Online Companion for

"A Piecewise-Diffusion Model of New-Product Demands"

Shun-Chen Niu
School of Management
The University of Texas at Dallas
P. O. Box 830688
Richardson, TX 75083-0688

August 9, 2005

This Online Companion is organized as follows. In Section 1, we give the details of proofs and supporting arguments that are left out of the main paper. In Section 2, we supply the details of a SAS implementation of the PDM. Throughout, the prefix P is used to refer to equations in the main paper (e.g., equation "(P-15)").

1 Proofs and Supporting Arguments

In Section 1.1, we prove Theorems 1–3. In Sections 1.2 and 1.3, we provide supporting arguments for (P-15) and (P-48), respectively.

1.1 The Theorems

Most of the arguments below are adapted from Sections 11.1 and 11.2 of Ethier and Kurtz (1986). The prefix EK will be used when we refer to specific results there (e.g., equation "(EK-2.18)").

Denote by $\hat{A}_m$ an SBM with specification $(m; \alpha, \beta)$, where $m \geq 2$ (which we assume throughout to avoid trivialities). We begin by putting the family $\{\hat{A}_m\}$ on the same probability space. This is based on the simple fact that if $\hat{X}$ and X are exponential random variables with respective rates λ and 1, then $\hat{X}$ can be defined via X as $\hat{X} = (1/\lambda)X$. Let $\hat{X}_{mj}$, $1 \leq j \leq m$, be the jth inter-adoption time in $\hat{A}_m$; and let $Y \equiv \{Y(t),\, t \geq 0\}$ be a standard Poisson process at rate 1. Then, the above fact implies that we can couple $\hat{A}_m$ and Y by defining

$$\hat{X}_{mj} \equiv \sup\left\{x \geq 0 : \int_0^x \lambda_{m,j-1}\, ds \leq Y_j\right\}, \tag{1}$$

where Y_j, $j \geq 1$, denotes the jth interevent time of Y. It follows from (1) and (P-5) that all of the $\hat{A}_m$s can be constructed via a common Y, with every $\hat{A}_m$ satisfying

$$\hat{A}_m(t) = Y\left(\int_0^t m\, b_m\left(\frac{\hat{A}_m(s)}{m}\right) ds\right), \qquad t \geq 0\,, \tag{2}$$

where $b_m(\cdot)$ is given by (P-12).

For the proofs of the theorems, we need several elementary bounds that involve $b_m(\cdot)$ and $b(\cdot)$.

Lemma 1 *For* $E = [0,\, 1]$,

$$\sup_{x \in E} b_m(x) \leq M_1 \equiv \frac{(\alpha + 2\beta)^2}{4\beta}\,. \tag{3}$$

Proof It is easily shown that $b_m(x)$, $x \in E$, is strictly concave and that the function achieves its maximum at $x^* = \max[0,\, (\beta_m - \alpha)/(2\beta_m)]$. Hence,

$$\sup_{x \in E} b_m(x) = (1 - x^*)(\alpha + \beta_m x^*) \leq \frac{(\alpha + \beta_m)^2}{4\beta_m}\,.$$

Note that, for $m \geq 2$, we have $\beta < \beta_m \leq 2\beta$. Therefore,

$$\frac{(\alpha + \beta_m)^2}{4\beta_m} < \frac{(\alpha + 2\beta)^2}{4\beta}\,;$$

and this establishes (3). □

Lemma 2 *For* $x \in E$,

$$|b_m(x) - b(x)| \leq \frac{\beta}{4(m-1)}\,. \tag{4}$$

Proof From (P-12) and (P-16), we have

$$b_m(x) - b(x) = (1 - x)x\,\frac{\beta}{m-1}\,.$$

Since $0 \leq (1 - x)x \leq 1/4$ for $x \in E$, (4) follows. □

Lemma 3 *For* $x, y \in E$,

$$|b(x) - b(y)| \leq M_2\, |x - y|\,, \tag{5}$$

where $M_2 \equiv \alpha + \beta$*; moreover,*

$$|b_m(x) - b(y)| \leq \frac{\beta}{4(m-1)} + M_2\, |x - y|\,. \tag{6}$$

Proof We only need to consider the case with $x \neq y$. The mean-value theorem implies that

$$b(x) - b(y) = b'(z)(x - y) \tag{7}$$

for some z between x and y. Since $b'(x) = -\alpha + \beta - 2\beta x$, we have $|b'(x)| \leq \alpha + \beta$ for all $x \in E$; and this, together with (7), yields (5). By combining (4), (5), and the triangle inequality, (6) follows immediately. □

We now prove Theorems 1–3.

