no larger than 0.02 , and $q_{1,1}$ not to exceed 10^9 . Estimates are comfortably within these domains, so they restrict the region for checking without excluding probable solutions.

To constrain the variables to lie in the required domains, we map to them by functions whose input variables are unconstrained (see Hamilton (1994, pp.146-8)). The functions are $\phi = 0.2\epsilon/(1 + |\epsilon|)$, and $\phi = k\lambda^2/(1 + \lambda^2)$ with the other parameters replacing λ for appropriate rescaling factors k . We then maximise the transformed functions with respect to the unconstrained variables using a Nelder-Mead algorithm. We start the algorithm from the parameter solutions in Givon *et al.* (1995). The solutions in the transformed parameters give solutions in the original parameters.

The Hessian for the maximised transformed function yields second derivative estimates of the standard errors for the transformed parameters. We can calculate standard error estimates for the original parameters by calculating the Hessian with respect to the non-transformed function. However, a flat likelihood function in a couple of the parameter directions and limits on accuracy for numerically calculated second derivatives meant that negative estimates of variance were occasionally produced in the non-transformed function (but never in the transformed function). So we use variance estimates calculated from the outer product of the score matrix at the original parameter values. The estimates were invariably positive.

The estimation was implemented in the R programming language (R Development Core Team, 2009) using the library packages MASS and numDeriv. We employed the Microsoft Excel add-in Excel2LaTeX to generate the tables. The code is available from the author's website¹.

¹http://ebasic.easily.co.uk/02E044/05304E/mpp_and_diffusion.html

4 Results

In this section we present our results. Subsection 4.1 gives estimates for our model and its restriction with constant marginal propensity to pirate, as in Givon *et al.* (1995). Subsection 4.2 looks at the model’s out of sample performance, subsection 4.3 examines estimates with alternative error specifications, and subsection 4.4 considers parameter estimates for a qualitatively similar but functionally different model.

4.1 Parameter estimates

Table 1: Parameter estimates for our pirate model and the $\epsilon = 0$ restricted model

	WP	WP	S	S
a	0.00146	0.00151	0.00155	0.00148
	0.00463	0.0048	0.00429	0.00262
b_1	0.109	0.229	0.00176	0.00000
	0.798	1.01	0.595	0.341
b_2	0.0888	0.0962	0.0509	0.114 **
	0.356	0.173	0.38	0.0547
α	0.163	0.124 **	0.224	0.101 **
	0.39	0.0596	1.36	0.045
σ^2	0.0108 ***	0.0107 ***	0.00379 ***	0.00403 ***
	0.00202	0.00203	0.000732	0.000756
ϵ	0.0226		0.0684	
	0.171		0.484	
AIC	1287.0	1285.1	1217.0	1218.8
MSE (%)	100	100	95	100

Standard deviations are shown below the coefficients. *** denotes a p-value of less than 0.01, ** of less than 0.05, * of less than 0.1. MSEs are expressed as percentages of the MSE for the corresponding restricted model. σ^2 is reported in units of 10^9 .

Table 1 shows parameter estimates for our model and its restriction to the Givon *et al.* (1995) model. In column one, our model is fitted to the word processor data. The a parameter equals 0.00146, which is higher than the 0.0002 rate reported in Givon *et al.* (1995). Our rates lie between the mean and median of estimates reported in the meta-analysis of Bass curves in Van den Bulte and Stremersch (2004), whereas the Givon *et al.*

(1995) estimate lies at the lower end of their range. Thus, we find that word processors were subject to external influence to a more usual extent than is found in Givon *et al.* (1995) (although the Van den Bulte and Stremersch (2004) data does not disaggregate their reported figure by annual, quarterly, and monthly frequency of calculation, so the sub-divided ranges may move closer to the Givon *et al.* (1995) figure). Thus, we find a higher effect of external influence on diffusion. The b_1 parameter equals 0.109, describing the internal influence on sales by legal owners. The value is lower than in Givon *et al.* (1995) where it is 0.135. Our estimate for the pirate internal influence parameter b_2 is 0.0888, again lower than in Givon *et al.* (1995) at 0.135 too. Our α parameter is 0.163, representing the proportion of internally influenced adopters who buy the legal software when few past pirate copies have been made. It is higher than in Givon *et al.* (1995) (0.144). The σ^2 variance parameter is 10,800,000, giving an implied standard deviation for legal and pirate acquisitions of 3,300 units per month. The ϵ parameter is 0.0226, indicating that the marginal propensity to pirate rises as the number of pirates rises. This is consistent with a hypothesis that increasing piracy prevalence makes it more viable or acceptable for new adopters to acquire pirate copies. The significance of all parameters is low, except for the error variance.

