Comparing parameter estimation methods for service-based diffusion models

Daniel Arthur
University of Surrey, CCSR / Dept of Economics
Guildford, Surrey, UK, GU2 7XH
d.arthur@surrey.ac.uk

Abstract

This paper compares methods for parameter estimation of diffusion models when seeking to extend these to service industry contexts from the traditional product focus. In the marketing science and economics literature, parameter estimation is dominated by econometric methods. This presents certain limitations as well as advantages compared to calibration in system dynamics modelling, which emphasises estimation of parameters by direct observation. But this poses a problem for industry or market-level diffusion models where deriving aggregate parameters observationally is impractical, especially for launches of new products or services which lack direct market knowledge. One solution is to use judgemental bootstrapping, entailing the estimation of parameters from an expert's forecast time series. Models parameterised this way can then be used as a basis for simulated structural experiments of proposed market architectures. Some interim results from three service industry case studies are presented.

Key words: aggregate diffusion models, estimation/calibration, models to predict / learn

Introduction

Two significant publications in the forecasting and marketing science fields have identified potential roles for system dynamics models that overlap with its more familiar modes of application. In the first publication that reviews demand forecasting in the telecommunication industry, Fildes (2002) identified the potential of simulation methods to help to structure the complexity of drivers of the adoption process when related to modelling new product launch situations. Standard forecasting methods typically employ price elasticity and econometric estimation for established markets, particularly for cross-sectional rather than longitudinal models. In contrast, Fildes identifies system dynamics (and indeed it was the only simulation method cited) for its potential to explore new markets. Beyond the normal use of system dynamics for policy modelling, he identifies an inherent applicability to forecasting applications, even though from its outset the system dynamics field has tended to eschew numerical forecasting in favour of behavioural prediction. He cites Lyneis (2000) who endorses the power of system dynamics to provide better short and mid-term forecasts than regression or trend extrapolation models.[1]

The second publication is also by Fildes (2003) and reviews a volume presenting the state of the art in new product diffusion models edited by well-established authors in quantitative marketing research, Mahajan et al (2000), updating a similar work by Mahajan and Wind (1986). Fildes notes two serious weaknesses in the more recent volume, namely the nearly complete neglect of the topic of validation and the other omission which is "less important but, perhaps, more fundamental": the failure to recognise that simulation models, usually using system dynamics, offer an alternative to Bass-type diffusion models. He claims the first omission is important since there is limited evidence that

[1] An aversion to forecasting in the system dynamics field was present at its outset (Forrester, 1961, Appendix K). However, the usual definition of forecasting, ie "predicting system condition at some future time" is contrasted with utility of "predicting the direction and degree of influence on system behaviour as a result of changing policy" (Forrester, 1996, 2003). Lyneis (2000) identifies that forecasting is both ubiquitous and necessary in business for estimating the approximate timing and impact of key business variables.

diffusion models 'work' or have been widely used in commercial settings. However, Mahajan et al (2000) do recognise that diffusion models can be used for analysing strategic decisions for product life-cycle dynamics and not just for pre-launch forecasting.

Given that the authors cited above represent significant voices in the forecasting and marketing science fields, this seems to offer system dynamics significant potential for a greater presence and impact. This is not to suggest that system dynamics has not been applied in those domains before; rather, there is a lack of awareness of what it has to offer, both in the area of strategy and policy for life-cycle analysis and in more detailed quantitative decision support roles. The latter should not be unfruitful areas of application since Lyneis (2000), Winch (1993), Homer (1996), Graham and Ariza (2001), Graham (2002), Graham et al (2002) all make a strong case for the value of such models that are comprehensively numerically parameterised. It is often detailed analysis that is influential in supporting major decisions. However, in doing so, the case will need to be made that system dynamics can effectively address predictive validation from a forecasting perspective and parameter estimation from an econometric perspective. Even though it will be highly beneficial to recognise the broader purview that system dynamics possesses in respect of model validation, there is no reason why these aspects cannot be encompassed.

For example, as Fildes (2002) notes, if system dynamics models are to be used to support important policy decisions, for instance, to determine the balance between spend on marketing spend as against capacity development, the neglect of predictive validation is "disingenuous", as the timescales in diffusion processes are critical both for understanding and decision making.

Thus there seems to be ample opportunity for system dynamics to be applied with some reward in new product diffusion dynamics, especially in successfully communicating its power in these different fields. A particular opportunity, identified in this paper, is to examine the potential of applying diffusion models to service markets, where industry structure on the supply side is much more uncertain and complex but still requires significant investment or strategic co-ordination. An example is in networked industries where there can be long value chains or important roles played by complementors, as identified in the case of investment decisions in vehicle telematics by General Motors (Barabba et al, 2002). Another opportunity is the development of the idea of a continuum of models from forecasting to policy analysis, along which there will be differing criteria for model credibility and demonstration of fitness-for-purpose.

A programme of research work is underway in which several of these issues are being addressed. However, as one starting point, a comparison of calibration of well-accepted models as against econometric estimation would be beneficial.

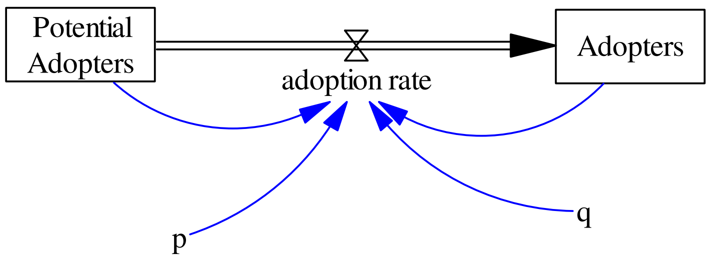
Diffusion models

There are two main traditions in for new product modelling: diffusion models which are typically used at an aggregate market level and conjoint-based or discrete choice models which are concerned with consumer choice at an individual level. The tradeoffs between differing levels of aggregation are discussed below. For the purposes of this paper, three widely accepted models of innovation diffusion demand are represented to represent the adoption of new products or technologies: the Bass, KS and EMM models. This follows the results of Zettelmeyer and Stoneman (1993/2002), who found these models to be the be the best performing in an econometrically-oriented comparison in terms of fit against historic data.

The standard Bass (1969) model defines the adoption rate by summing the diffusion from innovators and imitators. Bass added the external effect of innovators from the work of Fourt and Woodlock (1960) to the representation of epidemic growth of imitators proposed by (Mansfield, 1961).

$$\text{sales or adoption rate} = \frac{dN}{dt} = p(m-N) + q\frac{N}{m}(m-N)$$
$$= \left[p + q\frac{N}{m}\right](m-N) = pm + (q-p)N - \frac{q}{m}N^2 \tag{1}$$

where N is the current number of adopters, m is the market potential (or addressable market size), and p and q are the coefficients of innovation and imitation respectively and are usually reported in the units of year^{-1}.



