8-2013

An SIR Approach to Modeling Business Interactions in a Marketplace

Jessica R. Beiter

An SIR Approach to Modeling Business Interactions in a Marketplace

by

Jessica R. Beiter

Submitted to the School of Mathematical Sciences, College of Science
in partial fulfillment of the requirements for the degree of
Masters of Science in Applied and Computational Mathematics
at the

Rochester Institute of Technology

August 2013

Accepted by: David S. Ross, School of Mathematical Sciences
Accepted by: Darren A. Narayan, School of Mathematical Sciences
Accepted by: Matthew J. Hoffman, School of Mathematical Sciences

Abstract

As interest in online coupons from distributors like Groupon and Living Social have grown, eBay has proposed and built a marketplace. This marketplace is called eBay Local. It is a place where local merchants can post coupons that distributors can bid on the right to publish. Based on some initial data, we have built a model to fit and predict the growth of this marketplace. The influence of salesmen and organic growth convert potential merchants into active members of the marketplace posting their goods and services. We have modeled the recruitment and retention of businesses within the marketplace, based on interactions with businesses and monetary incentives. Our model has a structure similar to epidemiological models. Parameters are estimated based on initial data sets provided by eBay and numerical results are obtained using a fourth-order Runge-Kutta method coded in MATLAB specifically for this thesis. By adjusting the model and the parameters within reasonable values, the system displayed an accurate representation of the marketplace. Using the model we have found realistic conditions under which the system is optimized, creating a stable population of active businesses inexpensively.

Contents

List of Figures

List of Tables

1 The Problem

With the increasing reach and accessibility of the internet, many forms of communication have migrated from local papers to the web. An example of this migration is in online discounts and coupons that can be found for nearly every business. Groupon, one of the largest sources for these coupons, and other coupon distributors like Living Social and Daily Deals provide virtual meeting points for merchants and consumers. Here consumers peruse the discounted goods or services of the merchants. From the consumer's perspective, these deals found on bargain sites allow for a cheap experience to try out a new business or an incentive to return to an already experienced one.

For the merchants, these bargain sites can be viewed as a large source of potential customers. Often these sites have long lists of self-enrolling contacts; customers prepared and looking to make online purchases. By posting a deal on their products or services, merchants can greatly broaden their audience, bringing in new customers and increasing profits. When an online deal is successful, a merchant sees more returning customers [4].

There is a cost for these posted deals however. In order to encourage new customers to visit their places of business, merchants must first offer their wares at discounted rates. These coupons often offer a savings to the consumer at the expense of the merchant. The remaining profits can be further reduced by up to half by the distributors as a publishing fee. This can leave as little as 25% for the merchant to claim on its original product or service. To try to alleviate the cost of publishing the deals, eBay has proposed an online marketplace. This marketplace would be an online auction house where merchants could post a deal they would like published and the coupon distributors would bid on the right to sell the deal to the public. Currently, online coupons are only profitable for merchants approximately 67% of the time [4]. This marketplace would save the merchants money and make these deals more popular and accessible to consumers. The name for this proposed marketplace where local merchants sell discounted goods and services is eBay Local.

In order to build this marketplace of merchants and publishers, salesmen must first contact local merchants and encourage them to participate. Resources are limited however; so eBay would like to recruit businesses, as much as is possible, through word-of-mouth recruitment. We have been tasked with modeling the development of this marketplace and finding the optimal conditions in which it becomes self-sustaining with as little cost to eBay as possible. Model selection was based on the fact that eBay would like to promote organic growth based on interactions between abstaining and participating merchants. An epidemiology model that organizes the population by level of participation in the marketplace was selected.

2 An Introduction to the SIR Model

The basic SIR Model was developed in 1927 by W. O. Kermack and A. G. McKendrick [5] as a way to study the spread and potential reach of infectious disease. In this model, an individual can either be susceptible to, infected with, or recovered from a contagion, denoted by S, I, and R respectively. This results in a division of a fixed population into three separate compartments relating each to their connection to the contagion. The population begins with the majority of the individuals susceptible, a few already infected, and zero recovered individuals. Assuming a homogeneous mixing of the entire population, susceptible and infected individuals come into contact with each other with those susceptible transitioning to an infectious state at a rate of $\beta\left[\frac{1}{individuals \cdot time}\right]$. This means that the susceptible population decreases at a rate of $-\beta SI$ and the infected population increases at a rate of βSI. Individuals remain infected for a period of $\frac{1}{\gamma}\left[time\right]$ so the infected population decreases at a rate of $-\gamma I$ and the recovered population increases at a rate of γI. The SIR Model is represented by the following system where at all times, $S\left(t\right)+I\left(t\right)+R\left(t\right)$ equals the total population:

