

# The Value of a "Free" Customer

Sunil Gupta  
Carl F. Mela  
Jose M. Vidal-Sanz<sup>1</sup>

August 25, 2006

<sup>1</sup>Sunil Gupta (sgupta@hbs.edu) is a Professor of Business Administration, Harvard Business School, Soldiers Field, Boston, Massachusetts 02163. Carl F. Mela (mela@duke.edu) is a Professor of Marketing, The Fuqua School of Business, Duke University, Durham, North Carolina, 27708. José M. Vidal-Sanz (jvidal@emp.uc3m.es) is an Assistant Professor of Marketing, Universidad Carlos III, Calle Madrid 126, 28903 Getafe (Madrid), Spain. We thank seminar participants at the 2006 Marketing Dynamics Conference, 2006 Marketing Science Conference, and 2006 Yale Center for Customer Insights Conference for their comments.

## **Abstract: The Value of a "Free" Customer**

Central to a firm's growth and marketing policy is the revenue and profit potential of its customer assets. As a result, there has been a recent proliferation of work regarding customer lifetime value. However, extant research in this area is silent regarding how to assess the profitability of customers in a networked setting. In such settings, the presence of one type of customer can affect the value of another. Examples of such settings include job agencies (whose customers include both job seekers and listers), realtors (whose clients include home sellers and purchasers), and auction houses (whose customers include buyers and sellers). Customers such as buyers of an auction house pay no fees to the firm making their value difficult to compute. Yet these customers generate value to the firm because their presence attracts fee-paying sellers. In this paper we consider the value of a customer in these types of networked setting.

We compute the value of customers by developing a joint model of buyer and seller growth. This growth comes from three sources – marketing actions (price and advertising), direct network effects (e.g., buyer to buyer effects), and indirect network effects (e.g., buyer to seller effects). Using this growth model we concurrently solve the firm's problem of choosing optimal pricing and advertising subject to constraints on customer growth. By relaxing constraints on growth by one customer, we can then impute their lifetime value to the firm. We apply our model to data from an auction house.

Our results show that there are strong direct and indirect network effects present in our data. We find that in the most recent period buyers have a value of about \$550 and the sellers have a value of around \$500. We also find that our approach leads to estimates of firm value that are more accurate than models that fail to consider network effects. Finally, price and advertising elasticities are low (-0.16 and 0.006) and decrease over time as network effects become increasingly important.

Keywords: Customer Lifetime Value; CRM; Optimal Control; Dynamic Programming.

# 1 Introduction

Metrics of customer value are becoming more important as firms are increasingly compelled to justify the role of marketing investments on firm profitability (Marketing Science Institute 2006). A central metric for assessing the profitability of customers is customer lifetime value (CLV); the present value of all future profits generated by a customer (Kamakura et al. 2005). Using CLV a firm can rank order its customers or classify them into tiers based on their expected profitability. This allows firms to appropriately allocate resources across high versus low value customers (Reinartz and Kumar 2003, Rust, Lemon and Zeithaml 2004, Venkatesan and Kumar 2004). CLV can also be used for making customer acquisition decisions such that a firm does not spend more on acquiring a customer than the CLV of that customer (Gupta and Lehmann 2003, Gupta and Zeithaml 2006). It allows firms to balance their resources between customer acquisition and customer retention (Reinartz, Thomas and Kumar 2005). Recent studies also show that CLV can provide a link between customer value and firm value (Gupta, Lehmann and Stuart 2004).

Current models of CLV, however, omit an important element. Consider the case of Monster.com, an employment market place where job-seekers post their resumes and firms sign up to find potential employees. Monster provides this service free to job-seekers and obtains revenue by charging fees to the employers. A natural question arising from this business model is how much Monster should spend to acquire a job-seeker. Traditional models of CLV can not answer this question since job-seekers do not provide any direct revenue. In fact, if one includes the cost of maintaining resumes, the standard CLV for a job-seeker is negative. However, without job-seekers employers will not sign up, and without these firms Monster will have no revenues or profits. In other words, the value of job-seekers is through their *indirect network effect* on job listers.<sup>1</sup> This indirect network effect is not limited to employment services only (e.g., Monster, Hotjobs, Craigslist) but also extends to any exchange with multiple buyers and sellers (e.g. eBay, real estate).

The purpose of this study is to assess customer value when two parallel populations (e.g., buyers and sellers) interact and have strong 1) direct (within population) and 2) indirect (across populations) network effects. In these situations typically one set of customers (e.g., sellers) provide direct financial returns to the company. For example, sellers provide commissions to real estate

---

<sup>1</sup>Such markets are denoted as two-side matching markets in Sociology and Economics (Demange and Gale 1985). Our emphasis differ from this work inasmuch as we are not as concerned with matching specific employers with prospects as we are with understanding how the presence of one of these populations attracts the other.

agencies. However, firms must acquire and maintain the other set of customers (e.g., buyers). These customers are "free" as they do not provide any direct revenue. Our objective is to develop a model to assess the value of both types of customers.<sup>2</sup> This should enable us to answer the following questions:

- Are direct and indirect network effects sizable in practice? If so, this suggests the potential for large firms with strong network effects to dominate markets as the network grows.
- How much should a company spend to acquire new customers in the presence of these network effects? For example, how much should Monster spend on acquiring an additional job-seeker, how much should PayPal spend on acquiring a new account, how much should a dating service spend for a new client, how much should an auction house spend to acquire a new buyer? In other words, how much should a company spend on acquiring a "free" customer who does not provide any direct financial returns to the firm. This question is key to firms that operate in markets with strong direct and indirect network effects.
- How does the value of a customer change over time? Since the magnitude of network effects is likely to be time-dependent, customer value should also change over time. This suggests that the maximum amount of money that a firm should spend on a "free" customer changes over the life of a company. Once again, existing models of CLV do not address this dynamic aspect.
- How do we apportion value between buyers and sellers? In other words, how much of the value arising from the exchange between buyers and sellers accrues from each set of customers? Currently, firms have no metrics to apportion these revenues or profits. Some firms apportion all value to sellers since they generate the revenues. However, this clearly understates the value of buyers. Others use an arbitrary rule of thumb (e.g., 50-50) to split the profits between buyers and sellers. However, this implicitly assumes that both parties are equally important, which may not be necessarily correct. Finally, the network effects of buyers and sellers are likely to be different. This suggests that the proportion of value allocated between buyers and sellers may also change over time.

---

<sup>2</sup>Throughout the paper we will use the terms buyers and sellers for the two parallel population of customers.

- How should firms' marketing efforts change over time in the presence of network effects? How do customers' sensitivities to these marketing actions change over time? Marketing actions may be more critical in the early stages and network effects may dominate in later stages of a firm's life cycle. It is worth noting that extant CLV models fail to account for the role customer acquisitions play on subsequent marketing spending. As the market saturates, for example, optimal marketing spend may decrease even in the absence of network effects. Hence, there is an indirect link between customer acquisition and marketing spend, and this should logically affect customer valuation.
- Does the omission of network effects understate the value of customer base and hence the value of a firm? For example, Gupta et al. (2004) found that eBay's customer value is significantly lower than its market value. They speculate that this difference may be due to the omission of indirect network effects from their model.

While economists and marketing researchers have studied network effects in many contexts (e.g., Katz and Shapiro 1985, 1986, Neil, Kende and Rob 2000, Gupta, Jain and Sawhney 1999, Yao and Mela 2006), none of these studies examined the impact of these network effects on customer value. One recent study estimated the value of a lost customer by accounting for the word-of-mouth or direct network effects and found these effects to be very large (Hogan, Lemon and Libai 2004). However, we are not aware of any study that examines customer value when there are strong direct and indirect network effects. The indirect network effects are especially important in the buyer-seller contexts that we study here. Modeling these effects is crucial to assess the value of "free" customers who do not provide any direct revenue to the firm.

The paper proceeds as follows. We begin by developing a model that captures the growth of buyers and sellers from three sources – marketing actions (price and advertising), direct network effects or word-of-mouth, and indirect network effects. Next we define the firm's problem as an optimal control problem wherein the firm chooses its marketing actions to maximize its long run profits subject to the growth of these populations. These growth constraints imply costates or Lagrangian multipliers for the optimization problem yielding the incremental profits to a firm arising from an additional buyer or seller; that is, the lifetime value of that incremental seller or buyer to the firm. We apply our model to data obtained from an auction house, and estimate the

model using a Generalized Method of Moments based on both the growth and the Euler equations. This estimation approach explicitly accounts for endogeneity of marketing actions and increases the efficiency of the model estimates. We use the resulting parameter estimates to address the managerial questions highlighted above. We then conclude with limitations and next steps.

## 2 A Model of Customer Value in the Presence of Network Effects

### 2.1 Customer Growth and Network Effects

Consider two parallel populations of buyers and sellers. The acquisition of customers in each group can be captured by a "diffusion-type" model as follows:<sup>3</sup>

$$\dot{N}_t^B = \left( a(A_t) + b' \frac{N_t^B}{M^B} + g \frac{N_t^S}{M^S} \right) (M^B - N_t^B) - r(N_t^B) \quad (1)$$

$$\dot{N}_t^S = \left( \alpha(p_t) + \beta' \frac{N_t^S}{M^S} + \gamma \frac{N_t^B}{M^B} \right) (M^S - N_t^S) - \rho(N_t^S) \quad (2)$$

where  $N_t^B$  and  $N_t^S$  are the number of buyers and sellers at time  $t$ ,  $\dot{N}_t^B$  and  $\dot{N}_t^S$  denote the derivatives or the net customer acquisition rate of buyers and sellers,  $M^B$  and  $M^S$  are the potential market size of buyers and sellers,  $A_t$  is the advertising spending at time  $t$  and  $p_t$  is the price firm charges its sellers at time  $t$ . Some of the key characteristics of this system of equations are as follows:

1. The term  $a(A_t)$  recognizes that a firm can accelerate the growth of its buyers through buyer-targeted advertising. In our application of the e-auction house this takes the form of television and internet advertising. Consistent with prior literature (e.g., Horsky and Simon 1983), we further assume that  $a(A_t) = x + h \ln A_t$ . In other words, there are diminishing marginal returns from advertising and the coefficient  $h$  determines buyers' responsiveness to firm's advertising. Advertising is a decision variable for the firm and it can change over time. In our data (see Section 3), buyer-oriented advertising is the only marketing variable on the buyer side.
2. Similarly, the term  $\alpha(p_t)$  highlights the fact that the growth of sellers depends on the price the firm charges them. Once again, price is a decision variable and it can change over time.

---

<sup>3</sup>Similar models have been previously used in the context of international diffusion of products (Kumar and Krishnan 2002).

We assume that  $\alpha(p_t) = \phi - \theta \ln p_t$ . Here the parameter  $\theta$  indicates sellers' sensitivity to firm's pricing. In our data, price is the only marketing variable available on the seller side.