To allow for a time-dependent variation of the addressable market size, m can be replaced by m_t, where

$$m_t = m_0 + \frac{\gamma}{100} m_0 t \tag{2}$$

where γ is a growth parameter reflecting the percentage growth of the initial population m_0 per time period. Because the market size of some of the data tested was not constant, this modification was enabled for most subsequent models.

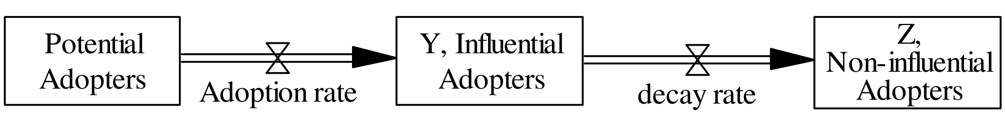
The EMM model (Easingwood, Mahajan and Muller, 1983) is a modification of the Bass model and is also termed the Non-Uniform Influence (NUI) model that allows the strength of the imitation effect to vary with the proportion of adopters via the parameter σ, and thereby exert a time-dependent effect:

$$q^* = q\left(\frac{N}{m}\right)^{\delta} \quad \text{hence} \quad \frac{dN}{dt} = \left(p + q\left(\frac{N}{m}\right)^{\sigma}(m-N)\right) \quad \text{where} \quad \sigma = \delta + 1 \tag{3}$$

The KS model (Karshenas and Stoneman, 1992) was developed as an attempt to integrate the behavioural basis of models in the Economics literature with the focus on forecasting performance and the epidemic tradition evident in the Marketing literature. It represents learning effects within the adopter population (and thus indirectly allows for some time-dependent behaviour) by splitting them into two groups, namely active and inactive adopters, ie those active in the imitation process and those who are not engaged in communicating to potential adopters. In epidemic terms, these groups are termed the 'infectives' and the 'recovered' populations. The rate of gain of the 'recovered' population is via the parameter α

$$\frac{dY}{dt} = \left[p + q\left(\frac{Y}{m_t}\right)\right].(m_t - N) - \alpha Y \tag{4}$$

$$\frac{dZ}{dt} = \alpha Y \tag{5}$$



The KS model is directly equivalent to the Susceptible-Infected-Recovered (SIR) model as widely used in the field of epidemiology.

Estimation approaches

The models above were all parameterised two approaches: econometric estimation and system dynamics calibration. The former implies fitting a pre-specified model (often a single equation, although systems of equations can now be handled) to historic time series using analytical or numerical techniques, invariably involving matrix theory and matrix inversions. Calibration is a somewhat broader term and implies that data may be derived from a variety of sources and brought to bear, as well as the possibility that structural changes to the equations may be made. However, it still implies fitting against historic data.

Econometric estimation from the adoption rate form of the equation with an additive error term; given the non-linearity of the models, Non-Linear Least Squares (NLS) estimation was used with the Eviews statistical/econometrics package. NLS was proposed by Srinivasan and Mason (1986) as a way of overcoming several limitations of Ordinary Least Squares estimation (Mahajan et al, 1986). Maximum Likelihood Estimation has also been proposed (Schmittlein and Mahajan, 1982). The limitations included a time interval bias from expressing a continuous differential equation in a discrete form (in equation (6), the approximation for dN/dt is evaluated at t whereas the right hand side is evaluated at t-1) and wrongly-signed values for p, q and m often arising from the quadratic equation solution. Also, standard errors are not available for deriving t-statistics as is common in econometric estimation[2]. Other techniques including simulataneous equation and stochastic models have been proposed in a more recent review by Putsis and Srinivasan (2000), although these authors admit that estimation errors may be swamped by specification errors. This concern is reflected by the system dynamics method's primary focus on structure.

The equations must be estimated in their discrete analogue form unless they can be expressed in a closed-form analytical solution as a function of time. This is only possible in the case of the Bass model. More complex models like the KS and EMM models cannot be expressed analytically. The discrete forms of the estimation equation are:

Bass:
$$\left.\frac{\Delta N}{\Delta t}\right|_t = \beta\left[p + q\frac{N_{t-1}}{m}\right](m - N_{t-1}) + \varepsilon_t \tag{6}$$

EMM:
$$\left.\frac{\Delta N}{\Delta t}\right|_t = \beta\left[p + q\left(\frac{N_{t-1}}{m}\right)^{\sigma}\right](m - N_{t-1}) + \varepsilon_t \tag{7}$$

KS:
$$\left.\frac{\Delta N}{\Delta t}\right|_t = \beta\left[p + q\frac{N_{t-1}}{m}\right](m - N_{t-1}) - \alpha N_{t-1} + \varepsilon_t \tag{8}$$

In the above equations, β represents a combination of exogenous economic variables, such as price, disposable income or interest rates that represent the the economic environment and act as modifiers to the take-up parameters p and q.

Product diffusion data is more often published as monthly or annual sales data. Sales data is often assumed to be synonymous with adoption data, but in reality sales data is often contaminated by repeat or replacement sales and often estimation studies are confined to the period of early sales only. In contrast, penetration data are desirable as they are not contaminated this way. In this case, since the discrete analogue estimation equations above (equations (6) to (8) require adoptions data (the flow rate), these must be derived from penetration data which can emphasise proportional measurement errors by differencing of large values over short time scales (eg monthly data).

[2] The t-statistic is not a reliable hypothesis test for the significance of model variables in the presence of measurement error (Mass and Senge, 1980; Sterman, 1984). A way of dealing with measurement error is Full Information Maximum Likelihood with Optimal Filtering (Peterson, 1980) or, more simply, using the behaviour anomaly test to see whether unreasonable model behaviour results from assuming a zero value for the parameter (Sterman, 1984).

System dynamics calibration allows more flexibility than the econometric procedures described above. First it allows the solution of the differential equations as a core part of its method and second, that means that any variable can be compared to data. Oliva (2003) has reviewed the techniques available to calibrate whole models to data and categorised them as Model Reference Optimisation (Lyneis and Pugh, 1996) and Full Information Maximum Likelihood using Optimal Filtering (FIMLOF, Peterson, 1980). MRO typically uses some form of minimisation of weighted least squares of the residuals between the data and the simulation values.