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta SI, \\
\frac{dI}{dt} &= \beta SI - \gamma I, \\
\frac{dR}{dt} &= \gamma I. \qquad (2.1)
\end{aligned}
$$

Essentially, the model describes how a population transitions from healthy to infected to recovered at rates β and γ. Here's an example of how the state of a population changes according to time with the following parameters:

Values	
S_0	500
I_0	1
R_0	0
β	0.001
γ	0.1

Table 1: Example values for the SIR model

The plot on the following page, Figure 1, is a visualization of how each of the population compartments change over time according to the values in Table 1 substituted into Equation 2.1.

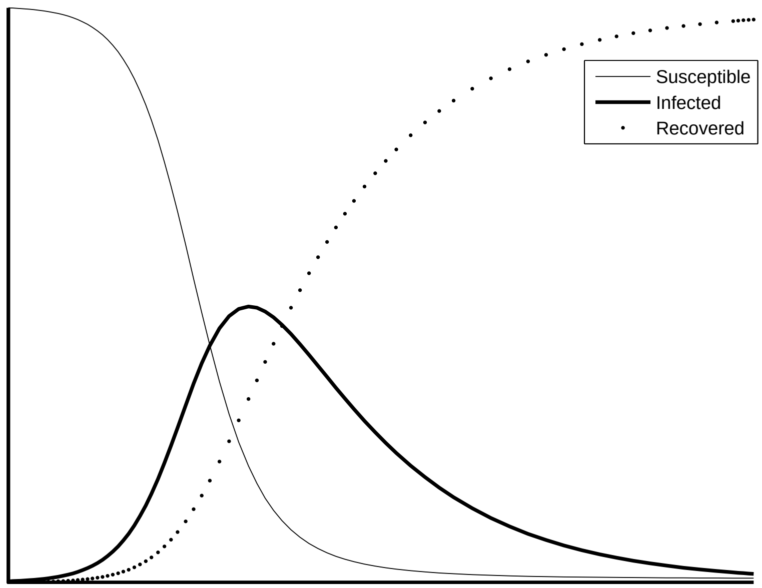

Figure 1: Graph of the SIR model

In our example, all members of the total population become infected and eventually recover causing the infection to slowly die out. Here we see the infection peaked when approximately half the population became infected but since the population is closed, no births or deaths, the equilibrium was reached when the number of infecteds became approximately zero.

If it were the case that the example disease in Figure 1 were lethal and death were the only recovered state, the population would be decimated. Fortunately, diseases rarely eliminate an entire susceptible population. Sometimes the goal is to reach an entire population though, not with disease but with an idea. The SIR epidemiology model has been studied and applied outside of its original biological context to the spread of information. There are some differences between disease and idea spread in the SIR model. One of the most important differences is the willingness to spread. “Ideas, unlike diseases, are usually beneficial and thus people’s behavior tends to maximize effective contacts” [2]. The human body has an immune system set up to fight infection but the mind, when open, collects infectious ideas, a distinct advantage for idea spreading when applying the SIR model.

Rumor propagation [8] and viral marketing [2] are two of the most heavily studied fields outside of biology to which variations of the SIR model are applied. If a marketer wished to spread an idea or product as efficiently as possible they would try to employ the method of viral marketing [7]. Viral marketing is done by initially hiring a few “infected” individuals and then wait for consumers to interact with each other to spread the idea “organically” or virally. This is the goal of eBay Local, to hire salesmen to first bring businesses into the online marketplace then have them communicate with other susceptible merchants to see the marketplace grow and thrive.

3 Building the SIR Model to be Applied to the Marketplace

We have developed a qualitative model to describe the pertinent interactions between members of the marketplace. The model takes into account various stages of growth a business can be at, different ways in which they can be contacted and encouraged to join eBay Local, and cost-per-acquisition estimates.