Proof of Theorem 1 It follows from (P-6) and (2) that the family $\{B_m\}$ can be put on the same probability space, with

$$B_m(t) = \frac{1}{m} Y\left(\int_0^t m\, b_m(B_m(s))\, ds\right), \qquad t \geq 0\,. \tag{8}$$

The proof of (P-14) is essentially the same as that of Theorem EK-2.1; the minor difference is that $b_m(\cdot)$ in (8) depends on m. The details are as follows.

Let $\tilde{Y}(t) = Y(t) - t$ and define $\tilde{Y} \equiv \{\tilde{Y}(t),\, t \geq 0\}$ (which is a martingale); then, (8) can be rewritten as

$$B_m(t) = \frac{1}{m} \tilde{Y}\left(\int_0^t m\, b_m(B_m(s))\, ds\right) + \int_0^t b_m(B_m(s))\, ds\,.$$

Next, observe that (P-3) is equivalent to

$$F_\infty(t) = \int_0^t b(F_\infty(s))\, ds, \qquad t \geq 0\,. \tag{9}$$

It follows that

$$B_m(t) - F_\infty(t) = \frac{1}{m} \tilde{Y}\left(\int_0^t m\, b_m(B_m(s))\, ds\right) + \int_0^t [b_m(B_m(s)) - b(F_\infty(s))]\, ds\,. \tag{10}$$

Now, define

$$\Delta_m(t) \equiv \sup_{0 \leq u \leq t} \left| \frac{1}{m} \tilde{Y}\left(\int_0^u m\, b_m(B_m(s))\, ds\right) \right|$$

and

$$\hat{\Delta}_m(t) \equiv \frac{\beta}{4(m-1)}\, t\,.$$

Then,

$$\left| \frac{1}{m} \tilde{Y}\left(\int_0^t m\, b_m(B_m(s))\, ds\right) \right| \leq \Delta_m(t)\,;$$

and furthermore, (6) implies that

$$\left| \int_0^t [b_m(B_m(s)) - b(F_\infty(s))]\, ds \right| \leq \hat{\Delta}_m(t) + \int_0^t M_2\, |B_m(s) - F_\infty(s)|\, ds\,.$$

Hence, from (10), we have

$$|B_m(t) - F_\infty(t)| \le [\Delta_m(t) + \hat{\Delta}_m(t)] + \int_0^t M_2 \, |B_m(s) - F_\infty(s)| \, ds \, .$$

Gronwall's inequality (Theorem EK-5.1, p. 498) then implies that

$$\sup_{0\le u\le t} |B_m(u) - F_\infty(u)| \le \sup_{0\le u\le t} [\Delta_m(u) + \hat{\Delta}_m(u)] e^{M_2 u} = [\Delta_m(t) + \hat{\Delta}_m(t)] e^{M_2 t}. \qquad (11)$$

Finally, from Lemma 1, we have

$$\Delta_m(t) \le \sup_{0\le u\le t} \left| \frac{1}{m} \tilde{Y}\left(m\, M_1 u\right) \right| = \sup_{0\le u\le M_1 t} \left| \frac{1}{m} \tilde{Y}\left(mu\right) \right| ,$$

implying (see (EK-2.5)) that $\lim_{m\to\infty} \Delta_m(t) = 0$ almost surely. Since we also have $\lim_{m\to\infty} \hat{\Delta}_m(t) = 0$, the proof is complete. □

Proof of Theorem 2 The proof of (P-15) and (P-17) is similar to that of Theorem EK-2.3; again, the minor difference is that $b_m(\cdot)$ in (8) depends on m. The details are as follows.

Multiplying both sides of (10) by $\sqrt{m}$ yields

$$V_m(t) = \frac{1}{\sqrt{m}} \tilde{Y}\left(\int_0^t m \, b_m(B_m(s)) \, ds \right) + \int_0^t \sqrt{m} [b_m(B_m(s)) - b(F_\infty(s))] \, ds \, . \qquad (12)$$

Define

$$U_m(t) \equiv \frac{1}{\sqrt{m}} \tilde{Y}\left(m \int_0^t b_m(B_m(s)) \, ds \right) ,$$

$$\Gamma_m(t) \equiv \sqrt{m} \int_0^t \left\{ [b(B_m(s)) - b(F_\infty(s))] - b'(F_\infty(s))[B_m(s) - F_\infty(s)] \right\} ds \, ,$$

and

$$\hat{\Gamma}_m(t) \equiv \sqrt{m} \int_0^t [b_m(B_m(s)) - b(B_m(s))] \, ds \, .$$

