Electronic Companion to
"Strategic New Product Media Planning under Emergent Channel Substitution and Synergy"

Vahideh Sadat Abedi, Oded Berman, Fred Feinberg, Dmitry Krass

Appendix A: Summary of Notation

Notation used in main model setup and analysis:

Notation	Definition
p	Product innovation rate
q	Product imitation rate
m	Potential market size
T	Media planning horizon
$F(t)$	Fraction of cumulative adoptions (market share) by time t
x_0	Fraction of adoptions (market share) at time 0
P	Product per unit price
θ	Profit discount rate
R	Set of marketing channels
$u_r(t)$	The overall influence of marketing effort in channel r on demand at time t (can incorporate impact of current and past advertising expenditures)
E_r	Total per capita investment in channel r over the media planning horizon, adding up to total investment of mE_r in channel r
$\Phi_r(E_r)$	Cumulative marketing effort in channel r on demand over the media planning horizon ($\Phi_r = \int_0^T u_r(t)dt$) given a total budget of mE_r to be spend in channel r. Also see equation 4.
K_r	Number of time-blocks that channel r spending can be changed over the media planning horizon $[0, T]$
τ_r	Length of each time block over which spending in channel r cannot be updated; $\tau_r = T/K_r$
a_{rk}	Total advertising spend in channel r over time block k
b_r	Maximum investing in channel r over each time block
$\phi_r^i(a)$	The measure of how much the baseline purchase rate of a non-adopter improves i blocks ahead when an investment of a is made over a given time block.
s^r	Investment in each channel r can impact current sales as well as the sales up to s^r blocks into the future.
γ_k	If the media planning horizon starts sometime after launch leading to some initial adoptions ($x_0 > 0$) and initial advertising spending, γ_k measures the overall effect of advertising expenditures prior to time 0 on demand in block k.
Φ_0	Constant capturing the overall effect of any potential advertising expenditures prior to

	time 0; $\Phi_0 = \tau_r \sum_{k=1}^{\min\{k,s_r\}} \gamma_k$
$F[t, \{a_r \mid E_r\}]$	[Alternative to $F(t)$] The market share captured by time t when the piecewise-constant advertising policy of $a_r(t)$ for $t \in [0, T]$ and $r \in R$ is used that sums up to E_r per capita for channel r over the planning horizon
$\Pi^D\{a_r \mid E_r\}$	The discounted profit (from DMP-DISC problem) when the advertising policy $\{a_r \mid E_r\}$ (as in definition of $F[t, \{a_r \mid E_r\}]$) is employed.
$\Pi^U\{a_r \mid E_r\}$	The undiscounted profit (from DMP problem) when the advertising policy $\{a_r \mid E_r\}$ (as in definition of $F[t, \{a_r \mid E_r\}]$) is employed.
C^0	Cumulative initial-investment marketing effort in all channels; $C^0 = \sum_{r \in R} \Phi_r(E_r^0)$;
C_{-s}^0	Cumulative initial-investment marketing effort in all channels except channel $s \in R$; $C_{-s}^0 = \sum_{r \in R - \{s\}} \Phi_r(E_r^0)$
E_r^0	Minimum feasible or allocated per capita investment in channel r, i.e., $E_r \geq E_r^0$
$E_s^*(C_{-s}^0)$	The optimal investment in channel s given cumulative effort in all other channels C_{-s}^0; this can be interpreted as the optimal "response" for channel s to C_{-s}^0
E_s^U	Upper bound on optimal investment in channel s
G_s^1 , G_s^2	Lower and upper market penetration thresholds – see equations (6) and (7)
R^A	A subset of channels that are "active" so that investments beyond the initial allocation of E^0 in them would be warranted
$R^{H/M}$	Subset of channels from R^A that have high leverage and medium momentum
$R^{H/L}$	Subset of channels from R^A that have high leverage and low momentum
$R^{M/M}$	Subset of channels from R^A that have medium leverage and medium momentum
$R^{M/L}$	Subset of channels from R^A that have medium leverage and low momentum
C_{-s}^{trans}	Threshold on C_{-s} when positive investment in channel s becomes feasible; $C_{-s}^{trans} = \inf\{C_{-s} \geq C_{-s}^0 \mid E_s^*(C_{-s}) > E_s^0\}$
C_{-s}^{LM}	Threshold on C_{-s} when channel s transitions to medium momentum from low momentum; $C_{-s}^{LM} = \inf\{C_{-s} \geq C_{-s}^0 \mid G[E_s^*(C_{-s}), C_{-s}] \geq G_s^1\}$
C_{-s}^{peak}	Threshold on C_{-s} when optimal channel s investment peaks; $C_{-s}^{peak} = \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^U)) \geq \frac{1}{2}(1 - \frac{p}{q})\}$
$C_{-s}^{\max}$	Threshold on C_{-s} when channel s transitions to high momentum and no further investment in that channel is warranted; $C_{-s}^{\max} = \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^0)) \geq G_s^2\}$

Additional notation used in empirical estimation and analysis of Camera case:

Notation	Definition
$S_r(k)$	*[Exponentially decaying impact of past advertising expenses:]* Stock of advertising goodwill in channel r by time block k: $u_r(t) = S_r(k)$ for $t \in [(k-1)\tau_r, k\tau_r)$
α_r	Channel r base effectiveness level
δ_r	Remembering rate of advertising goodwill after each time block
ρ_r	Advertising effectiveness exponent in order to impose diminishing return to advertising
$\Phi_r(E_r)$, β_r	Cumulative marketing effort in channel r on demand over the media planning horizon. Using exponentially decaying impact of past advertising, Section 3.4 derives this function as $\Phi_r(E_r) = \beta_r E_r^{\rho_r} + \Phi_r^0$.

Appendix B: Theorems & Proofs

Theorem 1 (<u>Section 3.2</u>): Assume that $\mathrm{E_r}$ dollars per target customer are pre-allocated to channel r for each $\mathrm{r} \in \mathrm{R}$. Then the optimal temporal plan of investing $\mathrm{E_r}$ in channel r in the DMP problem is obtained by maximizing the cumulative marketing effort in that channel over the horizon, which is independent of the plan for all the other channels and the word-of-mouth process. Therefore, the Tactical Planning Problem (TPP) can be stated as follows:

$$\Phi_r(E_r) = \max_{\substack{0\le a_{rk}\le b_r \\ k=1,\dots,K_r}} \tau_r \sum_{k=1}^{K_r} \sum_{i=0}^{\min\{k-1,s_r\}} \phi_r^i(a_{r,k-i}) + \Phi_{\mathrm{r}}^0 \qquad \textit{Subject to} \sum_{k=0}^{K_r} a_{rk} \le mE_r. \qquad \textit{(TPP)}$$

The resulting optimal cumulative effort, $\Phi_{\mathrm{r}}(\mathrm{E_r})$, is concave and non-decreasing in $\mathrm{E_r}$. In addition, the optimal plan of investing in channel r would be non-increasing over time, i.e. $\mathrm{a}_{\mathrm{rk}}^*$ is non-increasing in k.

Proof: First note that the differential equation in DMP problem has a closed form solution as follows:

$$F(T) = G\left(\sum_{r\in R} \tau_r \left(\sum_{k=1}^{K_r} \sum_{i=0}^{\min\{k-1,s_r\}} \phi_r^i(a_{r,k-i})\right) + \Phi_{\mathrm{r}}^0\right).$$

Now assume that the total budget allocated to each channel per target customer, E_r for all $r \in R$, is already decided. Then the objective function of the DMP problem can be written as

$$\max_{a_{rk}, r\in R, k=1,\dots,K_r} mP\left[G\left(\sum_{r\in R} \tau_r \left(\sum_{k=1}^{K_r} \sum_{i=0}^{\min\{k-1,s_r\}} \phi_r^i(a_{r,k-i})\right) + \Phi_{\mathrm{r}}^0\right) - x_0\right] - m\sum_{r\in R} E_r \ .$$

Since E_r values are pre-specified and $G(.)$ is an increasing function of its argument, the DMP problem is equivalent to the following:

$$\max_{\substack{0\le a_{rk}\le b_r, \\ r\in R, k=1,\dots,K_r}} \sum_{r\in R} \tau_r \left(\sum_{k=1}^{K_r} \sum_{i=0}^{\min\{k-1,s_r\}} \phi_r^i(a_{r,k-i})\right) + \Phi_{\mathrm{r}}^0 \qquad \text{subject to: } m\,E_r = \sum_{k=1}^{K_r} a_{rk} \text{ for all } r \in R.$$

However, note that the objective function and the constraints in this problem are separable in terms of each of the channels, meaning that it is enough to maximize with respect to the marketing expenses in each of the channels separately. Further, since all of the ϕ_r^i's are increasing functions, it is optimal to spend the whole budget mE_r in each channel r even if we set the constraint as $\sum_{k=1}^{K_r} a_{rk} \le mE_r$. Therefore, solving the TPP problem for all channels provide the optimal solution for the DMP problem when the budget in each channel is pre-determined.

To establish the properties of the TPP problem, we first rewrite the objective function, making the substitution that $\psi_r^j(a) = \sum_{i=0}^{\min\{s_r,k_r-j\}} \phi_r^i(a)$:

$$\sum_{k=1}^{K_r} \sum_{i=0}^{\min\{k-1,s_r\}} \phi_r^i(a_{r,k-i}) + \Phi_{\mathrm{r}}^0 =$$

$$\sum_{k=1}^{K_r} \sum_{j=\max\{k-s_r,1\}}^{k} \phi_r^{k-j}(a_{rj}) + \Phi_{\mathrm{r}}^0 = \sum_{j=1}^{K_r} \sum_{k=j}^{\min\{j+s_r,K_r\}} \phi_r^{k-j}(a_{rj}) + \Phi_{\mathrm{r}}^0 = \sum_{j=1}^{K_r} \psi_r^j(a_{rj}) + \Phi_{\mathrm{r}}^0.$$

As the $\psi_r^j(.)$ functions are concave, the TPP problem is an instance of a concave Knapsack problem. Therefore, based on Zipkin (1980), the optimal solution of this problem ($\Phi_r(E_r)$) is concave increasing in the level of the total budget E_r.

Now we establish the non-increasing property of the optimal temporal investment plan. First note that all the ψ_r^j are the same when $j \le K_r - s_r$, and involves less and less terms in the summation as j increases for $j > K_r - s_r$. Therefore, the derivative of $\psi_r^j(a)$, $D\psi_r^j(a)$, is non-increasing in j for any value of $a \in [0, b_r]$, in addition to being decreasing in a based on their concavity. Based on the KKT conditions and Zipkin (1980), there exists a number M^* that uniquely specifies the optimal solution a_{rj}^* to the TPP problem as follows:

$$D\psi_r^j(a_{rj}^*) > M^* \text{ iff } a_{rj}^* = b_r$$

$$D\psi_r^j(a_{rj}^*) = M^* \text{ iff } a_{rj}^* \in (0, b_r)$$

$$D\psi_r^j(a_{rj}^*) < M^* \text{ iff } a_{rj}^* = 0.$$

Given that each $D\psi_r^j(.)$ is strictly decreasing, we should have $a_{rj}^* > 0$ if and only if $D\psi_r^j(0) > M^*$. With the decreasing property of $D\psi_r^j$ in j, this means that there should be an integer k_0^* such that if $j \le k_0^*$, the optimal spending is positive ($a_{rj}^* > 0$), but the optimal spending drops to 0 for $j > k_0^*$. In addition, note that $a_{rj}^* = b_r$ if and only if $D\psi_r^j(0) > D\psi_r^j(b_r) > M^*$. Since this condition is a much stronger condition than $D\psi_r^j(0) > M^*$, it can only hold for some of $j \le k_0^*$. As $D\psi_r^j(b_r)$ is decreasing in j, there should be $k_1^* \le k_0^*$ such that for $j \le k_1^*$, the optimal spend is at the maximum level ($a_{rj}^* = b_r$), while for $k_1^* < j \le k_0^*$, we should have $D\psi_r^j(a_{rj}^*) = M^*$. In addition, for $k_1^* < j \le k_0^*$, the property that $D\psi_r^j(a_{rj}^*) = M^*$ translates to a_{rj}^* to be strictly decreasing in j. Putting it all together, this means that the optimal spending over time would be at the maximum level for $j \le k_1^*$, gradually drops toward 0 when $k_1^* < j \le k_0^*$, and remains at 0 for $j > k_0^*$, establishing the non-increasing property of optimal spend. ■

