



---

# Endogeneity and Exogeneity in Sales Response Functions

Wolfgang Polasek<sup>1</sup>

IHS, Stumpergasse 56, A-1060 Vienna, [polasek@ihs.ac.at](mailto:polasek@ihs.ac.at)

**Summary.** Endogeneity and exogeneity are topics that are mainly discussed in macroeconomics. We show that sales response function (SRF) are exposed to the same problem if we assume that the control variables in a SRF reflect behavioral reactions of the supply side. The supply side actions are covering a flexible marketing component which could interact with the sales responses if sales managers decide to react fast according to new market situations. A recent article of [3] suggested to use a class of production functions under constraints to estimate the sales responses that are subject to marketing strategies. In this paper we demonstrate this approach with a simple SRF(1) model that contains one endogenous variable. Such models can be extended by further exogenous variables leading to SRF-X models. The new modeling approach leads to a multivariate equation system and will be demonstrated using data from a pharma-marketing survey in German regions.

**Key words:** Sales response functions, stochastic derivative constraints, simultaneous estimation, MCMC, pharma-marketing, model choice.

## 1 Introduction

Kao et al. (2005) [3] have proposed a simultaneous estimation of sales in dependence of marketing inputs. The new idea behind this approach is that the (optimal) expenditures, which are inputs in the SRF, might depend on the current sales and should be estimated endogenously.

Polasek (2010) [5] has introduced a family of SRF(k) for a cross-sectional sample where the parameter k denotes the number of endogenous input variables (like marketing expenditures or sales related covariates). In this paper we show how to estimate a SRF(1) model that explains sales output by a Cobb-Douglas type of function of marketing variables.

The main point of the current approach is the extension of the traditional estimation of a SRF to a system of observations, because some input variables are jointly determined by the output. This problem was solved by system estimation in macroeconomics over the last decades. New is the assumption that

the endogeneity of the inputs stems from an implied (stochastic) optimality consideration, which allows the first partial derivative of the SRF to follow a normal distribution. We call this crucial behavioral assumption the stochastic partial derivative (SPD) assumption. More extensions of these modeling considerations lead to a model choice problem. The details of this model choice problem need to be more worked out in future, but we will concentrate in the next section on the model estimation part and we will use the marginal likelihood criterion for the Bayesian model choice.

If the first partial derivative of a response model is jointly determined with the dependent variable, then these 2 equations imply a simultaneous equation system, since the stochastic restrictions imply an endogeneity of the input variables. There are 3 main implications that constitute a SRF-SPD model:

1. Sales (demand) model: a (production) function of the input variables  $X$  explain  $\Rightarrow$  output sales variables  $y$  plus noise.
2. Behavioral (supply) model + SPD assumptions: results of marketing efforts are proxied by first derivatives and  $\Rightarrow$  follow a normal distribution.
3. Conditional on SPD and known SRF coefficients: The SPD assumption leads to  $\Rightarrow$  stochastic regressors.

In the next section we describe the SRF(1) model together with a Bayesian (MCMC) estimation procedure. Section 1 introduces the SRF(1)X(q) model (a notation similar to notations in multiple time series) to indicate that a SRF model does not need to have only endogenous variables, but can also include q exogenous covariates. Section 2 explains the estimation procedure and Section 3 discusses a regional sales response model example that involves a German pharma-marketing data set for the years 2008 and 2007. In a final section we conclude.

## 2 The SRF(1) model with endogeneity

We consider the simplest possible model to demonstrate the endogeneity effects in the SRF(1) sales response function with one input variable  $x$

$$y = \beta_0 x^{\beta_1} e^\epsilon, \quad (1)$$

where  $\epsilon$  is a  $N[0, \sigma_y^2]$  distributed error term. By taking logs for the n cross-sectional observations we get the following linear regression model

$$\ln y \sim N[\mu_y = X\beta, \sigma_y^2 I_n] \quad (2)$$

with the regression coefficients  $\beta = (\ln \beta_0, \beta_1)$  and the regressor matrix  $X = (1_n : \ln x)$  where  $1_n$  is a vector of ones and  $x$  is the cross-sectional decision variable that will influence the sales  $y$  (a  $n \times 1$  vector) in the n regions. Thus the model is of the type of a log linear production function as it is used in macro-economics. The new assumption is the stochastic partial derivative