3. The direct network or word-of-mouth effect for buyers and sellers is captured by the second term in equations (1) and (2). This formulation is consistent with the diffusion literature. Hogan, Lemon and Libai (2003) used a similar term to capture the direct network effect of losing a customer. The value of the parameters  $b'$  and  $\beta'$  indicate the strength of direct network effects. These may be positive as a result of word of mouth, or negative as a result of a crowding phenomenon where, all else equal, a buyer (or seller) prefers less competition.
4. As a firm acquires more buyers it becomes more attractive for sellers to join the firm as well. The reverse is also true – the more sellers a firm has, the more buyers it is likely to attract. This indirect network effect is captured by the third term in equations (1 and 2). The value of the parameters  $g$  and  $\gamma$  indicate the strength of indirect network effects.
5. Unlike the traditional diffusion models, in our context customers can also leave a firm. Customer defection is captured by the terms  $r(N_t^B)$  and  $\rho(N_t^S)$ . In particular, the number of customers who defect a firm is proportional to the number of current customers, and the parameters  $r$  and  $\rho$  denote the rate of defection. Some studies exogenously specify the defection rate (Gupta et al. 2004, Libai, Muller, and Peres 2006). Alternatively, one can assess the net impact of acquisition and defection by reformulating the above equations as follows:<sup>4</sup>

$$\dot{N}_t^B = \left( a(A_t) + b \frac{N_t^B}{M^B} + g \frac{N_t^S}{M^S} \right) (M^B - N_t^B) \quad (3)$$

$$\dot{N}_t^S = \left( \alpha(p_t) + \beta \frac{N_t^S}{M^S} + \gamma \frac{N_t^B}{M^B} \right) (M^S - N_t^S) \quad (4)$$

Here the parameter  $b$  and  $\beta$  capture the net effect of acquisition through word-of-mouth and defection through customer dissatisfaction. A negative value for these parameters will indicate that customer defection is greater than the word-of-mouth effect (or possibly that all else equal, sellers and buyers prefer less competition).

---

<sup>4</sup>In principle it is possible to estimate  $b'$  and  $r$  (and similarly  $\beta'$  and  $\rho$ ) separately. However, it is easy to show that when  $\frac{(M^B - N_t^B)}{M^B} \approx 1$ ,  $b = b' + r$ , thus making it difficult to estimate  $b'$  and  $r$  separately. In our data, estimates of  $\frac{(M^B - N_t^B)}{M^B}$  average about 0.9 and  $\frac{(M^S - N_t^S)}{M^S}$  averages about 0.85. Given these are close to 1.0, we were unable to separately identify  $b'$  and  $r$  in (1) and (2). Therefore we rely on the formulation given by (3) and (4) in our subsequent modeling discussion.

## 2.2 Optimal Marketing Policies and Customer Value

Equations (3) and (4) characterize the growth of buyers and sellers as a result of firm's actions (advertising and pricing), word-of-mouth or direct network effects as well as indirect network effects where an increase in the number of buyers makes it more attractive for sellers to join the firm and vice versa. The objective of the firm is to choose its advertising and pricing policies in such a fashion that they maximize its long run profits. More formally, the objective function of the company is:

$$\max_{\{p_t, A_t, N_t^B, N_t^S\}} \int_0^{\infty} (m_t(p_t) N_t^S - A_t) e^{-it} dt \quad (5)$$

subject to the growth equations (3) and (4) and initial conditions  $N_0^B = N_0^S = 0$ . Here the control variables are  $p_t$  and  $A_t$ , and the state variables are  $N_t^B$  and  $N_t^S$ . Further,  $m_t(p_t) = p_t - c_t$  is the margin,  $p_t$  is the price and  $c_t$  is marginal cost at time  $t$ . The discount rate for future profits is  $i$ .

We would like to highlight a few characteristics of equation (5). First, the profits for the company depend directly on the number of sellers and the price the firm charges them. If there is no indirect network effect of buyers on sellers, the number of buyers is irrelevant for profit maximization. In such a situation a firm has no reason to spend any money on buyer-oriented advertising and it has no way of assessing the long-term value of a buyer. This highlights the importance of indirect network effects in our context. Second, this formulation is consistent with the concepts of CLV and customer equity since it explicitly accounts for the long run profitability of current and future customers. Third, unlike most CLV models, this equation suggests that firm's actions (price and advertising) can influence customer growth and hence the overall value of the firm. In other words, price and advertising decisions are dynamic and endogenous. As a result, they are affected by customer acquisitions and should therefore be considered when computing CLV.

If we define  $G(t, N_t^B, N_t^S, p_t, A_t)$  to be the profit function (i.e., the terms within the integral) in equation (5),  $g_B(A, N^B, N^S)$  to be the acquisition rate of buyers as given by equation (3) and  $g_S(p, N^B, N^S)$  to be the acquisition rate of the sellers as indicated in equation (4), then the profit maximization of equation (5) becomes a standard optimal control problem whose solution can be determined by Pontriagyn's maximum principle (see Appendix A1). To solve this problem we form

the following Hamiltonian function

$$H(t, N^B, N^S, p, A, \lambda^B, \lambda^S) = G(t, N^B, N^S, p, A) + \lambda^B g_B(A, N^B, N^S) + \lambda^S g_S(p, N^B, N^S), \quad (6)$$

where  $\lambda^B, \lambda^S$  are auxiliary or costate variables. The solution to this optimal control problem satisfies the conditions indicated in Appendix A1.

The costates can be interpreted as Lagrangian multipliers. In the optimal solution,  $\lambda_t^B$  provides the customer value of an incremental buyer acquired at time  $t$  over an infinite horizon, i.e. the effect of an additional buyer on the long-term discounted profit of the company. Analogously,  $\lambda_t^S$  provides the customer value of an additional seller acquired at time  $t$ . Further, the value of an additional buyer or seller is allowed to vary over time. This intuitively makes sense as one might expect the network effects to vary over the lifecycle of the company. For example, in the early stages of a company, marketing actions may be more important to attract customers, while in the growth phase direct and indirect network effects may dominate. In other words, our model will allow us to find out the maximum amount of money a firm should spend to acquire a "free" buyer – something that traditional CLV models can not address. Further, we will also be able to show how this maximum acquisition cost varies over time.

## 2.3 Model Estimation

### 2.3.1 Customer Growth Model

For estimation purposes we make two adjustments to our model presented in equations (3) and (4). First, our data are in discrete time so the model formulation is changed accordingly. Second, an estimation error is included. These modifications are typical in empirical estimation of diffusion type models. Consequently, the customer growth model for estimation purposes becomes:

$$N_{t+1}^B - N_t^B = \left( x + h \ln A_t + b \frac{N_t^B}{M^B} + g \frac{N_t^S}{M^S} \right) (M^B - N_t^B) + e_t, \quad (7)$$

$$N_{t+1}^S - N_t^S = \left( \phi - \theta \ln P_t + \beta \frac{N_t^S}{M^S} + \gamma \frac{N_t^B}{M^B} \right) (M^S - N_t^S) + \varepsilon_t \quad (8)$$

In the past, these growth models have often been estimated using OLS, nonlinear least squares, seemingly unrelated regressions or similar approaches. In many cases the purpose of the model estimation was simply to assess the impact of a marketing variable on the diffusion process. For

example, Simon and Sebastian (1987) investigate the impact of advertising on the diffusion of new telephones in West Germany. These studies are descriptive in nature and do not attempt to provide optimal advertising or pricing policies that maximize firm's profits. In our context this is problematic as we seek to assess the impact of demand on firm profits in order to infer CLV. This necessitates an analysis of the firm's decisions of price and advertising, which we consider next.

### **2.3.2 Firm's Price and Advertising Decisions**

Studies that combine empirical estimation of diffusion-type models with analytical models of profit maximization and optimal marketing policies of the firm typically do it in two stages. In the first stage these studies ignore the optimal control problem of the firm and use the actual prices and advertising of the firm as exogenous variables when estimating the growth model of demand. In the second stage, they "plug-in" the parameter estimates of price and advertising in the optimal solutions of advertising and price to arrive at the optimal path for these decision variables and compare the optimal and actual values (e.g., Horsky and Simon 1983, Kalish 1985, Chintagunta and Vilcassim 1992, Chintagunta and Rao 1996). A third set of studies explore the dynamic policies in a purely theoretical fashion using the solutions of the optimal control problem and examining the comparative statics or using numerical illustrations (e.g., Feichtinger, Hartl and Sethi 1994, Thompson and Teng 1984, Horsky and Mate 1988, Dockner and Jorgensen 1988).

All these approaches have significant limitations in the context of CLV. Theoretical approaches provide directional results but are not very useful if the objective is to provide empirical estimates in a particular application. The empirical approaches used in the past have two major drawbacks. First, the estimation of the diffusion equations assumes price and advertising to be exogenous. If the objective of the firm is to find the optimal price and advertising policy then these marketing actions by definition depend on the customer growth. In other words, marketing activities should be endogenous. One ramification of this observation is that acquisition of a customer may increase the network effect which in turn may reduce the firm's marketing cost. As such, marketing spend should have an indirect affect on customer value. In principle this endogeneity issue can be solved by using an instrumental variable approach. However, this only solves half of the problem. The second drawback of the traditional approaches is that the first stage estimation of the diffusion model is completely divorced from the second-stage "plug-in" analysis of the firm's optimal control

problem. This second-stage analysis is necessary to make inferences about customer profitability as it yields information needed to compute the Lagrangians which yield the measures of buyer and seller CLV.

Therefore, we propose an estimation method that treats firm's actions (price and advertising) as endogenous and at the same time considers firm's objective function of maximizing long run profits. Specifically, the firm solves the following discrete time problem:

$$\begin{aligned} & \max_{\{p_t, A_t\}} E_0 \left[ \sum_{t=0}^{\infty} (1+i)^{-t} (N_t^S S p_t - A_t) \right] \\ & \text{s.t.} \\ N_{t+1}^B &= N_t^B + \left( x + h \ln A_t + b \frac{N_t^B}{M^B} + g \frac{N_t^S}{M^S} \right) (M^B - N_t^B) + e_t \\ N_{t+1}^S &= N_t^S + \left( \phi - \theta \ln p_t + \beta \frac{N_t^S}{M^S} + \gamma \frac{N_t^B}{M^B} \right) (M^S - N_t^S) + \varepsilon_t \end{aligned} \quad (9)$$

and subject to initial values  $N_0^S = 0$ ,  $N_0^B = 0$ . In compact notation, this problem can be expressed as

$$\begin{aligned} & \max_{\{p_t, A_t\}} E_0 \left[ \sum_{t=0}^{\infty} \delta^t G(N_t^B, N_t^S, p_t, A_t) \right] \\ N_{t+1}^B &= g^B(A_t, N_t^B, N_t^S, e_t) \\ N_{t+1}^S &= g^S(p_t, N_t^B, N_t^S, \varepsilon_t) \end{aligned} \quad (10)$$

where  $\delta = (1+i)^{-1}$ . The solution can be expressed by the first-order conditions associated with the following Lagrangian function

$$\begin{aligned} L(\{p_t, A_t, N_t^B, N_t^S, \lambda_t^B, \lambda_t^S\}) &= E_0 \left[ \sum_{t=0}^{\infty} \delta^t G(N_t^B, N_t^S, p_t, A_t) \right] \\ &+ \sum_{t=0}^{\infty} \lambda_{t+1}^B (N_{t+1}^B - g^B(A_t, N_t^B, N_t^S, e_t)) \\ &+ \sum_{t=0}^{\infty} \lambda_{t+1}^S (N_{t+1}^S - g^S(p_t, N_t^B, N_t^S, \varepsilon_t)). \end{aligned} \quad (11)$$

Equation (11) affords a natural metric for the CLV of buyers and sellers at time  $t$ ; these metrics are the Lagrangian multipliers  $\lambda_t^B$  and  $\lambda_t^S$ . These parameters represent the constraints on profits arising from not being able to add one customer at time  $t$ . Stated differently, the lifetime value to the firm arising from the growth of its customer base by one additional customer is given by  $\lambda_t^B$  and  $\lambda_t^S$ . Thus, estimates of the buyer and seller CLV are outputs of our model. We discuss the calculation of these parameters in more detail in Section 4.3.