In column two, we set the ϵ parameter to zero to see how the parameters and fit adjust compared with the model including it. The parameters on a , b_2 , and σ^2 change little, although the significance on b_2 increases whilst remaining low. The b_1 parameter rises to 0.229, with low significance. The α parameter drops to 0.124, and becomes significant at five percent. The Akaike Information Criterion selects the smaller model over the larger model, and the mean squared errors from the two models are almost identical indicating no in-sample predictive benefit from including the ϵ term.

Column three reports parameter estimates for our model applied to the spreadsheet data. The coefficient of external influence a is 0.00155, similar to that for word processors and compared to 0.00069 in Givon *et al.* (1995). The legal owners' influence parameter b_1 is negligible, and far below the external influence parameter b_2 of 0.0509. In Givon *et al.* (1995), the estimated parameters are larger and comparable at 0.0976 for the b_1 parameter and 0.104 for b_2 . Our α parameter is 0.224 compared with 0.121 in Givon *et al.* (1995). The error variance is 3,790,000 implies a standard deviation for monthly legal and pirate acquisitions of 1,900 units per month. The estimate for the ϵ parameter is 0.0684, so that the marginal propensity to pirate rises with the number of pirates. Except for the variance parameter, significance is low.

Column four shows the model when ϵ is excluded. The a , b_1 , and σ^2 parameters are similar to the model with it. The b_2 parameter rises to 0.114, and the α parameter drops to 0.101. Both are now significant at five percent. The Akaike Information Criterion selects our model with variable marginal propensity to pirate, and the model offers a non-trivial reduction in mean squared error.

The performance of our model vis-à-vis the restricted model is mixed. The ϵ parameter estimates are plausibly positive and low in value. With the word processor data, the extra variable offers no improvement in fit and weakens the significance on the α parameter, but not the other parameters. With the spreadsheet data, there are noticeable gains for the fit, but the parameter significance worsens perhaps indicating that qualitative behaviour of the larger model better describes the data but the functional form is not correctly specified. We examine these issues in the next subsections.

A further notable point is the insignificance of the legal internal influence in any specification. The estimates point to the only possibly statistically significant internal influence coming from past acquirers of pirate copies.

Figure 1 shows the predicted sales as generated by the Kalman filter at our estimated parameters. The top panel shows the fitted sales for word processors in our model (column one in table 1, with red dashes) and in the restricted model (column two, with green dots). Our model fits the data better across most of the period except towards its end where its predictions are lower than the restricted model, and far lower than the suddenly hiked sales. The extra parameter available in our model allows for better fitting but comes at a cost in that pirate sales dominate late in the period, if the optimal ϵ parameter is positive. The legal sales are lower as a result. In the bottom panel, we see the fitted sales for spreadsheets for our model (column three), and for the restricted model (column four). The fit is comparable between the two models for most of the period, but at the end of the period our model fits much better the sudden fall in sales. The reason for the relative fit is that the extra flexibility in our model allows for comparable fit quality over most of the curve, and then better fit to the late fall as pirate sales displace legal sales. The different directions of the sales shifts at the end of the period explains why the mean squared error for our model with the word processor data is the same as the restricted model, whereas it is much lower with the spreadsheet data.

Figure 2 shows the same graphs with pirate acquisitions included. The top panel shows forecast piracy for the word processor data. Our model, following the upper line with red dashes, shows pirate acquisitions rising quickly to account for most software acquisition. A similar path is shown for the restricted model, which is the upper line marked by green dots. Piracy incidence in our model is slightly lower than that in the restricted model at the start of the period but exceeds it later by another small amount, with the point of equality occurring quite early on at the end of 1988. Our model includes the extra term accelerating new pirate acquisitions at high rates of past pirate adoptions, so that early piracy tends to be lower and late piracy higher than in the absence of the extra coefficient.

The lower panel shows the corresponding curves for the spreadsheet data. Our model and the restricted model again show similar shapes. The restricted model has a shallower increase, so that our model again forecasts lower incidence of piracy when its prevalence is low, but greater incidence when its prevalence is high. Our model's forecasts exceed those

Figure 1: Predicted and actual sales for word processors (top) and spreadsheets (bottom). Red dashes are for the model with estimated ϵ , green dots are for the model with ϵ set at zero