(SPD) assumption, which implies an additional equation for the supply side behavior:

$$\partial y / \partial x = \beta_0 \beta_1 x^{\beta_1 - 1} e^\epsilon$$

Next, we make the assumption that the log derivatives across all units

$$\ln y_x = \log(\beta_0 \beta_1) + (\beta_1 - 1) \ln x \tag{3}$$

are normally distributed with parameters that can be estimated:

$$\ln y_x \sim N[\ln \lambda, \sigma_\lambda^2] \quad \text{or} \quad \ln y_x \sim N[\ln \lambda 1_n, \sigma_\lambda^2 I_n]. \tag{4}$$

This means that the sales responses  $y$  and the decision variable  $x$  might follow a restriction that allocates resources according to the first derivative of the SRF but the empirical observations across the  $n$  regions reveal some noise. These stochastic fluctuations across regions are captured by the mean response  $\lambda$ , and the variance  $\sigma_\lambda^2$  of the constraint measures the looseness or strength of this optimality behavior in the model and can be interpreted as a tightness parameter for the SPD restriction.

Adding the stochastic partial derivative (SPD) restrictions for the  $x$  regressor into the SRF model imposes the behavioral assumption that the marketing efforts are allocated in such a way that the marketing results (via the SRF derivative) should be equal across the regional units: Additionally, the above SPD assumption leads a normal distribution of the regressor  $x$ :

$$\ln x \sim N[\mu_x, \sigma_x^2 I_n] \tag{5}$$

where the parameters  $\theta_x = (\mu_x, \sigma_x^2)$  are determined by the SRF and SPD parameters. This approach actually implies the endogeneity of  $x$  in the SRF model and shows how the SPD assumption (i.e. the feedback from the sales to the marketing control variables) influences the inference process to yield a simultaneous estimation model.

To see how the SPD assumption translates to an assumption about the  $x$  variable, we re-write the exponent of the SPD density (4) and use the log derivative (3)

$$\begin{aligned} N[\ln \lambda, \sigma_\lambda^2] &\propto (\ln y_x - \ln \lambda)^2 / \sigma_\lambda^2 \\ &= (\log(\beta_0 \beta_1) + (\beta_1 - 1) \ln x - \ln \lambda)^2 / \sigma_\lambda^2 = \\ &= \left( \frac{\log(\beta_0 \beta_1 / \lambda)}{1 - \beta_1} - \ln x \right)^2 (\beta_1 - 1)^2 / \sigma_\lambda^2 \\ &\propto N[\ln x \mid \mu_x, \sigma_x^2 = \sigma_\lambda^2 / (\beta_1 - 1)^2] \end{aligned}$$

with the mean  $\mu_x = \frac{\log(\beta_0 \beta_1 / \lambda)}{1 - \beta_1}$  and variance  $\sigma_x^2 = \frac{\sigma_\lambda^2}{(\beta_1 - 1)^2}$  of  $\log(x)$ . (Note that the Jacobian for  $\ln x$  from  $\lambda$  is just  $1/|\beta_1 - 1|$ .)

Finally, we define the SRF(1)-SPD model by the following sequence of (conditional) normal densities:

**Definition 1.** *The SRF(1)-SPD model  $y = \beta_0 x^{\beta_1} e^u$  is defined as the following set of 3 log-normal densities:*

$$\begin{aligned} \ln y \mid \text{SRF}, \beta &\sim N[\ln \beta_0 + \beta_1 \ln x, \sigma_y^2] \\ \Rightarrow \ln x \mid \text{SPD}, \beta &\sim N[(\ln \beta_0 + \ln \beta_1 - \ln \lambda) / (\beta_1 - 1), \sigma_x^2] \\ \ln y_x \mid \text{SPD}, \theta_x &\sim N[\ln \lambda, \sigma_\lambda^2]. \end{aligned} \quad (6)$$

where 'SRF' stands for the functional form of the model and 'SPD' stands for the  $y_x$  is the first derivative of the SRF(1) model and with the parameters of the model given by  $\theta = (\theta_x, \theta_\lambda) = (\beta, \sigma_y^2, \lambda, \sigma_\lambda^2)$ . " $\Rightarrow$ " denotes the derived distribution for  $\ln x$ .