Theorem 2 (Section 3.3): The following error bounds can be obtained on discounted profit and total discounted sales of the discounted problem:

$$\begin{aligned}0 &\le \Pi^D\{a_r^D|E_r^D\} - \Pi^D\{a_r^D|E_r^U\} \le \Pi^D\{a_r^D|E_r^D\} - \Pi^D\{a_r^U|E_r^U\} \\ &\le mP\left[(1-e^{-\theta T})(F[T,\{a_r^U|E_r^U\}] - x_0) - \theta \int_0^T (F[t,\{a_r^U|E_r^U\}] - x_0)e^{-\theta t}dt\right],\end{aligned}$$

$$\begin{aligned}0 &\le \int_0^T \left(\frac{dF[t,\{a_r^D|E_r^U\}]}{dt} - \frac{dF[t,\{a_r^U|E_r^U\}]}{dt}\right)e^{-\theta t}dt \le \\ &\le (1-e^{-\theta T})(F[T,\{a_r^U|E_r^U\}] - x_0) - \theta \int_0^T (F[t,\{a_r^U|E_r^U\}] - x_0)e^{-\theta t}dt.\end{aligned}$$

Proof: First note that for a given investment plan $\{a_r|E_r\}$, $\Pi^D\{a_r|E_r\}$ can be written as follows after applying integration by parts:

$$\begin{aligned}\Pi^D\{a_r|E_r\} &= \left[\int_0^T \left(mP\frac{dF[t,\{a_r|E_r\}]}{dt}e^{-\theta t} - \sum_{r\in R} a_r(t)\right)dt\right] \\ &= mP\left[e^{-\theta T}(F[T,\{a_r|E_r\}] - x_0) + \theta \int_0^T (F[t,\{a_r|E_r\}] - x_0)e^{-\theta t}dt\right] - m\sum_{r\in R} E_r\end{aligned}$$

In the above expression, the profit depends on the cumulative adoption at any point of time rather than the derivative. A couple of notes are in place. First, cumulative adoptions can be written in closed form (i.e.,

$F[T, \{a_r|E_r\}] = G(\sum_{r\in R} \Phi_r))$, and therefore, the total market share adopted depends on the cumulative marketing effort up to that time. Second, the undiscounted tactical problem maximizes the cumulative marketing effort over the whole horizon based on the TPP, which means that the final market share adopted at time T based on the undiscounted policy would be no less than that of the discounted one if the same total budget is spent, i.e., $F[T, \{a_r^U|E_r\}] \geq F[T, \{a_r^D|E_r\}] \geq F[t, \{a_r^D|E_r\}]$. Therefore we have:

$$\Pi^D\{a_r^D|E_r\} \leq mP\left[e^{-\theta T}(F[T, \{a_r^D|E_r\}] - x_0) + \theta \int_0^T (F[T, \{a_r^D|E_r\}] - x_0)e^{-\theta t}dt\right] - m\sum_{r\in R} E_r$$
$$= mP(F[T, \{a_r^D|E_r\}] - x_0) - m\sum_{r\in R} E_r = \Pi^U\{a_r^D|E_r\} \leq \Pi^U\{a_r^U|E_r\} \qquad (I1)$$

Consequently, for a given strategic budget allocation $\{E_r\}$, if we implement the undiscounted optimal tactical plan instead of the optimal tactical plan for the discounted problem, we would have the following bounds:

$$0 \leq \Pi^D\{a_r^D|E_r\} - \Pi^D\{a_r^U|E_r\}$$
$$\leq \Pi^U\{a_r^U|E_r\} - mP\left(\int_0^T \frac{dF[t, \{a_r^U|E_r\}]}{dt} e^{-\theta t}dt\right) + m\sum_{r\in R} E_r$$
$$\leq \Pi^U\{a_r^U|E_r\} - mP\left[e^{-\theta T}(F[T, \{a_r^U|E_r\}] - x_0) + \theta \int_0^T (F[t, \{a_r^U|E_r\}] - x_0)e^{-\theta t}dt\right] + m\sum_{r\in R} E_r$$
$$= mP\left[\left(1 - e^{-\theta T}\right)(F[T, \{a_r^U|E_r\}] - x_0) - \theta \int_0^T (F[t, \{a_r^U|E_r\}] - x_0)e^{-\theta t}dt\right]. \quad (I2)$$

The above relationship then leads to the following regarding the difference in discounted sales if a_r^D or a_r^U are used at the tactical level given the strategic budget allocation $\{E_r\}$:

$$0 \leq \int_0^T \left(\frac{dF[t, \{a_r^D|E_r\}]}{dt} - \frac{dF[t, \{a_r^U|E_r\}]}{dt}\right) e^{-\theta t}dt$$
$$\leq \left(1 - e^{-\theta T}\right)(F[T, \{a_r^U|E_r\}] - x_0) - \theta \int_0^T (F[t, \{a_r^U|E_r\}] - x_0)e^{-\theta t}dt.$$

Now we would like to incorporate the effect of optimal strategic budget allocation rather than using a predetermined set of values. Using budget allocation $\{E_r^D\}$ in the set of inequalities of (I1) and noting that budget allocation $\{E_r^U\}$ optimizes the profit of the undiscounted DMP problem, we would get:

$$\Pi^D\{a_r^D|E_r^D\} \leq \Pi^U\{a_r^D|E_r^D\} \leq \Pi^U\{a_r^U|E_r^D\} \leq \Pi^U\{a_r^U|E_r^U\}.$$

If now the strategic budget allocation is done on the basis of undiscounted problem (either when only strategic undiscounted policy is used, or when both tactical and strategic discounted policies are used), the loss in profit would be as follows when (I2) is applied:

$$0 \leq \Pi^D\{a_r^D|E_r^D\} - \Pi^D\{a_r^D|E_r^U\} \leq \Pi^D\{a_r^D|E_r^D\} - \Pi^D\{a_r^U|E_r^U\} \leq \Pi^D\{a_r^D|E_r^U\} - \Pi^D\{a_r^U|E_r^U\}$$
$$\leq mP\left[\left(1 - e^{-\theta T}\right)(F[T, \{a_r^U|E_r^U\}] - x_0) - \theta \int_0^T (F[t, \{a_r^U|E_r^U\}] - x_0)e^{-\theta t}dt\right]. \blacksquare$$

Theorem 3 (Section 4.1.1): The optimal investment in channel s is bounded from above by E_s^U:

$$E_s^U = \begin{cases} E_s^0 & if\, \Phi_s'(E_s^0) < \frac{4q}{P(q+p)^2}, \\ sup_{E_s \geq E_s^0}\left[\Phi_s'(E_s) \geq \frac{4q}{P(q+p)^2}\right] & otherwise \end{cases}.$$

Proof: By the definition of E_s^U, for any investment in channel s in the interval (E_s^U, ∞), we should have $\Phi_s'(E_s) < \frac{4q}{P(q+p)^2}$. Then the following inequalities hold, where $G = G(\Phi_s(E_s) + C_{-s})$ and C_{-s} is the cumulative effort from expenditures in all channels except s:

$$\frac{\partial \Pi}{\partial E_s}([E_s, C_{-s}]) = m[P(1-G)(p+qG)\Phi_s'(E_s) - 1]$$

$$< m\left[P(1-G)(p+qG)\frac{4q}{P(q+p)^2} - 1\right] = \frac{-m}{(q+p)^2}[-4q(1-G)(p+qG) + (q+p)^2]$$

$$= \frac{-m}{(q+p)^2}[4q^2G^2 - 4q(q-p)G + (q-p)^2] = \frac{-m}{(q+p)^2}\big(2qG - (q-p)\big)^2 \leq 0.$$

Therefore profit Π decreases in E_s or remains the same for $E_s \in (E_s^U, \infty)$ and an investment larger than E_s^U cannot provide higher profit.■

Theorem 4 (Section 4.1.2): Assume that channel s is not a low-leverage channel, i.e. $E_s^U > E_s^0$. Then the threshold level G_s^2 is an upper bound on the market penetration level that can be achieved with investment vector $[E_s^*(C_{-s}^0), C_{-s}^0]$, in which an optimal investment is made in channel s while keeping the expenditures in all other channels at their initial levels. In addition, one of the following cases must hold for the optimal investment $E_s^*(C_{-s}^0)$ in channel s:

"High Momentum": when $G(C^0) \geq G_s^2$, demand adoption with current expenditure E^0 is sufficiently high and any further investment in channel s is not profitable, i.e. $E_s^*(C_{-s}^0) = E_s^0$.

"Medium Momentum": when $G_s^1 < G(C^0) < G_s^2$, it is always optimal to increase the investment in channel s, i.e. $E_s^*(C_{-s}^0) \in (E_s^0, E_s^U]$.

"Low Momentum": when $G(C^0) \leq G_s^1$, a small increase in expenditure in channel s results in a profit loss, i.e., there exists $E_s^{min} > E_s^0$ such that $\Pi[E_s, E_{-s}^0] < \Pi[E_s^0, E_{-s}^0]$ for all $E_s \in (E_s^0, E_s^{min}]$. However, a larger investment $E_s^* \in (E_s^{min}, E_s^U]$ may be profitable.

Proof: Consider an investment vector $E = [E_s, C_{-s}]$ that spends E_s in channel s and achieves a cumulative marketing effort of C_{-s} from all the other channels. Then the derivative of profit with respect to E_s can be written as:

$$\frac{\partial \Pi}{\partial E_s}[E_s, C_{-s}] = m(PG'(\Phi_s(E_s) + C_{-s})\Phi_s'(E_s) - 1) \quad \text{(A1)}$$

$$= m\Big(P\big(1 - G(\Phi_s(E_s) + C_{-s})\big)(p + qG(\Phi_s(E_s) + C_{-s}))\Phi_s'(E_s) - 1\Big)$$

$$= -mH_s^1(E_s; C_{-s})H_s^2(E_s; C_{-s}),$$

where the two functions H_s^1 and H_s^2 are defined as follows:

$$\begin{cases} H_s^1(E_s; C_{-s}) = G(\Phi_s(E_s) + C_{-s}) - \left[\frac{1}{2}(1 - \frac{p}{q}) - \frac{1}{2q}\sqrt{(q+p)^2 - 4q\frac{1}{P\Phi_s'(E_s)}}\right] \\ H_s^2(E_s; C_{-s}) = G(\Phi_s(E_s) + C_{-s}) - \left[\frac{1}{2}(1 - \frac{p}{q}) + \frac{1}{2q}\sqrt{(q+p)^2 - 4q\frac{1}{P\Phi_s'(E_s)}}\right] \end{cases}. \tag{A2}$$

Note that the functions H_s^1 and H_s^2 are well-defined, as channel s is not a low-leverage channel. Also, the last equality in (A1) is obtained by observing that $\frac{\partial \Pi}{\partial E_s}$ is a quadratic equation with respect to $G(\Phi_s(E_s) + C_{-s})$. These definitions also imply that, when $E = E^0$, $\frac{\partial \Pi}{\partial E_s}(E^0)$ reduces to:

$$\frac{\partial \Pi}{\partial E_s}(E^0) = -m(G(C^0) - G_s^1)(G(C^0) - G_s^2). \tag{A3}$$

Now, assume that the allocated investment vector E^0 is made. Then the optimal response of channel s, $E_s^*(C_{-s}^0)$, is either to maintain the current investment E_s^0 or increase it to a finite interior local maximizer, $E_s^{int} \in (E_s^0, E_s^U]$. If E_s^{int} exists, it should set $\frac{\partial \Pi}{\partial E_s}[E_s^{int}, C_{-s}]$ to 0, meaning that at least one of H_s^1 or H_s^2 should be zero at E_s^{int} as well.

First, we establish that G_s^2 is an upper bound on the optimal demand fraction that can optimally be adopted. To do this, note that if E_s^{int} exists and is a root of H_s^2, it results in a higher level of demand adoption than when it is a root of H_s^1 based on the definition of H_s^1 and H_s^2 functions. Therefore, we show the claim for when E_s^{int} is a root of H_s^2. As $\Phi_s'(E_s^0) \geq \Phi_s'(E_s^{int})$, we should have $G_s^2 \geq G\big(C_{-s}^0 + \Phi_s'(E_s^{int})\big)$, establishing the claim.

We now proceed with the investigation of each case, while bearing in mind that $\frac{\partial \Pi}{\partial E_s}(E^0)$ is a quadratic equation with respect to $G(C^0)$ (based on (A3)), having two possible roots, G_s^1 and G_s^2.