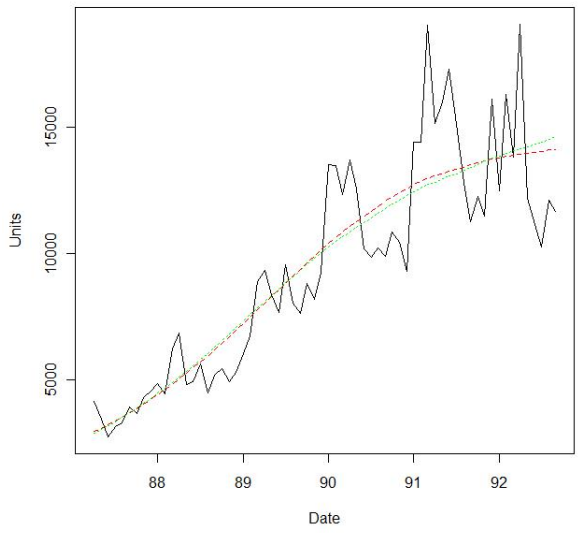
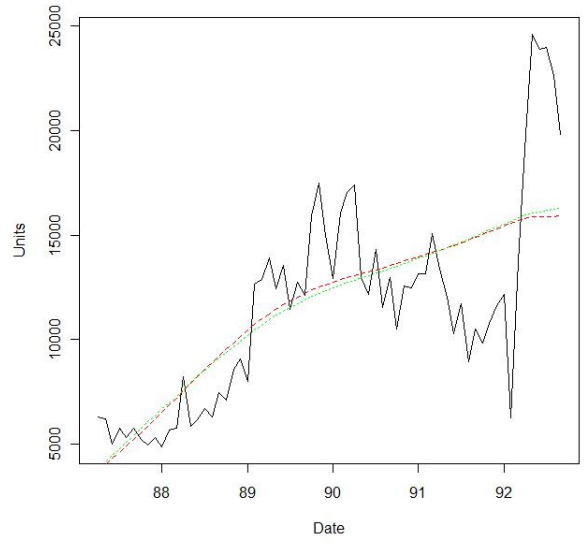
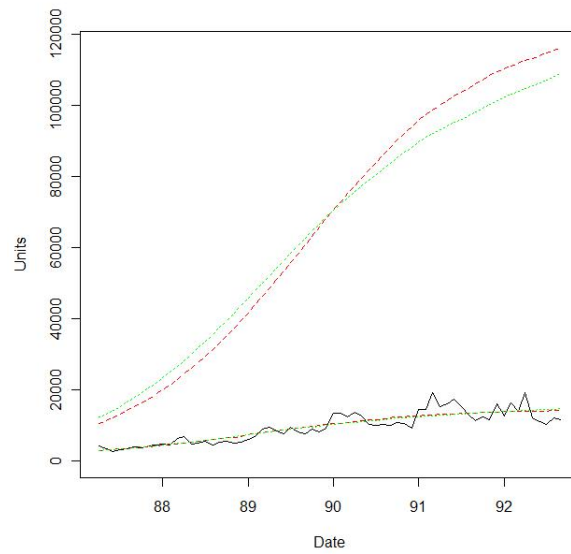
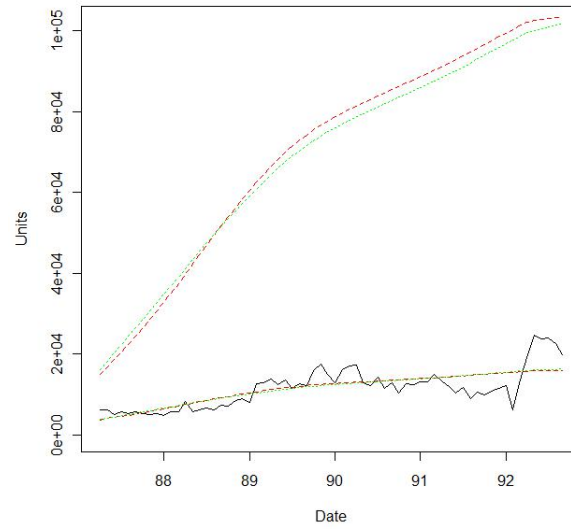


Figure 2: Pirate acquisitions, predicted sales, and actual sales for word processors (top) and spreadsheets (bottom). Red dashes are for the model with estimated ϵ , green dots are for the model with ϵ set at zero



of the restricted model from early 1990, and the gap is moderately large at high piracy prevalence. The larger gap for spreadsheets than word processors is due to the higher estimated ϵ coefficient for the spreadsheet data.

4.2 Out of sample performance

This section compares the predictive performance of our model with that of the restricted model. To do so, we re-estimated our model using data from the first 58 periods, retaining the last ten periods for assessing out of sample fit. For the word processor data, the out of sample period is marked by a large sales jump, while for the spreadsheet data the period saw a possible sales growth slowdown.