### 2.3.3 The Joint System of Estimation Equations

To obtain our estimation equations we define the optimal value function of the firm as:

$$V(N_0^S, N_0^B, e_0, \varepsilon_0) = \max_{\{p_t, A_t\}} E_0 \left[ \sum_{t=0}^{\infty} \delta^t G(N_t^B, N_t^S, p_t, A_t) \right]. \quad (12)$$

Then, the Jacobi-Bellman condition states that the solution satisfies for each integer  $t \geq 0$ ,

$$V(N_t^B, N_t^S, e_t, \varepsilon_t) = \max_{p_t, A_t} \{ G(N_t^B, N_t^S, p_t, A_t) + \delta E_t [V(N_{t+1}^S, N_{t+1}^B, e_{t+1}, \varepsilon_{t+1})] \}. \quad (13)$$

From the Jacobi-Bellman dynamic programming condition we obtain the following Euler equations (for details on the derivation see Appendix A2),

$$E_t \left[ \left( \frac{\frac{N_t^S S}{\theta(M^S - N_t^S)/p_t}}{\frac{1}{h(M^B - N_t^B)/A_t}} \right) + (1+i)^{-1} \left( \begin{pmatrix} 0 \\ Sp_{t+1} \end{pmatrix} + \begin{pmatrix} D_{t+1}^{11} & D_{t+1}^{12} \\ D_{t+1}^{21} & D_{t+1}^{22} \end{pmatrix} \begin{pmatrix} \frac{N_{t+1}^S S}{(-\theta(M^S - N_{t+1}^S)/p_{t+1})} \\ \frac{-1}{h(M^B - N_{t+1}^B)/A_{t+1}} \end{pmatrix} \right) \right] = 0, \quad (14)$$

where the  $D$  terms are defined in Appendix A2. The system of conditional moments in equation (14), one for price and the other for advertising, can be denoted by

$$E_t [U(\Omega_0, N_t^B, N_t^S, p_t, A_t, N_{t+1}^B, N_{t+1}^S, p_{t+1}, A_{t+1})] = 0, \quad (15)$$

where  $\Omega_0$  denotes the true parameter vector. This implies that for any instrument  $Z_t$  predetermined at time  $t$ , the unconditional expectations are null, i.e.

$$E [U(\Omega_0, N_t^B, N_t^S, p_t, A_t, N_{t+1}^B, N_{t+1}^S, p_{t+1}, A_{t+1}) Z_t] = 0, \quad (16)$$

In addition we have the two moment conditions associated with the dynamics of the state variables (the growth of buyers and sellers), yielding:

$$E \left[ \left( N_{t+1}^B - N_t^B - \left( x + h \ln A_t + b \frac{N_t^S}{M^S} + g \frac{N_t^B}{M^B} \right) (M^B - N_t^B) \right) Z_t \right] = 0, \quad (17)$$

$$E \left[ \left( N_{t+1}^S - N_t^S - \left( \phi - \theta \ln p_t + \beta \frac{N_t^B}{M^B} + \gamma \frac{N_t^S}{M^S} \right) (M^S - N_t^S) \right) Z_t \right] = 0, \quad (18)$$

where  $Z_t$  includes all the regressors of the model and its lags, as well as some transformations (see Section (2.3.4) for a description these instruments). It is these four equations (price and advertising paths, buyer and seller growth models) crossed with each instrument that form the basis of our estimation equations.

Note, if we ignore the firm's optimization problem and simply estimate the diffusion models, we get equations (17) and (18) corresponding to the buyers and sellers. This is the typical approach of most empirical studies in the past. However, by explicitly incorporating the optimal control problem of the firm we have additional equations as given by (16). These additional equations provide structure to the problem and help to identify the model parameters in the empirical estimation. We use Generalized Method of Moments for parameter estimation (details on the GMM estimation are provided in Appendix A3).

### 2.3.4 Discount Rate and Instruments

We assume firm's monthly discount rate  $i = 0.015$ . This monthly discount rate of 1.5% translates into an approximate annual discount rate of 20%. Given the limited number of variables, we use two lags for the instruments. The use of lags for instruments is common and as we shall show, these appear to be good instruments (e.g., Kadiyali, Chintagunta, and Vilcassim 2000). Specifically,  $Z_t$  is a  $9 \times 1$  vector

$$Z_t = \left( 1, N_{t-1}^S, N_{t-2}^S, N_{t-1}^B, N_{t-2}^B, \log p_{t-1}, \log p_{t-2}, \log A_{t-1}, \log A_{t-2} \right)'$$

and since we have 4 equations, we use 36 moment equations and 10 parameters (with 26 degrees of freedom). These instruments are tested via Hansen's (1982) test of overidentifying restrictions, or  $J$ -statistic (see Appendix A3).

### 3 Application

The application of our model requires information on the number of sellers and buyers over time as well as the marketing expenditures invested in the acquisition of these customers. An anonymous auctions house provided monthly data on these quantities for a large market between February 1, 2001 and March 1, 2005. The firm obtains revenues from sellers who list items on its web site for auction. These revenues are obtained from a listing fee, some promotional fees and a commission on the sales proceeds; these are combined into an overall margin value by the firm and this is measure used in our application. If  $S$  is the average annual gross merchandise sold per seller, and  $p_t$  is the percentage commission charged by the auction house at time  $t$ , then firm's annual margin from each seller is  $Sp_t$ . The marginal costs in this business are close to zero and therefore we exclude them from our analysis. The buyers provide no direct revenue to the firm. Not surprisingly the firm has a greater interest in acquiring and maintaining its sellers even though it recognizes that it needs to have buyers for its auctions. While sellers push the firm to acquire more and more buyers, the firm is not sure of how much money to value these buyers – which is the central question of our research.

The auction house spends money on TV and Internet advertising to attract buyers. The advertising data were compiled quarterly which we converted to monthly data by dividing by three. In addition, the auction firm provided information on margins and total transacted volume. To protect the confidentiality of the firm, we are unable to report the specific data means, but in Figure 1 we present information regarding the number of customers and the marketing expenditures over time, normalized so that the maximum value of each series is one.

The upper left and right panels of Figure 1 indicate rapid growth in the number of buyers and sellers. There are approximately 4.6 buyers for each seller. The pricing data depicted in the lower left panel indicates that take rates (margins) are increasing slightly over time, a sign of increasing pricing power which we speculate may arise from the growth of the buyer-seller network. This trend suggests it is desirable to account for pricing power and the endogeneity of pricing when computing the value of a customer. Moreover, the modest change in margins indicate that the change in seller value over time is not so large that prices change considerably. Advertising, as indicated by the lower right panel in Figure 1 also shows an increase over time which may be due to firm's increasing concern of attracting new buyers over time or it may be due to reduced advertising

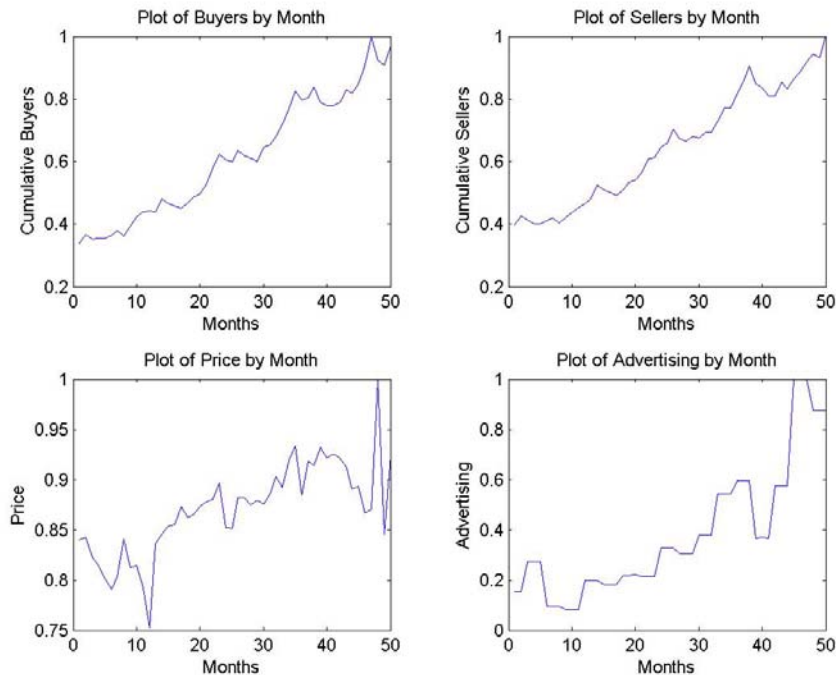


Figure 1: Changes in Marketing and Demand Over Time

sensitivity in the market prompting the firm to spend more to achieve the same results as before. In any event, increased spending on advertising is warranted only if buyer value is commensurate with this investment. These data serve as the basis of our empirical application.

## 4 Results

### 4.1 Model Fit

The overall model fit is given by the statistic  $J = T \cdot Q_{\widehat{W}}(\widehat{\Omega})$ , where  $Q_{\widehat{W}}(\widehat{\Omega})$  is defined in Appendix A3. The  $J$  statistics is distributed  $\chi^2_{r-k}$  where  $r$  is the number of equations and  $k$  the number of parameters. For our empirical application  $J = 9.89$ ,  $r = 36$  and  $k = 10$ . Therefore with  $Prob[\chi^2(26) > J] = 0.9982$  we accept the overidentifying moment conditions. In other words, the moment conditions are close to zero and the instruments are orthogonal to the error. The following figures depict the fit of the model to the data. In Figure 2 we present the model fit based on a one-step ahead forecasts for period  $t$  based on (7) and (8), and the estimation residuals. In Figure 3, we use actual  $\{A_t, p_t\}$  to generate a recursive forecast from period 0, and compare this

forecast with actual data. We conclude that the model provides a reasonably good fit to the data, as indicated by Figure 3.<sup>5</sup> The top panel of this figure provides one step ahead forecasts (using data up to period  $t - 1$  to forecast demand in  $t$ ) and the lower panel is a recursive formula using only information at  $t = 0$ .