Case 1 (High Momentum): The condition $G(C^0) \geq G_s^2$ implies that $\frac{\partial \Pi}{\partial E_s}(E^0) \leq 0$ and therefore profit is non-increasing at the point $E_s = E_s^0$. If no interior local maximizer ($E_s^{int} > E_s^0$) exists, no local minimum would exist as well, especially as $\lim_{E_s \to \infty} \Pi = -\infty$. Therefore, $\frac{\partial \Pi}{\partial E_s}(E^0)$ remains to be non-positive for any level of $E_s \geq 0$, and Π is non-increasing in E_s. This implies that not investing further in channel s is optimal. If $E_s^{int} > E_s^0$ exists, we have $G(C^0) \geq G_s^2 \geq G\big(\Phi_s(E_s^{int}) + C_{-s}^0\big)$ based on the upper bounding property of G_s^2 for the optimal demand fraction adopted. Since $E_s^{int} > E_s^0$, additional marketing cost has been incurred but lower or equal market penetration is obtained. Therefore $\Pi[E_s^{int}, C_{-s}^0] \leq \Pi[E_s^0, C_{-s}^0]$ and not investing further in channel s is optimal.

Case 2 (Medium Momentum): The condition $G_s^1 < G(C^0) < G_s^2$ implies that $\frac{\partial \Pi}{\partial E_s}(E^0) > 0$ which means that profit can be increased by at least infinitesimally investing more in channel s. Therefore maintaining the initial investment of E_s^0 is not optimal. In addition, as $\lim_{E_s \to \infty} \Pi = -\infty$, it should be the case that E_s^{int} exists and $E_s^*(E^0) = E_s^{int}$.

Case 3 (Low Momentum): When $G(C^0) \leq G_s^1$, we have $\frac{\partial \Pi}{\partial E_s}(E^0) \leq 0$; so increasing channel s investment by an arbitrarily small level would hinder profitability. If $E_s^{int} > E_s^0$ does not exist, no local minimum exists as well (because $\lim_{E_s \to \infty} \Pi = -\infty$) and the profit function is decreasing in E_s. Therefore, offering no additional channel s investment is optimal. If E_s^{int} exists, the profitability of offering E_s^{int} is dependent on whether it provides higher profit in comparison to maintaining the current investment E_s^0, that is, E_s^{int} is the optimal investment in channel s if $P(G(C^0) - \mathrm{x}_0) - E_s^0 < P\big(G\big(\Phi_s\big(E_s^{int}\big) + C_{-s}^0\big) - \mathrm{x}_0\big) - E_s^{int}$.■

Theorem 5 (Section 4.1.4): If there is a channel $r \in R$ such that its effectiveness always dominates that of channel s (that is, $\Phi_s'(E_s) < \Phi_r'(E_r)$ for all $E_s \in [E_s^0, E_s^U]$ and all $E_r \in [E_r^0, E_r^U]$) then it is never optimal to invest in channel s beyond E_s^0.

Proof: Let E_s^*, E_r^* be the investments in channel s and r in the optimal marketing resource allocation E^* resulting in cumulative marketing effort C^*. Based on the optimality of channel r investment, we should have $\frac{\partial\Pi}{\partial E_r}(E^*) \le 0$ (If optimal channel r investment is non-zero, we have $\frac{\partial\Pi}{\partial E_r}(E^*) = 0$ and if it is zero, $\frac{\partial\Pi}{\partial E_r}(E^*) \le 0$). Assume the contrary, that it is optimal to invest higher than E_s^0 in channel s, i.e., $E_s^* > E_s^0$. Then, using the expression $\frac{\partial\Pi}{\partial E_r}(E^*) = m[PG'(C^*)\Phi_r'(E_r^*) - 1]$, we have:

$$\frac{\partial\Pi}{\partial E_s}(E^*) = m[PG'(C^*)\Phi_s'(E_s^*) - 1] = \left[\frac{\partial\Pi}{\partial E_r}(E^*) + m\right]\frac{\Phi_s'(E_s^*)}{\Phi_r'(E_t^*)} - m$$

$$= \frac{\partial\Pi}{\partial E_r}(E^*)\frac{\Phi_s'(E_s^*)}{\Phi_r'(E_r^*)} + m\left(\frac{\Phi_s'(E_s^*)}{\Phi_r'(E_r^*)} - 1\right) \le m\left(\frac{\Phi_s'(E_s^*)}{\Phi_r'(E_r^*)} - 1\right) < 0.$$

The last inequality follows from the fact that the effectiveness of channel r dominates that of channel s. As $E_s^* > E_s^0$, $\frac{\partial\Pi}{\partial E_s}(E^*) < 0$ means that we can strictly increase the total profit by decreasing the level of channel s investment, which is in contradiction with optimality of E^*.■

Theorem 6 (Section 4.2): If $(\Phi_s^{-1}(x))'$ is convex in x for all $x \ge 0$ and $s \in R^A$, then $PG(x+z) - \Phi_s^{-1}(x)$ is concave or S-shaped in x.

Proof: The function $PG(x+z) - \Phi_s^{-1}(x)$ is concave or S-shaped when its derivative is either always decreasing, always increasing, or it increases up to some point and then decreases. This property is equivalent to the derivative being quasi-concave (as the derivative is defined over real numbers). The derivative can be written as follows:

$$\pi(x) \;=\; PG'(x+z) - \frac{1}{\Phi_s'\left(\Phi_s^{-1}(x)\right)} \;=\; \frac{P(p+q)^2}{p}\frac{e^{-(p+q)(x+z)}}{(1+(q/p)\,e^{-(p+q)(x+z)})^2} - (\Phi_s^{-1}(x))'.$$

Note that we can write $\pi(x)$ as a composition of two functions, that is, $\pi(x) = \bar{\pi} \circ y(x)$ where $\bar{\pi}(u) = \frac{P(p+q)^2}{p}\frac{u}{(1+q/p\,u)^2} - (\Phi_s^{-1}(x))'|_{x=\frac{-\ln(u)}{p+q}-z}$ for $u \in (0,1]$ and $y(x) = e^{-(p+q)(x+z)}$. Now we claim that the composition of a real-valued quasiconcave function with a real-valued decreasing function remains quasiconcave; as a result if $\bar{\pi}(u)$ is quasiconcave, $\pi(x)$ would be quasiconcave as well, noting that $y(.)$ is a monotone decreasing function. To show this, note that if $\bar{\pi}(u)$ is quasiconcave, one of the following holds: Either $\bar{\pi}(u)$ is monotone increasing or monotone decreasing over the range of the function $y(.)$, in which case the function $\pi(x) = \bar{\pi} \circ y(x)$ would be monotone decreasing or monotone increasing respectively. Otherwise, if $\bar{\pi}(u)$ increases and then decreases over the range of u in the range of function $y(.)$, it has a unique maximizer $u^* = y(x^*)$ such that for $u < u^*$ we have $\bar{\pi}(u)$ increasing, and for $u > u^*$ it is decreasing. Therefore, for two levels of $x_1 < x_2$ in the domain of $y(.)$ we can make the following deductions which shows that $\pi(x)$ would have an increasing-decreasing pattern:

$$x_1 < x_2 < x^* \;\Rightarrow\; y(x_1) > y(x_2) > u^* \;\Rightarrow\; \bar{\pi} \circ y(x_1) < \bar{\pi} \circ y(x_2) \;\Rightarrow\; \pi \text{ is increasing}$$

$$x^* < x_1 < x_2 \;\Rightarrow\; u^* > y(x_1) > y(x_2) \;\Rightarrow\; \bar{\pi} \circ y(x_1) > \bar{\pi} \circ y(x_2) \;\Rightarrow\; \pi \text{ is decreasing}$$

Now the only issue remaining is to setup conditions that guarantee quasiconcavity of $\bar{\pi}(u)$, with one such condition being $\bar{\pi}(u)$ to be concave. Note that the first term of $\bar{\pi}(u)$ is concave as its derivative

$\frac{P(p+q)^2}{p(1+q/pu)^4}\left(1-(\frac{q}{p}u)^2\right)$ is decreasing in u. Consequently, $\bar{\pi}(u)$ is concave (and hence quasi-concave) if $(\Phi_s^{-1}(x))'|_{x=\frac{-\ln(u)}{p+q}-z}$ is convex. As $\frac{-\ln(u)}{p+q}-z$ is convex in u and $(\Phi_s^{-1}(x))'$ is increasing, this function would be convex if $(\Phi_s^{-1}(x))'$ is convex in x. ■

Lemma 1 Assume that channel s has either medium leverage ($E_s^0 < E_s^U < \infty$) or high leverage ($E_s^U = \infty$). Also let the functions H_s^1 and H_s^2 be defined as in (A2), and the investment level E_s^{int} be denoted as follows:

$$E_s^{int}(C_{-s}^0) = \begin{cases} \text{The single root of } H_s^2(.;C_{-s}^0) & \text{If } E_s^U = \infty \text{ or } G\left(C_{-s}^0 + \Phi_s(E_s^U)\right) \geq \frac{1}{2}\left(1-\frac{p}{q}\right), \text{i.e.}, H_s^2(E_s^U, C_{-s}^0) \geq 0 \\ \text{The larger of the two possible roots of } H_s^1(.;C_{-s}^0) & \text{Otherwise} \end{cases}$$

If $G_s^1 < G(C^0) < G_s^2$ (medium momentum channel), $E_s^{int} \in (E_s^0, E_s^U]$ is well defined and represents the optimal investment of magnitude at least E_s^0. But if $G(C^0) \leq G_s^1$ (low momentum channel) and $E_s^{int} \in (E_s^0, E_s^U]$ exists, it represents the optimal investment in channel s provided that it generates a larger profit than the current investment E_s^0.

Proof: We would like to find the optimal expenditure in channel s from one of the following two (equivalent) optimization problems:

$$\max_{E_s \geq 0} \Pi = m(PG(\Phi_s(E_s) + C_{-s}^0) - E_s) \qquad \max_{x \geq 0} \overline{\Pi} = m(PG(x + C_{-s}^0) - \Phi_s^{-1}(x)).$$

These two problems are equivalent, as the objective of the second one is the composition of the first objective with the monotone increasing function $\Phi_s^{-1}(x)$. Because of the assumptions of Section 4.2 , either Π or $\overline{\Pi}$ is S-shaped with respect to its argument and can have at most two local minima and maxima, with at most one of them being a local maximum. If there is only one such point, then it should be a local maximum because increasingly large solutions would result in negative profit. As the two optimization problems are equivalent, if only $\overline{\Pi}$ is S-shaped, it would imply that Π would have at most one local maximum as well. Therefore, from now on we focus on the objective function Π having at most one local maximum and one local minimum. Also note that, as shown in the proof of the categorization of Section 4.1.2, $\frac{\partial \Pi}{\partial E_s}\left(E_s^{int}, E_{-s}^0\right) = -mH_s^1\left(E_s^{int}; C_{-s}^0\right)H_s^2\left(E_s^{int}; C_{-s}^0\right)$, with the local maximizer E_s^{int} setting at least one of H_s^1 or H_s^2 to zero.

Consider the case that E_s^{int} is a root of H_s^2. Based on the definition of H_s^2, we should have $G\left(\Phi_s(E_s^{int}) + C_{-s}^0\right) \geq \frac{1}{2}(1-\frac{p}{q})$ which is the range of demand adoption for which G, and consequently Π is concave. Therefore, if H_s^2 has a root, it is the local maximum. In addition, $H_s^2(E_s; C_{-s}^0)$ is increasing in E_s based on concavity and monotonicity of Φ_s, so it can have at most one root. Further, note that $H_s^2(E_s^0; C_{-s}^0) = G(C^0) - G_s^2 < 0$ and H_s^2 is an increasing function of E_s. As the optimal solution can be no greater than E_s^U, the existence of a root for $H_s^2(E_s; C_{-s}^0)$ depends on the sign of $H_s^2(E_s^U; C_{-s}^0)$, that is, $H_s^2(E_s^U; C_{-s}^0)$ has a root if and only if it is non-negative.

