# Stochastic Optimal Control for Online Seller under Reputational Mechanisms

Milan Bradonij'

Matthew Causley

Albert Cohen

*Article*

# Stochastic Optimal Control for Online Seller under Reputational Mechanisms

**Milan Bradonjić [1],\*, Matthew Causley [2] and Albert Cohen [3]**

[1] Mathematics of Networks and Systems, Bell Labs, 600 Mountain Avenue, Murray Hill, NJ 07974, USA

[2] Department of Mathematics, Kettering University, Flint, MI 48504, USA; E-Mail: mcausley@kettering.edu

[3] Department of Statistics and Probability and Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA; E-Mail: albert@math.msu.edu

**\*** Author to whom correspondence should be addressed; E-Mail: milan@research.bell-labs.com; Tel.: +1-908-582-2661.

**Abstract:** In this work we propose and analyze a model which addresses the pulsing behavior of sellers in an online auction (store). This pulsing behavior is observed when sellers switch between advertising and processing states. We assert that a seller switches her state in order to maximize her profit, and further that this switch can be identified through the seller's reputation. We show that for each seller there is an optimal reputation, *i.e.*, the reputation at which the seller should switch her state in order to maximize her total profit. We design a stochastic behavioral model for an online seller, which incorporates the dynamics of resource allocation and reputation. The design of the model is optimized by using a stochastic advertising model from [1] and used effectively in the Stochastic Optimal Control of Advertising [2]. This model of reputation is combined with the effect of online reputation on sales price empirically verified in [3]. We derive the Hamilton-Jacobi-Bellman (HJB) differential equation, whose solution relates optimal wealth level to a seller's reputation. We formulate both a full model, as well as a reduced model with fewer parameters, both of which have the same qualitative description of the optimal seller behavior. Coincidentally, the reduced model has a closed form analytical solution that we construct.

## 1. Introduction

In this work, we consider sellers in an online market like Amazon without a bidding mechanism. The setting may include concurrent online auctions on a site like eBay, for a product or service with only a "buy-it-now" option, under the assumption that buyers can compare reputation among online platforms. In this online environment, the buyer and seller have a certain amount of anonymity, which has implications on the fairness of the auction [4]. One of the only counter measures to ensure fairness is the use of feedback forums, such as Amazon, eBay, and Yahoo!, at online auction sites. The feedback provided by customers provides a ranking system for the sellers, which in turn rewards fair sellers and penalizes unfair or unreliable sellers [5–7]. Since it is in the sellers best interest to maximize their profits, it follows that a seller will seek to use reputation as a means to optimize their decision making [8,9]; e.g., for the optimal bidding strategies in sequential auctions, see [10].

Online auction sellers can be broadly categorized by two types of behavior: those focused on advertising and customer service, which includes following up with past customers; and those who instead focus only on processing existing orders. While the latter behavior will likely lower the seller reputation on average, this immediate increase in wealth may offset the long-term damage from a lower reputation. With this in mind, it is compelling to ask whether there exists an optimal long term strategy, in which a seller attempts to maximize her profit by achieving a fixed reputation.

We propose a simple behavioral model to study the relationship between a seller's wealth and reputation in the auctions environment with only the "buy-it-now" option. Following the standard Nerlove-Arrow construction [11], which was extended to a stochastic setting by Sethi [1], the wealth and growth of the seller are described by continuous stochastic processes. Note that our model is a continuous approximation of a discrete auctioning process. The discrete model of the process may be less tractable but still very interesting and useful. We propose an optimal infinite horizon strategy, and incorporate the wealth-growth model of Mink and Seifert [3], based on empirical studies of ebay sellers. Also, we use the model for reputation proposed by Sethi [1] and Rao [12] which has been adopted by many others in optimal advertising models, such as Raman's recent work in Boundary Value Problems in the Stochastic Optimal Control of Advertising [2]. This work incorporates the idea that humans do not multi task well, but rather switch, see [13]. In other words, switching costs do not factor in for an individual agent, but only for large online retailers who have marketing department reorganization costs. Additionally, the model in this work allows all sellers to have a little extra capacity/money for either shipping or advertising (e.g., writing a letter or similar individual campaign), which is rather small and not unlimited.

The rest of the paper is laid out as follows. In Section 2 we define the stochastic Nerlove-Arrow model, along with the growth rate due to Mink and Seifert. Using the principal of optimality we derive the corresponding Hamilton Jacobi Bellman (HJB) boundary value problem satisfied by the value function. In Section 3, we propose several important modifications to the existing methodology, and in doing so

derive a new model which is normalized with respect to the seller reputation, and has a linear growth rate. In making these modifications, we also find a reduced version of our new model, which is governed by a geometric Brownian motion. We prove that both models admit unique piecewise continuous solutions, and are qualitatively similar. However the reduced model, surprisingly, has a closed form piecewise defined solution. This allows us to prove that pulsing behavior is an optimal strategy for sellers and also serves as a benchmark to our numerical solution of the full model in Section 4, which does not have a closed form analytic solution. Our results are compelling, and warrant further investigation for the finite horizon case, which we discuss with concluding remarks in Section 5.

## 2. Explicit Resource Allocation Mechanism

Consider a seller in an online marketplace as a triple $(W, R, \mu)$ representing her wealth, reputation, and excess rate, respectively. Here, the excess rate is extra effort that can be allocated to either expedite payment *or* increase her reputation. In this setting, the reputation $R$ is a positive number reflecting customer satisfaction. As an initial approach, we consider only the effects of reputation on the optimal way to grow her wealth, and leave the more general model with shipping costs for future work. Also, in our mathematical formulation, we consider the reputation mechanism first suggested by the work of Nerlove and Arrow [11] and later generalized to stochastic settings in [1,2,12].

The seller is one of many competing in the online marketplace, and has her transactions verified and processed automatically via the marketplace. All sellers competing in this marketplace have an expected speed for mailing out completed orders to be able to sell in the marketplace, with some time allowed for delayed shipping before the marketplace intervenes. In return, the marketplace guarantees the transaction. Payment is modeled to be released to the seller once the purchased item is mailed to the buyer.

However, the seller can choose to go beyond these expected standards. For example, she may prepare packages and drive to the post office earlier than expected to expedite the payment from the marketplace. Or she may spend time and money communicating with buyers in a bid to increase their favorable ratings. This may include sending a small unexpected gift in the package to be mailed out or spending extra time to individually craft and send emails out to purchasers. Our model assumes that at most only one of these extra actions can happen with noticeable return at any given time. The seller may also decide to do nothing extra, and in this case $\mu = 0$.

For example, if she is focused on getting her payment released, then she is not able to focus on creative ways to engage her buyers post sale. We also assume that her extra capacity to do either extra action is limited by absolute minimum mailing rates and her pool of resources she is willing to contribute to going beyond expected standards. By choosing to shift up her resources from promotion to expedited mailing and back, the seller can influence her wealth and reputation levels via the excess rate $\mu$. Positive $\mu$ corresponds to a promotional state, while negative $\mu$ corresponds to an expedited mailing state.

We note that if our seller is in fact a large company with many such agents that can work on either processing or sales, then the company can choose to shift the ratio of workers who work in processing or sales, and this model is the subject of future work. The work we present here is limited to a one agent seller. We note here that our model also assumes no cost for the individual to switch behavior from

promotion to expedited mailing. In a large company, one would expect such costs to arise in restricting the size of advertising and shipping departments. One related paper that addresses such switching costs, albeit in a real options investment setting, is the paper by Duckworth *et al.* [14].

We utilize a continuous time formulation for wealth and reputation, as well as control of these processes. The seller will be determined to behave optimally if she maximizes her expected present value of total earnings, until her reputational value reaches $R = 0$. This produces a fully consistent formulation of both the value function $V$, and the strategy to achieve this optimal value.