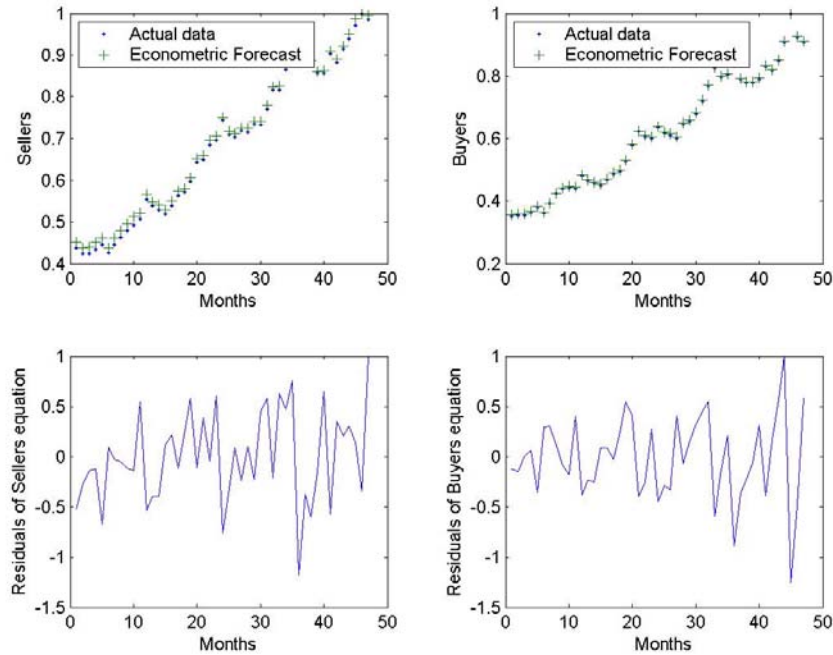


Figure 2: Model Fit and Residuals

## 4.2 Parameter Estimates

Table 1 presents the parameter estimates of our model along with the t-statistics.

---

<sup>5</sup>As the request of the sponsoring firm, the figure is normalized to disguise the data such that the maximum number of customers is scaled to one.

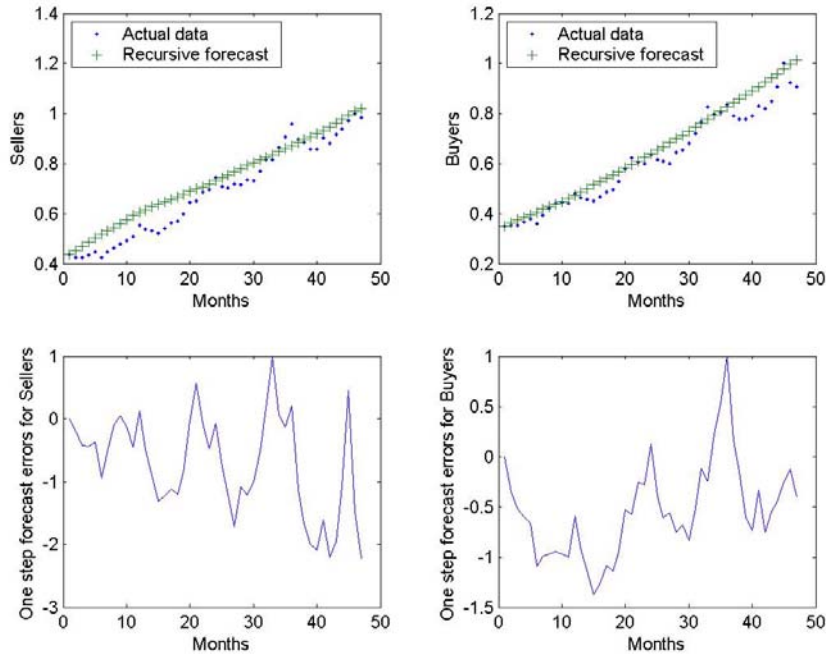


Figure 3: Recursive Forecast and Errors

	Parameter	Estimate	t-stat
<b>Buyers Equation</b>			
Intercept	$x$	-0.004	-4.26
Advertising	$h$	0.0003	5.90
Direct Network Effect of Buyers	$b$	-0.011	-2.45
Indirect Network Effect of Sellers	$g$	0.019	4.30
Potential Market Size of Buyers (million)	$M^B$	50.72	13.29
<b>Sellers Equation</b>			
Intercept	$\phi$	-0.037	-2.90
Price	$\theta$	0.015	2.96
Direct Network Effect of Sellers	$\beta$	-0.066	-2.95
Indirect Network Effect of Buyers	$\gamma$	0.095	2.13
Potential Market Size of Sellers (million)	$M^S$	7.50	4.29

Table 1: Parameter Estimates

All parameters are statistically significant. Advertising has a significantly positive impact on the acquisition of buyers. Price has a significant negative impact on sellers growth (recall  $\alpha(p_t) =$

$\phi - \theta \ln p_t$ , so price parameter  $\theta$  is expected to be positive). The parameters  $b$  and  $\beta$  are negative. As discussed earlier, these parameter captures the net effect of word-of-mouth as well as customer defection. The negative sign suggests that customer defection is outpacing the growth from word-of-mouth effect. This defection may also be a result of "crowding" where customers prefer to list/buy when fewer competitors are in the system.

The indirect network effects for both buyers and sellers are positive and very strong. This suggests that the more buyers we have in the system the more sellers are attracted to the firm and vice versa. Further, the indirect effect of buyers on sellers (0.095) is five times the indirect effect of sellers on buyers (0.019). In other words, even though buyers do not provide any direct revenue to the firm, they may be more critical for its growth. We will return to this issue when we translate these into the value of buyers and sellers. In sum, in response to the first question we raised in the introduction of this paper, we find that there are significant network effects present in our empirical application.

The market potential for buyers is estimated to be about 50.7 million and the number of sellers is estimated to be about 7.5 million. Though we can not reveal the specifics regarding the current market size in order to protect the confidentiality of the firm, based on our market potential estimates the market penetration of both populations is on the order of 1/5 the estimated market potential. This indicates that there are significant growth opportunities for this firm that should be reflected in its overall customer and firm value.

### **4.3 The Value of a Customer**

In this section, we address the three pertinent managerial questions we asked in the beginning of this paper.

1. What is the value of a buyer or a seller? In other words, what is the maximum the firm should spend on acquiring a buyer or a seller. This estimate is especially difficult for buyers who do not provide any direct revenue or profit to the firm.
2. How can the firm apportion the value of transactions between buyers and sellers?
3. How do these values change over time?

As indicated earlier, the value of a buyer is given by the Lagrangian multiplier of the optimal control problem. In Appendix A4 we develop an analog in the discrete time context so we can apply it to our data. In particular, we show that the shadow prices for buyers and sellers are given by

$$\begin{pmatrix} \lambda_{t+1}^B \\ \lambda_{t+1}^S \end{pmatrix} = -(1+i)^{-t} \begin{pmatrix} 0 \\ Sp_t \end{pmatrix} - (1+i)^{-t} \begin{pmatrix} D_t^{11} & D_t^{12} \\ D_t^{21} & D_t^{22} \end{pmatrix} \begin{pmatrix} \frac{N_t^S S}{(-\theta(M^S - N_t^S)/pt)} \\ \frac{-1}{h(M^B - N_t^B)/A_t} \end{pmatrix}. \quad (19)$$

Using this equation and the model parameters we estimate the current value of a buyer as approximately \$550. Note that the entire value of buyers is derived from their indirect network effect on the growth of sellers. Traditional models of CLV that do not account for these network effects are unable to estimate the value of these "free" customers.

The value of a seller at the current time is about \$500. Surprisingly the value of a seller is slightly less than that of a buyer. This is counter-intuitive for at least two reasons. Although the firm believes that buyers are important for its growth, its revenues are derived directly from sellers. Therefore, intuitively it makes sense to assume that the paying customers or sellers are more important for firm's profits. The second reason that supports the firm's intuition is the fact that there are approximately 4.6 buyers for each seller. Since each transaction requires a buyer and a seller, it is reasonable to argue that the value of a seller should be at least 4.6 times the value of a buyer. Our model results go against this intuition and suggest that the value of a seller is almost equal to or slightly less than the value of a buyer.

What explains this counter-intuitive result? First, Table 1 indicates that the parameter value for the indirect network effect of buyers on sellers growth (0.095) is almost five times the parameter value for the indirect network effect of sellers on buyers growth (0.019). Second, the net effect of word-of-mouth and attrition on sellers is  $-0.066$ , suggesting a natural tendency to attrite. Third, as indicated by  $\theta = 0.015$ , historical prices have had a negative impact on sellers' growth. Fourth, the negative intercept for seller's equation ( $\phi = -0.037$ ) suggests that there is no "organic" or "natural" growth of the seller population. In other words, except for the indirect network effect of buyers, all other factors are working against the growth of sellers. This makes the buyers even more critical for the overall growth and profitability of the firm. In the end, even though the firm has 4.6 buyers for each seller, the indirect network effect of a buyer is almost five times the indirect network effect of a seller. The net result of all these factors is such that the value of a seller is

almost equal to (or slightly less than) the value of a buyer.

Does the buyer and seller value change over time? Figure 4 shows these values over the time frame of our data. As expected these values change significantly over time since the networks effects vary over the lifecycle of the firm. Several important insights emerge from these results. First, the value of buyers and sellers has been growing steadily over time. In the beginning when the firm has limited number or no customers, marketing actions are the primary source of driving traffic. In other words in the introductory phase of the firm's lifecycle, marketing actions are likely to be more critical than the network effects. As the number of customers grow, the network effects become more important. Because the firm in our application is still in the growth phase of its lifecycle, it is not surprising to find that the network effects contribute more and more over time. Eventually, the impact of network effects will diminish. As the firm reaches the market potential for its customers, the value of an additional buyer or seller will approach zero. In other words, the value of a buyer and a seller is likely to follow a inverted U-shaped curve. Second, in the early time periods of our data, the value of a buyer is almost twice the value of a seller. Towards the end of our data period, the two values are almost equal. In other words, buyers are relatively more valuable in the early stages. This is due to the empirical result that the indirect network effect of buyers is stronger than that of sellers.