As with a typical epidemiological model, there are three main categories to organize the businesses. There are those that are potential members of the marketplace; waiting to have a seller or current member reach out to them, active members that have joined through either the work of a salesman or organic growth, and members that were involved then decided to drop out of the program. When being removed from the active compartment, merchants can return to the potential pool or drop out. Instead of the epidemiological notation of S, I, and R to denote our business categories, we use P, A, and D for Potential, Active, and Dropped out.

Businesses have been further broken down into categories based on their ages. This stratification of the businesses is done because of differing failure rates depending on the age of the business. Across all industries, approximately 66% of businesses were still operating two years after opening their doors and 44% after four years [6]. Based on these numbers, it may be of interest for eBay Local to target different businesses based on their age. Newer businesses that have been selling their goods and services for two years or less are denoted in the model with a subscript of 2. The businesses past the uncertainty of the first two years but still newer than five years have a subscript of 1. Finally, established businesses that have been a part of the community for five years or more are have a subscript of 0. Subscripts are applied throughout the model with respect to the age of the potential, active, or dropped out merchants along with the rates at which each category of business are expected to fail in a predicted economy.

Funding for salesmen is tracked in the system. This is to ensure that the marketplace is built in an economical fashion with a set budget, tracked by S in the model. The above definitions have been listed in Table 2.

Variables	
$P_2(t)$	Number of newly established potential businesses at time t [businesses]
$P_1(t)$	Number of moderately established potential businesses at time t [businesses]
$P_0(t)$	Number of well established potential businesses at time t [businesses]
$A_2(t)$	Number of newly established active businesses at time t [businesses]
$A_1(t)$	Number of moderately established active businesses at time t [businesses]
$A_0(t)$	Number of well established active businesses at time t [businesses]
$D_2(t)$	Number of newly established dropped out businesses at time t [businesses]
$D_1(t)$	Number of moderately established dropped out businesses at time t [businesses]
$D_0(t)$	Number of well established dropped out businesses at time t [businesses]
$S(t)$	Total funds spent on eBay Local at time t [dollars]

Table 2: Population breakdown for the SIR model

By considering merchant age as in Table 2, we are implying a non-fixed number of businesses in the marketplace unlike Equation 2.1. In order to maintain a constant total population with the above stratification and businesses failing, we introduce the parameter $\mu\,[businesses]$ to add newly created businesses to the model. Rates of failure for each business age compartment are represented by $\theta\left[\frac{businesses}{businesses \cdot time}\right]$ with appropriate subscripts. Merchants successfully age from newly established to moderately established at a rate of $\beta\left[\frac{businesses}{businesses \cdot time}\right]$ and there to well established at a rate of $\kappa\left[\frac{businesses}{businesses \cdot time}\right]$.

Parameters have also been defined to track the number of salesmen and the estimated cost of a new business joining. The number of salesmen contacting merchants at a given time is denoted by $\delta\,[salesmen]$ and their effectiveness, or rate at which they convince new merchants to join is $\lambda\left[\frac{1}{salesman \cdot time}\right]$. Each business brought into the marketplace by a salesman is an expense for eBay Local, who must pay the salesman for his work, $\alpha\,[dollars]$ is the cost-per-acquisition.

Merchants may also join the marketplace if influenced by other, already participating/active, businesses, without a salesman influence, at a rate $\gamma\left[\frac{1}{businesses \cdot time}\right]$. To officially join the marketplace and be considered an active business, a merchant must post a coupon for distributors to bid on. If a new coupon is not posted within 60 days of the most recent coupon, a business is removed from its appropriate A compartment. When leaving the marketplace, a merchant may either remove himself for good if he had a poor experience and be a dropped out business or return at a later time, in which case he are once again considered a potential business. Merchants become inactive at a rate of $\sigma\left[\frac{businesses}{businesses \cdot time}\right]$ and of those that are no longer active, some return to the potential pool at a rate of $\eta\left[\frac{businesses}{businesses \cdot time}\right]$ and are retrieved. If a merchant refuses to join the marketplace ever, he transitions from the potential compartment directly to the dropped out compartment. Merchants refuse at a rate of $\xi\left[\frac{businesses}{businesses \cdot time}\right]$.