Then, (12) can be rewritten as

$$V_m(t) = \hat{U}_m(t) + \int_0^t b'(F_\infty(s)) V_m(s) \, ds \, , \qquad (13)$$

where $\hat{U}_m(t) \equiv U_m(t) + \Gamma_m(t) + \hat{\Gamma}_m(t)$. Observe further that (13) is equivalent to

$$V_m(t) = \int_0^t \exp\left(\int_s^t b'(F_\infty(y))\, dy\right) d\hat{U}_m(s) \tag{14}$$

(see (EK-2.20); the "differential" of (14) is identical to that of (13)), and hence, upon integration by parts, to

$$V_m(t) = \hat{U}_m(t) + \int_0^t \exp\left(\int_s^t b'(F_\infty(y))\, dy\right) b'(F_\infty(s))\, \hat{U}_m(s)\, ds\,. \tag{15}$$

Now, since $\tilde{Y}$ has stationary and independent increments, we have, from the ordinary central limit theorem,

$$\frac{1}{\sqrt{m}}\tilde{Y}\left(m\int_{t_1}^{t_2} b(F_\infty(s))\, ds\right) \Rightarrow U(t_2) - U(t_1)\,, \qquad 0 \le t_1 \le t_2 < \infty\,,$$

where

$$U(t) \equiv W\left(\int_0^t b(F_\infty(s))\, ds\right);$$

hence, (6) and Theorem 1 imply (see, e.g., Billingsley 1999, Theorem 3.1, p. 27) that $\{U_m(t),\, t \ge 0\} \Rightarrow \{U(t),\, t \ge 0\}$ in $D[0,\infty)$. The relations (7) and $|b'(x) - b'(y)| = 2\beta|x-y|$ for $x, y \in E$, together with (11), imply that $\lim_{m\to\infty} \sup_{0\le u\le t} |\Gamma_m(u)| = 0$ almost surely (see (EK-2.25)). Furthermore, Lemma 2 implies that $|\hat{\Gamma}_m(t)| \le \sqrt{m}\,\hat{\Delta}_m(t)$ and hence $\lim_{m\to\infty} \sup_{0\le u\le t} |\hat{\Gamma}_m(u)| = 0$ almost surely. It then follows from (15) (by the continuous-mapping theorem) that $V_m \Rightarrow V$, where V satisfies

$$V(t) = U(t) + \int_0^t \exp\left(\int_s^t b'(F_\infty(y))\, dy\right) b'(F_\infty(s))\, U(s)\, ds\,. \tag{16}$$

Since (16) is equivalent (similar to (14), but using Itô's formula) to

$$V(t) = \int_0^t \exp\left(\int_s^t b'(F_\infty(y))\, dy\right) dU(s)\,,$$

we have established (P-17). Finally, since (16) is also equivalent (similar to (13)) to

$$V(t) = U(t) + \int_0^t b'(F_\infty(s))V(s)\, ds\,,$$

which is the integral form of (P-15), the proof of the theorem is complete. □

Proof of Theorem 3 From (P-17), we have $E[V(t)] = 0$ and, by Itô's isometry,

$$\psi(t) = E\left\{[V(t)]^2\right\} = \int_0^t \left[\exp\left(\int_s^t b'(F_\infty(y))\,dy\right)\sqrt{b(F_\infty(s))}\right]^2 ds\,. \tag{17}$$

Using the relations

$$\int_0^t [1 - F_\infty(y)]\,dy = \frac{1}{\beta}\,\ln\left[\frac{\alpha+\beta}{\alpha + \beta\,e^{-(\alpha+\beta)t}}\right]$$

and

$$b(F_\infty(t)) = f_\infty(t) = \frac{[(\alpha+\beta)^2/\alpha]\,e^{-(\alpha+\beta)t}}{[1 + (\beta/\alpha)\,e^{-(\alpha+\beta)t}]^2}\,,$$

(17) evaluates straightforwardly to (P-18) and (P-19) after considerable algebra.