Generating a fit between data and a model is rather easy in the case of a monotonically rising sigmoid curve that is so typical of market penetration data. Oliva (2003) warns that automated calibration procedures to achieve such fits can give practitioners false confidence in their models. Indeed Sterman (1984) describes it as a weak test, although it is a necessary one, and is but one of the many other tests required to build confidence, and he presents as a modified list of tests from Forrester and Senge (1980).

Aggregate vs individual-level models

Diffusion models have traditionally been put forward to represent aggregate market response of the adoption of new products or technologies. There are several ways of dealing with the limitations inherent in assuming aggregate behaviour. Four different positions have been proposed in a typology of dynamic sales models by Roberts and Lattin (2000) are shown in Figure 1.

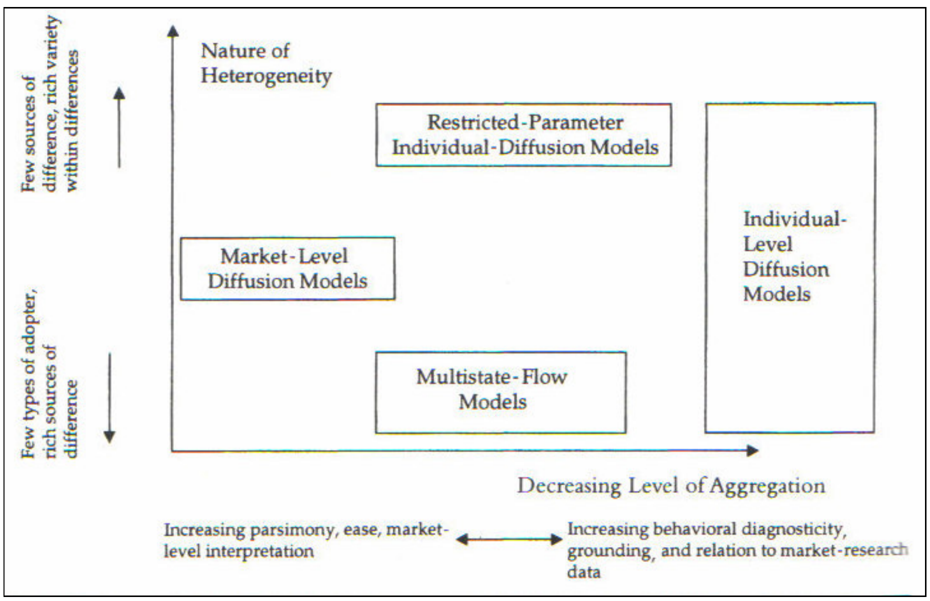


Figure 1: A typology of aggregation in models for sales of new products

At the two extremes are aggregated market-level models and individual-level models. Aggregate-level models have as their objective the understanding of overall market development and its response to managerial and environmental variables. Individual-level models start from classical utility and attitude models from economics and psychology and attempt to represent changes in expected utility over time. Discrete choice theory then provides a method to transform these utilities to probabilities of purchase and thus expected market shares. Multi-state flow models and restricted-parameter individual models fall in the continuum between the two extremes of aggregation. Multi-state models segment the market into a number of behavioural states and then observe the flows between them. System dynamics models fall more naturally into this category. Restricted parameter individual models retain the richness and theoretical rigour of individual-level models but consumers are allowed to be heterogeneous only with respect to a small number of parameters.

Parameter Estimation of Aggregate Diffusion Models for new products or services

The system dynamics approach has emphasised the importance of estimating parameters from 'below the level of aggregation' (Graham, 1980), that is, from direct knowledge of parameters at a disaggregate level, such as interviews, direct observation, literature etc, rather than from historic time series. All data sources are deemed relevant, whether from the mental, written or numerical databases (Forrester, 1980).

However, when one cannot observe the parameters typically used in Bass-type models, or variants, the question remains how these models should be parameterised. Several options are presented in Table 1.

Table 1: Approaches for parameterising models for new markets

	Method	Comments
Judgment or observation based		
1	Informed guesstimate – update as necessary	Useful as first-cut scenario testing
2	Expert judgement or direct observation of parameter values ('below level of aggregation' – Graham, 1980) At the extreme of disaggregation, using consumer choice theory such as conjoint analysis	Require currency of observations or expert judgement for likely situation at operational level / decision points The reliability of upscaling from a small sample of customer survey results to the market level is a query
Combined judgment and estimation		
3	By estimation from expert judgement forecasts (judgmental bootstrapping from time series forecast data)	Expert judgement forecasts used as a baseline for testing alternative scenarios, or to compare against published parameter values. † ‡
Formal estimation approaches		
4	By formal estimation from historic data for the application in question; possibly updating in the light of new data	Past data may not be relevant for future – only relevant for established markets / situations where structural stability prevails. † Bayesian or full-information methods such as Kalman Filtering to update with new data as it becomes available – relevant for ongoing planning eg for production management (peak sales/time to peak).
5	By using analogues (in the absence of historic data) – from published parameter values or by estimation from raw time series	Validity or relevance of analogue? Choose analogues at level of structure (ie expected market behaviour, not similarity of product / service) † ‡

† When estimating from existing data, for all logistic / S-growth penetration curves, need the point of inflection in the data to ensure the stability of estimates (Mahajan, Muller and Bass 1990)
‡ Ideally, time series data will include the point of inflection

Estimating parameters (p and q) from historic time-series of analogues products (or more simply, consulting published data where this has already been done) entails the question of how one chooses relevant analogues. Similarity in expected market behaviour is a better basis for selecting an analogue than product similarities (Lilien et al, 2000) and Thomas (1985) suggests five criteria to determine suitability: environmental context, market structure, buyer behaviour, marketing-mix strategies and characteristics of the innovation.

Market research can be done at the level of individual behaviour and aggregated across different groups which are deemed to be internally homogenous. However the reliability of such extrapolation to the market level is questionable with small samples. Whilst some variables are estimated at the individual level (eg strength of prior beliefs, utility of the innovation), others are calculated at the aggregate or segment level (eg factor scores, importance weights) (Roberts and Lattin, 2000).

Regarding the approach 3 (estimation of parameters from expert judgement forecasts) in Table 1, this is clearly not an independently derived forecast to serve as a comparison against the expert judgement forecasts themselves, although the parameters derived from this process can be compared against published data to check they are within feasible ranges. This is what would be required in any 'triangulation' exercise, commonly used in industry for market forecasting, which collects several independently sourced forecasts from which an average will probably be taken, and any outlier forecasts may be discounted.