#### **4.4 Network Effects and Firm Value**

How do network effects influence the value of the customer base and hence the value of the firm? Hogan, Lemon and Libai (2003) incorporate direct network effects through a diffusion model to examine the value of a lost customer in online banking. They find that the CLV of a customer without direct network effects is about \$208. However, direct network effects can be as large as \$650 in the early stages of the diffusion process. Our work augments this research by considering the indirect network effects that are critical in a buyer-seller situation. In one setting wherein these effects might be considerable, Gupta et al. (2004) estimated the CLV at eBay. However, they did not consider indirect network effects and found that their estimate of customer value for eBay was significantly lower than the market value of the firm. They suggested, "...eBay is an auction exchange, and thus there may be significant network externalities that are not captured by the traditional diffusion model. Furthermore, eBay's business includes both buyers and sellers ... it

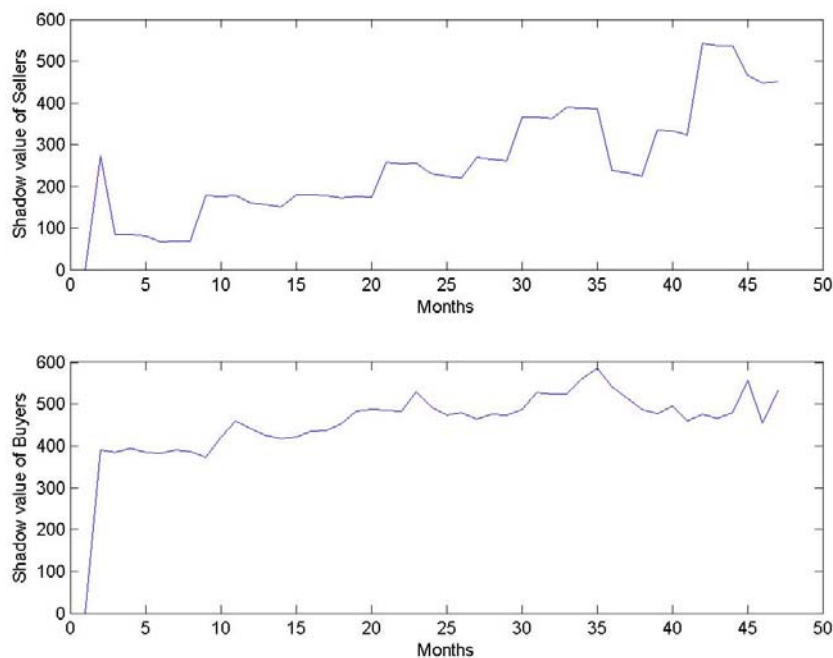


Figure 4: Value of a Buyer and a Seller Over Time

may be important to model buyers and sellers separately and then construct a model of interaction between them," (page 14). This is precisely what we have done in this paper.

Traditional models that do not include direct or indirect network effects estimate CLV based on annual margin and customer retention rate. Gupta and Lehmann (2003, 2005) suggest a simple formulation to estimate CLV as follows.

$$CLV = m \frac{r}{(1 + i - r)}$$

where  $m$  is the annual margin,  $r$  is the retention rate and  $i$  is the discount rate.

If we use this approach, the value is derived only from sellers since buyers do not generate any direct revenues or profits. Using an annual discount rate of 20%, retention rate of 80% as used by Gupta et al. (2004), our estimates for market potential for sellers and the current margin of sellers (which we can't disclose due to confidentiality reasons), we estimate the firm value using our data to be 2.5% of its market capitalization in the period considered. In contrast, when we account for direct and indirect network effects the value of a seller is about \$500, and the value of

a buyer is about \$550. Given the estimated market potential of 7.5 million sellers and 50.7 million buyers, a rough estimate of the customer value of this firm is  $\$500 \cdot 7.5 + \$550 \cdot 50.7 = \$31.7$  billion, which is closer to 80% of the market capitalization of this firm in the period considered.<sup>6</sup> Given these data are from only one geographic market of the firm (albeit the largest), this valuation seems reasonable and far better than one afforded by the traditional CLV approach that ignores network effects. Interestingly, buyers account for almost 90% of the total customer value. While we acknowledge that these are very crude estimates of market value, nonetheless they show that ignoring network effects can underestimate the firm value significantly.

#### 4.5 Price and Advertising Elasticities

Given the presence of strong network effects, how do the effectiveness of marketing actions change over time? To address this question, we estimate the price and advertising elasticities with respect to customer growth. In our model, advertising influences the growth of buyers and price affects the sellers' growth. The derivative of these marketing actions with respect to the number of buyers or sellers is given by the following equations.

$$\begin{aligned} \frac{\partial}{\partial A_t} g^B (A_t, N_t^B, N_t^S, e_t) &= \frac{h}{A_t} (M^B - N_t^B) \\ \frac{\partial}{\partial p_t} g^S (p_t, N_t^B, N_t^S, \varepsilon_t) &= -\frac{\theta}{p_t} (M^S - N_t^S) \end{aligned}$$

These equations capture the contemporaneous effects of advertising and price on the growth of buyers and sellers. Figure 5 shows the trajectory of the advertising and price elasticity over time.

Two interesting results emerge from this figure. First, both the advertising and the price elasticities are decreasing over time. In other words, as firms acquire more customers over time, network effects become increasingly important. This diminishes the impact of advertising and price on customer growth. Second, even in the early time periods sellers' price elasticities (-0.16) are small compared to the average price elasticity of -1.6 found by Tellis (1988) or -1.4 found by Bijmolt et al. (2004). This suggests that sellers have limited power given the auction house in the presence of strong network effects. Third, the advertising elasticity for buyers is a maximum of 0.006 in

---

<sup>6</sup>We recognize that for both approaches, strictly speaking, we should forecast the number of customers in the future as well their future value in each time period and use appropriate discounting. While this is straightforward in the traditional CLV model as shown by Gupta et al. (2004), it is a complex dynamic problem when direct and indirect network effects are included. We leave the exact solution for future research. Our purpose here is to get a rough estimation to show the significant differences that arise when we account for network effects.

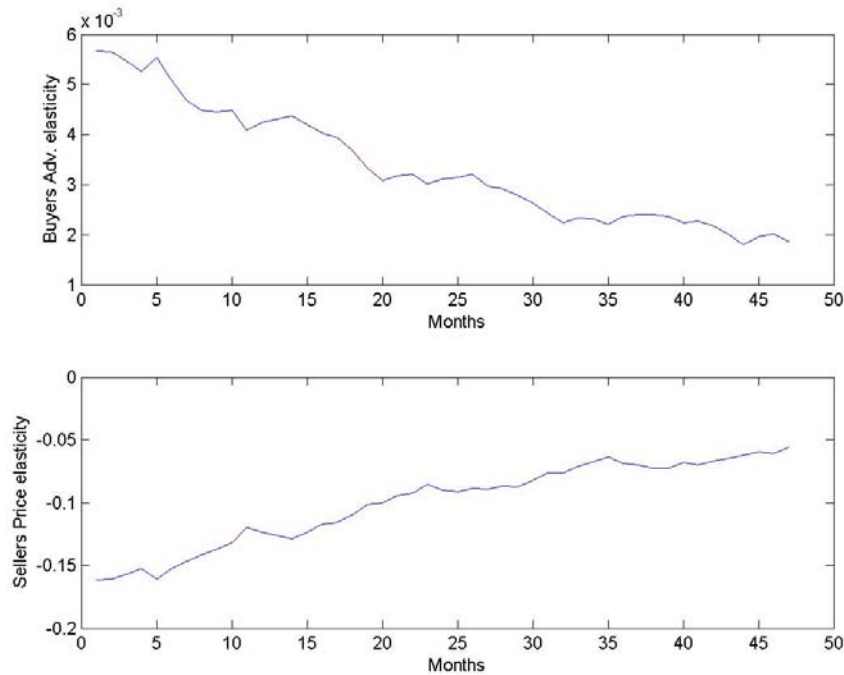


Figure 5: Advertising and Price Elasticity Over Time

the early periods and rapidly goes down over time. In a meta analysis, Lodish et al. (1995) found that the ad elasticities range from 0.05 for established products to 0.26 for new products. The low advertising elasticities in our application again suggest that in our application network effects are more pronounced and effective in acquiring customers than traditional advertising. With the changing landscape of advertising, the growth of new media and the explosion of viral marketing, the advertising elasticities in our study may be more reflective of the current status and effectiveness of advertising.

## 5 Conclusions

Customer profitability is a central consideration to many firms. Though research in this area is burgeoning, little work to date addresses the issue of customer valuation in the context of direct and indirect network effects in a dynamic setting. This limitation is palpable because network effects exist in many contexts including sellers and buyers for auction houses, job seekers and job providers on job sites, real estate listings and buyers on listing services, and so forth. In these contexts,

traditional methods of customer valuation that are based on summing an individual customer's cash flows are misleading, as the interaction of populations is relevant to customer valuation. Even customers that provide no direct source of revenue add value to the firm. For example, a job seeker does not pay for a job, but an increase in the number of job seekers makes a job listing more valuable to an employer, and hence the listing agency. Therefore, CLV is not solely a function of the cash flows generated by a customer, but also the effect this customer has on attracting other customers. In addition extant methods of valuing customers do not consider the role that customer acquisitions play in affecting marketing expenditures. This can also affect CLV, especially in the presence of network effects, as firms can reduce their marketing expenses as network grows and the diminution in costs further amplifies customer value.

We address these problems by offering a new approach to assess customer value that is predicated upon the incremental profits to the firm over an infinite horizon as a result of adding another buyer or seller to the firm's portfolio of customers. To do this, we begin by developing a diffusion-type model of the growth in the firm's buyer and seller customer populations over time. In this model, the size of each customer population is a function of a firm's marketing, direct network effects, and indirect network effects. We then use this system of growth equations, coupled with margin data and marketing expenditures, to determine firm profits over time. The resulting first order conditions for marketing spend over time imply an optimal price and advertising path. This leads to four equations which we estimate via GMM: the seller growth model, the buyer growth model, the advertising spend over time, and the pricing over time. Using the parameter estimates from this model, it is possible to compute the Lagrange multipliers arising from the constraints that the seller and buyer growth models place on a firms' infinite horizon profits. The Lagrange multiplier for buyers (providers) is a natural measure of the marginal impact of an additional buyer or seller on a firm's net discounted sum of future profits, and hence provides the CLV of each type of customer.

The presence of network effects led to a number of questions we set forth at the start of this paper. We reprise these questions here and our findings pertaining to each in the considered application (and note that the method developed in this paper could afford answers to these questions in other applications as well):

- Are direct and indirect network effects present in practice? We find strong evidence of the presence of network effects, and find the network effects of buyers on seller are nearly six

times the effect of seller on buyers.

- How much should a company spend to acquire new customers in the presence of these network effects? In the considered context, we find that each type of customer provides the firm with roughly \$500-550 over the lifetime of the customer.
- How does the value of a customer change over time? The value of the customers is increasing over time as the network builds, and we expect the network to reach a point wherein an additional customer no longer enhances the effect of the network.
- How do we apportion value between buyers and sellers? Though there exist 4.6 seller for each buyer, we find each to have roughly equal value. This counterintuitive effect is explained in part by the greater indirect network effect associate with buyers and a greater tendency of seller to attrite.
- How should firms' marketing efforts change over time in the presence of network effects? How do customers' sensitivities to these marketing actions change over time? As the network effect become stronger, marketing plays less of a role in attracting buyers and sellers. Hence, elasticities decrease as well as the optimal spend.
- Does the omission of network effects understate the value of customer base and hence the value of a firm? We find that traditional CLV methods capture only 2% of the firms' value whereas the approach indicated herein affords a good approximation of firm value.