To prove (P-20), we apply a dominated-convergence argument. The easily-established inequality $(x+y)^2 \le 2(x^2+y^2)$ and (12) imply that

$$|V_m(t)|^2 \le \frac{2}{m}\left|\tilde{Y}\left(\int_0^t m\,b_m(B_m(s))\,ds\right)\right|^2 + 2m\left|\int_0^t [b_m(B_m(s)) - b(F_\infty(s))]\,ds\right|^2. \tag{18}$$

From the Cauchy-Schwarz inequality and (6), we have

$$\begin{aligned} & \left|\int_0^t [b_m(B_m(s)) - b(F_\infty(s))]\,ds\right|^2 \\ \le\ & \int_0^t 1^2\,ds \int_0^t |b_m(B_m(s)) - b(F_\infty(s))|^2\,ds \\ \le\ & 2\left[\frac{\beta}{4(m-1)}\right]^2 t^2 + \int_0^t 2\,T(M_2)^2\,|B_m(s)) - F_\infty(s)|^2\,ds\,, \end{aligned} \tag{19}$$

where T is any constant satisfying $t \le T$. Define

$$\Lambda_m(t) \equiv \frac{2}{m}\left|\tilde{Y}\left(\int_0^t m\,b_m(B_m(s))\,ds\right)\right|^2$$

and

$$\hat{\Lambda}_m(t) \equiv \frac{m\beta^2}{4(m-1)^2}\,t^2.$$

Then, it follows from (18) and (19) that

$$|V_m(t)|^2 \le [\Lambda_m(t) + \hat{\Lambda}_m(t)] + \int_0^t 4\,T(M_2)^2\,|V_m(s)|^2\,ds\,,$$

and hence, from Gronwall's inequality,

$$|V_m(t)|^2 \le [\Lambda_m(t) + \hat{\Lambda}_m(t)] \exp\left[4\,T(M_2)^2 t\right]. \tag{20}$$

Observe that $\Lambda_m(t) = 2|U_m(t)|^2 \Rightarrow 2|U(t)|^2$ and that $\lim_{m\to\infty} \hat{\Lambda}_m(t) = 0$ deterministically. Furthermore, we have

$$E[\Lambda_m(t)] = 2\,E\left[\int_0^t b_m(B_m(s))\,ds\right];$$

hence, Lemma 1, (6), and Theorem 1 together imply that

$$\lim_{m\to\infty} E[\Lambda_m(t)] = 2\int_0^t b(F_\infty(s))\,ds = 2\,E\{[U(t)]^2\}.$$

Since $|V_m(t)|^2 \Rightarrow |V(t)|^2$ (Theorem 2), (20) and the weak-convergence version of the dominated convergence theorem (Theorem EK-1.2, p. 492) now yield that $\lim_{m\to\infty} E\{[V_m(t)]^2\} = E\{[V(t)]^2\}$. Finally, since a similar argument applied to $|V_m(t)|$ also yields that $\lim_{m\to\infty} E[V_m(t)] = 0 = E[V(t)]$, the proof of (P-20) is complete. □

1.2 Intuitive Explanation of (P-15)

Suppose m is large and, with t fixed, consider the increment $V_m(t+h) - V_m(t)$, where $h > 0$. Let $Y_m(h) \equiv A_m(t+h) - A_m(t)$; then, for given $B_m(t)$, the stochastic component of this increment is

$$\sqrt{m}\,B_m(t+h) - \sqrt{m}\,B_m(t) = \frac{Y_m(h)}{\sqrt{m}}. \tag{21}$$

When h is small, $Y_m(h)$ is approximately distributed as a Poisson random variable with mean and variance $mb_m(B_m(t))h$. Furthermore, since this approximating Poisson random variable can be regarded as the sum of m i.i.d. Poisson random variables (each with mean $b_m(B_m(t))h$), $Y_m(h)/\sqrt{m}$ is approximately normal. Now, suppose $B_m(t) - F_\infty(t)$ is small, which applies almost surely when m is large (Theorem 1). Then, from Taylor's formula, we have

$$b_m(B_m(t)) \approx b_m(F_\infty(t)) + b'_m(F_\infty(t))[B_m(t) - F_\infty(t)];$$

and this suggests that

$$
\begin{aligned}
E\left[\frac{Y_m(h)}{\sqrt{m}}\right] &\approx [\sqrt{m}\, b_m(F_\infty(t)) + b_m'(F_\infty(t))V_m(t)]h \\
&\approx [\sqrt{m}\, b(F_\infty(t)) + b'(F_\infty(t))V_m(t)]h \qquad (22)
\end{aligned}
$$

and, similarly,

$$
\mathrm{Var}\left[\frac{Y_m(h)}{\sqrt{m}}\right] \approx \{b(F_\infty(t)) + b'(F_\infty(t))[B_m(t) - F_\infty(t)]\}h \approx b(F_\infty(t))h\,.
$$

It follows that in the neighborhood of the origin, the process $\{\sqrt{m}\,[B_m(t+u) - B_m(t)],\ u \geq 0\}$ would evolve like a Wiener process with drift $\sqrt{m}\, b(F_\infty(t)) + b'(F_\infty(t))V_m(t)$ and variance rate $b(F_\infty(t))$. Finally, note that the deterministic component $\sqrt{m}\, F_\infty(t)$ of $V_m(t)$ grows at rate $\sqrt{m}\, b(F_\infty(t))$ (see (9)). Since this growth offsets the first term in (22), we see that the "net" drift of V_m at time t is $b'(F_\infty(t))V_m(t)$. Letting m go to infinity now leads to (P-15).