Estimating parameters from expert judgement forecasts is called judgemental bootstrapping. This is a procedure normally used on cross-sectional data to infer experts' rules by induction to a mathematical model so as to make expert forecasting processes more reliable. However, it is less commonly used on time series (Armstrong, 2001, p184)

Diffusion of services compared with products

The Bass-type diffusion models have almost entirely been applied to consumer products, whether high or low value. The question is, how applicable are diffusion models to services or, at least, to services based on an installed (diffused) product base? Service industries have several different characteristics, some of which are:

- services diffuse in a layered adoption process in social context (ie a service often diffuses on top of an existing product or infrastructure)
- complex and qualitative supply side factors are very important, such as multiple complementary services, customer service quality etc
- continuous technical development takes place, more than in discrete products or their replacements

The models tested below represent an effort to calibrate various diffusion models in technology service markets. However, a fuller treatment of the supply-side structures and likely network effects is outside the scope of the current paper, although a diffusion model as a core component of larger models has been demonstrated (see eg the regulatory/competition model, Graham and Godfrey, 2005).

Data used for calibration / estimation tests

To assess the applicability of diffusion models to services, analogues which possessed sufficient data histories were sought. In the context of exploring market situations where complementary services support an installed product base in the telecoms industry, wireless data services in the mobile telecoms were thought to be relevant as the environmental context and market structures are similar. The three analogues chosen are mobile phone diffusion in the UK (1984 – 2004), internet broadband diffusion in the UK (2001 – 2005) and the growth of the subscriber base for the "i-mode" mobile phone internet access service in Japan (1997 – 2005).

However, it is noted that some services are actually service *enablers* rather than services per se, where the service is based on embedded hardware in a product. Consequently, the development of broadband internet access, for which some monthly penetration and price data for the UK were available, was considered relevant. The development of the i-mode by operator NTT Docomo was thought to be a relevant analogue in terms of the successful management of the complexity of the supply side. The business strategy enabling i-mode's success is claimed to be the symbiosis of the roles of hardware, infrastructure and content providers and the steps taken to ensure that all parties develop and benefit mutually from organic growth (Tee, 2005; NTT Docomo, 2006).

However a consistent feature among the analogues is the complementarity – the symbiosis between hardware and the complementary content and services [internet was well developed content by the time broadband became popular].

The enable a comparison against the econometrically oriented study of Zettelmeyer and Stoneman (1993/2002), which considers the diffusion of camcorders, CD players and cars, the economic variables of price, personal disposable income, and interest rates were included as economic factors. In equations (6) to (8) above, this was done by multiplying the p and q factors in the diffusion models above by a multiplier β where $\beta = P_{t-i}^{a_1} YD_{t-j}^{a_2} R_{t-k}^{a_3}$ where P is the price of the product, YD is personal disposable income and R the interest rate. A priori expectations from economic theory are that $a_1 < 0$, $a_2 > 0$, $a_3 < 0$, and $i, j,$ and k stand for lags that can be chosen individually for each data set. The prices and economic data were normalised to unity for the beginning observation. The sources used for compiling the time series data are shown in Table 2.

Table 2: Data sources for calibration experiments

1. UK Broadband adoption (2000 – 2005, monthly data)
    a. UK National Office of Statistics: Index of internet connectivity [for penetration data] (ONS, 2006a)
    b. Ofcom: The Communications Market 2004 [for prices, based on April 2004 values] (Ofcom, 2004)
2. UK mobile telephone adoption (1984 – 2004, annual data)
    a. OECD Telecommunications Database 2003 (OECD, 2003) [for revenue and penetration data]
    b. OECD Communications Outlook 2005 (OECD, 2005) [for revenue and penetration data]
    c. Ofcom: The Communications Market 2004 (Ofcom, 2004) [for revenue data]
    d. Euromonitor International Global Market Information database (Euromonitor.2006) [for additional penetration data].
    e. Office for National Statistics (ONS, 2006b) [for Retail Price Index data]
3. i-mode adoption (1997 - 2005, mainly monthly data), NTT DoCoMo website, Japan (NTT DoCoMo, 2006).
4. Office for National Statistics for personal disposable income (YD) and Interest rate (R) data[3] (ONS, 2006c)

Penetration data are generally preferable as this avoids contamination with multiple or repeat sales. However, for estimation purposes using Eviews, the penetration data must be transformed by first-differencing into discrete sales data, which inevitably increases the proportional measurement errors in the derived sales.

Results and discussion

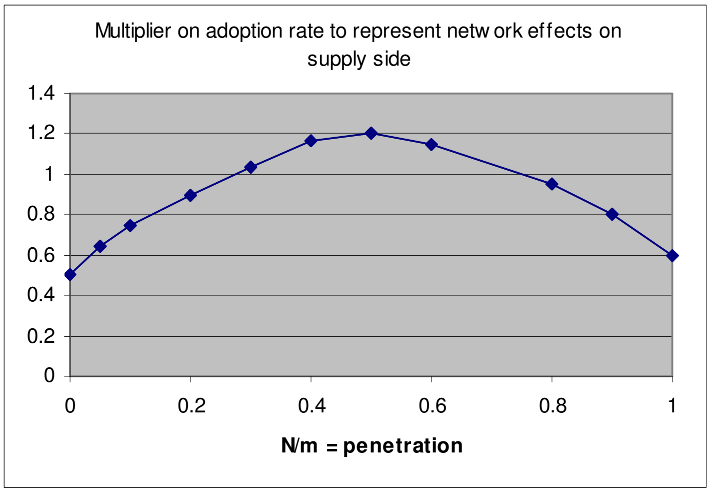
Some work-in-progress results are presented in the Appendix. The results show that it is easy to obtain a fit for the stock of adopters against cumulative subscribers or penetration data. The very high R^2 values and low Durbin Watson statistics – suggesting autocorrelation of residuals – are a result of accumulating the errors in the process that actually generates the stock of adopters. Given that it is intuitively appealing, and possible to reproduce the sigmoid time history using system dynamics calibration, is this wrong? The answer is no, if one is simply seeking a set of parameters to calibrate a model. If one wishes to do hypothesis tests or to check the normality of residuals (for the purpose of constructing confidence intervals on the parameters), then it is better to calibrate the adoption rate against the data. The system dynamics perspective that integration is a better way of seeing the world, ie that a stock of adopters result from accumulating sales, rather than sales being the differential of the adopters, would support the view that it is the adoption rate which is the key variable, that is the adoption rate is "causing" the stock of adoptions rather than vice versa.

1. *Calibration against stock or flow data?* Calibrating against the adoption rate yields similar parameter estimates and usually better statistics for normality and homoscedasticity.