Given the nascent state of customer valuation research in the context of network effects, there remain many potential directions in which the foregoing analysis can be extended. First, as richer data become available in more contexts, our analysis can be further generalized. One extension of particular interest is to forecast future expenses, demand and customer value. This would require the resolution of a stochastic dynamic program to solve the infinite horizon problem. In sum, we hope this initial foray into customer valuation in the context of network effects leads to additional insights that will be useful to firms who are concerned with managing their customer portfolio in a networked economy.

## References

- [1] ASCHER, U. M., R. M. M. MATTHEIJ, R. D. RUSSELL (1995), *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, SIAM's Classics in Applied Mathematics, 13, SIAM, Philadelphia.
- [2] BERTSEKAS D. P. (1995), *Dynamic Programming and Optimal Control*, Vols.1 and 2, Athena Scientific, Belmont, MA.
- [3] BIJMOLT, TAMMO H.A., HARALD J. VAN HEERDE, AND RIK G.M. PIETERS (2005), "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42 (May), 141-156.
- [4] CHAMBERLAIN, GARY S. (1987), "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics*, 34, 305-334.
- [5] CHIANG, A. C. (1992): *Elements of Dynamic Optimization*. McGraw Hill, New York.
- [6] CHINTAGUNTA, PRADEEP AND NAUFEL VILCASSIM (1992), "An Empirical Investigation of Advertising Strategies in a Dynamic Duopoly," *Management Science*, 38 (9), 1230-1244.
- [7] CHINTAGUNTA, PRADEEP AND VITHALA RAO (1996), "Pricing Strategies in a Dynamic Duopoly: A Differential Game Model," *Management Science*, 42 (11), 1501-1514.
- [8] DEMANGE, GABRIELLE AND DAVID GALE (1985), "The Strategy Structure of Two-Sided Matching Markets," *Econometrica*, 53, 4 (July), 873-888.
- [9] DOCKNER, ENGELBERT AND STEFFEN JORGENSEN (1988), "Optimal Pricing Strategies for New Products in Dynamic Oligopolies," *Marketing Science*, 7 (4), 315-334.
- [10] FEICHTINGER, GUSTAV, RICHARD HARTL AND SURESH SETHI (1994), "Dynamic Optimal Control Models in Advertising: Recent Developments," *Management Science*, 40 (2), 195-226.
- [11] GUPTA, SACHIN, DIPAK JAIN AND MOHANBIR SAWHNEY (1999), "Modeling the Evolution of Markets with Indirect Network Externalities: An Application to Digital Television," *Marketing Science*, 18 (3), 396-416.
- [12] GUPTA, SUNIL AND DONALD R. LEHMANN (2003), "Customers as Assets," *Journal of Interactive Marketing*, 17, 1(Winter), 9-24.
- [13] Gupta, Sunil and Donald R. Lehmann (2005), *Managing Customers as Investments: The Strategic Value of Customers in the Long Run*, Wharton School Publishing, Upper Saddle River, New Jersey.
- [14] GUPTA, SUNIL, DONALD R. LEHMANN AND JENNIFER AMES STUART (2004), "Valuing Customers," *Journal of Marketing Research*, 41, 1, 7-18.
- [15] GUPTA, SUNIL AND VALARIE ZEITHAML (2006), "Customer Metrics and Their Impact on Financial Performance," *Marketing Science*, forthcoming.

- [16] HANSEN, LARS PETER (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 4. (July), 1029-1054.
- [17] HOGAN, JOHN, KATHERINE LEMON AND BARAK LIBAI (2003), "What is the True Value of a Lost Customer?" *Journal of Service Research*, 5, 3, February, 196-208.
- [18] HORSKY, DAN AND LEONARD S. SIMON (1983), "Advertising and the Diffusion of New Products," *Marketing Science*, 2, 1(Winter), 1-17.
- [19] HORSKY, DAN AND KARL MATE (1988), "Dynamic Advertising Strategies of Competing Durable Good Producers," *Marketing Science*, 7, 4(Fall), 356-367.
- [20] KADIYALI, VRINDA, PRADEEP CHINTAGUNTA, AND NAFUEL VILCASSIM (2000), "Manufacturer-Retailer Channel Interactions and Implications for Channel Power: An Empirical Investigation of Pricing in a Local Market," *Marketing Science*, 19, 2 (Spring), 127-148.
- [21] KALISH, SHLOMO (1985), "A New Product Adoption Model with Price, Advertising and Uncertainty," *Management Science*, 31, 12, 1569-1585.
- [22] KAMAKURA, WAGNER A, CARL F. MELA, ASIM ANSARI, ANAND BODAPATI, PETE FADER, RAGHURAM IYENGAR, PRASAD NAIK SCOTT NESLIN, BAOHONG SUN, PETER VERHOEF, MICHEL WEDEL, AND RON WILCOX (2005), "Choice Models and Customer Relationship Management," *Marketing Letters*, 16, 3/4, 279-291.
- [23] KATZ, M. AND C. SHAPIRO (1985), "Network Externalities, Competition, and Compatibility", *American Economic Review*, 75, 3, 424-440.
- [24] KATZ, M. AND C. SHAPIRO (1986), "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, 94, 22-41.
- [25] KUMAR, V. AND TRICHY KRISHNAN (2002), "Multinational Diffusion Models: An Alternative Framework," *Marketing Science*, 21, 3(Summer), 318-330.
- [26] LIBAI, BARAK, EITAN MULLER AND RENANA PERES (2006), "The Effect of Customer Attrition on Service Growth and Equity," working paper, Tel-Aviv University.
- [27] LIONS, J. L. (1971), *Optimal Control of Systems Governed by Partial Differential Equations*, Springer Verlag, New York.
- [28] LODISH, LEONARD M., MAGID ABRAHAM, STUART KALMENSON, JEANNE LIVELSBERGER, BETH LUBETKIN, BRUCE RICHARDSON, AND MARY ELLEN STEVENS (1995), "How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Cable T.V. Advertising Experiments," *Journal of Marketing Research*, 32, 2(May), 125-139.
- [29] MARKETING SCIENCE INSTITUTE (2006), "2006-2008 Research Priorities," Research Report, Boston, Massachusetts.
- [30] NEIL, G., M. KENDE AND R. ROB (2000), "The Dynamics of Technological Adoption in Hardware/Software Systems: The Case of Compact Disc Players," *The RAND Journal of Economics*, 31 (1), 43-61.

- [31] NEWHEY, WHITNEY K. AND DANIEL. MCFADDEN (1994), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, vol. iv, ed. by R. F. Engle and D. L. McFadden, pp. 2111-2245, Amsterdam: Elsevier.
- [32] NEWHEY, WHITNEY K. AND KENNETH D. WEST (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 3(May), 703-708.
- [33] PONTRYAGIN, L. S.; V. BOLTYANSKII; R. V. GAMKRELIDZE AND E. F. MISHENKO (1964), *The Mathematical Theory of Optimal Processes*, McMillan, New York.
- [34] REINARTZ, WERNER AND V. KUMAR (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67, 1(January), 77-99.
- [35] REINARTZ, WERNER, JACQUELYN THOMAS AND V. KUMAR (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69, 1 (January), 63-79.
- [36] RUST, ROLAND, KATHERINE LEMON AND VALARIE ZEITHAML (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68, 1(January), 109-126.
- [37] SEIERSTAD, A. & SYDSAETER, K. (1987), *Optimal Control Theory with Economic Applications*, Advanced Textbooks in Economics, (C. J. Bliss and M. D. Intriligator Editors ), Vol. 24. North-Holland, New York.
- [38] SIMON, HERMANN AND KARL-HEINZ SEBASTIAN (1987), "Diffusion and Advertising: The German Telephone Campaign," *Management Science*, 33, 4, 451-466.
- [39] TELLIS, GERARD. J. (1988), "The Price Elasticity of Selective Demand: A Meta-analysis of Econometric Models of Sales," *Journal of Marketing Research*, 25, 331-341.
- [40] THOMPSON, GERALD AND JINN-TSAIR TENG (1984), "Optimal Pricing and Advertising Policies for New Product Oligopoly Models," *Marketing Science*, 3, 2(Spring), 148-168.
- [41] VENKATESAN, RAJKUMAR AND V. KUMAR (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68, 4(November), 106-125.
- [42] YAO, SONG AND CARL F. MELA (2006), "Online Auction Demand," Working Paper, Duke University.

## Appendix

### A1 Optimal Control Solution

Equation (6) gives the following Hamiltonian function

$$H(t, N^B, N^S, p, A, \lambda^B, \lambda^S) = G(t, N^B, N^S, p, A) + \lambda^B g_B(A, N^B, N^S) + \lambda^S g_S(p, N^B, N^S), \quad (\text{A-1})$$

where  $\lambda^B, \lambda^S$  are auxiliary or costate variables. The solution to this optimal control problem satisfies the following conditions,

$$\begin{aligned} 0 &= \frac{\partial}{\partial p} H = \frac{\partial}{\partial p} G(t, N_t^B, N_t^S, p_t, A_t) + \lambda_t^S \frac{\partial}{\partial p} g_S(p_t, N_t^B, N_t^S), \\ 0 &= \frac{\partial}{\partial A} H = \frac{\partial}{\partial A} G(t, N_t^B, N_t^S, p_t, A_t) + \lambda_t^B \frac{\partial}{\partial A} g_B(A_t, N_t^B, N_t^S), \\ \dot{N}_t^B &= \frac{\partial}{\partial \lambda^B} H = g_B(A_t, N_t^B, N_t^S), \\ \dot{N}_t^S &= \frac{\partial}{\partial \lambda^S} H = g_S(p_t, N_t^B, N_t^S), \\ \dot{\lambda}_t^B &= -\frac{\partial}{\partial N^B} H = -\frac{\partial}{\partial N^B} G(t, N_t^B, N_t^S, p_t, A_t) - \lambda_t^S \frac{\partial}{\partial N^B} g_S(p_t, N_t^B, N_t^S) \\ &\quad - \lambda_t^B \frac{\partial}{\partial N^B} g_B(A_t, N_t^B, N_t^S), \\ \dot{\lambda}_t^S &= -\frac{\partial}{\partial N^S} H = -\frac{\partial}{\partial N^S} G(t, N_t^B, N_t^S, p_t, A_t) - \lambda_t^S \frac{\partial}{\partial N^S} g_S(p_t, N_t^B, N_t^S) \\ &\quad - \lambda_t^B \frac{\partial}{\partial N^S} g_B(A_t, N_t^B, N_t^S), \end{aligned}$$

together with the initial conditions  $N_0^B = N_0^S = 0$ , and the transversality conditions

$$\lim_{T \rightarrow \infty} \lambda_T^B N_T^B = 0, \quad \lim_{T \rightarrow \infty} \lambda_T^S N_T^S = 0,$$

(see e.g., Pontryagin et al. 1964, Lions 1971, Seierstad and Sydsaeter 1987, Chiang 1992 and Bertsekas 1995).