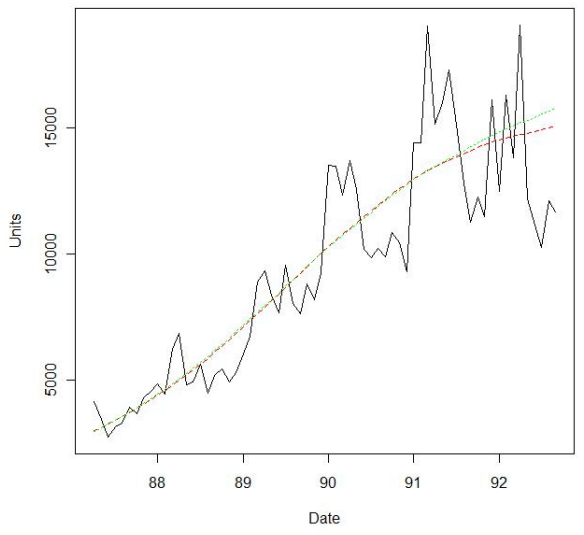
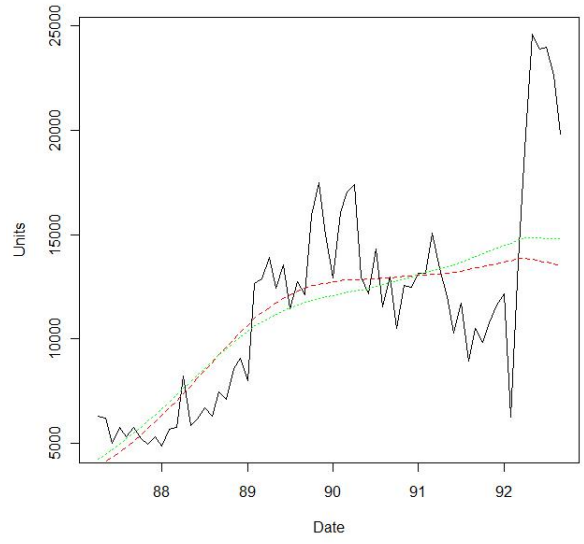
Table 2: Parameter estimates in-sample and MSEs out of sample

	WP	WP	S	S
a	0.00183	0.00242	0.00164	0.00161
	0.00187	0.00179	0.00336	0.00207
b_1	0.00252	0.00000	0.00075	0.00001
	0.239	0.194	0.587	0.305
b_2	0.0594	0.145 ***	0.0556	0.106 *
	0.128	0.044	0.406	0.055
α	0.291	0.105 ***	0.199	0.105 **
	0.449	0.0219	1.17	0.0411
σ^2	0.00413 ***	0.0053 ***	0.00309 ***	0.00314 ***
	0.000263	0.000419	0.000473	0.000477
ϵ	0.0843		0.056	
	0.136		0.462	
MSE, o.o.s. (%)	119	100	83	100

Standard deviations are shown below the coefficients. *** denotes a p-value of less than 0.01, ** of less than 0.05, * of less than 0.1. MSEs out of sample are expressed as percentages of the MSE out of sample for the corresponding restricted model. σ^2 is reported in units of 10^9 .

The results are in table 2. Column one shows the estimates for our model applied to the word processor data. The a parameter of external influence is 0.00183, a little higher than the estimate over the whole range. The b_1 internal legal influence parameter is economically and statistically inconsequential. The b_2 internal pirate influence parameter is 0.0594, slightly below the whole domain estimate. The legal share parameter α is 0.291,

Figure 3: Predicted and actual sales for word processors (top) and spreadsheets (bottom), with ten out-of-sample periods. Red dashes are for the model with estimated ϵ , green dots are for the model with ϵ set at zero



far above the whole period value. The variance estimate σ^2 is 0.00413, under half of the full period estimate that includes the sudden sales growth. The estimate of the marginal propensity to pirate parameter ϵ is 0.0843 compared with the full sample ϵ estimate of 0.0226. The parameter significance is generally low except on σ^2 , and very low on b_1 .

Column two shows similar parameters for the restricted model estimated on the word processor data, without ϵ . The pirate internal influence parameter is fifty percent higher than its full period estimate. The legal share parameter is a little lower. Both these parameters become significant at one percent, unlike the full period estimate or the estimate for the model with variable propensity to pirate.

Our model has a much higher out of sample MSE than the restricted model. Including the variable marginal propensity to pirate worsens fit in this case. The positive ϵ coefficient results in the share of pirate acquisitions rising over time and the share of legal acquisitions falling, which leads to a far better fit than the restricted model over the in sample period (figure 3, top panel). However, the curvature in our predicted sales curve means they lie below the predicted sales of the restricted model after the unprecedented sales jump in the out of sample period.