[3] YD Real households' disposable income per head, CVM SA; R: selected UK retail banks' base rate

2. *The effect of exogenous economic variables.* The economic variables added to the quality of the fit in the case of UK mobiles, since diffusion has seemed to be strongly driven by price decreases. The other economic variables remained relatively static over their time history. Generally, the fit of models is good without including decision variables (eg price) and environmental economic variables (Bass, Krishnan and Jain, 1994) and adding these variables often does not improve forecast performance (Bottomley and Fildes, 1998).

3. *Network effects on supply side.* A 'table function' multiplier, to represent the effect of network effects on the supply-side, was found to improve the fit (as measured by log likelihood) and is an independent confirmation of what is known from other sources (ie literature and management reports) to be the case. The relationship was found to be applicable for both i-mode and UK mobile phone adoption but led to no improvements for the broadband case. It could alternatively be seen as a time dependence (or, more strictly, a penetration-dependence) of the demand-side p and q coefficients.



4. *Lack of identifiability of system.* The weakness of trying to estimate both structure and parameters from historic time series data is that different model specifications and parameter sets can fit the data equally well. This is a strong point in favour of the system dynamics method where parameters, and particularly structure should be elicited from managerial or 'operator' knowledge

5. *Endorsement of the 'hand calibration' method.* As implemented in Vensim, and especially using the Synthesim capability within Vensim, it is easy to get a picture of what changes different parameters make by changing the simulation interactively and thus learning about the model's behaviour. In conjunction with the calibration routines, reasonable bounds on parameters can be defined. In contrast, the automatic calibration procedures in Eviews sometimes resulted in parameters that artificially tracked signals in the exogenous data, yielding extreme and incorrectly signed parameters.

6. *Calculation of confidence intervals.* The likelihood ratio method of calculating confidence intervals that is implemented in Vensim is based on the assumptions of normality, independence and homoscedasticity of residuals. Where these conditions do not apply, this method cannot be used. An alternative method called bootstrapping, which relaxes these assumptions, is mentioned by Dogan (2004) but this seems rather cumbersome at present.

Further work planned

Further work is planned in this research to:

1. undertake judgemental bootstrapping experiments from expert judgement forecasts to serve as a basis for later model experimentation

2. supply additional statistics, eg the Theil statistics (as recommended by Sterman) could be added to the diagnostic indicators in the results.
3. perform forecast tests, by the usual procedure of splitting historic data into two portions and carrying out 'ex post' forecasting. Papers testing differing models in the forecasting literature often perform forecast tests, and not just fitting against historic data, so 1-step ahead forecast tests could be added.
4. examine supply side structures with the involvement of real operator experience in developing potential business models – ie not just relying on calibration
5. develop practical ideas on model validation and how this varies across the continuum of models from policy modelling to 'predictive' modelling

Conclusions

Calibration of system dynamics models against historic time series data in the end is very similar to econometric estimation. Given that the marketing and economics literature is dominated by econometric estimation methods, a higher recognition for system dynamics and its approaches to parameter estimation would be aided by connecting better with econometric procedures and quoting diagnostic indicators.

This is particularly true in accepting that parameter estimation by direct observation, which is part of good system dynamics practice, is not really feasible for aggregated market level models. A major weakness of econometric estimation is that one does not have the same opportunity to explore manual simulation and calibration methods, which aid model understanding. However, econometric packages have a greater variety of tools to explore diagnostics of the fit against the data. For example, knowing whether the residuals pass a test for normality and heteroscedasticity is valuable, if not to perform statistical hypothesis testing, but to check whether confidence bounds can be reliably stated using the simple likelihood ratio procedure available in Vensim.

In many practical situations, a broad estimate of parameter values will be all that is needed in diffusion models, since the overall behaviour will not alter, especially if one is comparing alternative structures for a future market. This paper has proposed that one way of incorporating expert judgement into models of new markets is through judgemental bootstrapping, ie calibrating models against expert forecasts as a basis for subsequent structural experiments.

System dynamics have a potentially strong role in forecasting applications, even though the community has tended to reject the use of models to predict rather than learn (de Geus, 1992). At the least, forecasting practitioners will want to know whether system dynamics models can provide managerial guidance to the assessment of new market entry strategy. Marketing science practitioners will be interested in techniques that allow market strategy to be conceptualised and quantified, not just pre-launch forecasting. However, different modes of modelling represent poles of a continuum and therefore some of the criteria by which forecasting models are judged need to be incorporated in system dynamics practice.

However, returning to the 'policy modelling' that is a more mainstream domain for the system dynamics method, there is wide opportunity to use diffusion models as a basis for lifecycle management strategies or identifying the best supply-side configurations for firms in complex service delivery chains. Initial evidence has been adduced that positive network effects on the supply side act to accelerate adoption with increasing penetration factors.

Acknowledgements
The author would like to acknowledge the contribution of Professor Barry Evans, Head of the Centre for Communications Systems Research as lead supervisor of this research work and particularly of

Professors Paul Levine and Neil Rickman, Department of Economics for their guidance and suggestions on the econometric analysis.

References

Armstrong JS (2001). Judgmental bootstrapping: inferring experts' rules for forecasting. In: Principles of Forecasting: A handbook for researchers and practitioners. Armstrong (ed). Kluwer Academic Publishers, p 171-192.

Barabba V, Huber C, Cooke F, Pudar N, Smith J and Paich M (2002). A Multimethod Approach for Creating New Business Models: The General Motors OnStar Project. Interfaces, Jan/Feb 2002, Vol. 32(1), 20-34.

Barlas Y (1996). Formal aspects of model validity and validation in system dynamics. System Dynamics Review 12(3), 183-210.

Bass FM (1969). A new product growth model for consumer durables. *Management Science.* **13**(3): 203–223.

Bass FM, Krishnan TV and Jain D (1994). Why the Bass model fits without decision variables. Marketing Science 13 (Summer), 204-223.

Bottomley PA and Fildes R (1998). The role of prices in models of innovation diffusion. Journal of Forecasting, 17, 539-555.

Dogan G (2004). Confidence interval estimation in system dynamics models: bootstrapping vs likelihood ratio method. Proceedings of the International System Dynamics Conference, Oxford. Proceedings of the International System Dynamics Conference, Oxford, Kennedy M, Winch GW, Langer RS, Rowe JI and Yanni JM (Eds).

Easingwood CJ, Mahajan V and Muller E (1983). A nonuniform influence innovation diffusion model of new product acceptance. Marketing Science 2(3-Summer), 273-295.

Euromonitor (2006). Euromonitor International Global Market Information database. Euromonitor Plc. www.euromonitor.com

Fildes R (2002). Telecommunications demand forecasting – a review. International Journal of Forecasting 18, 489-522.