When $E_s^U = \infty$, we need to evaluate $\lim_{E_s\to\infty} H_s^2(E_s; C_{-s}^0)$ to see whether H_s^2 has a root or not. To do this, note that when $E_s^U = \infty$, $\lim_{E_s\to\infty} \Phi_s{}'(E_s) > \frac{4q}{P(q+p)^2}$, which means that $\lim_{E_s\to\infty} \Phi_s(E_s) = \infty$. Substituting these quantities into $\lim_{E_s\to\infty} H_s^2(E_s; C_{-s}^0)$ we have:

$$\lim_{E_s\to\infty} H_s^2(E_s; C_{-s}^0) = G\left(C_{-s}^0 + \lim_{E_s\to\infty}\Phi_s(E_s)\right) - \left[\frac{1}{2}(1-\frac{p}{q}) + \frac{1}{2q}\sqrt{(q+p)^2 - 4q\frac{1}{P\lim_{E_s\to\infty}\Phi_s'(E_s)}}\right]$$

$$\geq 1 - \left[\frac{1}{2}\left(1-\frac{p}{q}\right) + \frac{1}{2q}\sqrt{(q+p)^2 - 4q\frac{1}{P\,\Phi_s'(E_s^0)}}\right] = 1 - G_s^2 \geq 0$$

Therefore we can conclude that, when $E_s^U = \infty$, H_s^2 always has a root determining the optimal channel s investment. But when $E_s^U < \infty$, we have $\Phi_s'(E_s^U) = \frac{4q}{P(q+p)^2}$, which allows us to simplify $H_s^2(E_s^U; C_{-s}^0) = G(C_{-s}^0 + \Phi_s(E_s^U)) - \frac{1}{2}(1-\frac{p}{q})$. Therefore H_s^2 has a root if and only if $G(C_{-s}^0 + \Phi_s(E_s^U)) \geq \frac{1}{2}(1-\frac{p}{q})$.

Alternatively if E_s^{int} is not a root of H_s^2, we should have $H_s^2(E_s^U; C_{-s}^0) < 0$, and a local maximizer should be the root of $H_s^1(.; C_{-s}^0)$ if it exists (which we know it does for medium momentum channel). As H_s^2 does not have a root, H_s^1 can have at most two roots. Note that we cannot have two local maximizers without having a local minimum. Therefore, one of the roots of H_s^1 should be a local maximum which should happen at the concave portion of Π, and the other should be a local minimum located where Π is convex. The S-shaped property of the objective function forces the convex portion of Π to locate for smaller values of E_s than for the concave portion, meaning that the interior maximizer should be the larger of the possible two roots of H_s^1, and it represents the optimal solution if it provides larger profit than investing E_s^0. ■

Lemma 2 If channel s has low momentum, along with medium or high leverage, with respect to the initial spend vector E^0, there is a level of C_{-s}, C_{-s}^{trans}, such that if $C_{-s} < C_{-s}^{trans}$, it is optimal to maintain the current channel s investment E_s^0, but when $C_{-s} \geq C_{-s}^{trans}$, while the channel remains low momentum, it is optimal to increase the level of channel s investment beyond E_s^0.

Proof: Let the interval $[C_{-s}^0, \bar{C}_{-s}]$ represent the largest interval of C_{-s} levels over which channel s maintains its low-momentum category. Therefore, for all $C_{-s} \in [C_{-s}^0, \bar{C}_{-s}]$, we have $G([E_s^0, C_{-s}]) \leq G([E_s^*(C_{-s}), C_{-s}]) \leq G_s^1$, which means that, based on low momentum categorization, it is optimal to increase the level of channel s investment when two conditions hold: (1) an interior maximizer for channel s investment larger than E_s^0 exists; and (2) this interior maximizer provides greater profit than maintaining the current investment E_s^0. As a result, we first investigate over which part of the interval $[C_{-s}^0, \bar{C}_{-s}]$ an interior solution exists; then over this part, we check whether it is profitable to increase channel s investment compared to the current investment E_s^0.

We first build the C_{-s} level, C_{-s}^{exist}, such that for $C_{-s} \geq C_{-s}^{exist}$ an interior maximizer exists, but none exists for $C_{-s} < C_{-s}^{exist}$. If $E_s^U = \infty$, an interior maximizer always exists and is obtained from a root of H_s^2 based on Lemma 1; therefore, we set $C_{-s}^{exist} = C_{-s}^0$. But now focus on the case that $E_s^0 \leq E_s^U < \infty$, in which case E_s^U is independent of C_{-s} and would remain unchanged as C_{-s} changes. Also recall that the interior maximizer (if it exists) is either a root of H_s^1 or H_s^2. Lemma 1 establishes that in this case, if $H_s^2(E_s^U, C_{-s}) \geq 0$, then H_s^2 would have a root that coincides with the single interior maximizer of profit. However, note that H_s^2 is increasing in C_{-s}, and as a result, a level of C_{-s} called $C_{-s}^1 \in [C_{-s}^0, \bar{C}_{-s}]$ exists such that $H_s^2(E_s^U, C_{-s}) < 0$ for $C_{-s} \in [C_{-s}^0, C_{-s}^1)$, but $H_s^2(E_s^U, C_{-s}) \geq 0$ for $[C_{-s}^1, \bar{C}_{-s}]$, in which case an interior maximizer exists.

Now we focus on the interval $[C_{-s}^0, C_{-s}^1)$ to see in which subinterval an interior maximizer does not exist. Again based on Lemma 1, if an interior maximizer exists over this interval, it should be a root of H_s^1. We claim that there is a C_{-s} level called $C_{-s}^{exist} \in [C_{-s}^0, C_{-s}^1)$ such that for $C_{-s} \in [C_{-s}^0, C_{-s}^{exist})$, H_s^1 does not have a root (and hence no interior maximizer exists), but for $C_{-s} \in [C_{-s}^{exist}, C_{-s}^1)$, H_s^1 has at least one root and an interior maximizer exists. To see this, let $\tilde{E}_s(C_{-s})$ be the smallest root of $H_s^1(.; C_{-s})$ if it exists (resulting in $H_s^1(\tilde{E}_s(C_{-s}); C_{-s}) = 0$). Note that H_s^1 is increasing in C_{-s}, and as the channel has low momentum, $H_s^1(E_s^0, C_{-s}) =$

$G(C^0) - G_s^1 \leq 0$. If for a given C_{-s} level $c \in [C_{-s}^0, C_{-s}^1)$ a root for H_s^1 exists (i.e., $\tilde{E}_s(c)$ exists), then for all $C_{-s} \in [c, C_{-s}^1)$ we should have $H_s^1(\tilde{E}_s(c), C_{-s}) \geq 0$ while $H_s^1(E_s^0, C_{-s}) \leq 0$. Therefore, a root for $H_s^1(., C_{-s})$ should exist such that $\tilde{E}_s(C_{-s}) \in [E_s^0, \tilde{E}_s(c)]$ which is either the interior maximizer or another larger root for H_s^1 exists that locally maximizes profit. The continuity of H_s^1 implies that a level of $C_{-s}^{exist} \in [C_{-s}^0, C_{-s}^1]$ should exist to be the infimum of all C_{-s} levels for which a root for H_s^1 exists.

When $C_{-s} \in [C_{-s}^0, C_{-s}^{exist})$, neither of H_s^1 and H_s^2 has a root, so the profit function should be decreasing and the optimal channel s investment would be E_s^0. However, for $C_{-s} \in [C_{-s}^{exist}, \bar{C}_{-s}]$, for which an interior maximizer exists, we need to look at when the profit of increasing channel s investment to the interior maximizer, $E_s^{int}(C_{-s})$, exceeds that of maintaining the initial investment E_s^0 in that channel. Define the difference of these two profits per capita as

$$\begin{aligned} h(C_{-s}) &= \frac{1}{m}\left[\Pi\left(E_s^{int}(C_{-s}), C_{-s}\right) - \Pi(E_s^0, C_{-s})\right] \\ &= PG(\Phi_s(E_s^{int}(C_{-s})) + C_{-s}) - E_s^{int}(C_{-s}) - PG(\Phi_s(E_s^0) + C_{-s}) + E_s^0, \end{aligned}$$

where with a little abuse of notation $\Pi(E_s, C_{-s})$ is the profit obtained by investing E_s in channel s and a vector of E_{-s} in all channels except s resulting in cumulative marketing effort of C_{-s} from all channels except s. We claim that $h(C_{-s})$ is increasing in C_{-s} over $[C_{-s}^{exist}, \bar{C}_{-s}]$, which would imply that there is a level of C_{-s} called $C_{-s}^{trans} \in [C_{-s}^{exist}, \bar{C}_{-s}]$ such that for $C_{-s} \in [C_{-s}^{exist}, C_{-s}^{trans}]$ offering the interior maximizer is never optimal, but it remains to be optimal for all C_{-s} levels in $[C_{-s}^{trans}, \bar{C}_{-s}]$. To see this, note that Π is twice continuously differentiable and therefore $E_s^{int}(C_{-s})$, for $C_{-s} \geq C_{-s}^{exist}$, is continuously differentiable in C_{-s} as well. Also, the derivatives of Π at the two points $E_s^{int}(C_{-s})$ and E_s^0 have the following properties by their definition and the fact that channel s has low momentum:

$$G'[E_s^{int}(C_{-s}), C_{-s}]\Phi_s'\left(E_s^{int}(C_{-s})\right) = 1 \qquad PG'[E_s^0, C_{-s}]\Phi_s'(E_s^0) - 1 < 0.$$

Therefore, the derivative of $h(.)$ is as follows:

$$\begin{aligned} h'(C_{-s}) &= PG'[E_s^{int}(C_{-s}), C_{-s}]\left[\Phi_s'\left(E_s^{int}(C_{-s})\right)\frac{\partial E_s^{int}(C_{-s})}{\partial C_{-s}} + 1\right] - \frac{\partial E_s^{int}(C_{-s})}{\partial C_{-s}} - PG'[E_s^0, C_{-s}] \\ &> \left(\frac{1}{\Phi_s'(E_s^{int}(C_{-s}))}\right)\left[\Phi_s'\left(E_s^{int}(C_{-s})\right)\frac{\partial E_s^{int}(C_{-s})}{\partial C_{-s}} + 1\right] - \frac{\partial E_s^{int}(C_{-s})}{\partial C_{-s}} - \frac{1}{\Phi_s'(E_s^0)} \\ &= \frac{1}{\Phi_s'(E_s^{int}(C_{-s}))} - \frac{1}{\Phi_s'(E_s^0)} \geq 0, \end{aligned}$$

with the last inequality following from the concavity of Φ_s'. As a result, $h(.)$ is strictly increasing, which establishes the claim. ■

Lemma 3 For a given spend vector $[E_s^0, C_{-s}^1]$, let channel s have medium or high leverage such that $G[E_s^0, C_{-s}^1] < G_s^2$ (low or medium momentum). Then, if it is optimal to increase channel s investment, and

a) the optimal expenditure is obtained from the root of H_s^2, an arbitrarily small increase in C_{-s} results in reduction in the optimal channel s investment.

b) the optimal expenditure is obtained from a root of H_s^1, an arbitrarily small increase in C_{-s} results in an increase in the optimal channel s investment.

Proof: a) In this case, we should have $H_s^2(E_s^U, C_{-s}^1) \geq 0$. As $H_s^2(E_s^U, C_{-s})$ is increasing in C_{-s}, we can find an arbitrarily small increase in the level of cumulative marketing effort for all channels except s, $C_{-s}^2 > C_{-s}^1$, such that it would remain optimal to increase channel s investment (by continuity of the profit function), $H_s^2(E_s^U, C_{-s}^2) \geq 0$, and we still have $G[E_s^0, C_{-s}^2] < G_s^2$. Also let $E_s^{int}(C_{-s}^1)$ and $E_s^{int}(C_{-s}^2)$ be the corresponding roots of H_s^2 (which correspond to the optimal channel s investments) for cumulative marketing efforts of C_{-s}^1 and C_{-s}^2 respectively. As H_s^2 is increasing in E_s and $E_s^{int}(C_{-s}^1)$ is the root of $H_s^2(., C_{-s}^1)$, we should have $H_s^2(E_s, C_{-s}^1) \geq 0$ for all $E_s \geq E_s^{int}(C_{-s}^1)$. Therefore, for all $E_s \geq E_s^{int}(C_{-s}^1)$ we should have $H_s^2(E_s, C_{-s}^2) \geq H_s^2(E_s, C_{-s}^1) \geq 0$, implying that $H_s^2(., C_{-s}^2)$ cannot have a root over the interval $[E_s^{int}(C_{-s}^1), E_s^U]$. In other words, we should have the root of $H_s^2(., C_{-s}^2)$ such that $E_s^{int}(C_{-s}^2) \leq E_s^{int}(C_{-s}^1)$, establishing the decreasing property.

b) In this case, since H_s^1 has a root, H_s^2 should not have a root as $H_s^2(E_s^U, C_{-s}^1) < 0$ by Lemma 1. Now we consider what happens in each of the two cases that $G[E_s^0, C_{-s}^1]$ is above or below G_s^1.