Column three has the parameter estimates for our model using the spreadsheet data. The external influence, internal legal influence, and legal share parameter are all close to their full period estimates. The internal pirate influence parameter is even smaller than for the full sample, and the error variance parameter is moderately lower. The marginal propensity to pirate parameter is also similar to the value for the whole period, and is positive. Except the variance parameter, significance is low.

The spreadsheet parameter estimates with the ϵ parameter set to zero are in column four. The estimated volatility is twenty percent lower, but otherwise the parameters are similar to their full period estimates in size and significance.

The mean squared error in spreadsheet sales projection for our model is only 83 percent of the MSE for the restricted model. In our model, the positive ϵ parameter means that the share of new piracy rises as the number of past pirates increases, so that legal sales' share declines. Legal sales are lower in our model than in the restricted model during the sample period, tracking the decline in actual sales (figure 3, bottom panel). The ϵ parameter is quite high at 0.056, so that the gap is quite large which accounts for the size of the MSE forecast gain.

The non-additive functional form meant that we could not verify the error correlation assumptions necessary to run Clark-West tests (Clark and West, 2007) of the extra variable offering no predictive gains. When the tests were run under uncertain assumption validity, the null of no gains was rejected at ten percent for spreadsheets (assuming no autocorrelation in the data). Significance fell a little on allowing for first order autocorrelation.

Our model seems to capture behaviour over periods without large sales jumps better than the restricted model. However, the restricted model offers more robust predictions

when faced with shocks biased against the general direction of movement. It is unclear, based on the available data, whether such shocks are inherent to the system and occur frequently. If they are and do, then more extensive misspecifications may be preferable to less misspecified models if the partial misspecification decreases predictive accuracy after the shock.

4.3 Alternative error specifications

Our model in equation 2 describes the errors in the legal sales and pirate acquisition series by a bivariate normal distribution. In order to achieve parameter identification, our base estimations restricted the variance-covariance matrix \mathbf{Q} to be a multiple of the identity matrix. In this section, we compare our estimates under other error specifications.

We consider four \mathbf{Q} specifications, applied to the word processor and spreadsheet series in turn. The first specification puts the variance for the larger pirate series to be twice the variance for the legal series, so $q_{1,1} = \sigma^2$ for some constant σ^2 , $q_{2,2} = 2\sigma^2$, and $q_{1,2} = q_{2,1} = 0$. The second specification allows for positive correlation between the legal sales and pirate acquisitions, setting $q_{1,1} = q_{2,2} = \sigma^2$ and $q_{1,2} = q_{2,1} = \sigma^2/2$. Thirdly, we specify negative correlation between the two series, so $q_{1,1} = q_{2,2} = \sigma^2$ and $q_{1,2} = q_{2,1} = -\sigma^2/2$. In the fourth specification, heteroscedastic errors are allowed so $q_{1,1} = q_{2,2} = N_t\sigma^2$ and $q_{1,2} = q_{2,1} = 0$, recalling that N_t is market capacity at time t , which we express in millions.

Our model was re-estimated with each of these variance specifications. The coefficient estimates are given in Table 3. The word processor results are presented in the first four columns. With doubled pirate variance in column one, the coefficients are similar to the base estimates in value and significance. The marginal propensity to pirate rises a little more quickly. The AIC and MSE are no different. For positive correlation in column two, the estimated effect of legal internal influence is much larger ($b_1 = 0.334$) than in the base model. The marginal propensity to pirate declines as the number of pirate copies rises ($\epsilon = -0.0147$), unlike in the base estimates. The positive correlation provides a route by which pirate acquisitions influence legal sales, and which can therefore account for the diminished importance of the direct influence of piracy on diffusion in this estimation. The AIC is higher for this specification and the mean squared errors are the same.

Column three shows the estimates with negative correlation. Parameter estimates for external influence and variance within series are similar to the base estimate. However, the estimates of the legal internal influence and the legal share parameter are much larger, at 0.381 and 0.226 respectively. Thus, for small times t the proportion of past legal adopters who induce new legal adoptions is $0.381 \times 0.226/N_t = 0.086/N_t$. The estimated effect of pirate internal influence is negligible. Taking also into account the number of new adoptions induced by past legal adoptions and the size of the external influence parameter, the early legal diffusion looks like a standard univariate Bass diffusion. The marginal propensity to