Fildes R (2003). Book review: New Product Diffusion Models: Mahajan V, Muller E and Wind Y (Eds) (2000). International Journal of Forecasting 19, 327-328.

Ford A and Flynn H (2005). Statistical screening of system dynamics models. System Dynamics Review, 21(4), 273-303.

Fourt LA and Woodlock JW (1960). Early prediction of market success for new grocery products. Journal of Marketing 25(2): 31-38.

Forrester JW (1961). Industrial Dynamics. MIT Press, Cambridge, MA.

Forrester JW (1980). System dynamics: future opportunities. In: System Dynamics, TIMS Studies in the Management Sciences, Legasto AA Jr (ed). North-Holland, New York, 7-22.

Forrester JW (1980) and Senge PM (1980). Tests for building confidence in system dynamics models. In: System Dynamics, TIMS Studies in the Management Sciences, Legasto AA Jr (ed). North-Holland, New York, 209-28.

Forrester JW (1996). National Model Response to Criticisms of the Economics Profession by Economists. D-memo D-4664-1. System Dynamics Group D-memo collection, MIT.

Forrester JW (2003). Economic theory for the new millennium. Plenary Address. Proceedings of the 21st International System Dynamics Society conference, New York. Eds: Eberlein RL, Diker VG, Langer RS and Rowe JI.

de Geus AP (1992). Modelling to predict or learn? European Journal of Operational Research 59, 1-5. Reprinted in Modeling for Learning Organizations (1994) Morecroft JDW and Sterman JD (Eds), Productivity Press, Portland, OR.

Graham AK (1980). Parameter estimation in system dynamics modeling. In: Elements of the system dynamics method. Randers, J (ed), Productivity Press, Cambridge, Mass, (originally MIT Press), p143-161.

Graham AK and Ariza CA (2001). Dynamic, hard and strategic questions: Using optimization to answer a marketing resource allocation question. Proceedings of the 2001 International System Dynamics Conference. Atlanta, Georgia.

Graham AK (2002). On positioning system dynamics as an applied science of strategy, or: SD is scientific. We haven't said so and we should. Proceedings of the International System Dynamics Conference, Palermo.

Graham AK, Moore J and Choi CY (2002). How robust are conclusions from a complex, calibrated model, really? A project management model benchmark using fit-constrained Monte Carlo analysis. Proceedings of the International System Dynamics Society Conference, Palermo.

Graham AK and Godfrey J (2005). Achieving win-win in a regulatory dispute: managing 3G competition. Proceedings of the International System Dynamics Conference, Boston.

Homer JB (1996). Why we iterate: scientific modeling in theory and practice. System Dynamics Review 12(1), 1-20.

Karshenas M and Stoneman P (1992). A flexible model of technological diffusion incorporating economic factors with an application to the spread of colour television ownership in the UK. Journal of Forecasting, 11(7), 577-601.
Lilien GL, Rangaswamy A and van den Bulte, C (2000). Diffusion models: managerial applications and software. In: Mahajan V, Muller E and Wind Y (2000) (eds). New product diffusion models. International series in quantitative marketing. Kluwer Academic Publishers, Boston.
Lyneis JM and Pugh, III, AL (1996). Automated vs 'Hand' Callibration of System Dynamics Models: An Experiment with a Simple Project Model. International System Dynamics Conference, Cambridge, Massachusetts, System Dynamics Society.
Lyneis J (2000). System dynamics for market forecasting and structural analysis. System Dynamics Review 16(1), 3-25.
Mahajan V, Mason CH and Srinivasan V (1986). An evaluation of estimation procedures for new product diffusion models. In: Innovation diffusion models of new product acceptance. Vol 5 of Series on Econometrics and Management Sciences. Mahajan and Wind (Eds). Ballinger Publishing Company, Cambridge, Mass.
Mahajan V, Muller E and Bass FM (1990). New product diffusion models in marketing: a review and directions for research. *Journal of Marketing 54(1), 1-26.*
Mahajan V, Muller E and Wind Y (2000) (eds). New product diffusion models. International series in quantitative marketing. Kluwer Academic Publishers, Boston.
Mahajan V and Wind Y (1986). Innovation diffusion models of new product acceptance – a re-examination. In: Innovation diffusion models of new product acceptance. Vol 5 of Series on Econometrics and Management Sciences. Mahajan and Wind (Eds). Ballinger Publishing Company, Cambridge, Mass.
Mansfield E. (1961). Technical change and the role of imitation. Econometrica 29(4): 741-766.
Mass NJ and Senge PM (1980). Alternative tests for selecting model variables. In: Elements of the system dynamics method, Randers, J (ed), Productivity Press, Cambridge, Mass, (originally MIT Press), p. 205-225.
NTT Docomo (2006). NTT DoCoMo website and Investor Relations website data. http://www.nttdocomo.co.jp/english/corporate/investor_relations/top_e.html
OECD (2003). Telecommunications Database Vol 2003 release 01 - *OECD Organisation for Economic Co-operation and Development.*. Database Edition (ISSN 1608-1315). Available from *http://new.SourceOECD.org/database/telecom*
OECD (2005). OECD Communications Outlook. Information and Communications Technologies. Organisation for Economic Co-operation and Development. OECD Publishing. ISBN 92-64-00950-7.
Ofcom (2004). The Communications Market 2004 (August). http://www.ofcom.org.uk/research/cm/
Oliva R (2003). Model calibration as a testing strategy for system dynamics models. European Journal of Operational Research 151, 552-568.
ONS (2006a). Index of Internet Connectivity. Survey of UK Internet Service Providers. http://www.statistics.gov.uk/statbase/Product.asp?vlnk=8251
ONS (2006b). Retail Prices Index RP02. Office for National Statistics. London. http://www.statistics.gov.uk/downloads/theme_economy/Rp02.pdf
ONS (2006c). Statbase, ~Economic Trends. Office for National Statistics. London.http://www.statistics.gov.uk/statbase/
Peterson DW (1980). Statistical tools for system dynamics. In: The elements of the system dynamics method, Randers J (Ed). Productivity Press, Cambridge, Mass.
Putsis Jr WP and Srinivasan V (2000). Estimation techniques for macro diffusion models. In: Mahajan V, Muller E and Wind Y (2000) (eds). New product diffusion models. International series in quantitative marketing. Kluwer Academic Publishers, Boston.
Roberts JH and Lattin JM (2000). Disaggregate-level diffusion models. In: Mahajan V, Muller E and Wind Y (2000) (eds). New product diffusion models. International series in quantitative marketing. Kluwer Academic Publishers, Boston.
Srinivasan V and Mason CH (1986). Non-linear least squares estimation of new product diffusion models. Marketing Science 5(2), 169-178.
Sterman JD (1984). Appropriate summary statistics for evaluating the historical fit of system dynamics models. Dynamica 10(2), Winter, 51-66.
Tee R (2005). Different directions in the mobile internet: analysing mobile internet services in Japan and Europe. In Mobile World: past, present and future. Hamill L and Lasen A (Eds). Springer.
Thomas RJ (1985). Estimating market growth for new products: an analogical diffusion models approach. *Journal of Product Innovation Management*, 2 (March), 45-55.
Winch GW (1993). Consensus building in the planning process: benefits from a 'hard' modeling approach. System Dynamics Review 9(3): 287-300.