1.3 Informal Support for (P-48)

To avoid duplication, we will only consider the actual-history SBM; moreover, we will ignore the D_is so as to focus on the essential.

With $a_0 = 0$, a_j replaced by A_j, and $\pi_i = 1$ in (P-26)–(P-29), we have, for all $1 \leq i \leq n$,

$$
\alpha_i = \alpha + \frac{\beta}{m-1} \sum_{j=1}^{i-1} A_j
$$

and

$$
\beta_i = \frac{m - \sum_{j=1}^{i-1} A_j - 1}{m-1}\, \beta\,.
$$

It follows that

$$
\alpha_i + \beta_i = \alpha + \beta \qquad (23)
$$

and

$$
\frac{\beta_i}{\alpha_i} = \frac{\beta}{\alpha}\, \frac{1 - \sum_{j=1}^{i-1} A_j/(m-1)}{1 + (\beta/\alpha)\sum_{j=1}^{i-1} A_j/(m-1)}\,. \qquad (24)
$$

Now, with $t = t_{i-1}$ in (P-6), we have $B_m(t_{i-1}) = \sum_{j=1}^{i-1} A_j/m$; and hence Theorem 1 suggests that (almost surely)

$$\sum_{j=1}^{i-1} A_j \approx mF_\infty(t_{i-1}) \tag{25}$$

for sufficiently large m. With (25) in (24), we then have

$$\frac{\beta_i}{\alpha_i} \approx \frac{\beta}{\alpha}\,\frac{1-F_\infty(t_{i-1})}{1+(\beta/\alpha)\,F_\infty(t_{i-1})} = \frac{\beta}{\alpha}\,e^{-(\alpha+\beta)t_{i-1}}, \tag{26}$$

where the equality follows from substituting (P-8). Next, from (P-37) (or (P-11)), we have the conditional approximation:

$$E\left[A_i \mid \sum_{j=1}^{i-1} A_j\right] \approx \left(m - \sum_{j=1}^{i-1} A_j\right) \frac{1-e^{-(\alpha_i+\beta_i)}}{1+(\beta_i/\alpha_i)\,e^{-(\alpha_i+\beta_i)}}\,;$$

and, upon substitution of (25), (23), and (26), this yields

$$\begin{aligned} E\left[A_i \mid \sum_{j=1}^{i-1} A_j\right] &\approx m\left[1-F_\infty(t_{i-1})\right] \frac{1-e^{-(\alpha+\beta)}}{1+(\beta/\alpha)\,e^{-(\alpha+\beta)(t_{i-1}+1)}} \\ &= m\left[F_\infty(t_i)-F_\infty(t_{i-1})\right], \end{aligned} \tag{27}$$

where the equality follows from the easily-established relation

$$\frac{F_\infty(t+h)-F_\infty(t)}{1-F_\infty(t)} = \frac{1-e^{-(\alpha+\beta)h}}{1+(\beta/\alpha)\,e^{-(\alpha+\beta)(t+h)}}$$

(for any $t \geq 0$ and $h \geq 0$). Finally, since the right-hand side of (27) is independent of $\sum_{j=1}^{i-1} A_j$, we obtain

$$E[A_i] = E\left\{E\left[A_i \mid \sum_{j=1}^{i-1} A_j\right]\right\} \approx m\left[F_\infty(t_i)-F_\infty(t_{i-1})\right], \tag{28}$$

which is (P-48).

2 SAS Implementation of PDM

In this section, we provide the data and program files that are used to produce parameter estimates for the Room Air Conditioners (RACs). Estimation results summarized in the last two entries in Table 1 of the paper, for the actual-history and the expected-history versions of PDM-F, are obtained by running these programs.

The data file and the programs are listed in Section 2.1. Explanations of the programs are provided in Section 2.2.

A brief summary of the PDM is given in Section 4.3.3 of the paper. Before looking over the details of the programs, the reader should review that section.