Table 3: Parameter estimates with alternative error specifications

	WP	WP	WP	WP	S	S	S	S
	Err. $\times 2$	Pos. cor.	Neg. cor.	Het. err.	Err. $\times 2$	Pos. cor.	Neg. cor.	Het. err.
a	0.00139	0.00163	0.00123	0.00217	0.00163	0.00161	0.00157	0.00169
	0.0046	0.00579	0.00701	0.00163	0.00493	0.00443	0.00407	0.00183
b_1	0.0911	0.334	0.381	0.00007	0.000104	0.000989	0.000366	0.00232
	0.786	1.59	1.49	0.293	1.36	0.719	0.495	0.367
b_2	0.0864	0.101	0.000769	0.0518	0.103	0.0597	0.044	0.0524
	0.386	0.378	0.0999	0.176	0.831	0.458	0.318	0.28
α	0.175	0.103	0.226	0.317	0.109	0.188	0.259	0.207
	0.483	0.178	0.472	0.74	0.724	1.18	1.52	0.896
σ^2	0.0108 ***	0.0108 ***	0.0103 ***	0.00242 ***	0.004 ***	0.00385 ***	0.00375 ***	0.000936 ***
	0.00198	0.00209	0.00206	0.000491	0.000813	0.000754	0.000716	0.000213
ϵ	0.0283	-0.0147	0.0454	0.0924	0.00776	0.0535	0.0826	0.0624
	0.202	0.116	0.216	0.224	0.454	0.477	0.494	0.354
AIC	1287.0	1287.5	1286.4	1261.1	1220.4	1217.9	1216.1	1197.3
MSE (%)	100	100	98	104	105	101	99	101

Standard deviations are shown below the coefficients. *** denotes a p-value of less than 0.01, ** of less than 0.05, * of less than 0.1. MSEs are expressed as percentages of the MSE for the base estimations (columns one and three of table 1). σ^2 is reported in units of 10^9 .

pirate rises much more quickly in this specification than in the base model (ϵ equals 0.0454 compared with 0.0226). The AIC and MSE are lower. In this specification, piracy acts primarily to displace legal diffusion without promoting it.

Column four shows the parameters for the model with heterogeneous errors. Compared with the base model parameters, the external diffusion parameter is moderately larger, as is the estimate of the series error variance. The legal internal influence coefficient is negligible, and the pirate internal influence coefficient is much the same. The marginal propensity to pirate rises much more quickly in this specification ($\epsilon = 0.0924$). The AIC is lower, but the MSE increases by a larger percentage.

Parameter estimates for the spreadsheet data are shown in columns five to eight. The parameters for the double pirate error specification are shown in column five. The external influence and legal internal influence parameters are much the same. The variance coefficient is similar at 0.004. The pirate internal influence parameter is almost doubled at 0.103, while the legal share parameter is halved at 0.109. The rise in marginal propensity to pirate is positive and low, with ϵ at 0.00776 compared with 0.0684 in the base specification. Significance is generally low, the AIC is higher and the MSE increases by five percent.

In column six, the results are shown for the positive correlation error specification. The parameter estimates are not too dissimilar from the base specification. The AIC and MSE are higher. The parameters for the negative correlation error specification are shown in column seven. The parameters are similar to the base specification, with the marginal propensity to pirate parameter a bit larger. The AIC and MSE are both a little lower. The parameters for the heterogeneous error specification are in column eight. They are again similar to the base specification. The AIC improves by two percent and the MSE worsens by one percent.

In summary, the alternative error specifications broadly support an increasing marginal propensity to pirate. For both the word processor and spreadsheet data, the negative correlation specifications reduce the AIC and the MSE. For the word processor data, the resulting model shows piracy growing rapidly and making little contribution to legal diffusion. Parameter significance is generally low.

4.4 An alternative model

4.4.1 The alternative specification

We noted earlier that our model gives plausible values for the ϵ parameter and some improvements in fit, but with low parameter significance. In this section, an alternative specification is examined that exhibits the same broad type of qualitative behaviour. The aim is to see if better fitting parameters can be produced, or at least verification of the qualitative outcomes.

The alternative specification makes the legal share α decline as the number of cumulative pirate adopters rises. The share of new internally influenced adopters who buy is revised to $\alpha/(1 + \max(Y_t, 1)^\epsilon)$, where Y_t is the number of pirates. The remaining share, $1 - \alpha/(1 + \max(Y_t, 1)^\epsilon)$, acquires a pirate copy. We omit the piracy multiplier $\max(Y_t, 1)^\epsilon$ described in the model in section 2. In the new model, piracy has no accelerating effect unlike the earlier model, and instead only displaces legal sales.

Algebraically, the model is

$$\begin{aligned} dX_t &= \left(\left[a + \frac{\alpha}{1 + \max(Y_t, 1)^\epsilon} \frac{b_1 X_t + b_2 Y_t}{N_t} \right] (N_t - X_t - Y_t) \right) dt + dw_1 \\ dY_t &= \left(\left(1 - \frac{\alpha}{1 + \max(Y_t, 1)^\epsilon} \right) \frac{b_1 X_t + b_2 Y_t}{N_t} \right) (N_t - X_t - Y_t) dt + dw_2 \end{aligned} \quad (31)$$

where $d\mathbf{w} = (dw_1, dw_2) \sim N(0, dt^2 \mathbf{Q})$, and $\mathbf{Q} = \sigma^2 \mathbf{I}(2)$.