An alternative way of writing the SRF(1) model is:

$$\begin{aligned} \text{Demand : } & N[\ln y \mid \mu_y, \sigma_y^2], \quad \mu_y = \ln \beta_0 + \beta_1 \ln x; \\ \Rightarrow \text{Supply : } & N[\ln x \mid \mu_x, \sigma_x^2], \quad \mu_x = \frac{\log(\beta_0 \beta_1 / \lambda)}{1 - \beta_1}; \\ \text{Behavior : } & N[\ln y_x \mid \ln \lambda, \sigma_\lambda^2]. \end{aligned}$$

It is important to note that while we assume independence between the 3 densities in Definition 1, the parameters of the  $x$  distribution are not independent and are functions of the parameters of the demand and behavioral equation.

The supply side of the model reacts to the market by 'targeting' the first derivative of the SRF model to get approximately equal results across all regions. Making the assumption that the first derivative of the SRF model with respect to the control variable  $x$  is the marketing target of the supply side has 2 implication for the modeling process: (1) the control variable  $x$  is for the SRF estimation and therefore a simultaneous system for  $x$  and  $y$  has to be estimated. (2) The parameters of the  $x$ -density depend on the functional form and the parameters of the behavioral equation,  $\theta_\lambda$ . Because the results of marketing efforts on  $y$  can not be observed directly, this variable is a latent variable and has to be proxied by the first derivative that depends on the functional form. Thus, the behavioral equation has to be included into the model and is part of the simultaneous estimation process.

For statistical inference we can estimate the parameter vector  $\theta$  of the system by maximum likelihood or by MCMC, assuming a prior density given by  $p(\theta)$ . In the next section we outline the MCMC procedure.

## 2.1 MCMC estimation in the SRF(1) model

The MCMC estimation of the SRF(1) model requires the likelihood function for  $\mathcal{D} = (\ln y, \ln x)$  given by

$$l(\mathcal{D} \mid \theta) = N[\ln y \mid \mu_y, \sigma_y^2 I_n] N[\ln x \mid \mu_x, \sigma_x^2 I_n] * J \quad (7)$$

and  $J$  being the Jacobian from the transformation of  $\theta_\lambda$  to  $\theta_x$ .

$$\mu_y = \ln \beta_0 + \beta_1 \ln x \text{ and } \mu_x = (\ln \beta_0 + \ln \beta_1 - \mu_\lambda) / (\beta_1 - 1)$$

The prior density is assumed to be

$$p(\theta) = N[\beta | \beta_*, H_*] N[\lambda | \lambda_*, \sigma_{\lambda_*}^2] \prod_{j \in \{y, \lambda\}} \Gamma[\sigma_j^{-2} | \sigma_{j_*}^2 n_{j_*} / 2, n_{j_*} / 2]. \quad (8)$$

Then the posterior distribution for  $\theta$  is given by  $p(\theta | \mathcal{D}) \propto l(\mathcal{D} | \theta) p(\theta)$ .

**Theorem 1 (MCMC in the SRF(1)-SPD model).**

The MCMC iteration in the SRF(1)-SPD model with the likelihood function (7) and the prior density (8) takes the following draws of the full conditional distributions (fcd):

1. Starting values: set  $\beta = \beta_{OLS}$  and  $\lambda = 0$
2. Draw  $\lambda$  from  $p(\lambda | \lambda_{**}, \tau_{**}^2)$
3. Draw  $\beta$  by a Metropolis step (see [2]) from  $p[\beta | \mathbf{b}_*, \mathbf{H}_*] l(\theta | y)$
4. Draw  $\sigma_y^{-2}$  from  $\Gamma[\sigma_y^{-2} | s_{y**}^2, n_{y**}]$
5. Draw  $\sigma_\lambda^{-2}$  from  $\Gamma[\sigma_\lambda^{-2} | s_{\lambda**}^2, n_{\lambda**}]$
6. Repeat until convergence.

The proof can be found in Polasek (2010) [5]. The marginal likelihood of model  $\mathcal{M}$  is computed by the Newton-Raftery (1994) [4] formula

$$\hat{m}(\mathbf{y} | \mathcal{M})^{-1} = \frac{1}{n_{rep}} \sum_{j=1}^{n_{rep}} \left( \sum_{i=1}^n \ln l(\mathcal{D}_i | \mathcal{M}, \theta_{(j)}) \right)^{-1} \quad \text{or} \quad \hat{m}_\alpha = \bar{m}_\alpha (l(\mathcal{D} | \mathcal{M}, \theta_{(j)})^{-1})$$

where  $\mathcal{D}_i$  is the i-th data observation and with the likelihood given in (7) and  $\hat{m}_\alpha$  is the  $\alpha$ -trimmed mean to avoid outliers.