## A2 The Euler Equation

To solve the dynamic problem indicated by (9) we begin by defining the optimal value function: for an arbitrary initial point  $(N_0^S, N_0^B, e_0, \varepsilon_0)$  this function is given by:

$$V(N_0^S, N_0^B, e_0, \varepsilon_0) = \max_{\{p_t, A_t\}} E_0 \left[ \sum_{t=0}^{\infty} \delta^t G(N_t^B, N_t^S, p_t, A_t) \right]. \quad (\text{A-2})$$

Then, the Jacobi-Bellman condition states that the solution satisfies the following recursion for each period  $t \geq 0$ ,

$$V(N_t^B, N_t^S, e_t, \varepsilon_t) = \max_{p_t, A_t} \left\{ G(N_t^B, N_t^S, p_t, A_t) + \delta E_t [V(N_{t+1}^S, N_{t+1}^B, e_{t+1}, \varepsilon_{t+1})] \right\}. \quad (\text{A-3})$$

Incorporating the state equations into (A-3) leads to

$$\begin{aligned} & V(N_t^B, N_t^S, e_t, \varepsilon_t) \\ = & \max_{p_t, A_t} \left\{ G(N_t^B, N_t^S, p_t, A_t) + \delta E_t [V(g^S(A_t, N_t^B, N_t^S, e_t), g^B(p_t, N_t^B, N_t^S, \varepsilon_t), e_{t+1}, \varepsilon_{t+1}))] \right\}. \end{aligned}$$

Therefore, the first order conditions associated to the right hand side optimization problem are satisfied, which are

$$\frac{\partial}{\partial p_t} G(N_t^B, N_t^S, p_t, A_t) + \delta E_t \left[ \frac{\partial}{\partial N_{t+1}^S} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}) \right] \frac{\partial}{\partial p_t} g^S(p_t, N_t^B, N_t^S, \varepsilon_t) \quad =(\text{A04})$$

$$\frac{\partial}{\partial A_t} G(N_t^B, N_t^S, p_t, A_t) + \delta E_t \left[ \frac{\partial}{\partial N_{t+1}^B} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}) \right] \frac{\partial}{\partial A_t} g^B(A_t, N_t^B, N_t^S, e_t) \quad =(\text{A05})$$

Using the envelope theorem, it can be proved that

$$\begin{aligned} & \left( \frac{\partial}{\partial N_t^B} V(N_t^B, N_t^S, e_t, \varepsilon_t) \right) = \left( \frac{\partial}{\partial N_t^B} G(N_t^B, N_t^S, p_t, A_t) \right) + \delta \\ & \times \left( \begin{array}{cc} \frac{\partial}{\partial N_t^B} g^B(p_t, N_t^B, N_t^S, \varepsilon_t) & \frac{\partial}{\partial N_t^B} g^S(A_t, N_t^B, N_t^S, e_t) \\ \frac{\partial}{\partial N_t^S} g^B(p_t, N_t^B, N_t^S, \varepsilon_t) & \frac{\partial}{\partial N_t^S} g^S(A_t, N_t^B, N_t^S, e_t) \end{array} \right) E_t \left[ \left( \begin{array}{c} \frac{\partial}{\partial N_{t+1}^B} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}) \\ \frac{\partial}{\partial N_{t+1}^S} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}) \end{array} \right) \right] \quad (\text{A-6}) \end{aligned}$$

Substituting  $E_t \left[ \frac{\partial}{\partial N_{t+1}^B} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}) \right]$  and  $E_t \left[ \frac{\partial}{\partial N_{t+1}^S} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}) \right]$  from (A-4) and (A-5) into (A-6),

$$\begin{aligned} & \left( \frac{\partial}{\partial N_t^B} V(N_t^B, N_t^S, e_t, \varepsilon_t) \right) = \left( \frac{\partial}{\partial N_t^B} G(N_t^B, N_t^S, p_t, A_t) \right) \\ & + \left( \frac{\partial}{\partial N_t^S} V(N_t^B, N_t^S, e_t, \varepsilon_t) \right) = \left( \frac{\partial}{\partial N_t^S} G(N_t^B, N_t^S, p_t, A_t) \right) \\ & + \left( \frac{\partial}{\partial N_t^B} g^B(p_t, N_t^B, N_t^S, \varepsilon_t) \quad \frac{\partial}{\partial N_t^B} g^S(A_t, N_t^B, N_t^S, e_t) \right) \left( \frac{\frac{\partial G(N_t^B, N_t^S, p_t, A_t)/\partial p_t}{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)/\partial p_t}}{\frac{\partial G(N_t^B, N_t^S, p_t, A_t)/\partial A_t}{\partial g^B(A_t, N_t^B, N_t^S, e_t)/\partial A_t}} \right) \end{aligned} \quad (\text{A-7})$$

updating the resulting condition and combining it with the first order conditions (A-4), (A-5) yields the system,

$$\begin{aligned} & \left( \frac{\frac{\partial G(N_t^B, N_t^S, p_t, A_t)/\partial p_t}{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)/\partial p_t}}{\frac{\partial G(N_t^B, N_t^S, p_t, A_t)/\partial A_t}{\partial g^B(A_t, N_t^B, N_t^S, e_t)/\partial A_t}} \right) = \delta E_t \left[ \left( \frac{\partial}{\partial N_{t+1}^B} G(N_{t+1}^B, N_{t+1}^S, p_{t+1}, A_{t+1}) \right) \right. \\ & \left. + \left( \frac{\partial}{\partial N_{t+1}^S} G(N_{t+1}^B, N_{t+1}^S, p_{t+1}, A_{t+1}) \right) \right. \\ & \left. + \left( \frac{\partial}{\partial N_{t+1}^B} g^B(p_{t+1}, N_{t+1}^B, N_{t+1}^S, \varepsilon_{t+1}) \quad \frac{\partial}{\partial N_{t+1}^B} g^S(A_{t+1}, N_{t+1}^B, N_{t+1}^S, e_{t+1}) \right) \right. \\ & \left. + \left( \frac{\partial}{\partial N_{t+1}^S} g^B(p_{t+1}, N_{t+1}^B, N_{t+1}^S, \varepsilon_{t+1}) \quad \frac{\partial}{\partial N_{t+1}^S} g^S(A_{t+1}, N_{t+1}^B, N_{t+1}^S, e_{t+1}) \right) \right] \\ & \cdot \left( \frac{\frac{\partial G(N_{t+1}^B, N_{t+1}^S, p_{t+1}, A_{t+1})/\partial p_{t+1}}{\partial g^S(p_{t+1}, N_{t+1}^B, N_{t+1}^S, \varepsilon_{t+1})/\partial p_{t+1}}}{\frac{\partial G(N_{t+1}^B, N_{t+1}^S, p_{t+1}, A_{t+1})/\partial A_{t+1}}{\partial g^B(A_{t+1}, N_{t+1}^B, N_{t+1}^S, e_{t+1})/\partial A_{t+1}}} \right) \end{aligned}$$

This expression is the Euler equations system. Note that the left hand side can be introduced in the conditional expectation with a sign change. Computing the partial derivatives, we obtain the expression (14).

$$E_t \left[ \left( \frac{\frac{N_t^S S}{\theta(M^S - N_t^S)/p_t}}{\frac{1}{h(M^B - N_t^B)/A_t}} \right) + (1+i)^{-1} \left( \left( \begin{array}{c} 0 \\ S p_{t+1} \end{array} \right) + \left( \begin{array}{cc} D_{t+1}^{11} & D_{t+1}^{12} \\ D_{t+1}^{21} & D_{t+1}^{22} \end{array} \right) \left( \frac{\frac{N_{t+1}^S S}{(-\theta(M^S - N_{t+1}^S)/p_{t+1})}}{\frac{-1}{h(M^B - N_{t+1}^B)/A_{t+1}}} \right) \right) \right] = 0,$$

where

$$\begin{aligned}
D_{t+1}^{11} &= 1 + \left( \frac{b}{M^B} \right) (M^B - N_{t+1}^B) - \left( x + h \ln A_{t+1} + b \frac{N_{t+1}^B}{M^B} + g \frac{N_{t+1}^S}{M^S} \right) \\
&= 1 + b - x - h \ln A_{t+1} - 2b \frac{N_{t+1}^B}{M^B} - g \frac{N_{t+1}^S}{M^S} \\
D_{t+1}^{12} &= \frac{\gamma}{M^B} (M^S - N_{t+1}^S) \\
D_{t+1}^{21} &= \frac{g}{M^S} (M^B - N_{t+1}^B) \\
D_{t+1}^{22} &= 1 + \frac{\beta}{M^S} (M^S - N_{t+1}^S) - \left( \phi - \theta \ln P_{t+1} + \beta \frac{N_{t+1}^S}{M^S} + \gamma \frac{N_{t+1}^B}{M^B} \right) - \rho \\
&= 1 + \beta - \phi + \theta \ln p_{t+1} - 2\beta \frac{N_{t+1}^S}{M^S} - \gamma \frac{N_{t+1}^B}{M^B}
\end{aligned}$$

### A3 GMM Estimation

Equations (16), (17) and (18) yield a set of moment conditions. To simplify the notation, we express these moment equations as  $E[m(W_t, \Omega)] = 0$ , where  $\Omega$  denotes the set of all parameters. If this system of equations is just identified then one can use the method of moments. When the system is overidentified as a result of adding more instrument conditions, then generalized method of moments or GMM is used. In this approach we estimate the parameter vector by minimizing the sum of squares of the differences between the population moments and the sample moments, using the variance of the moments as a metric. Specifically, GMM estimates  $\Omega$  by minimizing

$$Q_W(\Omega) = \left( \frac{1}{T} \sum_{t=1}^T m(W_t, \Omega) \right)' W^{-1} \left( \frac{1}{T} \sum_{t=1}^T m(W_t, \Omega) \right), \quad (\text{A-8})$$

where  $W$  is a positive definite weight matrix. While the researcher cannot make the moment conditions exactly equal to zero, s/he can choose parameters such that (A-8) is close to zero. The choice of  $W$  impacts what it means to be close and the variance of the estimate  $\hat{\Omega}$  depends on this chosen matrix. It is always possible to choose  $W = I$ . However, this will, in general, lead to inefficient estimates. The optimal  $W$  weights the moment conditions such that those conditions with a high degree of variance get weighted less, thus affecting the minimization routine less. The

estimate  $\widehat{\Omega}$  with minimum variance is obtained for the limit variance covariance matrix,

$$\begin{aligned} W_{\Omega_0} &= \lim_{T \rightarrow \infty} T \cdot E \left[ \left( \frac{1}{T} \sum_{t=1}^T m(W_t, \Omega_0) \right) \left( \frac{1}{T} \sum_{t=1}^T m(W_t, \Omega_0) \right)' \right] \\ &= \sum_{t=-\infty}^{\infty} E [m(W_0, \Omega_0) m(W_t, \Omega_0)'], \end{aligned} \quad (\text{A-9})$$

i.e.,  $W_{\Omega_0}$  is  $2\pi$  times the spectral density matrix for  $\{m(W_t, \theta_0)\}$  at frequency zero. This matrix depends on  $\Omega_0$ , and therefore the optimal GMM is unfeasible.