2.1 The Data File and Programs

2.1.1 Data File for RACs

The suggested name for the data file below is "RACs_Data.txt". Before running the programs, data in this file need to be imported into SAS. Only the last three columns are used in the programs; their headings should not be changed, as they are the variable names used in the programs. Sales is in thousands of units; price is in dollars; and advertising spending is in millions of dollars.

```
Year   Time   Sale   Price   Adv
1949   1      96     410     0
1950   2      195    370     0.615
1951   3      238    365     1.198
1952   4      365    388     3.196
1953   5      1045   335     5.34
1954   6      1230   341     14.372
1955   7      1270   320     9.391
1956   8      1828   293     13.61
1957   9      1586   310     16.785
1958   10     1673   279     9.238
1959   11     1660   269     5.863
1960   12     1580   275     3.923
1961   13     1500   259     1.493
```

2.1.2 Program for the Expected-History PDM-F

The suggested file name for the program below is "RACs_PDM-F_EH.sas". The first couple of lines set the title, call the Nonlinear-Programming Procedure (nlp), invoke the input data file, and set running conditions. The remaining lines have been numbered, which are referenced in Section 2.2 below. These line numbers should be removed in the program file.

```
       title '--- Expected-History PDM-F for RACs ---'  ;

proc nlp data=RACs_Data maxit=500 outest=mle tech=quanew out=vars
              MSING=1e-36 GCONV=1e-12 FDH=central FDInt=obj
              covariance=1 pcov phes  ;

01     max loglik  ;
02
03     parms beta=20, pi=0.005, delta=40, eta=6.5,
04           gamma_p=0.009, gamma_b=0.38, pi_m=0.04  ;
05
06     bounds 0 < beta <= 150  ;
07     bounds 0 < pi <= 1  ;
08     bounds 0 < delta <= 200  ;
09     bounds 0 < eta <= 10  ;
10     bounds 0 < gamma_p <= 10  ;
11     bounds 0 < gamma_b <= 10  ;
12     bounds 0 < pi_m <= 1  ;
13
14     m = 53291  ;
15     alpha_0 = 0  ;
16     S_0 = 744  ;
17     Pr_Ref = 410  ;
18
19     iad_i = exp(-gamma_p*Adv)  ;
20     arg_i = (-log((1-pi/pi_m)*iad_i))*((Price/Pr_Ref)**(-eta))  ;
```

```
pi_i = pi_m*(1-exp(-arg_i))  ;

if _obs_ = 1 then
   do  ;
      part_i = max(m-S_0,0)*pi_i  ;
      m_i = part_i+S_0  ;
      CAdv_i = Adv  ;
      alpha_i = alpha_0+S_0*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
      beta_i = max(part_i-1,0)*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
      sum_i = alpha_i+beta_i  ;
      rho_i = beta_i/alpha_i  ;
      var_i = (1+rho_i)*exp(-2*sum_i)*(exp(sum_i)-1
              +2*rho_i*sum_i+(rho_i**2)
              *(1-exp(-sum_i)))/(1+rho_i*exp(-sum_i))**4  ;
      F_i = (1-exp(-sum_i))/(1+rho_i*exp(-sum_i))  ;
      ES_i = part_i*F_i  ;
      CES_i = ES_i+S_0  ;
   end  ;

if _obs_ > 1 then
   do  ;
      part_i = max(m-CES_i,0)*pi_i  ;
      m_i = part_i+CES_i  ;
      CAdv_i = Adv+CAdv_i  ;
      alpha_i = alpha_0+CES_i*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
      beta_i = max(part_i-1,0)*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
      sum_i = alpha_i+beta_i  ;
      rho_i = beta_i/alpha_i  ;
      var_i = (1+rho_i)*exp(-2*sum_i)*(exp(sum_i)-1
              +2*rho_i*sum_i+(rho_i**2)
              *(1-exp(-sum_i)))/(1+rho_i*exp(-sum_i))**4  ;
      F_i = (1-exp(-sum_i))/(1+rho_i*exp(-sum_i))  ;
```

```
        ES_i = part_i*F_i  ;
        CES_i = ES_i+CES_i  ;
     end  ;

retain CAdv_i  ;
retain CES_i  ;

sigma_i = (part_i*var_i+delta**2)**(1/2)  ;
loglik = -log(sigma_i)-(1/2)*((Sale-ES_i)/sigma_i)**2  ;

run  ;
```