### 3 Application: Sales response in pharma marketing

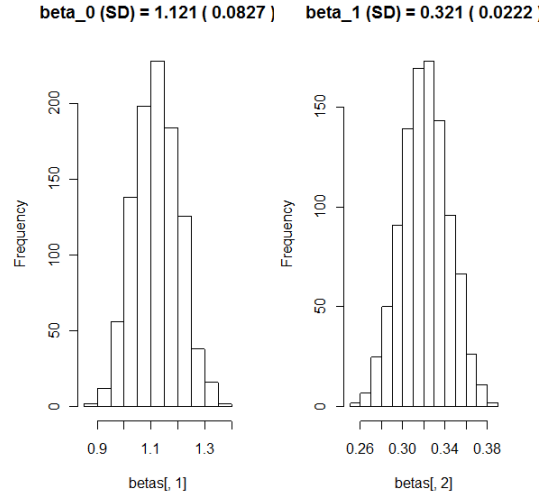
We apply the SRF(1) model with and without exogenous variables to regional pharma sales in Germany for the year 2008. The model explains the regional whole sales of the product M11 by the total sales  $U2008$  of the company across all regions. The posterior means (and SD) of the SRF(1) model for  $M11_{pc} = \ln(U2008M11/Pop)$  by  $U2008_{pc}$  (pc: per capita) is:

$$M11_{pc} = 1.121 + 0.321 U2008_{pc} \quad R^2 = 0.1502$$

(0.0827) (0.0222)  $\sigma_y = 1.2581$

with  $\lambda = 27.8(1.678)$  and  $\sigma_\lambda = 1.26(0.021)$ ; acceptance rate = 50.8%. The OLS fit is quite similar to the posterior mean and the F-statistic ( $df = 1, 1898$ ) = 335.4101 is highly significant (p-value = 0).

OLS	Estimate	Std.Err	t-value	Pr(> t )	R <sup>2</sup> =0.1502
Intercept	1.1237	0.0673	16.6984	0	
X	0.3222	0.0176	18.3142	0	

**Fig. 1.** The beta coefficients of visits and sales U2008 per capita (pc)

### 3.1 Exogenous variables: A SRF(1) model for regional pharma sales

We extend the SRF(1) model for the sales of the M11 product  $M11_{pc} = \ln(U2008M11/Pop)$  by including 2 more regressors: the number of visits for the related product  $V13_{pc} = \ln(visitsM13/Pop)$  and the purchasing power potential  $\ln(PPP)$ . The posterior mean (and SD) of the Bayesian regression estimation are:

$$M11_{pc} = 1.152 + 0.22 U2008_{pc} + 0.19 V13_{pc} + 0.19 \ln(PPP)$$

(0.2507)    (0.0251)                      (0.0239)                      (0.0359)

with  $\lambda = 22.478(2.268)$  and  $\sigma_\lambda = 1.225(0.020)$  (acceptance rate = 51%).

The OLS fit is (with  $R^2 = 0.1967$ ,  $F(df = 3, 1896) = 154.76$ , p-value=0) and  $\sigma_y = 1.2238$ :

$$M11_{pc} = 1.1510 + 0.2197 U2008_{pc} + 0.1893 V13_{pc} + 0.0962 \ln(PPP)$$

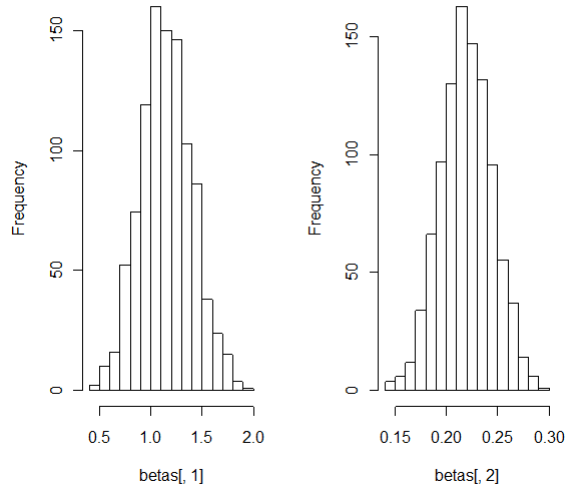
(0.1987)    (0.0197)                      (0.0191)                      (0.0283)