If $G_s^1 \leq G[E_s^0, C_{-s}^1] \leq G_s^2$ (medium momentum channel), we know that profit increases as the level of channel s investment increases infinitesimally from E_s^0 (i.e. $\frac{\partial \Pi}{\partial E_s}(E_s^0, E_{-s}^1) \geq 0$ where E_{-s}^1 is the vector of expenditures in all channels except s resulting in C_{-s}^1). Therefore, as E_s increases, the first point setting $\frac{\partial \Pi}{\partial E_s}(., E_{-s}^1)$ to zero should be a local maximizer. From the other hand, the assumptions in Section 4.2 guarantee that the profit function can have at most one local minimum in addition to the local maximum. However, if a local minimum exists (which should be after the local maximum), the property of the profit function to be such that $\lim_{E_s \to \infty} \Pi(E_s, E_{-s}^1) = -\infty$ cannot be satisfied. Therefore, in this case, the profit function has a single critical point which is a local maximum. This consequently means that H_s^1 has only one root as H_s^2 does not have any root. In addition, since $H_s^1(E_s^0; C_{-s}^1) = G(C^1) - G_s^1 > 0$, $H_s^1(E_s; C_{-s}^1) > 0$ for all $E_s \in [E_s^0, E_s^{int}(C_{-s}^1))$, where $E_s^{int}(C_{-s}^1)$ is the root of $H_s^1(., C_{-s}^1)$. Now consider an arbitrarily small increase in the level of cumulative marketing effort for all channels except s, $C_{-s}^2 > C_{-s}^1$, such that with C_{-s}^2 we still have $G[E_s^0, C_{-s}^2] \leq G_s^2$ and H_s^1 continues to have a single root denoted by $E_s^{int}(C_{-s}^2)$ (achievable by continuity of the profit function and H_s^1). As H_s^1 is increasing in C_{-s}, we should have $H_s^1(E_s; C_{-s}^2) > 0$ for all $E_s \in [E_s^0, E_s^{int}(C_{-s}^1))$. Therefore, the root of $H_s^1(.; C_{-s}^2)$ cannot happen over the interval $[E_s^0, E_s^{int}(C_{-s}^1))$, implying that $E_s^{int}(C_{-s}^2) \geq E_s^{int}(C_{-s}^1)$.

But if $G[E_s^0, C_{-s}^1] < G_s^1$ (low momentum channel), we know that profit decreases as the level of channel s investment increases infinitesimally from E_s^0 (i.e. $\frac{\partial \Pi}{\partial E_s}(E_s^0, E_{-s}^1) \leq 0$ where E_{-s}^1 is the vector of expenditures in all channels except s resulting in C_{-s}^1). Therefore, as E_s increases, the first point setting $\frac{\partial \Pi}{\partial E_s}(., E_{-s}^1)$ to zero should be a local minimizer. From the other hand, as $\lim_{E_s \to \infty} \Pi(E_s, E_{-s}^1) = -\infty$ along with the two conditions laid out in Section 4.2, we can conclude that the profit function has exactly one local maximum and one local minimum. This consequently means that H_s^1 has exactly two roots (as H_s^2 does not have any) with the larger one representing the local maximizer $E_s^{int}(C_{-s}^1)$. In addition, as $H_s^1(E_s^0; C_{-s}^1) = G(C^1) - G_s^1 \leq 0$, H_s^1 is positive between its two roots. Now consider an arbitrarily small increase in the level of cumulative marketing effort for all channels except s, $C_{-s}^2 > C_{-s}^1$, such that with C_{-s}^2 we still have $G[E_s^0, C_{-s}^2] < G_s^1$ and H_s^1 continues to have two roots with the larger one denoted by $E_s^{int}(C_{-s}^2)$ again, achievable by continuity of the profit function and H_s^1). Let e be a level of channel s investment that is less than both $E_s^{int}(C_{-s}^1)$ and $E_s^{int}(C_{-s}^2)$, which makes both $H_s^1(e; C_{-s}^1) > 0$ and $H_s^1(e; C_{-s}^2) > 0$; this level of channel s investment exists by continuity of H_s^1 and the fact that C_{-s}^2 is only infinitesimally larger than C_{-s}^1. With this construction, we would then have $H_s^1(E_s; C_{-s}^1) \geq 0$ for $E_s \in [e, E_s^{int}(C_{-s}^1)]$ and $H_s^1(E_s; C_{-s}^2) \geq 0$ for $E_s \in [e, E_s^{int}(C_{-s}^2)]$. Now look at the interval $[e, E_s^{int}(C_{-s}^1)]$; as H_s^1 is increasing in C_{-s}, we should have $0 \leq H_s^1(E_s; C_{-s}^1) < H_s^1(E_s; C_{-s}^2)$ for all $E_s \in [e, E_s^{int}(C_{-s}^1))$. Therefore, the root of $H_s^1(.; C_{-s}^2)$ cannot happen over the interval $[e, E_s^{int}(C_{-s}^1))$, implying that $E_s^{int}(C_{-s}^2) \geq E_s^{int}(C_{-s}^1)$, establishing the increasing property. ■

Recall definition of H_s^1 and H_s^2 functions as well as the critical quantities C_{-s}^{trans}, C_{-s}^{LM}, C_{-s}^{peak} and C_{-s}^{max}:

$$\begin{cases} H_s^1(E_s; C_{-s}) = G(\Phi_s(E_s) + C_{-s}) - \left[\frac{1}{2}(1 - \frac{p}{q}) - \frac{1}{2q}\sqrt{(q+p)^2 - 4q\frac{1}{P\Phi_s'(E_s)}}\right] \\ H_s^2(E_s; C_{-s}) = G(\Phi_s(E_s) + C_{-s}) - \left[\frac{1}{2}(1 - \frac{p}{q}) + \frac{1}{2q}\sqrt{(q+p)^2 - 4q\frac{1}{P\Phi_s'(E_s)}}\right] \end{cases}$$

$$\begin{aligned} C_{-s}^{trans} &= \inf\{C_{-s} \geq C_{-s}^0 \mid E_s^*(C_{-s}) > E_s^0\} \\ C_{-s}^{LM} &= \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^0)) \geq G_s^1\} \\ C_{-s}^{peak} &= \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^U)) \geq \frac{1}{2}(1 - \frac{p}{q})\} \\ C_{-s}^{\max} &= \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^0)) \geq G_s^2\} \end{aligned}$$

Theorem 7 (Section 4.2.1): Suppose the classification of channel s corresponding to the initial spend vector E^0 is $R^{M/L}$. The quantities $C_{-s}^{trans}, C_{-s}^{LM}, C_{-s}^{peak}$ and C_{-s}^{max} must exist with $C_{-s}^0 \leq C_{-s}^{trans} \leq C_{-s}^{peak} < C_{-s}^{max}$ and $C_{-s}^{trans} \leq C_{-s}^{LM} < C_{-s}^{max}$. As C_{-s} increases to C_{-s}^{LM}, channel s transitions to medium momentum category, and as it further increases to C_{-s}^{max}, it transitions to the high momentum category. Moreover,

1. $E_s^*(C_{-s}) = E_s^0$ for $C_{-s} \in [C_{-s}^0, C_{-s}^{trans})$
2. $E_s^*(C_{-s})$ is increasing smoothly in C_{-s}, for $C_{-s} \in C_{-s}^{trans}, C_{-s}^{peak}]$ reaching a maximum value of E_s^U at C_{-s}^{peak}. Therefore in this range, the interaction of s with others is dominantly synergistic.
3. $E_s^*(C_{-s})$ is decreasing smoothly in C_{-s} for $C_{-s} \in C_{-s}^{peak}, C_{-s}^{max}]$ reaching its minimum value of E_s^0 at C_{-s}^{max}. Therefore in this range, the interaction of s with others is dominantly substitutive.
4. $E_s^*(C_{-s}) = E_s^0 \; for \; C_{-s} \geq C_{-s}^{max}$.

Proof: First note that as channel s has a low momentum, we should have $G[E_s^0, C_{-s}^0] \leq G_s^1$. As C_{-s} increases, $G[E_s^0, C_{-s}]$ increases accordingly, while the two threshold levels G_s^1 and G_s^2 remain unchanged. Therefore, the two quantities $C_{-s}^0 \leq C_{-s}^{LM} \leq C_{-s}^{max}$ must exist by their definition and the increasing property of $G(.)$. Also, with increase in C_{-s}, channel s transitions from low to medium momentum when C_{-s} reaches C_{-s}^{LM}, and to high momentum when C_{-s} reaches C_{-s}^{max} by the definition of C_{-s}^{LM} and C_{-s}^{max}.

Based on Lemma 2 and the initial low momentum category of channel s, C_{-s}^{trans} exists in range $[C_{-s}^0, C_{-s}^{max}]$, and it is optimal to increase investment in channel s more than E_s^0 for $C_{-s} \in (C_{-s}^{trans}, C_{-s}^{\max})$. Also, Lemma 1 (and its proof) specifies that the optimal level of increase in investment depends on the sign of $H_s^2(E_s^U, C_{-s})$, specially as here we have $E_s^0 < E_s^U < \infty$; if $H_s^2(E_s^U, C_{-s}) \geq 0$, the optimal investment is found from the root of H_s^2, but otherwise, it is obtained from the larger root of H_s^1. But note that $H_s^2(E_s^U, C_{-s})$ is increasing in C_{-s}. Therefore, if for some level of C_{-s}, $H_s^2(E_s^U, C_{-s})$ is non-negative, it would remain to be so for larger levels of C_{-s}. By definition, the level C_{-s}^{peak} essentially represents the level of C_{-s} which determines when $H_s^2(E_s^U, C_{-s})$ changes sign and sets $H_s^2(E_s^U, C_{-s}^{peak}) = G\big(C_{-s}^{peak} + \Phi_s(E_s^U)\big) - \frac{1}{2}(1 - \frac{p}{q}) = 0$. Based on the continuity of H_s^2, C_{-s}^{peak} should exist in the range $[C_{-s}^{trans}, C_{-s}^{\max}]$. Therefore when $C_{-s}^{\max} > C_{-s} \geq C_{-s}^{peak}$, the optimal channel s investment is obtained from the root of H_s^2 and based on Lemma 3(a), as the level of C_{-s} increases, the optimal channel s investment decreases. With increase in C_{-s}, $H_s^2(E_s^0, C_{-s}) = G(\Phi_s(E_s^0) + C_{-s}) - G_s^2$ increases as well until $H_s^2(E_s^0, C_{-s}) > 0$ and H_s^2 does not have a root any more which coincides with channel s switching to high momentum making the offering of E_s^0 optimal. However, when $C_{-s}^{trans} < C_{-s} < C_{-s}^{peak}$, the optimal channel s spend is obtained from a root of H_s^1 and based on Lemma 3(b), the optimal spend is increasing in C_{-s}.

Lastly, we need to show that C_{-s}^{LM} is in the range $[C_{-s}^{trans}, C_{-s}^{max}]$. Note that $C_{-s}^{LM} \geq C_{-s}^{trans}$ since if channel s has medium momentum for a level of C_{-s}, it is optimal to increase investment in channel s to more than E_s^0, and the existence of such an interior optimal solution means that it should exist in the first place, establishing that the transition to medium momentum has to happen on or after the case that the interior optimal solution exists. Also this transition has to happen before the channel switches to medium momentum category by the continuity of $G[E_s^0, C_{-s}]$ in C_{-s}, which establishes the claim. ■

Theorem 8 (<u>Section 4.2.2</u>): Suppose the classification of channel s corresponding to the initial spend vector E^0 is $R^{H/L}$. Let C_{-s}^{trans}, C_{-s}^{LM}, and C_{-s}^{max} be defined as in Theorem 11. These quantities must exist with $C_{-s}^{0} \leq C_{-s}^{trans} = C_{-s}^{peak} \leq C_{-s}^{LM} < C_{-s}^{max}$. As C_{-s} increases to C_{-s}^{trans}, channel s transitions to medium momentum category, and as it further increases to C_{-s}^{max}, it transitions to the high momentum category. Moreover,

1. $E_s^*(C_{-s}) = E_s^0$ for $C_{-s} \in [C_{-s}^0, C_{-s}^{trans})$.
2. $E_s^*(C_{-s})$ is maximized at C_{-s}^{trans}, and decreases smoothly for $C_{-s} \in [C_{-s}^{trans}, C_{-s}^{max}]$ until reaching its minimum value of E_s^0 at C_{-s}^{max}.
3. $E_s^*(C_{-s}) = E_s^0$ for $C_{-s} \geq C_{-s}^{max}$.