## 2.1. Mathematical Formulation

Following the framework in [15,16], reputation $R$ is observed in $\mathcal{O} := (0, \infty)$ and the control $\mu$ is constrained to $U := [-\epsilon, \epsilon]$. Since this process will be stochastic, we introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and corresponding reference probability systems $\nu = ((\Omega, \{\mathcal{F}_s\}, \mathbb{P}), B)$ where $\mathcal{F}_s \subset \mathcal{F}$ and $B$ is an $\mathcal{F}_s$-adapted Brownian motion. The evolution of wealth and reputation is hence modeled by a controlled two-state Markov process $(W_\cdot^{\mu_\cdot}, R_\cdot^{\mu_\cdot})$, where wealth $W$ grows at a rate proportional to a function $h(R)$ that links reputation to revenue per sale, subject to a controlled extra processing rate $\mu$ per unit time chosen from the admissible class $A_\nu$,

$$\mathcal{A}_\nu := \left\{ \mu \in \mathcal{F}_s \mid \mu \text{ is } \mathcal{F}_s\text{-p.m.}, \mu_\cdot \in U \text{ on } [0, \infty), \mathbb{E}\left[ \int_0^{\tau_\mathcal{O}} e^{-\rho s} \left| (1 - \mu_s) h(R_s^{(\mu)}) \right| \mathrm{d}s \right] < \infty \right\} \quad (1)$$

Here, p.m. denotes *progressively measurable*, $\rho$ is the constant discount factor to account for the time value of money, and $\tau_\mathcal{O}$ is the first exit time of reputation $R$ from $\mathcal{O}$, or $\infty$ if $R \in \mathcal{O}$ for all $s \geq 0$. We set $\infty$ as an absorbing state for reputation. Symbolically,

$$\tau_\mathcal{O} := \inf \left\{ t \geq 0 \mid R_t^{(\mu)} \notin \mathcal{O} \right\} \quad (2)$$

Formally, our two state controlled Markov process is

$$\begin{aligned} dW_t^{(\mu)} &= (1 - \mu_t) h(R_t^{(\mu)}) \, \mathrm{dt} \\ dR_t^{(\mu)} &= (\mu_t - \kappa R_t^{(\mu)}) \, \mathrm{dt} + \sigma \, \mathrm{dB_t} \end{aligned}$$

where $\kappa$ is a proportionality constant which accounts for mean reversion [11] and $\sigma$ is the constant volatilty accounting for random effects to reputation.

The major difficulty is in defining the growth rate $h(R)$. However, an *explicit* mechanism has been proposed by Mink and Seifert [3]. There the authors not only propose, but empirically justify a growth rate of

$$\begin{aligned} h(R) &= A + C \left( 1 - \frac{1}{\ln (e + R)} \right) \\ &= A \left[ 1 + \frac{C}{A} \left( 1 - \frac{1}{\ln (e + R)} \right) \right] \end{aligned} \quad (3)$$

where $A$ relates to the inherent value of the object for sale and $C$ is a parameter to be fitted. This is accomplished by obtaining data using an auction robot and then computing a single regression, which

gives $C = 2.50$ in Equation (3). To the best of our knowledge, the model in [3] is among the first to give an explicit relationship between reputation and price. We also note that the parameter $\frac{C}{A}$ represents the maximal reputational effect of a seller as a fraction of inherent value $A$, and is what we choose in the numerical examples below. The model also suggests a multiple regression formula where other factors, such as shipping costs and whether a "buy-it-now" price is offered, are considered as well, in which case $C = 1.93$. The authors in [3] also comment that highly experienced sellers have higher feedback scores and design the auction more favorably, which reflects to their higher revenue. In fact, they show that the coefficient attributed to shipping costs is larger than one, implying that customers put a high value on shipping when deciding on their bids, and that savvy agents take this into consideration. Notice that since $h$ is bounded on $\mathcal{O}$, our admissible class reduces to

$$\mathcal{A}_\nu := \{\mu \in \mathcal{F}_s \mid \mu \text{ is } \mathcal{F}_s\text{-p.m.}, \mu_\cdot \in U \text{ on } [0, \infty)\}. \tag{4}$$

Finally, the work in [3] posits that the horizon does not affect the revenue stream as much as the shipping cost and reputation factors, and so we consider an infinite horizon model here.

## 2.2. Hamilton-Jacobi-Bellman Formulation

With the stochastic dynamics for reputation proposed in [1], a growth rate model for reputational effect on sales [3], and an infinite horizon, we expect that switching would depend only on the current reputational state. We now seek a twice-continuously differentiable, polynomially growing function $V \in C^2[0, \infty) \cap C_p[0, \infty)$ which is a candidate solution of the optimal control problem

$$
\begin{aligned}
\bar{V}(R) &= \sup_\nu \sup_{\mu \in \mathcal{A}_\nu} \mathbb{E}\left[\int_0^{\tau_0} e^{-\rho s}(1 - \mu_s)h(R_s^{(\mu)})\,\mathrm{d}s \mid R_0 = R\right] \\
&= \sup_\nu \sup_{\mu \in \mathcal{A}_\nu} \mathbb{E}\left[\int_0^{\tau_\mathcal{O}} e^{-\rho s}(1 - \mu_s)h(R_s^{(\mu)})\,\mathrm{d}s + 1_{\{\tau_\mathcal{O} < \infty\}} e^{-\rho \tau_\mathcal{O}} \frac{1 + \epsilon}{\rho} h(\infty) 1_{\left\{R_{\tau_\mathcal{O}}^{(\mu)} = \infty\right\}} \mid R_0 = R\right] \\
h(R) &= A + C\left(1 - \frac{1}{\ln(e + R)}\right) \\
\tau_0 &:= \inf\{t \geq 0 \mid R_t = 0\}
\end{aligned}
\tag{5}
$$

One approach to finding $\bar{V}$ would be to solve Equation (5) directly. For example, by the definition of $\bar{V}$, it follows that

$$
\begin{aligned}
\bar{V}(0) &= 0 \\
\bar{V}(\infty) &= \sup_\nu \sup_{\mu \in \mathcal{A}_\nu} \mathbb{E}\left[\int_0^\infty e^{-\rho s}(1 - \mu_s)h(R_s^{(\mu)})\,\mathrm{d}s \mid R_0 = \infty\right] \\
&= \frac{1 + \epsilon}{\rho} h(\infty)
\end{aligned}
\tag{6}
$$

However, we shall instead apply the principle of optimality, and in doing so arrive at the following nonlinear Hamilton-Jacobi-Bellman (HJB) boundary value problem (suppressing the explicit dependence of $R$ on $\mu$)

$$0 = \max_{-\epsilon \leq \mu \leq \epsilon} \left[ (\mu - \kappa R)\frac{\partial V}{\partial R} + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial R^2} + (1 - \mu)h(R) - \rho V \right]$$

$$V(0) = 0$$

$$V(\infty) = \frac{1 + \epsilon}{\rho}h(\infty) \tag{7}$$

$$h(R) = A + C\left(1 - \frac{1}{\ln(e + R)}\right) = A + C\frac{\ln\left(1 + \frac{R}{e}\right)}{1 + \ln\left(1 + \frac{R}{e}\right)},$$

which simplifies to

$$\rho V = h(R) - \kappa R\frac{\partial V}{\partial R} + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial R^2} + \epsilon\left|\frac{\partial V}{\partial R} - h(R)\right|$$

$$V(0) = 0$$

$$V(\infty) = \frac{1 + \epsilon}{\rho}(A + C) \tag{8}$$

$$h(R) = A + C\frac{\ln\left(1 + \frac{R}{e}\right)}{1 + \ln\left(1 + \frac{R}{e}\right)}$$

We therefore instead solve the HJB problem, which is justified by the following standard theorem (Corollary IV.5.1 in [15]):

**Theorem 1.** *Let* $V \in C^2(\mathcal{O}) \cap C_b(\bar{\mathcal{O}})$ *be a twice-continuously differentiable and bounded solution to an associated HJB equation of a control problem*

$$V_{\mathrm{PM}}(x) := \inf_{\nu}\inf_{A_\nu}\mathbb{E}\left[\int_0^{\tau_{\mathcal{O}}} e^{-\rho s}G(R_s^\mu, \mu_s)\,ds\right] \tag{9}$$

*for a process* $R$ *whose SDE has drift and volatility that are Lipschitz in* $\mu$ *and* $R$, *and a* $G(R, \mu)$ *that is continuous and polynomially growing for all* $R \in \mathcal{O}$ *and continuous for all* $\mu \in U$. *Moreover, assume either that* $\beta > 0$ *or that* $\tau_{\mathcal{O}} < \infty$ *with probability* 1 *for every admissible progressively measurable control process* $\mu$. *Then* $V(x) = V_{\mathrm{PM}}(x)$ *for an optimal control* $u^*(s) \in \operatorname{argmin}[L_R[V] + G(R, u)]$, *where* $L$ *is the generator of the process* $R$.

We now have our verification theorem:

**Theorem 2.** *Let* $V \in C^2(\mathcal{O}) \cap C_b(\bar{\mathcal{O}})$ *be a twice-continuously differentiable and bounded solution to the boundary value problem Equation* (8). *Then* $V$ *is an optimal solution to the optimal control problem Equation* (5).