We repeat the state space representation in section 3.2.2. Now the vector \mathbf{f} is given by

$$f_1 = \left[a + \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon} \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \quad (32)$$

$$f_2 = \left[\left(1 - \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon} \right) \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \quad (33)$$

and the components of the matrix \mathbf{F} are given by the following expressions:

$$\begin{aligned}
\mathbf{F}_{1,1} &= \left[\frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon} \frac{b_1}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \\
&\quad - \left[a + \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon} \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right]
\end{aligned} \tag{34}$$

$$\begin{aligned}
\mathbf{F}_{1,2} &= \left[-\epsilon \max(Y_{t-1}, 1)^{\epsilon-1} \frac{\alpha}{(1 + \max(Y_{t-1}, 1)^\epsilon)^2} \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right. \\
&\quad \left. + \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon} \frac{b_2}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \\
&\quad - \left[a + \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon} \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right]
\end{aligned} \tag{35}$$

$$\begin{aligned}
\mathbf{F}_{2,1} &= \left[\left(1 - \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon}\right) \frac{b_1}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \\
&\quad - \left[\left(1 - \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon}\right) \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right]
\end{aligned} \tag{36}$$

$$\begin{aligned}
\mathbf{F}_{2,2} &= \left[\epsilon \max(Y_{t-1}, 1)^{\epsilon-1} \frac{\alpha}{(1 + \max(Y_{t-1}, 1)^\epsilon)^2} \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right. \\
&\quad \left. + \left(1 - \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon}\right) \frac{b_2}{N_t} \right] (N_t - X_{t-1} - Y_{t-1}) \\
&\quad - \left[\left(1 - \frac{\alpha}{1 + \max(Y_{t-1}, 1)^\epsilon}\right) \frac{b_1 X_{t-1} + b_2 Y_{t-1}}{N_t} \right]
\end{aligned} \tag{37}$$

4.4.2 Results for the alternative specification

Table 4: Parameter estimates for the alternative model form

	WP	WP	S	S
a	0.0016	0.00167	0.00115	0.00135
	0.0039	0.00391	0.00525	0.00271
b_1	0.167	0.283	0.000853	0.00000
	1.27	1.25	0.932	0.551
b_2	0.11	0.0857	0.112	0.113
	0.237	0.229	0.18	0.0854
α	0.246	0.245 ***	0.33	0.208 **
	0.628	0.0947	2.43	0.0903
σ^2	0.0108 ***	0.0107 ***	0.00403 ***	0.00407 ***
	0.00209	0.00203	0.000827	0.000752
ϵ	0.000467		0.0503	
	0.326		0.664	
AIC	1287.0	1285.0	1220.8	1219.4
MSE (%)	100	100	99	100

Standard deviations are shown below the coefficients. *** denotes an p-value of less than 0.01, ** of less than 0.05, * of less than 0.1. MSEs are expressed as percentages of the MSE for the corresponding restricted model. σ^2 is reported in units of 10^9 .

Table 4 shows the results of estimation for the alternative model. In column one, we see the model fitted to the word processor data. The external influence parameter is 0.0016, comparable with our estimate for our main model. The legal internal influence parameter is 0.167, higher than the main model estimate but comparable with Givon *et al.* (1995). The same is true for the pirate internal influence parameter. The α parameter is larger than for the main model, but the two are not directly comparable. In our model and for low values of the ϵ parameter, the α parameter is twice the legal share. Error variance estimates are unchanged. The ϵ parameter is positive, indicating the pirate share of internally influenced adoption rises as the number of pirate copies rises. Parameter significance is low except for variance. Column two fixes the pirate share, and produces almost the same estimates as column two of table 1 whose specification is identical. Differences arise from slight variations in numerical convergence. Including the variable marginal propensity to pirate raises the AIC and does not change the MSE.

Column three reports the fitted parameters for the spreadsheet data. The external

influence parameter is comparable with that in the main model, as is the negligible legal internal influence coefficient and the error variance estimate. The pirate internal influence parameter is 0.112, which is twice as high as for the main model and comparable with Givon *et al.* (1995). The legal share is 0.165 after adjustment for comparison with the earlier work, making it lower than in the main model but a little higher than in Givon *et al.* (1995). The ϵ parameter is moderately positive, indicating rising marginal propensity to pirate. Its significance, as for the other parameters except error variance, is low. As shown in column four, the model with ϵ set to zero has much the same parameters as the constrained main model. Its AIC is slightly lower, but has higher MSE than the unconstrained model in column three.