Proof: Similar to the proof above for Medium-Leverage Channels, with increase in C_{-s}, channel s transitions from low to medium to high momentum, with the transitions happening at C_{-s}^{LM} and C_{-s}^{max} by their definition. Now based on Lemma 2 and the momentum-based categorization, the level of C_{-s}^{trans} exists and it is optimal to increase investment in channel s for $C_{-s} \in (C_{-s}^{trans}, C_{-s}^{\max})$ but not before. However, for C_{-s} in the range $(C_{-s}^{trans}, C_{-s}^{\max})$, we know that according to Lemma 2, especially that $E_s^U = \infty$, an interior maximizer exists and is obtained from the root of H_s^2. Based on Lemma 3(a), we can conclude that until when $C_{-s} \in (C_{-s}^{trans}, C_{-s}^{\max})$, the optimal channel s investment is decreasing in C_{-s}, but for $C_{-s} > C_{-s}^{\max}$ it is optimal to suffice to the initial investment level E_s^0 given that in that range the channel has high momentum. Given that over this range, the maximum channel s spending happens at C_{-s}^{trans}, by the definition of C_{-s}^{peak} we should have $C_{-s}^{peak} = C_{-s}^{trans}$.

Next, we investigate how the channel transitions between momentum categories. Similar to the proof for medium leverage channels above and by the continuity of $G[E_s^0, C_{-s}]$ in C_{-s}, the transition from medium to high momentum happens at $C_{-s}^{\max}$, and the transition from low to medium momentum occurs on or after C_{-s}^{trans} ($C_{-s}^{LM} \geq C_{-s}^{trans}$). ■

Appendix C: Bayesian Estimation for Camera Data Application & Results

Here we present the method for extracting the full posterior distribution of the 12 parameters of the model in our empirical application to camera sales. Channel 1 is expenditures via free-standing inserts (FSI) in flyers; Channel 2 is radio. Specifically, we have $\phi_r^i(a) = \alpha_r\, \delta_r^i\, a^{\rho_r}$ for months $i = 0,1, ...,$ channels $r = 1,2$, leading to six parameters for relative channel effectiveness α_r, exponent ρ_r, and advertising remembering rate δ_r. We also must estimate the diffusion parameters p, q and m. Lastly, because advertising data start several months after the initial product offering, we estimate three additional parameters: x_0 (initial adoption proportion), and $\{S_1(0), S_2(0)\}$ (initial advertising goodwill levels).

Estimation results will be presented in the following order for the 12 estimated quantities:

Parameter	Definition
p	Product innovation rate
q	Product imitation rate
m	Potential market size
$\{\alpha_1, \alpha_2\}$	Effectiveness in Channel {1,2}
$\{\rho_1, \rho_2\}$	Exponent in Advertising function for Channel {1,2}
$\{\delta_1, \delta_2\}$	Remember rate for Channel {1,2}
x_0	Fraction of adoptions (market share) at time 0
$\{S_1(0), S_2(0)\}$	Initial Ad level for Channel {1,2}

The model described in Section 3.4 for Sales gives rise to a (log-)likelihood function for the 12 parameters listed above. A target posterior is obtained by adding in log-priors. Because we do not wish to impose conjugacy, we choose relatively flexible priors that conform to the domains of each of the parameters, e.g., Beta for those bounded on the unit interval, gamma for non-negative parameters, etc. In all cases, we aim for relatively non-informative priors, avoiding those that place too much posterior mass either in the center or edges for (half-)bounded parameters, and whose first moments do not stray too far from the empirical values reported in Van den Bulte & Stremersch (2004). We empirically tune the priors to ensure that the measure of model fit, the posterior standard deviation – which is calculated in the Bayesian fashion at each pass of the sampler – is minimized. The results we present, with lowest posterior standard deviation, corresponded to the least informative prior settings of approximately 20 sets run.

The sampler was set to use a M-H proposal based on the given log-posterior, with each likelihood evaluation coded into MATLAB to compute quickly. All model results are based on 1 million draws burn-in and 1 million draws for inference, thinned by 50, with a random walk Metropolis-Hastings proposal and pre-tuned stepsizes that were held constant across all models. Specific prior settings and stepsizes were as follows, and we note in passing that estimation results were not especially sensitive to the latter:

Parameter	Prior	Step Size for M-H
$\{p, q\}$	$\text{Beta}(1.5, 1.5)$	0.001
m	$\text{Gamma}(14.0, 500)$	50
$\{\alpha_1, \alpha_2\}$	$\text{Gamma}(1.5, 1.5)$	0.001
$\{\rho_1, \rho_2\}$	$\text{Beta}(1.5, 1.5)$	0.010
$\{\delta_1, \delta_2\}$	$\text{Beta}(1.5, 1.5)$	0.010
x_0	$\text{Beta}(1.5, 3.0)$	0.010
$\{S_1(0), S_2(0)\}$	$\text{Gamma}(2.0, 0.1)$	0.010

We note in passing that the prior for m is set to have a data-specific mean of 7,000, but high variance, so is relatively noninformative across the possible range of values of m.

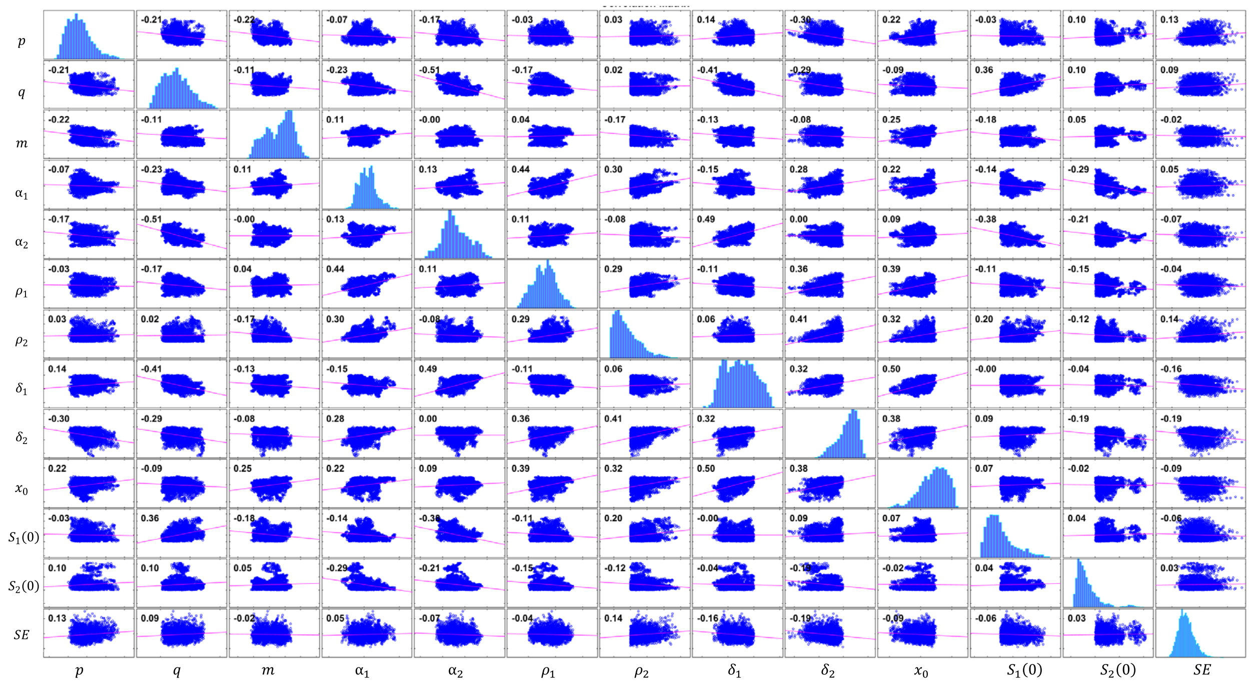
Estimation Results

The sampler appeared to converge well, with Effective Sample Sizes above 4000 for all parameters, nearly equal summary statistics for the first and second half of the draws used for inference, and split chain Gelman-Rubin statistics near 1. Moreover, as shown (far) below, empirical correlation between parameter pairs were not extreme, with all squared correlations well below 0.2, nearly unimodal marginal densities, and an essentially bell-shaped log-SE histogram.

Summary statistics for the parameters themselves were given as follows, with mean, standard error, effective sample size, lower (0.025) median (0.500), and upper (0.975) quantiles for the posterior credible intervals:

Parameter	Mean	StdErr	ESS	Lower	Median	Upper
p	0.0391	0.0172	4037	0.0118	0.0362	0.0738
q	0.5445	0.0489	4004	0.4782	0.5376	0.6471
m	6738.9	1963.7	4007	3593.3	6918.9	10225.9
α_1	0.0124	0.0081	4122	0.0030	0.0103	0.0371
α_2	0.0081	0.0058	4228	0.0010	0.0069	0.0223
ρ_1	0.3777	0.1493	4048	0.0990	0.3816	0.6507
ρ_2	0.2722	0.1684	4036	0.0545	0.2330	0.6952
δ_1	0.4038	0.1497	4050	0.1232	0.4068	0.7065
δ_2	0.3602	0.1632	4041	0.0745	0.3697	0.6354
x_0	0.0808	0.0458	4297	0.0119	0.0758	0.1837
$S_1(0)$	1.5781	0.5436	4004	0.6735	1.6771	2.3481
$S_2(0)$	1.4068	0.3332	4010	0.6309	1.5054	1.8763
$StdErr$	16.0890	2.5794	5615	12.1743	15.8109	21.7919

Correlation Plot for 12 estimated parameters and model StdErr



Reference

Van den Bulte, C., S. Stremersch. (2004). Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Science*, *23*(4), 530-544.

Appendix D: Impact of Changes in Demand Response to Advertising on Channel Interactions – The Case of Camera Sales

In this section, we return to the case studied in Sections 3.4 and 4.3, and explore the dependence of nature of channel interaction on the nature of demand response to advertising. Specifically, we explore sensitivity of channel interactions with respect to two factors: the exponent term ρ_2 representing how easily Channel 2 can substitute investment in the remaining Channel 1, and inclusion of explicit interaction term in demand response to advertising to explicitly enforce substitution or synergy (as has been done in some existing research such as Naik & Raman (2003)) in addition to the natural arising of these factors in our setup.

Sensitivity of Channel Interactions to ρ_2

Our cumulative marketing effort functions (8) resemble the well-known family of utility functions with constant elasticity of substitution (CES) in economics, where ρ_s is used as the measure of how easily the investment in channel s substitutes with others. In our context for Channel 2, ρ_2 can be interpreted as the strength of the substitutive interaction of Channel 2 with the remaining Channel 1: as ρ_2 increases, the tendency for Channel 2 to substitute investment in Channel 1 becomes stronger; in fact, at $\rho_2 = 1$ the function Φ_2 becomes linear and the interaction of Channel 2 with Channel 1 becomes entirely substitutive. In Section 4.3, we illustrated the results based on the estimated value of $\rho_2 = 0.2722$, but in this section, we explore how the change in the value of ρ_2 impacts the nature of interaction between the two channels.

In the sensitivity analysis, simply increasing ρ_2 in (8) is not very informative, as it creates two counter-effects: on the one hand, the strength of the substitution of Channel 2 with Channel 1 increases; but on the other, the marketing effectiveness of Channel 2 also increases, making comparisons across different ρ_2 values problematic. In order to isolate the substitution effect, we re-scale the effectiveness coefficient β_2 so that investing 1% of possible revenue per target customer in Channel 2 (i.e., setting $E_2 = 0.01$) achieves the same level of effectiveness for any ρ_2 compared to when it is set to its estimated value of $\rho = 0.2722$. Note that the investment level of $E_2 = 0.01$ is set to be higher than the typical level of investments in Channel 2 for various values of ρ_2, and it is in fact a bit larger than maximum level of Channel 2 investment ($E_2^U = 0.0098$) when $\rho = 0.2722$. The optimal expenditure $E_2^*(E_1)$ on Channel 2 with respect to E_1 is depicted in Figure 6 for six representative values of ρ_2 along with the estimated value. This figure shows the interaction pattern for two settings that was considered in Section 4.3: the regular setting in which the media planning horizon starts with ending conditions after 28 months, as well setting as if the media planning horizon had started 28 months ago.