The posterior distribution of the parameters are displayed in the next 2 figures:

We see that the exogenous variables have improved the fit but the  $R^2$  is still low, despite the significant coefficients. This is because the underlying dispersion across regions is quite high, but positive slopes are present. Not all exogenous variables fit nicely with an SRF(1)-X model. The reasons for the high volatility of the SRF in dependence of covariates have to be further

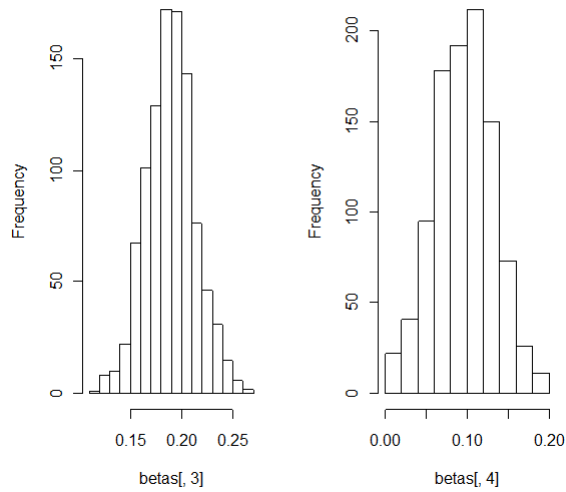
**Fig. 2.** SRF(1)X(2): Betas(intercept, U2008pc) on sales U2008M11pc

**beta\_0 (SD) = 1.152 ( 0.2507 )    beta\_1 (SD) = 0.22 ( 0.0251 )**



**Fig. 3.** SRF(1)X(2): Betas(visitsM13pc,PPP) on sales U2008M11pc

**beta\_2 (SD) = 0.19 ( 0.0239 )    beta\_3 (SD) = 0.096 ( 0.0359 )**



explored. If covariates can be found that reduces this volatility more likely accurate forecasts on the effectiveness of varying the marketing control variable  $x$  can be made. More on how regional heterogeneity can be modeled can be found in Polasek (2010) [5].

## 4 Summary

In this paper we have proposed the SRF(1) model family with 1 endogenous variable stochastic derivative constraints, as it was suggested by Kao et al. (2005) [3]. The MCMC estimation of the model is quite straightforward, despite the fact that the constraints lead to a simultaneous sales model. The class of SRF models with stochastic partial derivative (SPD) constraints is quite flexible and allows the inclusion of exogenous variables denoted by SRF(1)-X models. Future research will concentrate on the model selection procedure and on the problem, if model choice can help to decide statistically what variables in a SRF model are endogenous or exogenous. We have demonstrated this approach by a regional pharmaceutical sale models for Germany. In Polasek (2010) [5] it was shown that a whole SRF(p)-X(q) model family can be estimated by MCMC and the SRF approach can be extended to a spatial sales response model that takes the neighborhood structure of the observations or other regional variations into account (see Baier and Polasek, 2009) [1].

## References

1. Baier, D. and Polasek W. (2009), Marketing and Regional Sales: Evaluation of Expenditure Strategies by Spatial Sales Response Functions, *to appear in GFKL 2009 proceedings*.
2. Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm, *The American Statistician*, 49, 327-335.
3. Kao L.-J., Chiu C.-C., Gilbride T.J., Otter T., and Allenby G.M. Evaluating the Effectiveness of Marketing Expenditures. *Working Paper*, Ohio State University, Fisher College of Business, 2005.
4. Newton, M.A. and Raftery A.E. (1994). Approximate Bayesian Inference with the Weighted Likelihood Bootstrap (with discussion), *J. Royal Statistical Soc., Series B*, 56, 3-48.
5. Polasek, W. (2010). Sales Response Functions (SRF) with Stochastic Derivative Constraints. *Working Paper*, Institute of Advanced Studies, Wien, 2010.
6. Rossi P.E., Allenby G.M., and McCulloch R. *Bayesian Statistics and Marketing*. John Wiley and Sons, New York, 2005.