In summary, the alternative model also identifies a rising marginal propensity to pirate in the word processor and spreadsheet data. The model performance is not quite as good as for the main model. Perhaps the acceleration of pirate sales due to piracy, which is present in the main model but missing from this specification, captures an aspect of the data generating process.

5 Conclusion

This paper has examined diffusion of computer software in the presence of piracy. It has generally found that the marginal propensity to pirate rises with the number of past pirate copies. It has found that piracy is responsible for most of the internally influenced diffusion in the spreadsheet market in the period under examination. In the word processor market, an error specification with negative correlation outperformed other specifications and indicated that only legal sales were responsible for internally influenced diffusion.

A number of avenues for future work are suggested. One of them follows from noting that a rising propensity to pirate alters the timing of welfare and profit emergence. Further work could examine their dynamics and strategic behaviour undertaken by legal sellers in order to manage piracy.

Including the marginal propensity to pirate in the pirate diffusion model offers gains in fit and assumption plausibility. However, they were not entirely functionally convincing, with low parameter significance possibly indicating lack of parsimony. Better specifications of the deterministic components of the model could be sought.

Our model's predictive performance was mixed. It underperformed the restricted model after a large shock contrary to the general sales curvature. Further work could clarify whether these shocks form error corrections to drifts away from the restricted model, or are not systemically related to the models here. In the latter case, analysing the frequency of shocks and their direction would help to clarify the probability and severity of predictive underperformance.

More general specification of stochastic components has allowed us to strengthen Givon

et al. (1995)'s and Haruvy *et al.* (2004)'s findings on spreadsheets and contradict them on word processors. For the latter point, it is conceivable that the difference between the results is that our allowance of negative error correlation strips out a source of interaction which is forced to be included in the deterministic components of their model. Further work could further distinguish between deterministic and stochastic interactions. It could also allow for serial correlation, for example in a revised state space formulation.

Our work finds contrasting effects of past piracy on the word processor and spreadsheet market. Qualitative studies could examine the reasons for the difference. Conceivably it arises because of the presence of a dominant but declining word processor product (Word Perfect) over the period without an equivalent in the spreadsheet market, or due to different corporate strategies by market leaders.

References

- Bass, F. (1969) A New Product Growth Model for Consumer Durables, *Management Science*, **15**, 215–27.
- Basu, A., Basu, A. and Batra, R. (1995) Modeling the Response Pattern to Direct Marketing Campaigns, *Journal of Marketing Research*, **32**, 204–12.
- Clark, T. and West, K. (2007) Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics*, **138**, 291–311.
- Givon, M., Mahajan, V. and Muller, E. (1995) Software Piracy: Estimation of Lost Sales and the Impact on Software Diffusion, *Journal of Marketing*, **59**, 29–37.
- Givon, M., Mahajan, V. and Muller, E. (1997) Assessing the Relationship between the User-Based Market Share and Unit Sales-Based Market Share for Pirated Software Brands in Competitive Markets, *Technological Forecasting and Social Change*, **55**, 131–44.
- Hamilton, J. (1994) *Time Series Analysis*, Princeton University Press.
- Haruvy, E., Mahajan, V. and Prasad, A. (2004) The Effect of Piracy on the Market Penetration of Subscription Software, *Journal of Business*, **77**, S81–107.
- Jain, D. and Rao, R. (1990) Effect of Price on the Demand for Durables: Modeling, Estimation, and Findings, *Journal of Business & Economic Statistics*, **8**, 163–70.
- Liu, Y., Cheng, H., Tang, Q. and Eryarsoy, E. (2011) Optimal software pricing in the presence of piracy and word-of-mouth effect, *Decision Support Systems*, **51**, 99–107.

- Prasad, A. and Mahajan, V. (2003) How many pirates should a software firm tolerate? An analysis of piracy protection on the diffusion of software, *International Journal of Research in Marketing*, **20**, 337–53.
- R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>.
- Schmittlein, D. and Mahajan, V. (1982) Maximum Likelihood Estimation for an Innovation Diffusion Model of New Product Acceptance, *Marketing Science*, **1**, 57–78.
- Srinivasan, V. and Mason, C. (1986) Technical Note - Nonlinear Least Squares Estimation of New Product Diffusion Models, *Marketing Science*, **5**, 169–78.
- Van den Bulte, C. and Stremersch, S. (2004) Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test, *Marketing Science*, **23**, 530–44.
- Xie, J., Song, X., Sirbu, M. and Wang, Q. (1997) Kalman Filter Estimation of New Product Diffusion Models, *Journal of Marketing Research*, **34**, 378–93.