Zettelmeyer F and Stoneman P (1993). Testing alternative models of new product diffusion. In: Chapter 10: Stoneman P (2002). The economics of technological diffusion. Blackwell, Oxford.

Notes

1. Importance of calibration and a statistical fit criterion see: Reichelt, K. S., et al. (1996). Calibration Statistics: Selecting a Statistic and Setting a Standard. 1996 International System Dynamics Conference, Cambridge, Massachusetts, System Dynamics Society.
2. 'Positive economics' = ability of model to predict without need for explanation or the need for the model structure to be consistent with reality

APPENDIX – Interim results for Eviews estimation / Vensim calibration

Table 3: i-mode calibration

i-mode penetration (mainly monthly) data: 22 Feb 1999 - 31 Dec 2005

					Parameters						Diagnostic statistics										
#	model	soft-ware	data	Comments	p 1/yr	q 1/yr	m [-]	γ % / yr	σ [-]	α %/yr	SSR	T	k	SE	R^2	$(R\text{-bar})^2$	LL	DW	J-B Norm JB, p	B-G LM12 F, p	White Heteros Obs*R2, p
1	Bass	E	"	constant m	0.09613	1.031	34.11	0.00	1	0	31.5641	29	3	1.1018	0.9931	0.9926	-42.38	0.2209	2.386	21.078	0.4920
	se			N formul'n	0.01182	0.091	0.46	–	–	–									0.303	0.000	0.7819
	t-ratio				8.13008	11.277	73.60	–	–	–											
2	Bass	E	delta subscri	const m	0.0636	1.595	38.25				4.266	77	3	0.2401	0.7471	0.7403	2.13	0.9134	0.471	10.003	3.7482
	se			NLS formul'	0.00869	0.106	1.66												0.790	0.000	0.5862
	t-ratio				7.31668	15.075	22.99														
3	Bass	E	subcribers	const m	0.10162	0.985	43.69				135.722	78	3	1.3452	0.9916	0.9914	-132.28	0.0409	4.662	394.020	9.0965
	se			N formul'n	0.00713	0.053	0.35												0.097	0.000	0.0106
	t-ratio				14.2505	18.671	124.22														
4	Bass	E		const m	0.1466	0.868	44.50				755.933	77	3	3.1961	0.6566	0.6473	-197.20	0.6538	4.112	12.331	12.0421
	se			sales	0.02404	0.091	0.66												0.128	0.000	0.0024
	t-ratio				6.09742	9.585	67.06														
5	Bass	E		ord sales	0.10258	1.928	26.66	11.46			284.660	77	4	1.9747	0.8707	0.8654	-159.60	1.624	11.400	4.225	6.8586
	se				0.02668	0.116	0.78	0.87											0.003	0.000	0.2314
	t-ratio				3.8457	16.623	33.99	13.13													
6	EMM	E		ord sales	0.0315	1.654	28.02	10.22	0.7518		260.532	77	5	1.9022	0.8816	0.8751	-156.19	1.7866	34.237	3.771	6.9859
	se				0.0404	0.148	1.14	1.10	0.0923										0.000	0.000	0.2217
	t-ratio				0.77955	11.176	24.56	9.29	8.1446												
7	Bass	V	subscribers		0.06106	2.028	26.44	11.77			9.041	78	4	0.3495	0.9994	0.9994	-26.63	0.1782			
	95% CI lower				0.05748	1.985	26.26	11.53													
	95% CI upper				0.06494	2.072	26.62	12.01													
8	Bass	V	sales	const m	0.07381	1.452	40.05				574.556	77	3	2.7864	0.7442	0.7373	-186.64	0.8758	0.133		6.21
																			0.936		0.04
9	Bass	V	sales data		0.04756	2.221	25.00	12.71			260.244	77	4	1.8881	0.8818	0.8769	-156.14	1.8837	27.219		9.9493
																			0.000		0.0069
10	EMM	V	sales data		0.02591	2.059	25.50	12.04	0.89168	0	258.021	77	5	1.8930	0.8828	0.8763	-155.81	1.9025	8.738		
			for subscribers data		0.04464	1.959	26.49	11.77	0.92365												
11	KS	V	sales data		0.0277	3.436	22.56	21.54		93.2145	240.258	77	5	1.8267	0.8908	0.8848	-153.07	2.0533	8.565		

Key
E = Eviews; V = Vensim; SSR = sum of squared residuals; T = number of observations, k = number of estimated coefficients; SE = standard error of the regression; R^2 = coefficient of determination; $(R\text{-bar})^2$ = is R^2 adjusted for degrees of freedom; LL = log likelihood; DW = Durbin Watson statistic; JB = Jarque-Bera test for normality (JB statistic, p-value); B-G = Breusch-Godfrey Lagrange Multiplier (LMj) test for serial correlation to the jth degree [F-statistic, p value]; White = White's test for heteroscedasticity [Obs*R^2 statistic, p value];