**Proof.** By Corollary IV.5.1 in [15], since

- $(i)$ the drift and volatility are (uniformly) Lipschitz in $R$ and $\mu$,
- $(ii)$ $\rho > 0$, and

- (*iii*) we impose an additional boundary condition on the solution as $R \to \infty$ as our indefinite horizon is now $[0, \tau_{\mathcal{O}}) = [0, \tau_0) = [0, \infty)$.

it follows directly that our bounded classical solution $V = \bar{V}$ and the control

$$\mu^*(R) \in \underset{\mu \in [-\epsilon, \epsilon]}{\mathrm{argmax}} \left[ (\mu - \kappa R) \frac{\partial V}{\partial R} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial R^2} + (1 - \mu) h(R) - \rho V \right] \tag{10}$$

leads to an optimal, stationary Markov control $\mu^*(R_s^{(\mu^*)})$. $\quad\square$

**Remark 1.** We note here that it is sufficient to require polynomial growth on $V$, not bounded growth. However, as $h$ is bounded above by $A + C$ on $\mathcal{O}$, we can restrict our attention to bounded $V$. Furthermore, an upper bound (and in fact limit as $R \to \infty$) for $V$ is the value obtained by a seller earning maximum value $A + C$ on her items on the time interval $[0, \infty)$ while always processing orders.

## 3. Market Share Based Pricing Mechanism

We observe that under its current definition, the reputation $R \in \mathcal{O}$. In this section, we employ the mapping

$$Y := f(R) = \frac{R}{1 + R} \tag{11}$$

which results in a normalized *market share* reputation $Y \in \mathcal{Q} := (0, 1)$. This is consistent with prior assumptions [3] that the value function $h(R)$ be concave, and bounded. We therefore also simplify (3) accordingly, replacing logarithms with a rational function

$$\begin{aligned} \bar{h}(R) &= A + C \frac{R}{1 + R} \\ &= A \left[ 1 + \frac{C}{A} \frac{R}{1 + R} \right] \end{aligned} \tag{12}$$

This modification, when combined with the mapping from a reputational score to a market share score, $R \to Y$, produces a more intuitive growth rate, $\tilde{h}(Y)$ which grows linearly with respect to market share reputation. Recall that $\infty$ is an absorbing state for $R$; consequently, we have $dY_t \equiv 0$ if $Y_t \geq 1$. For $Y_t < 1$, it follows from Ito calculus that $dY_t = f'(R)dR_t + \frac{1}{2}f''(R)dR_t dR_t$ and so

$$\begin{aligned} dY_t &= (1 - Y)^2 \left[ \left( \mu - \kappa \frac{Y}{1 - Y} \right) dt + \sigma\, dB_t \right] + \frac{1}{2}(-2(1 - Y)^3)\sigma^2\, dt \\ &= \left[ \mu(1 - Y)^2 - \kappa Y(1 - Y) - \sigma^2(1 - Y)^3 \right] dt + \sigma(1 - Y)^2\, dB_t \\ \tilde{h}(Y) &= A + CY \\ \tilde{V}(y) &= \sup_{\nu} \sup_{\mu \in \mathcal{A}_\nu} \mathbb{E} \left[ \int_0^{\tau_0} e^{-\rho s}(1 - \mu_s)\tilde{h}(Y_s)\, ds \mid Y_0 = y \right] \end{aligned} \tag{13}$$

Note that a function that is bounded for market share $Y \in (0, 1)$ is also correspondingly bounded for all reputation $R \in (0, \infty)$. We therefore study the market share model below, confident that our results will hold for the full reputation model by virtue of the inverse map of Equation (11).

### 3.1. Rescaled HJB Model

After market share rescaling, the HJB now takes a form which is in fact degenerate at both endpoints $y = 0, 1$:

$$-\frac{\sigma^2}{2}(1-y)^4 V''(y) + \rho V = \tilde{h}(y) - \left[\kappa y(1-y) + \sigma^2(1-y)^3\right] V'(y) + \epsilon \left|(1-y)^2 V'(y) - \tilde{h}(y)\right|$$

$$V(0) = 0, \qquad V(1) = \frac{1+\epsilon}{\rho}\tilde{h}(1)$$

(14)

In theory, a closed form solution for this boundary value problem can be formally constructed piecewise, to the left and right of the special point $y^*$, for which $\tilde{h}(y^*) = (1 - y^*)^2 V'(y^*)$. We shall refer to $y^*$ as the switching point below, since it is the reputational value at which the seller switches from advertising to processing. In practice, we shall resort to numerical computation to approximate this solution, and in doing so estimate the value $y^*$.

**Theorem 3.** *If $V^{**} \in C^2(\mathcal{Q})$ is a monotonically increasing solution to the HJB problem Equation* (14), *then $V^{**} = \tilde{V}$.*

**Proof.** Enforcing the boundary condition on $V^{**}(1)$ enforces bounded growth on our monotonically increasing $V^{**}$. By the inverse map of Equation (11), the result follows as a corollary to Theorem 2. Note that $\mathcal{Q} = (0, 1)$ is open, but $y = 1$ is not reached in finite time from anywhere in $\mathcal{Q}$ and the diffusion $Y$ is in fact absorbed for all time in state $Y = 1$. Hence, the exact value of $V^{**}(1)$ imposed. $\square$

### 3.2. Reduced Model

With the previous mapping, note that the market (reputation) share $Y$ is absorbed at $Y = 1$, which corresponds to $R \to \infty$. Moreover, as in the Stochastic Nerlove evolution, we have that the probability that $Y_t < 0$ for some positive $t$ is greater than 0, and the drift term is a third-order polynomial in $Y$. Based on these observations, we propose a reduced model for *Stochastic Reputation Share*, expressed by the following stochastic differential equation and associated stochastic optimal control problem

$$\check{V}(y) = \sup_\nu \sup_{\mu \in \mathcal{A}_\nu} \mathbb{E}\left[\int_0^{\tau_0} e^{-\rho s}(1 - \mu_s)(A + CY_s)\,\mathrm{d}s \mid Y_0 = y\right]$$

$$\frac{dY_t}{1 - Y_t} = \mu\,\mathrm{dt} + \sigma\,\mathrm{dB_t}$$

(15)

This leads to a corresponding nonlinear HJB boundary value problem,

$$0 = \max_{-\epsilon \leq \mu \leq \epsilon}\left[\mu(1-y)V'(y) + \frac{\sigma^2}{2}(1-y)^2 V''(y) + (1-\mu)(A + Cy) - \rho V\right]$$

$$V(0) = 0$$

$$V(1) = \sup_\nu \sup_{\mu \in \mathcal{A}_\nu} \mathbb{E}\left[\int_0^\infty e^{-\rho s}(1-\mu_s)\tilde{h}(1)\,\mathrm{ds}\right] = \rho^{-1}(1+\epsilon)\tilde{h}(1)$$

(16)

which can be recast as

$$-\frac{\sigma^2}{2}(1-y)^2 V''(y) + \rho V = \tilde{h}(y) + \epsilon \left| (1-y)V'(y) - \tilde{h}(y) \right|$$
$$V(0) = 0 \tag{17}$$
$$V(1) = \frac{1+\epsilon}{\rho}\tilde{h}(1)$$

One of the most attractive features of the reduced model (17) is that it has a closed form, piecewise defined analytic solution, which we derive in the Appendix. Here we state the following theorem.

**Theorem 4.** *There exists a solution* $V^* \in C^2(\mathcal{Q}) \cap C_b(\bar{\mathcal{Q}})$ *to the reduced model* (17). *Furthermore,* $V^* = \check{V}$, *and is given by the piecewise solution*

$$V^*(y) = \begin{cases} V_\ell(y), & 0 \leq y \leq y^* \\ V_r(y), & y^* \leq y \leq 1 \end{cases} \tag{18}$$

*where*

$$V_\ell = c_1(1-y)^{\gamma^\ell_-} + c_2(1-y)^{\gamma^\ell_+} + \alpha_\ell + \beta_\ell y \tag{19}$$
$$V_r = c_3(1-y)^{\gamma^r_+} + \alpha_r + \beta_r y \tag{20}$$

*with constants defined in the Appendix.*

**Proof.** The construction of this piecewise solution is found in the Appendix. The solution constructed is monotonically increasing in $y$. Enforcing the boundary condition on $V^*(1)$ uniformly bounds $V^*$ on $\mathcal{Q}$. The proof of equality $V^* = \check{V}$ then follows from Corollary IV.5.1 in [15], as in our Theorem 2 above. $\square$

**Remark 2.** Since the reduced model has a known closed form solution, we can measure the error made in constructing numerical solutions in Section 4, which in turn acts as a benchmark for the code we use to study the full model. It is uncommon to find a closed form solution for most optimal control problems.