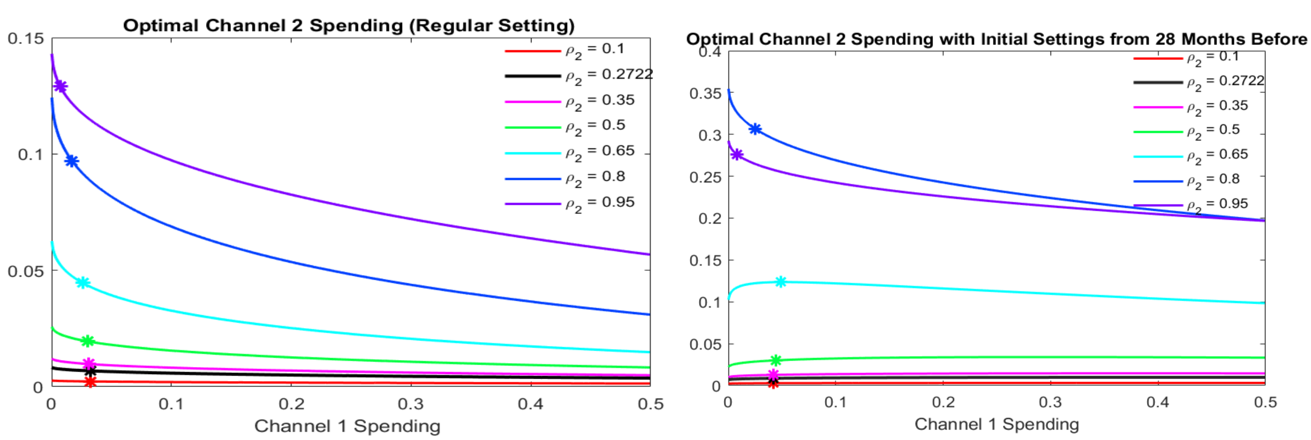


Figure 1: Interaction pattern of Channel 2 with Channel 1 for various values of ρ_2 – with settings after the initial 28 months (left panel), and from before the initial 28 months (right panel); * - corresponding optimal spend plan

A number of key observations emerge as ρ_2 increases from 0.1 to 0.95. In the left graph of Figure 6, the interaction of the two channels is mainly substitutive for $\rho_2 = 0.2722$ as observed in Section 4.3, and with stronger substitutive tendency of Channel 2, the main substitutive pattern is maintained for larger values of ρ_2 as well. In the right graph where both synergistic and substitutive interaction can be observed for $\rho_2 = 0.2722$, both types of interaction can still be observed for lower values of ρ_2, but with larger values of ρ_2 (0.8 and 0.95), the synergy region disappears entirely as the substitutive strength of Channel 2 increases. Moreover, both graphs reveal that the optimal Channel 2 spend curves are flatter for lower values of ρ_2 and become increasingly more peaked at higher values. Here flatter curves correspond to a larger region where it is optimal to spend above allocated amounts on both channels; in the right graph, the range where the spend curve is increasing (indicating a dominantly synergistic interaction) is also larger for smaller values of ρ_2. This is quite intuitive: as ρ_2 decreases, it becomes harder for Channel 2 to substitute Channel 1, leading to more synergistic interactions.

It is also interesting to observe the impact on the optimal spend plan with change in ρ_2 (the * points on the graphs). In the left figure that the interaction is mainly substitutive, Channel 1 optimal spending (i.e., the x-coordinate of the $*$ points) is decreasing in ρ_2 and the optimal Channel 2 spending is increasing. This indicates that with when Channel 2 can more easily substitute Channel 1, it is beneficial to contribute more on this channel and cut back on the other. In the right graph, however, the change in the optimal spend plans is non-monotonic in ρ_2, and for small values of ρ_2 they fall on the synergistic region while for larger values they are mainly in the substitutive region of interaction. This suggests that factors determining the optimal spend level are relatively complex even in this case with only two channels.

Inclusion of explicit interaction term in demand response to advertising

As discussed earlier, our demand adoption model is relatively general and allows for both substitutive and synergistic interaction between channels. However, substitution and synergy, unlike prior literature, are not deliberatively "built into" the model, but arises naturally from the native GBM setting, in a way not anticipated by prior literature. In this section, we illustrate for the camera sales case of sections 3.4 and 4.3 as to who the explicit addition of substitution or synergy into the model would impact the nature of channel interactions. To capture this, we follow the form of inclusion of such effects in Naik & Raman (2003) and Naik & Peters (2009) through addition of an explicit interaction term in the form of demand response to advertising. Therefore, the total market adoption over the media planning horizon in (2) is replaced with the following:

$$F(T) = G(\Phi_1(E_1) + \Phi_2(E_2) + \omega\Phi_1(E_1)\Phi_2(E_2)).$$

The above demand form reduces to the one considered in this paper when $\omega = 0$, it explicitly enforces substitutive interaction between channels when $\omega < 0$, and it explicitly induces synergy between channels when $\omega > 0$. However, notice that addition of ω also influences the level of total cumulative marketing effort and demand adoption. To isolate the impact of explicit enforcement of substitution and synergy, we re-scale the effectiveness coefficients β_1 and β_2 so that investing 10% of per-target-customer revenue into each of the channels (i.e., setting $E_1 = E_2 = 0.1$) achieves the same level of effectiveness for any ω compared to when $\omega = 0$. The optimal expenditure $E_2^*(E_1)$ on Channel 2 with respect to E_1 is depicted in Figure 7 for various representative values of ω for two cases similar to above: the regular setting in which the media planning horizon starts with ending conditions after 28 months, as well setting as if the media planning horizon had started 28 months ago.

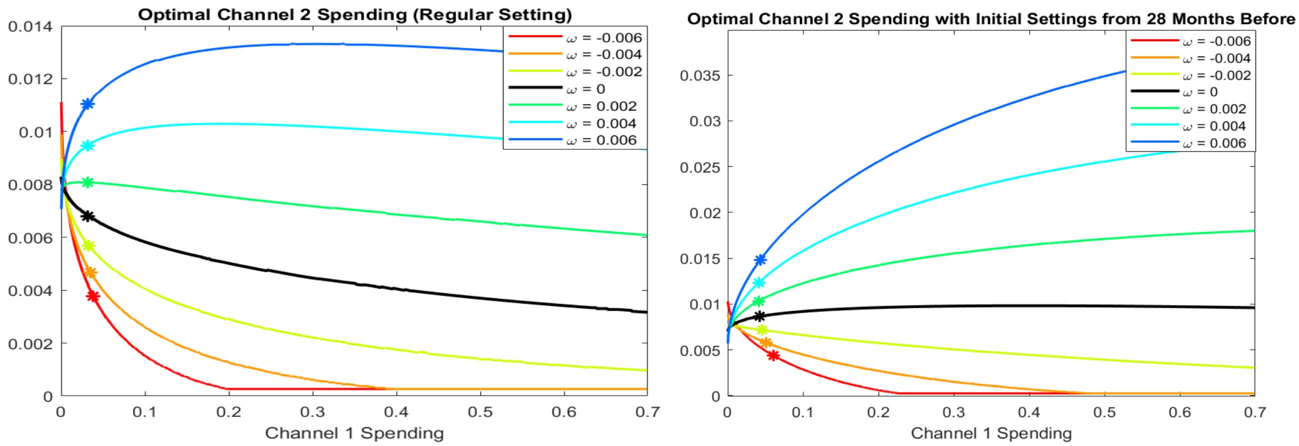


Figure 2: Interaction pattern of Channel 2 with Channel 1 for various values of ω – with settings after the initial 28 months (left panel), and from before the initial 28 months (right panel); * - corresponding optimal spend plan

Figure 7 reveals that when an explicit substitutive or synergistic interaction is imposed on the nature of demand response to advertising, the optimal pattern of interaction between channels would turn out to be more biased towards that enforced pattern. In both right and left panels of this figure it can be seen that when $\omega > 0$ (building synergy into the model), the increasing trend of Channel 2 investment happens for a larger range of Channel 1 spending and the Channel 2 investment increases more rapidly. In contrast, when $\omega < 0$ (building substitution into the model), the decline in Channel 2 spending happens more sharply and for a larger range of Channel 1 spending, until Channel 1 spending falls to the minimum allocated Channel 2 spending of $E_2^0 = 0.0002$. Interestingly, even when either of the two forces are build into the model, the other type of interaction is not entirely eliminated. For example, when $\omega = 0.004$ in the left panel, synergistic interaction is moderately built into the model, but still a declining pattern of optimal Channel 2 investment can be observed that is indicative of an overall substitutive interaction between channels in the media plan. In contrast, when $\omega = -0.002$ in the right panel, channels are forced to moderately substitute one another, but still for low values of Channel 1 spending, the optimal Channel 2 investment is increasing, suggesting a synergistic pattern of interaction. In summary, this analysis highlights that even when intrinsic substitutive or synergistic tendencies between channels are built into the model, the need for analyzing resulting “media planning interactions” between channels remain, as carried out in this paper. These media planning interactions in essence represents the overall outcome of various forces in play such as customer word of mouth, intrinsically assumed channel interactions, and allocated/prior advertising expenditures.

Appendix E: Sensitivity of Optimal Allocation to Relative Channel Cost

The analysis in Section 4 provided typologies to guide judicious managerial action in different channel settings, but held aside the impact of changes in channel costs; here, we examine the role of such cost changes on the optimal allocation of marketing expenditures. Recall that the quantity E_s represents the total expenditure on channel s per target customer and $\Phi(E_s)$ is the resulting cumulative effectiveness over the time horizon T. While we have used the same monetary yardstick for both and applied it across all channels, in practice marketers typically use channel-specific measures of "currency" when deciding on the level of usage of a particular channel. For example, the standard measure of currency for television is GRP (gross rating points); a marketing plan will usually specify the number of GRPs that should be purchased for a particular TV category (local, national, specialty, etc.) during a given time period. Of course, the cost of one GRP may be quite different depending on the channel and program type. Similarly, the common currency for online media is "impressions", for print media it is the "subscriber base", etc. Thus, to differentiate channel-specific currency unit from the actual dollar amount required to acquire it, we will introduce scaling coefficients ε_s for $s \in R$, and assume that $\varepsilon_s E_s$ dollars are required per target customer to purchase E_s units of channel-specific currency for channel s, which will result in the cumulative effectiveness of $\Phi(E_s)$. A small (large) value of ε_s implies that channel s is inexpensive (expensive) relative to the other channels. We thus treat ε_s as the "relative cost" of channel s. This leads to the following modification of the Marketing Effectiveness Allocation (MEA) problem:

$$\max_{E_r, r\in R} \Pi = mP\left[G\left(\sum_{r\in R} \Phi_r(E_r)\right) - x_0\right] - m\sum_{r\in R} \varepsilon_r E_r, \text{ s.t. } \quad E_r \in [0, b_r T], \; r \in R \quad \text{(Modified MEA)}$$

Obviously, when $\varepsilon_s = 1$ for all $s \in R$, the problem above reduces to the original (MEA) problem. The above revised problem also resembles the Lagrangian relaxation of the original problem when budgetary constraints on the total spend of some or all channels are added to the model. Therefore, ε_s can alternatively be interpreted as the "shadow price" of the spend in channel s when budgetary limitations are present.

Replicating the analysis for the MEA problem, it is straightforward to revise the channel classifications and the definitions of the quantities $E_s^U(\varepsilon_s)$, $C_{-s}^{trans}(\varepsilon_s)$, $C_{-s}^{peak}(\varepsilon_s)$, and $C_{-s}^{max}(\varepsilon_s)$ (which now depend on ε_s) for the Modified MEA problem. We first analyze the effect of increasing the relative cost ε_s on these quantities to understand how it impacts the channel interactions. **Theorem 9** below characterizes the solution of Modified MEA as follows:

Cost Effects for Modified MEA: As ε_s increases, $E_s^U(\varepsilon_s)$ and $C_{-s}^{max}(\varepsilon_s)$ cannot increase, while $C_{-s}^{trans}(\varepsilon_s)$ cannot decrease. Moreover, if channel s has medium-leverage for some value of ε_s and remains so for an increase in ε_s, the value of $C_{-s}^{peak}(\varepsilon_s)$ cannot decrease.

This result states that, as channel s becomes relatively more expensive, its maximal spend E_s^U cannot increase, and stronger support from all other channels is required for s to reach its optimal spend. Moreover, the range $[C_{-s}^{trans}(\varepsilon), C_{-s}^{max}(\varepsilon)]$, where the optimal spend for channel s exceeds the allocated level E_s^0, is shrinking in ε_s, as it becomes more challenging for the firm to additionally spend in a more expensive channel. In addition, the interval where the interaction is dominantly substitutive is shrinking (for medium-leverage channels, this interval is $[C_{-s}^{peak}, C_{-s}^{\max}]$, while for high-leverage channels it is $[C_{-s}^{trans}, C_{-s}^{\max}]$). This means that the added cost burden of channel s limits the appropriate development of demand momentum, and leads the firm to rely more on the support from the other channels.