Table 4: UK mobile phone subscription calibration

UK mobile annual data 1984 - 2004

model	s/w	data	Comments	p	q	m	σ	α	[P] a1	[YD] a2	[R] a3	SSR	T	k	SE	R^2	$(\text{R-bar})^2$	LL	DW	JB Norm p	B-G LM1 F, p	White Heteros Obs*R2, p
				Parameters								Diagnostic statistics										
Long tail means low p and high q																						
Bass	V	% pen	with econ f	0.00089	0.08953	89.1628	1.000	0	-1.056	0.000	-0.013	69.65	20	6	2.23052	0.996889	0.99578	-40.9	1.72256	4.173487		
			adoption peak leads data by about 0.7 yr																	0.124091		
		adoption rate	with econ f	0.000144	0	85.0965	1.000	0	-3.465	0.000	-0.131	25.88	20	6	1.35958	0.972424	0.96258	-31.0	2.26823	1.494622		vis hetero
																				0.473638		
EMM	V	% pen	with econ f	0.001533	0.10341	87.9631	1.198	0	-0.915	0.756	0.000	63.19	20	7	2.20474	0.997177	0.99587	-39.9	1.94908	2.735555		
																				0.254672		
		adoption rate		as for Bass given that q=0																		
KS	V	% pen	with econ f	0.000988	0.12667	89.8475	1.000	23.6062	-1.068	0.000	-0.080	66.72	20	7	2.26549	0.99702	0.99564	-40.4	1.71629	4.337957		
																				0.114294		
Bass	V	% pen	econ+mult	0.003145	0.05584	90.4168	1.000	0	-1.199	0.057	0.000	37.31	20	6	1.63245	0.998334	0.99774	-34.6	1.98346	2.864735		
																				0.238743		
EMM	V	% pen	econ+mult	0.00333	0.06138	90.1289	1.057	0	-1.091	0.446	-0.001	37.17	20	7	1.69097	0.99834	0.99757	-34.6	2.05588	3.615716		
																				0.164005		
KS	V	% pen	econ+mult	results not improved on Bass ie. Alpha = 0																		
Bass	E	sales	no econ	0.0001	0.86528	88.5429																
				fixed	0.0935	2.51727																
					9.254	35.1741																
Bass	E	w econ		0.000252	0	91.328			-3.015		-0.492	19.42	20	7	1.22225	0.979306	0.96975	-28.1	1.89735	11.8883	2.765169	7.359731
		negative values of p or q		0.000104		2.21119			0.155		0.362									0.002621	0.117085	0.59972
EMM	E			2.421469		41.3028			-19.45		-1.362											
KS	E																					

Key

E = Eviews; V = Vensim; SSR = sum of squared residuals; T = number of observations, k = number of estimated coefficients; SE = standard error of the regression; R^2 = coefficient of determination; $(\text{R-bar})^2$ = is R^2 adjusted for degrees of freedom; LL = log likelihood; DW = Durbin Watson statistic; JB = Jarque-Bera test for normality (JB statistic, p-value); B-G = Breusch-Godfrey Lagrange Multiplier (LMj) test for serial correlation to the jth degree [F-statistic, p value]; White = White's test for heteroscedasticity [Obs*R^2 statistic, p value];

UK broadband residential subscriptions Jan 2000 - Dec 2005

				Parameters								Diagnostic statistics													
									[P]	[YD]	[R]											J-B		B-G	White
model	pkge	data	Commen	p	q	m	σ	α	a1	a2	a3	SSR	T	k	SE regr'	mean dv	%SE	R^2	$(\text{R-bar})^2$	LL	DW	Norm	Skew Kurt*	LM12 F, p	Heteros Obs*R2, p
			consideration made to lagging the economic variables but visual inspection did not suggest this would help																						
Bass	V	UK b'band pen %		0.0254	0.6899	100	1.000	0	-0.1777	0.213	0	17.70	60	6	0.57249	####	0.02510	0.9992	0.9992	-48.5	0.1487	1.003	-0.125		13.402
																						0.606	-0.619		0.00123
EMM	V			0.0257	0.7227	100	1.009	0	-0.1252	0.188	0	18.53	61	7	0.58584	####	0.02569	0.9992	0.9991	-50.2	0.1741	34.068			vis homo
																						0.000			
KS	V			0.0248	0.7306	100	1.000	3.74	-0.1404	0.403	0	18.63	61	7	0.58734	####	0.02575	0.9992	0.9991	-50.4	0.1735	33.405			vis homo
																						0.000			
			multiplier doesn't assist calibration																						
Bass	V	adoption rate	no econ	0.0207	0.8234	100						361.22	60	3	2.51737	####	0.19822	0.8917	0.8879	-139.0	1.2815	53.225	0.8825		0.572641
																						0.000	4.3927		0.751022
Bass	V	adoption rate		0.0232	0.719	100	1.000	0	-0.0076	1.346	0	362.25	61	6	2.56638			0.8963	0.8869	-140.9	1.2854	7.108			vis homo
																						0.029			
EMM	V	adoption rate		0.0166	0.5505	100	0.86269	0	−0.1525	1.901	0	358.13	61	7	2.57526			0.8975	0.8861	-140.5	1.3029	12.439			vis homo
																						0.002			
KS	V		same results as Bass, ie alpha = 0																						
Bass	E	sales	no econ	0.0217	0.7697	108.2						367.30	60	2	2.53849	####	0.19988	0.8899	0.8880	-139.5	1.2907	33.616	-0.775	1.1982	2.59103
				0.0053	0.0592	8.472																0.000	6.3233	0.3137	0.273757
				4.118	12.997	12.78																			
	E	m=100		0.02	0.828	100						375.49	60	2	2.54441	####	0.20035	0.8874	0.8855	-140.2	1.2635	44.884	-0.951	1.2436	2.487511
				0.0056	0.028																	0.000	6.7863	0.2841	0.288299
				3.5618	29.624																				
	E	slhs		0.0048	0.955	100						11.20	60	2	0.43937	-2.10	0.20900	0.8730	0.8708	-34.8	1.4401	######	-2.529	0.731	4.04062
				0.0023	0.0683																	0.000	15.234	0.7141	0.132614
				2.1135	13.991																				
	E	NLS method		0.0226	0.7491	110.9						2.43	60	3	0.20634	1.06	0.19496	0.8952	0.8916	11.1	1.3183	34.925	-0.712	1.2242	2.24962
				0.0022	0.062	9.662																0.000	6.4556	0.2969	0.324714
				10.492	12.078	11.48																			
				0.0213	0.8221	100																			
Bass	E	slhs	w econ	-0.0194	3.6989	100			−1.2345	−21.1	1.699	8.07	60	5	0.38295	-2.10	0.18216	0.9085	0.9019	-24.9	2.1628	340.20	-2.284		
																						0.000	13.73		

Key
E = Eviews; V = Vensim; SSR = sum of squared residuals; T = number of observations, k = number of estimated coefficients; SE = standard error of the regression; R^2 = coefficient of determination; $(\text{R-bar})^2$ = is R^2 adjusted for degrees of freedom; LL = log likelihood; DW = Durbin Watson statistic; JB = Jarque-Bera test for normality (JB statistic, p-value); B-G = Breusch-Godfrey Lagrange Multiplier (LMj) test for serial correlation to the jth degree [F-statistic, p value]; White = White's test for heteroscedasticity [Obs*R^2 statistic, p value];