## 4. Numerical Results

In this section, we construct numerical solutions of the market share scaled model Equation(14), as well as our proposed reduced model Equation (8). In Sections A and A.1, the piecewise analytic solution and a nonlinear equation for the reduced switching point $y^*$ are found. These will be used to validate the numerical results for the reduced model, which act as a benchmark for the full problem. Please note that in what follows, the symbol $y^*$ is used to denote the switching point in reputational share for both the reduced and full model and their corresponding ODEs.

### 4.1. Numerical Results for the Reduced Model

The boundary value problem Equation (8) is discretized using a standard finite difference scheme. Let $y_k = k/N_y$, and $V_k = V(y_k)$, for $k = 0, 1, \ldots, N_y$. Then we solve a linear system of $N_y + 1$ equations of the form $M\mathbf{v} = \mathbf{f}$, where

$$M\mathbf{v} \approx \rho V - \frac{\sigma^2}{2}(1-y)^2 V'', \qquad \text{and} \qquad \mathbf{f} \approx h + \epsilon \left| (1-y)V' - h \right|.$$

Since $\mathbf{f}$ depends on $V$, the solution is implicit, and therefore must be obtained by using a fixed point iteration $M\mathbf{v}^{(k+1)} = \mathbf{f}^{(k)}$. In our numerical experiments, we let $N_y = 1000$, and set the maximum iteration count at $K = 20$. Convergence is observed in all tested cases. In Figure 1, several numerical solutions $V(y)$ are shown, for fixed $A = 1$, $C = 0.15$, $\epsilon = 0.02$, and various values of $\rho = 0.1, 0.2, 0.5, 2.0$ and $\sigma = 0.2, 0.5, 1.0, 5.0$. The values of $A$ and $C$ are chosen via $\frac{C}{A} = 0.15$ to reflect a maximal 15% reputational premium for sellers above the inherent value $A$.
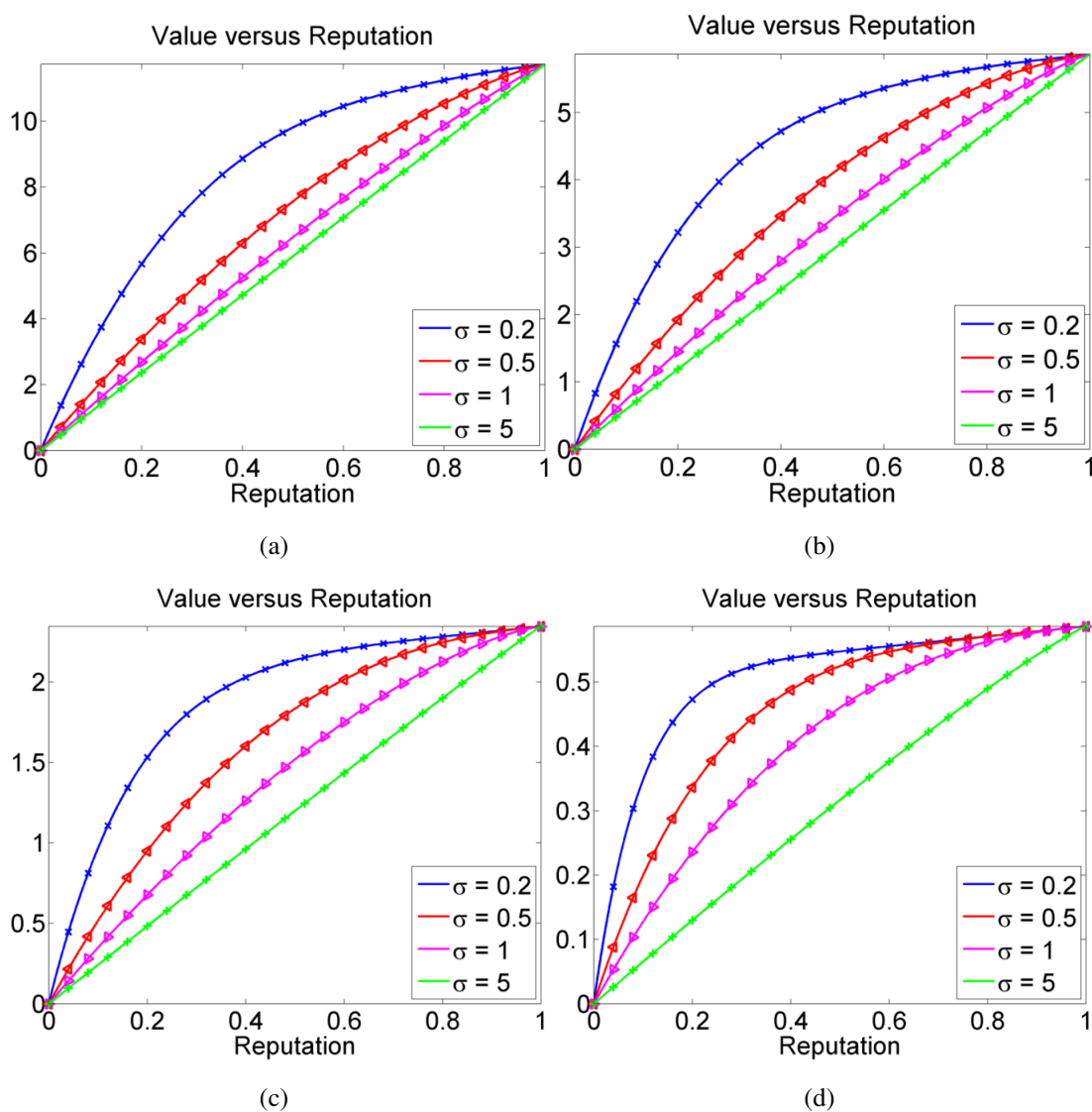


**Figure 1.** A plot of several numerical solutions $V(y)$ for the reduced boundary value problem Equation (17) are shown with $A = 1$, $C = 0.15$, and $\epsilon = 0.02$. (**a**) $\rho = 0.1$; (**b**) $\rho = 0.2$; (**c**) $\rho = 0.5$; (**d**) $\rho = 2.0$.

It follows that both $\rho$ and $\sigma$ have a dramatic effect not only the shape of the solution, but also, as is demonstrated in Figure 2, on the switching point $y^*$. These numerical results suggest that both the discount rate $\rho$ and the volatility $\sigma$ of a seller's reputation can "dramatically" affect her behavior, particularly the critical reputation at which she switches from the processing to advertising mode.
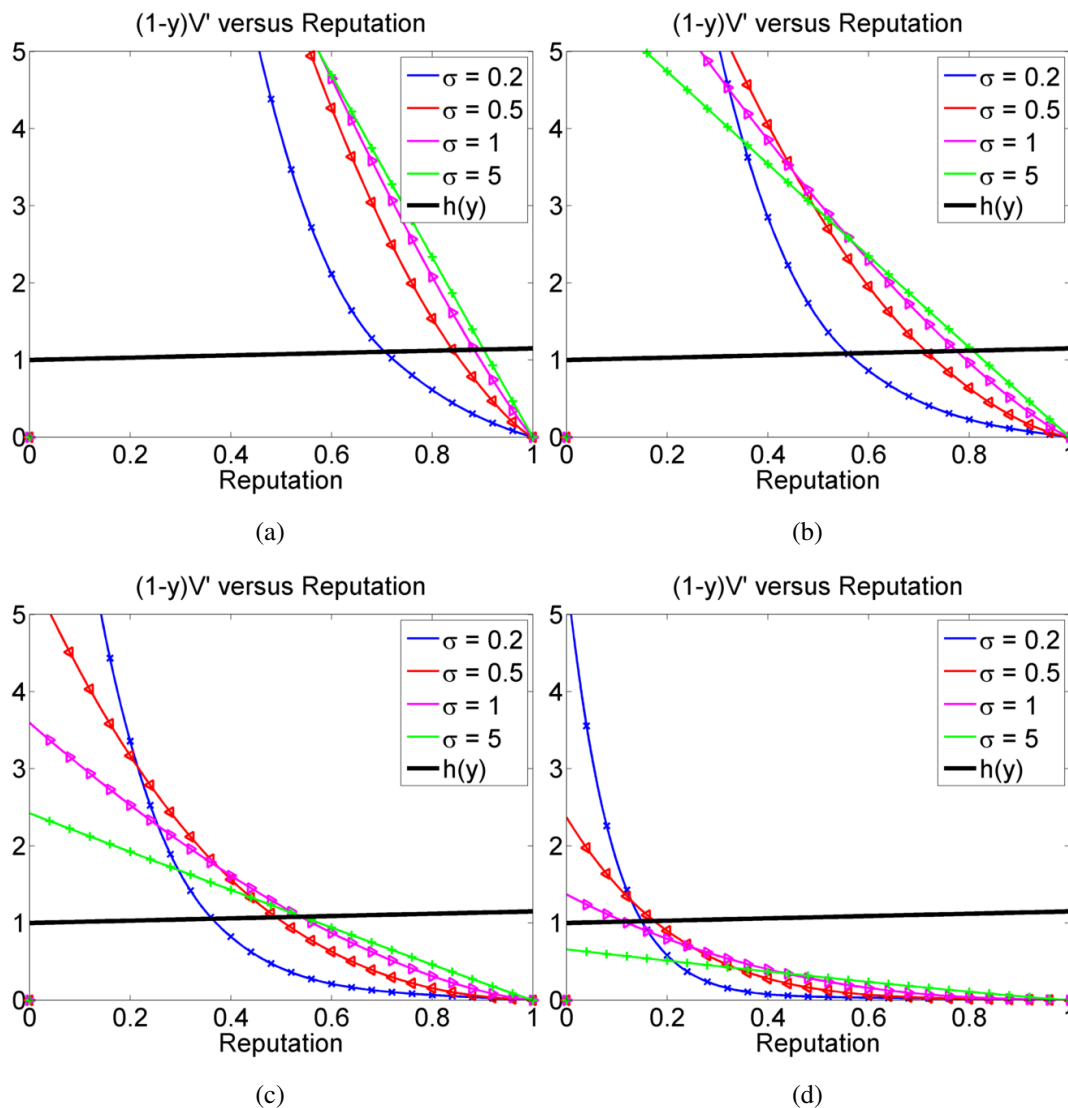