2.1.3 Program for the Actual-History PDM-F

The suggested file name for the program below is "RACs_PDM-F_AH.sas". Note again that line numbers should be removed in the program file.

```
       title '--- Actual-History PDM-F for RACs ---'  ;

proc nlp data=RACs_Data maxit=500 outest=mle tech=quanew out=vars
              MSING=1e-36 GCONV=1e-12 FDH=central FDInt=obj
              covariance=1 pcov phes  ;

max loglik  ;

parms beta=20, pi=0.005, delta=40, eta=6.5,
      gamma_p=0.009, gamma_b=0.38, pi_m=0.04  ;

bounds 0 < beta <= 150  ;
bounds 0 < pi <= 1  ;
bounds 0 < delta <= 200  ;
bounds 0 < eta <= 10  ;
```

```
bounds 0 < gamma_p <= 10  ;
bounds 0 < gamma_b <= 10  ;
bounds 0 < pi_m <= 1  ;

m = 53291  ;
alpha_0 = 0  ;
S_0 = 744  ;
Pr_Ref = 410  ;

iad_i = exp(-gamma_p*Adv)  ;
arg_i = (-log((1-pi/pi_m)*iad_i))*((Price/Pr_Ref)**(-eta))  ;
pi_i = pi_m*(1-exp(-arg_i))  ;

if _obs_ = 1 then
   do  ;
      part_i = max(m-S_0,0)*pi_i  ;
      m_i = part_i+S_0  ;
      CAdv_i = Adv  ;
      alpha_i = alpha_0+S_0*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
      beta_i = max(part_i-1,0)*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
      sum_i = alpha_i+beta_i  ;
      rho_i = beta_i/alpha_i  ;
      var_i = (1+rho_i)*exp(-2*sum_i)*(exp(sum_i)-1
              +2*rho_i*sum_i+(rho_i**2)
              *(1-exp(-sum_i)))/(1+rho_i*exp(-sum_i))**4  ;
      F_i = (1-exp(-sum_i))/(1+rho_i*exp(-sum_i))  ;
      ES_i = part_i*F_i  ;
      CAS_i = Sale+S_0  ;
   end  ;

if _obs_ > 1 then
   do  ;
```

```
42          part_i = max(m-CAS_i,0)*pi_i  ;
43          m_i = part_i+CAS_i  ;
44          CAdv_i = Adv+CAdv_i  ;
45          alpha_i = alpha_0+CAS_i*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
46          beta_i = max(part_i-1,0)*beta*(1+gamma_b*CAdv_i)/(m-1)  ;
47          sum_i = alpha_i+beta_i  ;
48          rho_i = beta_i/alpha_i  ;
49          var_i = (1+rho_i)*exp(-2*sum_i)*(exp(sum_i)-1
50                  +2*rho_i*sum_i+(rho_i**2)
51                  *(1-exp(-sum_i)))/(1+rho_i*exp(-sum_i))**4  ;
52          F_i = (1-exp(-sum_i))/(1+rho_i*exp(-sum_i))  ;
53          ES_i = part_i*F_i  ;
54          CAS_i = Sale+CAS_i  ;
55       end  ;
56
57    retain CAdv_i  ;
58    retain CAS_i  ;
59
60    sigma_i = (part_i*var_i+delta**2)**(1/2)  ;
61    loglik = -log(sigma_i)-(1/2)*((Sale-ES_i)/sigma_i)**2  ;
62
63    run  ;
```

2.2 Explanation of the Programs

2.2.1 Expected-History PDM-F

Line 01: Define the objective function. The variable loglik is to be maximized.

Lines 03–04: Specify starting values of the parameters β, π, δ, η, γ_p, γ_b, and π_m. Note that α is absent; this is because a pilot run with α as a parameter indicated that it is not significant. The specified starting values are based on the output of a separate Excel program, using the genetic-algorithm-based Evolutionary Solver add-in (Frontline Systems). Without a separate Excel program, one can guess by analogy. Good starting values reduce the necessary number of iterations, the maximum of which is set above at 500 (in this case). Using the logic provided in the SAS program, it is not difficult to construct a corresponding Excel program. One advantage of Excel is that it provides convenient graphing tools. This is very helpful when we wish to examine the host of output characteristics (4th paragraph in Section 6 of the paper) of the PDM in detail.

Lines 06–12: Set reasonable ranges (or search space) for the parameters. Except for π and π_m, these should be changed if the parameter estimates hit the specified bounds.

Lines 14–17: Specify input data m, α, a_0, and p_{ref}. For α, the name alpha_0 is used; and for a_0, we use S_0. The value of α is set to 0 because it is not significant. The setting 744 for S_0 comes from a pilot run; and this is discussed in Section 5 of the paper. The reference price p_{ref} is set to p_1, the average price in the first period.