The next result (see **Theorem 10** below) considers the impact of an increase in the relative cost of channel s on the optimal channel-level expenditures in the Modified MEA problem. We let the vector E^*, with components $E_r^* \geq E_r^0$ for $r \in R$, be an optimal solution to the MEA problem, and summarize the impact of an increase in the relative cost ε_s of channel s, while keeping all the other relative costs at their original level.

Optimal Investment for Modified MEA: The optimal investment E_s^* in channel s and the optimal profit Π^* are both non-increasing in ε_s. If $R^* = \{r \in R |\ E_r^* > E_r^0\}$ is the set of channels for which the optimal spends exceed allocated levels, then, for channels $s \neq r \in R^*$, optimal investments must satisfy the proportionality condition $\frac{\Phi_r'(E_r^*)}{\varepsilon_r} = \frac{\Phi_s'(E_s^*)}{\varepsilon_s}$.

Therefore, if ε_s increases so that the set R^* does not change, the optimal investment in channel r decreases if and only if this ratio increases in ε_s.

The Optimal Investment characterization for the modified MEA confirms the intuition that, when a channel becomes relatively more expensive, its usage cannot increase, resulting in a decline in the optimal profit. It then specifies that, when the firm (optimally) sticks to its media plan – that is, the choice of channels to further utilize does not change – the firm must balance marginal effectiveness of chosen channels proportional to their relative cost, measured by $\frac{\Phi_r'(E_r^*)}{\varepsilon_r}$. Note that this simplified condition does not involve the dynamics of demand penetration, reflected in the function $G(\cdot)$ and product price: if this measure increases for channel s, the optimal spend for any other channel in R^* must decrease (this follows from the concavity of Φ_r), assuming the set R^* does not change. However, whether $\frac{\Phi_s'(E_s^*)}{\varepsilon_s}$ increases or decreases with ε_s is not certain, and depends on the properties of Φ_s and the original optimal spend vector E^*. Thus, the impact of increase in ε_s on the optimal spend of other channels in R^* cannot be predicted in general.

Theorems & Proofs of Appendix E:

Theorem 9: As ε_s increases, $E_s^U(\varepsilon_s)$ and $C_{-s}^{max}(\varepsilon_s)$ cannot increase, while $C_{-s}^{trans}(\varepsilon_s)$ cannot decrease. Moreover, if channel s has medium-leverage for some value of ε_s and remain so for an increase in ε_s, the value of $C_{-s}^{peak}(\varepsilon_s)$ cannot decrease.

Proof: Adding the parameter ε_s and following our analysis of Section 4.1, we redefine E_s^U as follows which coincides with the ones in Section 4.1 when $\varepsilon_s = 1$.

$$E_s^U(\varepsilon_s) = \begin{cases} E_s^0 & \text{if } \Phi_s'(E_s^0) < \frac{4q\varepsilon_s}{P(q+p)^2} \\ sup\left\{E_s \geq 0 | \Phi_s'(E_s) \geq \frac{4q\varepsilon_s}{P(q+p)^2}\right\} & \text{otherwise} \end{cases}.$$

Note that the proofs of all earlier claims hold true trivially with these new definitions when ε_s is added to the optimization problem. From the above, it is easily verifiable that when ε_s increases, E_s^U does not increase as the set of E_s values for which the inequality $\Phi_s'(E_s) \geq \frac{4q\varepsilon_s}{P(q+p)^2}$ holds becomes smaller.

Next, $C_{-s}^{\max}(\varepsilon_s) = \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^0)) \geq G_s^2(\varepsilon_s)\}$ is non-increasing in ε_s as G_s^2 is decreasing in ε_s. Also we show that $C_{-s}^{trans}(\varepsilon_s) = \inf\{C_{-s} \geq C_{-s}^0 \mid E_s^*(C_{-s}) > E_s^0\}$ is non-decreasing in ε_s. If with an increase in ε_s, an interior maximizer continues to exist, we show that the feasibility of optimally increasing channel s investment decreases with an increase in ε_s. To check this, recall the function $h(C_{-s})$ in the proof of Lemma 2 which when positive (over the range of C_{-s} values for which an interior maximizer exists) indicated that the interior local maximizer is optimal. With the addition of ε_s, the function becomes:

$$h(C_{-s}) = PG(\Phi_s(E_s^{int}(C_{-s})) + C_{-s}) - PG(\Phi_s(E_s^0) + C_{-s}) - \varepsilon_s\big(E_s^{int}(C_{-s}) - E_s^0\big).$$

Taking derivative of $h(C_{-s})$ with respect to ε_s we have:

$$\frac{\partial h(C_{-s})}{\partial \varepsilon_s} = PG'\big(\Phi_s(E_s^{int}) + C_{-s}\big)\Phi_s'\big(E_s^{int}\big)\frac{\partial E_s^{int}}{\partial \varepsilon_s} - \big(E_s^{int}(C_{-s}) - E_s^0\big) - \varepsilon_s\frac{\partial E_s^{int}}{\partial \varepsilon_s}.$$

Since $E_s^{int}(C_{-s})$ is the interior maximizer and sets $\frac{\partial \Pi}{\partial E_s} = 0$, we have

$PG'\big(\Phi_s(E_s^{int}) + C_{-s}\big)\Phi_s'\big(E_s^{int}\big) = \varepsilon_s$. Putting these together, we find that $\frac{\partial h}{\partial \varepsilon_s} = -\big(E_s^{int}(C_{-s}) - E_s^0\big) < 0$. This derivation indicates that if for some level of C_{-s}, investing at the level of interior maximizer in channel s is not optimal, then with an increase in ε_s it remains to be so as well and the optimal channel s remains at E_s^0. Therefore, the level of C_{-s}^{trans} cannot decrease.

Lastly, if channel s has medium-leverage for some value of ε_s and remains so for an increase in ε_s, the value of $C_{-s}^{peak}(\varepsilon_s) = \inf\{C_{-s} \geq C_{-s}^0 \mid G(C_{-s} + \Phi_s(E_s^U(\varepsilon_s))) \geq \frac{1}{2}(1 - \frac{p}{q})\}$ is well defined for both levels of ε_s. Since E_s^U is non-increasing in ε_s, the set of C_{-s} values satisfying the condition inside the brackets shrinks, leading to C_{-s}^{peak} to potentially increase. ■

Theorem 10: The optimal investment $\mathrm{E_s^*}$ in channel s and the optimal profit Π^* are both non-increasing in ε_s. Also, let $\mathrm{R}^* = \{\mathrm{r} \in \mathrm{R} \mid \mathrm{E_r^*} > \mathrm{E_r^0}\}$ be the set of channels for which the optimal spends exceed the allocated levels. For channels $\mathrm{s} \neq \mathrm{r} \in \mathrm{R}^*$, optimal investments must satisfy

$$\frac{\Phi_r'(E_r^*)}{\varepsilon_r} = \frac{\Phi_s'(E_s^*)}{\varepsilon_s}.$$

Therefore, if ε_s increases so that the set R^* does not change, the optimal investment in channel r decreases if and only if this ratio increases in ε_s.

Proof: Define $E^{*,\varepsilon}$ to be the vector of the optimal marketing spends when the relative cost of channel s is ε and $C^{*,\varepsilon}$ be the resulting cumulative marketing effort from spending $E^{*,\varepsilon}$ in all channels. In addition, define the total profit of strategy $E^{*,\varepsilon}$ as $\Pi_\varepsilon(E^{*,\varepsilon})$.

The optimal profit decreases in ε_s: Let $\varepsilon_s^1 < \varepsilon_s^2$ be two levels of ε_s. By the optimality of E^{*,ε_s^1} and E^{*,ε_s^2} and the linearity of $\Pi_{\varepsilon_s}(.)$ in ε_s, we have $\Pi_{\varepsilon_s^1}(E^{*,\varepsilon_s^1}) \geq \Pi_{\varepsilon_s^1}(E^{*,\varepsilon_s^2}) \geq \Pi_{\varepsilon_s^2}(E^{*,\varepsilon_s^2})$, which establishes the claim.

The optimal channel s spend decreases in ε_s: Note that $\frac{\partial \Pi_{\varepsilon_s}}{\partial E_s} = m[PG'(\Phi_s(E_s) + C_{-s})\Phi_s'(E_s) - \varepsilon_s]$. If $\frac{\partial \Pi_{\varepsilon_s}}{\partial E_s}$ is always negative (for $E_s \in [E_s^0, E_s^U]$), we know that it is not optimal to increase the level of investment of channel

s beyond E_s^0. In this case, when ε_s increases, $\frac{\partial \Pi_{\varepsilon_s}}{\partial E_s}$ remains negative and therefore the same level of E_s^0 remains optimal. Therefore, the claim is trivially valid.

However, if for some range of E_s values this derivative is positive, it should take the value of zero for some level E_s^{int} as we know that $\lim_{E_s \to \infty} \frac{\partial \Pi_{\varepsilon_s}}{\partial E_s} < 0$. From the other hand, we know from the assumptions of Section 4.2 (Structure of Channel Interactions) and the proof of Lemma 1 that $\frac{\partial \Pi_{\varepsilon_s}}{\partial E_s}$ can have at most two roots, with the larger of the possible two roots representing a (unique) local maximizer. We denote this local maximizer as E_s^{int,ε_s} to show its dependence on ε_s. Therefore, this local maximizer has the property that $\frac{\partial \Pi_{\varepsilon_s}(E_s,E_{-s})}{\partial E_s} < 0$ for all $E_s > E_s^{int,\varepsilon_s}$.

Now consider two levels of relative cost $\varepsilon_s^1 < \varepsilon_s^2$. By the above argument and the fact that $\frac{\partial \Pi_{\varepsilon_s}}{\partial E_s}$ is decreasing in ε_s we have $\frac{\partial \Pi_{\varepsilon_s^2}(E_s,E_{-s})}{\partial E_s} < \frac{\partial \Pi_{\varepsilon_s^1}(E_s,E_{-s})}{\partial E_s} < 0$ for all $E_s > E_s^{int,\varepsilon_s^1}$.This means that at the level ε_s^2, $\frac{\partial \Pi_{\varepsilon_s^2}(E_s,E_{-s})}{\partial E_s}$ cannot have a root over the interval $[E_s^{int,\varepsilon_s^1}, \infty)$, so the new local maximizer at level ε_s^2, $E_s^{int,\varepsilon_s^2}$ (which is the larger root of $\frac{\partial \Pi_{\varepsilon_s^2}(E_s,E_{-s})}{\partial E_s}$), should be smaller than $E_s^{int,\varepsilon_s^1}$. Also as established in the proof of Cost Effects for Modified MEA, when ε_s increases, it takes a larger level of C_{-s} for the local maximizer to be globally optimal. But as here we have fixed the level of investment in other channels, we can conclude that if at the level ε_s^1, $E_s^{int,\varepsilon_s^1}$ represents the optimal level of channel s investment, at the level ε_s^2 either it is optimal to increase the level of channel s investment to $E_s^{int,\varepsilon_s^2}$ or maintain the current investment of E_s^0; but if at the level ε_s^1, it is more profitable to offer E_s^0 over $E_s^{int,\varepsilon_s^1}$, the same would hold for the level ε_s^2. In either case, the optimal level of channel s does not increase with an increase in ε_s.

Balancing the spends in the optimal solution: We know from the KKT necessary conditions that at the optimal solution, we should have $\frac{\partial \Pi_{e_s}}{\partial E_r} \leq 0$ for all $r \in R$, where the equality holds for all channels with positive additional spend, i.e. $\frac{\partial \Pi_{e_s}}{\partial E_r} = 0$ for all $r \in R^*$. Therefore, for each $r \in R^*$ we have $\frac{\partial \Pi_{e_s}}{\partial E_s}(E^{*,\varepsilon_s}) = PG'(C^{*,\varepsilon_s})\Phi_s'(E_s^{*,\varepsilon_s}) - \varepsilon_s = 0$ or equivalently $\frac{1}{PG'(C^{*,\varepsilon_s})} = \frac{\Phi_r'(E_r^{*,\varepsilon_s})}{\varepsilon_r}$ with the right hand side to be the same for all channels that belong to R^*. Consequently, for any pair of channels $s, r \in R^*$ we can write $\frac{\Phi_s'(E_s^{*,\varepsilon_s})}{\varepsilon s} = \frac{\Phi_r'(E_r^{*,\varepsilon_s})}{\varepsilon_r}$. ■