**Figure 2.** A plot of the quantity $(1-y)V'(y)$, where $V(y)$ is a numerical solution of the the reduced boundary value problem Equation (17). The value $y^*$ is defined as the intersection of these curves with $h(y)$ (the solid line). (**a**) $\rho = 0.1$; (**b**) $\rho = 0.2$; (**c**) $\rho = 0.5$; (**d**) $\rho = 2.0$.

### 4.2. Numerical Results for the Full Model

The same numerical discretization is employed to solve the full Nerlove-Arrow model Equation (14), where we solve a linear system of the form $M\mathbf{v} = \mathbf{f}$, with

$$M\mathbf{v} \approx \rho V + \left[\kappa y(1-y) + \sigma^2(1-y)^3\right] V'(y) - \frac{\sigma^2}{2}(1-y)^4 V''(y), \qquad \mathbf{f} \approx h + \epsilon \left|(1-y)^2 V' - h\right|$$

We first set $\kappa = 0$, and hold all remaining parameters fixed, and plot the results in Figures 3 and 4. The full solutions have more curvature than those of the reduced model, but nonetheless remain

monotone, and have a single unique switching point $y^*$. The additional curvature is due to the appearance of $V'$ terms in the differential operator, which now depend on $\sigma$, as well as $\kappa$, which we recall incorporates mean reversion. That is, independent of the sellers strategy, reputation will tend to decrease to a smaller amount of the market share, with constant rate $\kappa$.

In Figures 5 and 6, the same solutions are shown with $\kappa = 1$. Here it becomes apparent that both the rate of mean reversion, as well as the volatility will affect the seller's optimal strategy.

**Remark 3.** Our numerical results illustrate that a wide range of seller behaviors can be described by varying the parameters, and therefore that reputational value is a strong indicator of optimal seller behavior.
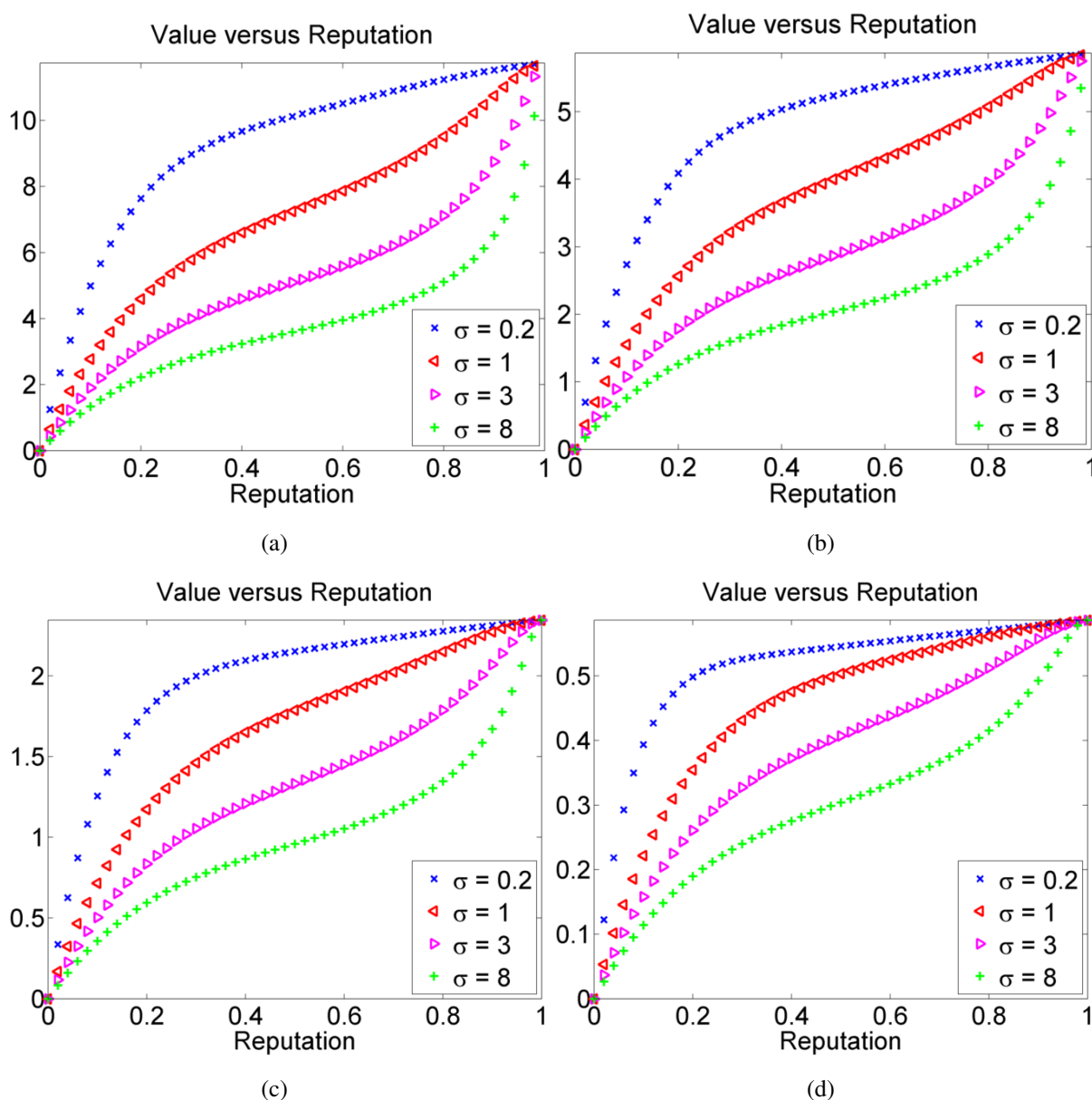