Lines 19–21: Compute the participation fraction π_i for the current period, using the parameter values in the current execution cycle. The formula is given in equation (P-57). The current price and advertising spending are supplied by the input data file.

Lines 23–38: In an execution cycle (or iteration), SAS processes the input data file one row, and hence one period, at a time, for a given set of parameter values. This section of the program applies only for the first period, i.e., only if _obs_, the current row index in the data file, equals 1. Having a separate section for the first period allows us to perform initialization tasks easily. The line-by-line explanation is as follows.

Line 25: Compute $(m - a_0)\pi_1$, the active number of adopters in period 1. The extra "max" operation is intended to guard against potential anomaly in some parameter scenarios; similar uses of this also appear elsewhere in the program. It is redundant unless there is a parameter problem.

Line 26: Compute m_1, which is the current level of market penetration, or market ceiling; see equation (P-49). The m_is, which offer valuable information, will appear in the final SAS output.

Line 27: Set up a variable for cumulative advertising spendings and initialize it.

Line 28: Compute $\hat{\alpha}_1$; see equation (P-26). Note that for period 1, $\hat{\alpha}_1 = \alpha_1$ and $\hat{\beta}_1 = \beta_1$. Moreover, β in (P-26) is modified by the effect of cumulative advertising spendings; see the scaling in (P-60).

Line 29: Compute $\hat{\beta}_1$; see equation (P-27).

Lines 30 and 31: Compute $\hat{\alpha}_1 + \hat{\beta}_1$ and $\hat{\beta}_1/\hat{\alpha}_1$. These are intermediate steps in computation.

Lines 32–34: Compute $\psi_1(t)$, with $t = 1$; see equations (P-18) and (P-19). Recall that we assume $t_i - t_{i-1} = 1$ for all i. The formula used here is an equivalent version that combines the two terms in (P-18) into a single expression, given explicitly by

$$\psi(t) = \frac{(1+\beta/\alpha)\, e^{-2(\alpha+\beta)t}}{[1+(\beta/\alpha)\, e^{-(\alpha+\beta)t}]^4} \left\{ e^{(\alpha+\beta)t} - 1 + 2\left(\frac{\beta}{\alpha}\right)(\alpha+\beta)t + \left(\frac{\beta}{\alpha}\right)^2 (1 - e^{-(\alpha+\beta)t}) \right\}.$$

This formula is more compact.

Line 35: Compute the last term in (P-44).

Line 36: Compute $\hat{s}_1$, the expected sales in the current period, using (P-44).

Line 37: Set up the variable CES_i to accumulate expected sales and initialize it.

Lines 40–55: This is the core section of the program; it applies for periods 2 through n. Since the logic is similar to that for period 1, we will only note the differences.

Lines 42 and 43: Replace a_0 by cumulative expected sales.

Line 44: Update cumulative advertising spendings.

Line 45: Compute $\hat{\alpha}_i$ using cumulative expected sales; see equation (P-31).

Line 46: Compute $\hat{\beta}_i$, using (P-32). Note again that β in both (P-31) and (P-32) has been modified according to (P-60).

Line 54: Update cumulative expected sales.

Lines 57 and 58: Retain the current values of CAdv_i and CES_i, which are needed for computations in the next period. Without these commands, the contents of CAdv_i and CES_i will be cleared as SAS moves on to the next period.

Line 60: Compute $\hat{\sigma}_i$, the standard deviation of the current S_i; see equation (P-45).

Line 61: Compute the contribution to loglik, the log-likelihood function, from the current period; see equation (P-47). SAS will accumulate loglik automatically as it moves through the periods. The current execution cycle ends when SAS reaches the end of the input data file. At that point, the specified NLP algorithm (Quasi-Newton in this case) will pick a new set of parameters for SAS to begin a new execution cycle. This continues until the specified termination criteria are met.

2.2.2 Actual-History PDM-F

The program for the actual-history version of PDM-F is essentially the same as that for the expected-history version. The only revisions appear in lines 37, 42, 43, 45, 54, and 58, where the variable CAS_i, for cumulative actual sales, replaces CES_i in the previous program. Note that with these revisions, equations (P-31), (P-32), (P-44), and (P-45) automatically become (P-28), (P-29), (P-39), and (P-40), respectively.

2.2.3 Other Versions or Products

The above programs can be easily modified to work with other versions of the PDM, as well as the SBM; see Section 4.3.3 in the paper. To run the programs for other products, simply replace the data file.