A feasible estimation is typically achieved using a two-step process. In the initial step, indicated (0), a positive definite matrix  $W^{(0)}$  matrix is chosen (e.g.  $W^{(0)} = I$ , the identity matrix), leading to a consistent initial set of parameter estimates  $\widehat{\Omega}^{(0)}$ . In the second step, we take the estimate  $\widehat{\Omega}^{(0)}$  and estimate the variance-covariance estimator using the Bartlett spectral density estimator of  $W_{\Omega_0}$ ,

$$\begin{aligned} \widehat{W}_{\Omega_0} &= C_0 + \sum_{l=1}^L \left(1 - \frac{l}{L}\right) (C_l + C_l'), \\ C_l &= T^{-1} \sum_{t=1}^{T-l} m(W_t, \widehat{\Omega}^{(0)}) m(W_{t+l}, \widehat{\Omega}^{(0)})' = C_{-l}', \end{aligned} \quad (\text{A-10})$$

as suggested by Newey and West (1987). A moderate number of lags  $L < T-1$  is usually considered, thus  $L$  increases slowly with the sample size, so that  $T/L \rightarrow \infty$ . This new weight matrix is then used to solve the problem (A-8), and the resulting estimation  $\widehat{\Omega}^{(1)}$  is an asymptotically efficient estimator of the true parameters. One can iterate over the weighting matrix  $M$  times (it does not affect the asymptotic distribution, but the accuracy for small samples usually increases), but the process converges sufficiently fast (e.g., in our context  $M < 5$ ).

It has also been shown that, for the optimal weight,  $\sqrt{T} (\widehat{\Omega} - \Omega)$  is asymptotically distributed  $N(0, (\Gamma' W_{\Omega}^{-1} \Gamma)^{-1})$  with  $\Gamma = E \left[ \frac{\partial}{\partial \Omega'} m(W_t, \Omega_0) \right]$ , which can be estimated by  $\Gamma = T^{-1} \sum_{t=1}^T \frac{\partial}{\partial \Omega'} m(W_t, \Omega_0)$ . In addition,  $T \cdot Q_{\widehat{W}}(\widehat{\Omega})$  is asymptotically distributed  $\chi_{r-k}^2$ , where  $r$  is the number of equations and  $k$  the number of parameters, and this can be used to test the overidentifying restrictions (see,

e.g., Hansen 1982, Chamberlain 1987 and Newey and McFadden 1994) to test the appropriateness of the instruments.

The estimation is computed in two steps. In the first step, a Newey-West matrix was evaluated at the starting values with  $L = 12$ , and in the second iteration we use Newey-West on the consistent first step estimator.

#### A4 Lagrange Multipliers

We begin by characterising the solution to company problem via the Lagrange functional (11), for each  $t \geq 0$ ,

$$\begin{aligned} \frac{\partial E_0 \left[ \sum_{s=t}^{\infty} \delta^s G(N_s^B, N_s^S, p_s, A_s) \right]}{\partial p_t} + \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial p_t} &= 0, \\ \frac{\partial E_0 \left[ \sum_{s=t}^{\infty} \delta^s G(N_s^B, N_s^S, p_s, A_s) \right]}{\partial A_t} + \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial A_t} &= 0, \\ \frac{\partial E_0 \left[ \sum_{s=t}^{\infty} \delta^s G(N_s^B, N_s^S, p_s, A_s) \right]}{\partial N_t^B} + \lambda_t^B - \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial N_t^B} \\ &\quad - \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial N_t^B} = 0, \\ \frac{\partial E_0 \left[ \sum_{s=t}^{\infty} \delta^s G(N_s^B, N_s^S, p_s, A_s) \right]}{\partial N_t^S} - \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial N_t^S} + \lambda_t^S \\ &\quad - \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial N_t^S} = 0. \end{aligned}$$

In the optimum  $E_0 \left[ \sum_{s=t}^{\infty} \delta^{(s-t)} G(N_s^B, N_s^S, p_s, A_s) \right] = V(N_t^S, N_t^B, e_t, \varepsilon_t)$ , therefore

$$E_0 \left[ \sum_{s=t}^{\infty} \delta^s G(N_s^B, N_s^S, p_s, A_s) \right] = \delta^t V(N_t^S, N_t^B, e_t, \varepsilon_t),$$

and the first order conditions for each time  $t$  are,

$$\delta^t \frac{\partial}{\partial p_t} V(N_t^B, N_t^S, e_t, \varepsilon_t) + \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial p_t} = 0,$$

$$\delta^t \frac{\partial}{\partial A_t} V(N_t^B, N_t^S, e_t, \varepsilon_t) + \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial A_t} = 0,$$

$$\delta^t \frac{\partial}{\partial N_t^B} V(N_t^B, N_t^S, e_t, \varepsilon_t) + \lambda_t^B - \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial N_t^B} - \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial N_t^B} = 0,$$

$$\delta^t \frac{\partial}{\partial N_t^S} V(N_t^B, N_t^S, e_t, \varepsilon_t) - \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial N_t^S} + \lambda_t^S - \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial N_t^S} = 0.$$

Therefore,

$$\lambda_{t+1}^B = -\delta^t \frac{\frac{\partial}{\partial A_t} V(N_t^B, N_t^S, e_t, \varepsilon_t)}{\frac{\partial}{\partial A_t} g^B(A_t, N_t^B, N_t^S, e_t)},$$

$$\lambda_{t+1}^S = -\delta^t \frac{\frac{\partial}{\partial p_t} V(N_t^B, N_t^S, e_t, \varepsilon_t)}{\frac{\partial}{\partial p_t} g^S(p_t, N_t^B, N_t^S, \varepsilon_t)},$$

also,

$$\lambda_t^B = \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial N_t^B} + \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial N_t^B} - \delta^t \frac{\partial V(N_t^B, N_t^S, e_t, \varepsilon_t)}{\partial N_t^B},$$

$$\lambda_t^S = \lambda_{t+1}^B \frac{\partial g^B(A_t, N_t^B, N_t^S, e_t)}{\partial N_t^S} + \lambda_{t+1}^S \frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}{\partial N_t^S} - \delta^t \frac{\partial V(N_t^B, N_t^S, e_t, \varepsilon_t)}{\partial N_t^S}.$$

Note that last two equations can be used to compute the multipliers backwards, recursively from the terminal transversality conditions. But in this paper we will use the two non-recursive

expressions for the Lagrange Multipliers,

$$\begin{aligned}
\lambda_{t+1}^B &= -\delta^t \frac{\frac{\partial}{\partial A_t} V(N_t^B, N_t^S, e_t, \varepsilon_t)}{\frac{\partial}{\partial A_t} g^B(A_t, N_t^B, N_t^S, e_t)} = -\delta^t \frac{\frac{\partial}{\partial A_t} V(g^B(A_t, N_t^B, N_t^S, e_t), g^S(p_t, N_t^B, N_t^S, \varepsilon_t), e_{t+1}, \varepsilon_{t+1})}{\frac{\partial}{\partial A_t} g^B(A_t, N_t^B, N_t^S, e_t)}} \\
&= -\delta^t \frac{\partial}{\partial N_{t+1}^B} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}). \\
\lambda_{t+1}^S &= -\delta^t \frac{\frac{\partial}{\partial p_t} V(N_t^B, N_t^S, e_t, \varepsilon_t)}{\frac{\partial}{\partial p_t} g^S(p_t, N_t^B, N_t^S, \varepsilon_t)} = -\delta^t \frac{\frac{\partial}{\partial p_t} V(g^B(A_t, N_t^B, N_t^S, e_t), g^S(p_t, N_t^B, N_t^S, \varepsilon_t), e_{t+1}, \varepsilon_{t+1})}{\frac{\partial}{\partial p_t} g^S(p_t, N_t^B, N_t^S, \varepsilon_t)}} \\
&= -\delta^t \frac{\partial}{\partial N_{t+1}^S} V(N_{t+1}^B, N_{t+1}^S, e_{t+1}, \varepsilon_{t+1}),
\end{aligned}$$

Using (A-7), we obtain that,

$$\begin{aligned}
\begin{pmatrix} \lambda_{t+1}^B \\ \lambda_{t+1}^S \end{pmatrix} &= -\delta^t \begin{pmatrix} \frac{\partial}{\partial N_t^B} V(N_t^B, N_t^S, e_t, \varepsilon_t) \\ \frac{\partial}{\partial N_t^S} V(N_t^B, N_t^S, e_t, \varepsilon_t) \end{pmatrix} = -\delta^t \begin{pmatrix} \frac{\partial}{\partial N_t^B} G(N_t^B, N_t^S, p_t, A_t) \\ \frac{\partial}{\partial N_t^S} G(N_t^B, N_t^S, p_t, A_t) \end{pmatrix} \\
&= -\delta^t \begin{pmatrix} \frac{\partial}{\partial N_t^B} g^B(p_t, N_t^B, N_t^S, \varepsilon_t) & \frac{\partial}{\partial N_t^B} g^S(A_t, N_t^B, N_t^S, e_t) \\ \frac{\partial}{\partial N_t^S} g^B(p_t, N_t^B, N_t^S, \varepsilon_t) & \frac{\partial}{\partial N_t^S} g^S(A_t, N_t^B, N_t^S, e_t) \end{pmatrix} \begin{pmatrix} \frac{\partial G(N_t^B, N_t^S, p_t, A_t) / \partial p_t}{\frac{\partial g^S(p_t, N_t^B, N_t^S, \varepsilon_t) / \partial p_t}} \\ \frac{\partial G(N_t^B, N_t^S, p_t, A_t) / \partial A_t}{\frac{\partial g^B(A_t, N_t^B, N_t^S, e_t) / \partial A_t}} \end{pmatrix}.
\end{aligned}$$

Computing the partial derivatives, we obtain the expression (19),

$$\begin{pmatrix} \lambda_{t+1}^B \\ \lambda_{t+1}^S \end{pmatrix} = -(1+i)^{-t} \begin{pmatrix} 0 \\ Sp_t \end{pmatrix} - (1+i)^{-t} \begin{pmatrix} D_t^{11} & D_t^{12} \\ D_t^{21} & D_t^{22} \end{pmatrix} \begin{pmatrix} N_t^S S \\ \frac{-1}{h(M^B - N_t^B)/A_t} \end{pmatrix}. \quad (\text{A-11})$$