**Figure 3.** A plot of several numerical solutions $V(y)$ for the full boundary value problem Equation (14) are shown with $A = 1$, $C = 0.15$, $\epsilon = 0.02$, and $\kappa = 0$. (**a**) $\rho = 0.1$; (**b**) $\rho = 0.2$; (**c**) $\rho = 0.5$; (**d**) $\rho = 2.0$.
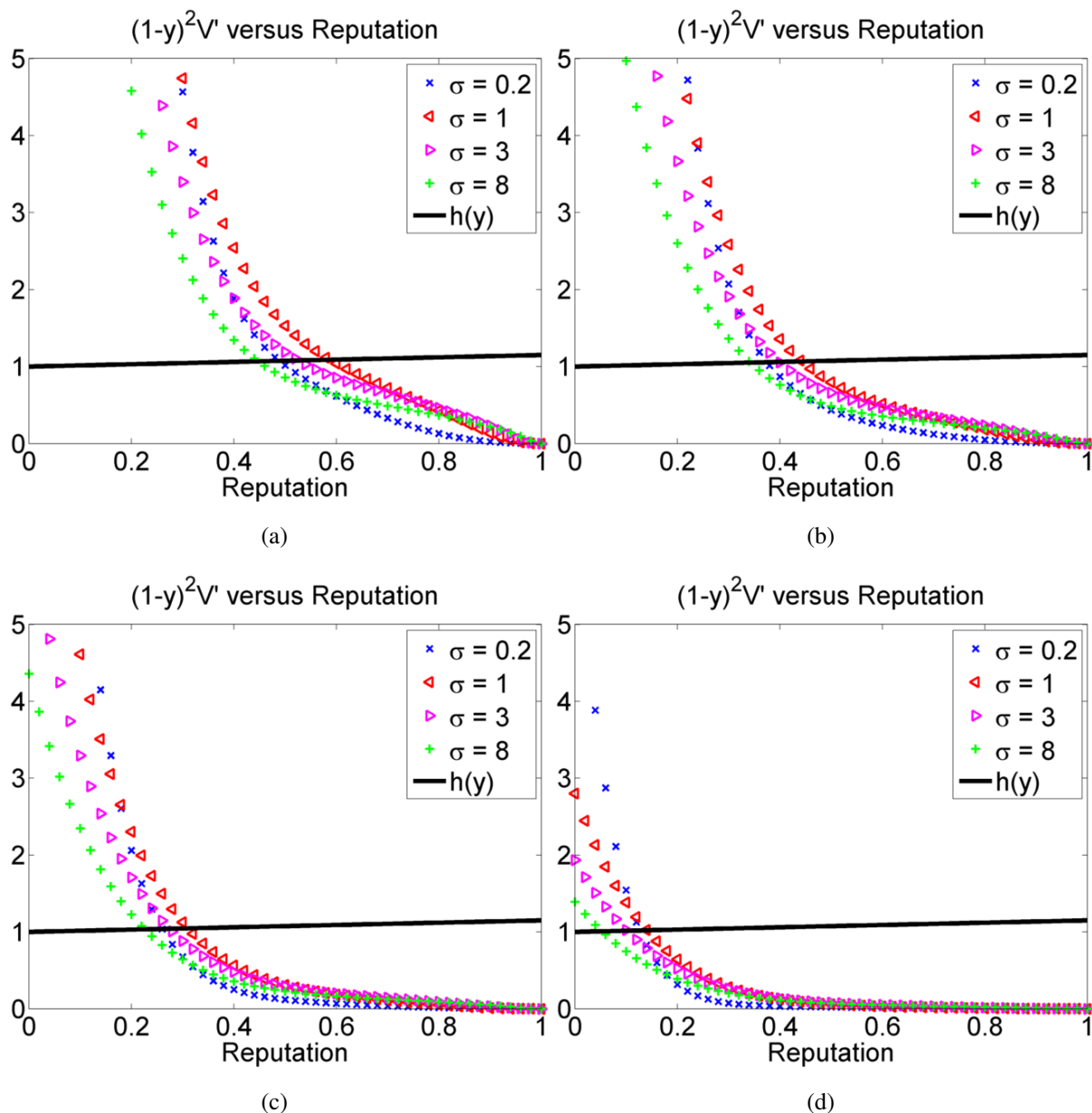
**Figure 4.** A plot of the quantity $(1 - y)^2 V'$, where $V$ is a numerical solution of the the full boundary value problem Equation (14). The value $y^*$ is defined as the intersection of these curves with $h(y)$ (black solid line). (**a**) $\rho = 0.1$; (**b**) $\rho = 0.2$; (**c**) $\rho = 0.5$; (**d**) $\rho = 2.0$.
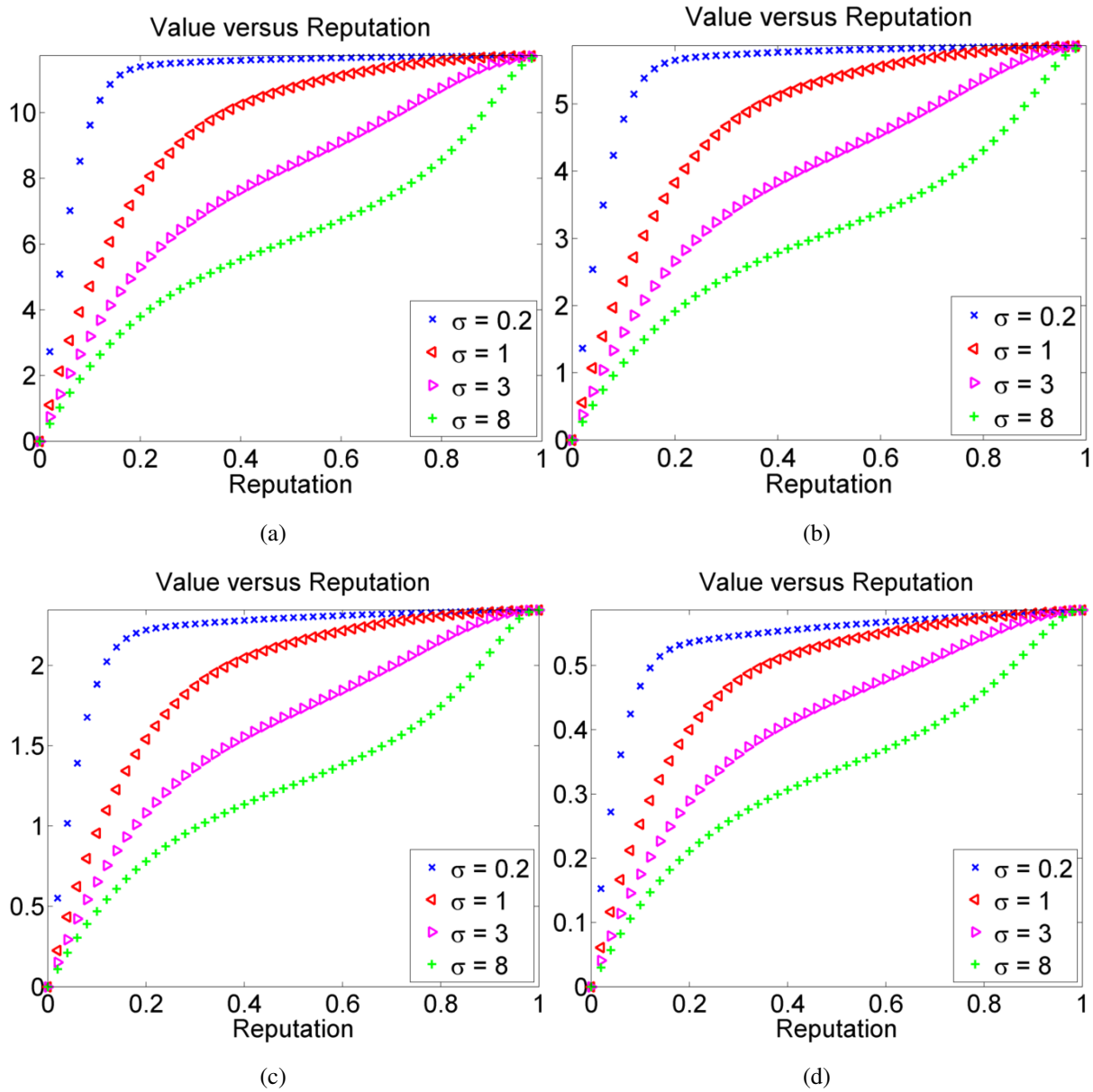
**Figure 5.** The same as Figure 3, but with $\kappa = 1$. (**a**) $\rho = 0.1$; (**b**) $\rho = 0.2$; (**c**) $\rho = 0.5$; (**d**) $\rho = 2.0$.
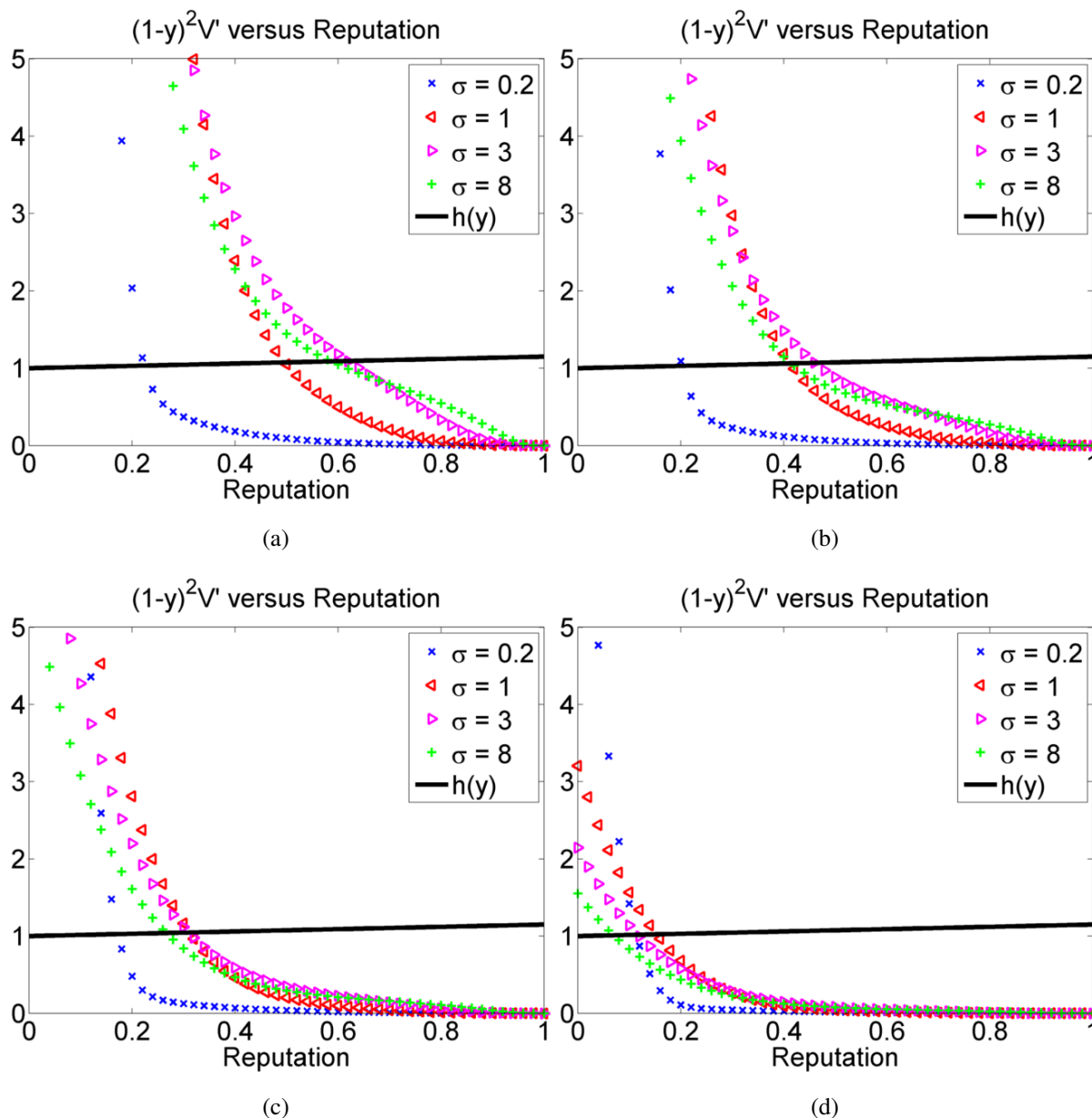
**Figure 6.** The same as Figure 4, but with $\kappa = 1$. (**a**) $\rho = 0.1$; (**b**) $\rho = 0.2$; (**c**) $\rho = 0.5$; (**d**) $\rho = 2.0$.

## 5. Conclusions and Future Work

In this work, we have proposed a model which describes the behavior of sellers participating in online auctions (markets). Our model is based on the premise that buyer feedback greatly impacts the sales rate, which motivates our assumption that the revenue per sale (or the price per unit) is solely dependent upon seller reputation. We assumed that a seller has a fixed amount of resources that must be allocated either to (i) advertising to new buyers or (ii) processing orders for current customers. In doing so, we were able to design an optimal selling strategy, wherein the seller switches their behavior when an optimal market share reputation is reached, which depends on statistical modeling parameters.

These modeling parameters were introduced through the wealth-reputation mechanisms in [11], which have been generalized to stochastic settings in [1,2,12]. We have additionally modified the

empirical model $h(R)$ found in [3], relating the price per unit to the reputation. Rather than viewing reputation $R$ as an unbounded quantity, we instead introduced a market scaled reputation $Y = R/(1+R)$, where $Y \in [0, 1)$, which reduces the price per unit to a simple linear model $h(Y) = A + CY$.

We then optimized the stochastic model over the control of the excess rate $\mu$ using the Hamilton-Jacobi-Bellman equation, yielding a deterministic equation relating the value per unit to the seller reputation. The resulting boundary value problem is then solved numerically, and we qualitatively see that the value of sellers goods increase monotonically with reputation, but that a unique optimal reputation $y = y^*$ determines when the seller should switch from advertising to processing. The numerical scheme was validated with a reduced model, which has a closed form piecewise analytic solution, and permits direct determination of $y^*$. The numerical results validate our modeling assumptions, and provide a framework for studying seller behavior based on seller reputation. Although the used techniques are standard, we believe that the optimal strategy presented both analytically and numerically has implications on the way reputational information can be used to predict the behavior of an individual seller in an online market.

This work can naturally be generalized in a variety of ways. For instance, the model could be developed in the case of finite horizon time $T$, since a more realistic assumption is that a seller strategy depends on both time and the current reputation state. Unlike our current infinite horizon model, the resulting HJB equation will lead to a time-dependent partial differential equation, which will be the subject of future work.

## Acknowledgments

## Author Contributions

The authors equally contributed.

## Conflicts of Interest

The authors declare no conflict of interest.

## A. Piecewise Analytic Solution of Reduced Model

We now construct the exact solution to the reduced HJB boundary value problem Equation (8), which incidentally can be found analytically. The principal difficulty that must be overcome is the appearance of an absolute value, which contains the unknown variable $V'$. We therefore consider a piecewise defined solution $V(y)$, which remains $C^2$ across the switching point $y^*$, defined by the vanishing of the expression in the absolute value

$$(1 - y^*)V'(y^*) = \tilde{h}(y^*) \tag{21}$$

We now separately consider the regions $y < y^*$ and $y > y^*$ in the reduced model Equation (8), and define the corresponding linear differential operators $L_\pm$ as

$$L_\pm[V] = \left( -\frac{\sigma^2}{2}(1-y)^2\frac{d^2}{dy^2} \pm \epsilon(1-y)\frac{d}{dy} + \rho \right) V \tag{22}$$

The boundary value problem can then be decomposed into two smaller problems for $V_\ell$ and $V_r$. The boundary conditions provide one condition for each of $V_\ell$ and $V_r$. The remaining two conditions are provided by enforcing $C^2$ smoothness of the solution at the switching point Equation (21). Hence we now formulate two (well-posed) boundary value problems:

$$L_-[V_\ell] = (1-\epsilon)\tilde{h}(y) \qquad 0 < y < y^* \tag{23}$$

$$V_\ell(0) = 0, \qquad (1-y^*)V_\ell'(y^*) = \tilde{h}(y^*) \tag{24}$$

and

$$L_+[V_r] = (1+\epsilon)\tilde{h}(y) \qquad y^* < y < 1 \tag{25}$$

$$V_r(1) = \frac{1+\epsilon}{\rho}\tilde{h}(1), \qquad (1-y^*)V_r'(y^*) = \tilde{h}(y^*) \tag{26}$$

Once solved, enforcing continuity uniquely defines the value of the switching point $y^*$:

$$V_\ell(y^*) = V_r(y^*) \tag{27}$$

### A.1. Constructing the Solution

We shall construct the solutions $V_\ell$ and $V_r$ separately first, although the approach for both solutions will be the same. Following standard methods, we first decompose the full solution into a homogeneous and particular solution, where the homogeneous part satisfies $L_\pm[u] = 0$. Since the operators $L_\pm$ are equi-dimensional, we seek solutions of the form

$$u = c(1-y)^\gamma \tag{28}$$

and the application of the differential operator yields

$$L_\pm[u] = c(1-y)^\gamma \left( -\frac{\sigma^2}{2}(\gamma^2 - \gamma) \mp \epsilon\gamma + \rho \right) = 0 \tag{29}$$

We set the parenthetical term of this expression to zero in order to solve for the admissible exponents $\gamma$

$$\gamma_\pm^\ell = \frac{1}{2} + \frac{\epsilon}{\sigma^2} \pm \sqrt{\left(\frac{1}{2} + \frac{\epsilon}{\sigma^2}\right)^2 + 2\frac{\rho}{\sigma^2}} \qquad \text{(from } L_-\text{)} \tag{30}$$

$$\gamma_\pm^r = \frac{1}{2} - \frac{\epsilon}{\sigma^2} \pm \sqrt{\left(\frac{1}{2} - \frac{\epsilon}{\sigma^2}\right)^2 + 2\frac{\rho}{\sigma^2}} \qquad \text{(from } L_+\text{)} \tag{31}$$

Because $\gamma_-^r < 0$ for $\rho > 0$, the homogeneous solution $(1-y)^{\gamma_-^r}$ will be unbounded as $y \to 1$, and so we exclude it from further consideration below. Furthermore, since $\tilde{h}(y) = A + Cy$ is a linear function, the particular solution will also be linear. We therefore have a general solution of the form

$$V_\ell = c_1(1-y)^{\gamma_-^\ell} + c_2(1-y)^{\gamma_+^\ell} + \alpha_\ell + \beta_\ell y \tag{32}$$

$$V_r = c_3(1-y)^{\gamma_+^r} + \alpha_r + \beta_r y \tag{33}$$

The particular solution is fixed by enforcing that the differential Equations (23) and (25) are satisfied, which yields

$$\alpha_\ell = \left(\frac{1-\epsilon}{\rho}\right)(A+C) - \left(\frac{1-\epsilon}{\rho+\epsilon}\right)C, \qquad \beta_\ell = \left(\frac{1-\epsilon}{\rho+\epsilon}\right)C \tag{34}$$

$$\alpha_r = \left(\frac{1+\epsilon}{\rho}\right)(A+C) - \left(\frac{1+\epsilon}{\rho-\epsilon}\right)C, \qquad \beta_r = \left(\frac{1+\epsilon}{\rho-\epsilon}\right)C \tag{35}$$

The homogeneous coefficients $c_1$ and $c_2$ are then found by applying the boundary conditions (24), and the resulting linear system

$$\begin{pmatrix} 1 & 1 \\ \gamma_-^\ell(1-y^*)^{\gamma_-^\ell} & \gamma_+^\ell(1-y^*)^{\gamma_+^\ell} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -\alpha_\ell \\ \beta_\ell(1-y^*) - h(y^*) \end{pmatrix}$$

is solved by

$$c_1 = -\frac{\alpha_\ell\gamma_+^\ell(1-y^*)^{\gamma_+^\ell} + \beta_\ell(1-y^*) - \tilde{h}(y^*)}{\gamma_+^\ell(1-y^*)^{\gamma_+^\ell} - \gamma_-^\ell(1-y^*)^{\gamma_-^\ell}} \tag{36}$$

$$c_2 = \frac{\beta_\ell(1-y^*) - \tilde{h}(y^*) + \alpha_\ell\gamma_-^\ell(1-y^*)^{\gamma_-^\ell}}{\gamma_+^\ell(1-y^*)^{\gamma_+^\ell} - \gamma_-^\ell(1-y^*)^{\gamma_-^\ell}} \tag{37}$$

The boundary condition Equation (26) at $y = 1$ is automatically satisfied by the particular solution, and so the final coefficient $c_3$ is determined by the condition at $y = y^*$,

$$c_3 = \frac{\beta_r(1-y^*) - \tilde{h}(y^*)}{\gamma_+^r(1-y^*)^{\gamma_+^r}} \tag{38}$$

Finally, having determined the general solutions $V_\ell$ and $V_r$, the solution $V$ is obtained by enforcing continuity. This also fixes the value $y^*$, which must now satisfy the nonlinear transcendental equation

$$c_1(1-y^*)^{\gamma_-^\ell} + c_2(1-y^*)^{\gamma_+^\ell} + \alpha_\ell + \beta_\ell y^* = c_3(1-y^*)^{\gamma_+^r} + \alpha_r + \beta_r y^* \tag{39}$$

By expanding the expression, it follows that

$$0 = \left((1-y^*)^{(\gamma_+^\ell-\gamma_-^\ell)} - 1\right)\left(\beta_\ell(1-y^*) - \tilde{h}(y^*)\right) - \alpha_\ell(\gamma_+^\ell - \gamma_-^\ell)(1-y^*)^{\gamma_+^\ell}$$

$$+ \left(\gamma_+^\ell(1-y^*)^{(\gamma_+^\ell-\gamma_-^\ell)} - \gamma_-^\ell\right)\left(\alpha_\ell - \alpha_r + (\beta_\ell - \beta_r)y^* - \frac{\beta_r(1-y^*) - \tilde{h}(y^*)}{\gamma_+^r}\right) \tag{40}$$

*A.2. Comparison With $\mu \equiv 0$.*

If the seller decides to take no action to enhance her reputation and long term wealth growth by letting $\mu \equiv 0$, then her value function solves

$$-\frac{\sigma^2}{2}(1-y)^2 V_0''(y) + \rho V_0(y) = \tilde{h}(y) \qquad 0 < y < 1$$

$$V_0(0) = 0 \tag{41}$$

$$V_0(1) = \frac{1}{\rho}\tilde{h}(1)$$

which leads to a solution

$$V_0(y) = \frac{\tilde{h}(y)}{\rho} - \frac{C}{\rho}(1-y)^{\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{2\rho}{\sigma^2}}} \tag{42}$$

It follows that for $V$ which solves $(3.15)$ that

$$\frac{A+C}{\rho}\epsilon = V(1) - V_0(1) \leq \sup_{0\leq y\leq 1}\left|V(y) - V_0(y)\right| \tag{43}$$

Hence, she enhances her overall expected wealth by order $\epsilon$ by pulsing strategies.

## References

1. Sethi, S.P. Deterministic and stochastic optimization of a dynamic advertising model. *Optim. Control Appl. Methods* **1983**, *4*, 179–184.
2. Raman, K. Boundary value problems in stochastic optimal control of advertising. *Automatica* **2006**, *42*, 1357–1362.
3. Mink, M.; Seifert, S. Reputation on eBay and its Impact on Sales Prices. In Proceedings of The Group Decision and Negotiation International Conference, Karlsruhe, Germany, 25–28 June 2006, pp. 253–255.
4. Beam, C.; Segev, A. *Auctions on the Internet: A Field Study*; Blackwell Publishers Ltd: Oxford, UK, 1998.
5. Bajari, P.; Hortacsu, A. Economic insights from internet auctions. *J. Econ. Lit.* **2004**, *42*, 457–486.
6. Resnick, P.; Zeckhauser, R.; Swanson, J.; Lockwood, K. The Value of Reputation on eBay: A Controlled Experiment. *Exp. Econ.* **2006**, *9*, 79–101.
7. Resnick, P.; Zeckhauser, R. Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. In *The Economics of the Internet and E-Commerce*; Baye, M.R., Ed.; Volume 11 of Advances in Applied Microeconomics. Elsevier Science: Amsterdam, The Netherlands, 2002; pp. 127–157.
8. Houser, D.; Wooders, J. Reputation in Auctions: Theory, and Evidence from eBay. *J. Econ. Manag. Strateg. Blackwell Publ. Vol.* **2006**, *15*, 353–369.
9. Dellarocas, C. Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard. *Inf. Syst. Res.* **2005**, *16*, 209–230.
10. Apt, K.R.; Markakis, E. Optimal Strategies in Sequential Bidding. In Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems, Budapest, Hungary, 10–15 May 2009; pp. 1189–1190.
11. Nerlove, M.; Arrow, K.J. Optimal Advertising Policy under Dynamic Conditions. *Economica* **1962**, *29*, 129–142.
12. Rao, R.C. Estimating Continuous Time Advertising-Sales Models. *Market. Sci.* **1986**, *5*, 125–142.
13. Levitin, D.J. *Extracted from The Organized Mind: Thinking Straight in the Age of Information Overload*; Dutton Penguin Random House: New York, NY, USA, 2014.
14. Duckworth, K.; Zervos, M. A Model for Investment Decisions with Switching Costs. *Ann. Appl. Probab.* **2001**, *11*, 239–260.

15. Fleming, W.; Soner, H.M. *Controlled Markov Processes and Viscosity Solutions*; Stochastic modelling and applied probability, Springer: New York, NY, USA, 2006.

16. Fleming, W.H.; Rishel, R.W. *Deterministic and Stochastic Optimal Control*; Springer-Verlag: Berlin, Germany; New York, NY, USA, 1975.