The parameters we have explained above have been collected in Table 3

Parameters	
μ	new business development rate $[\frac{businesses}{time}]$
β	newly to moderately established business maturation rate $[\frac{businesses}{businesses \cdot time}]$
κ	moderately to well established business maturation rate $[\frac{businesses}{businesses \cdot time}]$
θ_2	newly established business failure rate $[\frac{businesses}{businesses \cdot time}]$
θ_1	moderately established business failure rate $[\frac{businesses}{businesses \cdot time}]$
θ_0	well established business failure rate $[\frac{businesses}{businesses \cdot time}]$
δ	salesmen [salesmen]
α	cost per acquisition [dollars]
λ	business acquisition rate $[\frac{1}{salesman \cdot time}]$
γ	organic growth rate $[\frac{1}{businesses \cdot time}]$
σ	drop out rate $[\frac{businesses}{businesses \cdot time}]$
η	retrieval rate $[\frac{businesses}{businesses \cdot time}]$
ξ	refusal rate $[\frac{businesses}{businesses \cdot time}]$

Table 3: Transition parameters for the SIR model

All of the parameters in Table 3 are elements of the non-negative real space. Some of the parameters are more heavily restricted because they represent portions of the existing population. This is why σ, η, xi, β, κ, θ_2, θ_1, and θ_0 are all bounded by $[0, 1]$. The parameter for salesmen, δ is an element of the natural numbers because it represents individuals, then the rest; α, λ, γ, and μ are all in the range of $[0, \infty)$. With these restrictions in mind we will continue to build the SIR model for the marketplace.

The differential equations of potential, active, and dropped-out businesses with respect to time show an open system where each possible interaction with a business occurs, including those between new and old, potential and active, and potential and salesmen. The last equation relates the depletion of funds to the growth of the marketplace for salesmen with respect to time. Once the funds for the salesmen have been used up, we remove them from the equation, at which point organic growth must take over in order for the marketplace to continue to grow. The system of differential equations brings all the variables and parameters together and is defined in Equation 3.1,

$$
\begin{aligned}
\frac{dP_2}{dt} &= \mu - \beta P_2(t) - \lambda\delta P_2(t) - \gamma P_2(t)\sum_{j=0}^{2} A_j(t) + \eta\sigma A_2(t) - \xi P_2(t) - \theta_2 P_2(t), \\
\frac{dP_1}{dt} &= \beta P_2(t) - \kappa P_1(t) - \lambda\delta P_1(t) - \gamma P_1(t)\sum_{j=0}^{2} A_j(t) + \eta\sigma A_1(t) - \xi P_1(t) - \theta_1 P_1(t), \\
\frac{dP_0}{dt} &= \kappa P_1(t) - \lambda\delta P_0(t) - \gamma P_0(t)\sum_{j=0}^{2} A_j(t) + \eta\sigma A_0(t) - \xi P_0(t) - \theta_0 P_0(t), \\
\frac{dA_2}{dt} &= -\beta A_2(t) + \lambda\delta P_2(t) + \gamma P_2(t)\sum_{j=0}^{2} A_j(t) - \sigma A_2(t) - \theta_2 A_2(t), \\
\frac{dA_1}{dt} &= \beta A_2(t) - \kappa A_1(t) + \lambda\delta P_1(t) + \gamma P_1(t)\sum_{j=0}^{2} A_j(t) - \sigma A_1(t) - \theta_1 A_1(t), \\
\frac{dA_0}{dt} &= \kappa A_1 + \lambda\delta P_0(t) + \gamma P_0(t)\sum_{j=0}^{2} A_j(t) - \sigma A_0(t) - \theta_0 A_0(t), \\
\frac{dD_2}{dt} &= -\beta D_2(t) + (1-\eta)\sigma A_2(t) + \xi P_2(t) - \theta_2 D_2(t), \\
\frac{dD_1}{dt} &= \beta D_2(t) - \kappa D_1(t) + (1-\eta)\sigma A_1(t) + \xi P_1(t) - \theta_1 D_1(t), \\
\frac{dD_0}{dt} &= \kappa D_1(t) + (1-\eta)\sigma A_0(t) + \xi P_0(t) - \theta_0 D_0(t), \\
\frac{dS}{dt} &= -\alpha\lambda\delta\sum_{j=0}^{2} P_j(t). \qquad (3.1)
\end{aligned}
$$

This system of equations is what we will use to model the joining and abstaining of small businesses in the online marketplace proposed by eBay. Estimates for each of the parameters and compartment populations will come later.

4 Analyzing the Model

In order to better understand the behavior of the eBay Local system and how best to optimize the results, it is important to analyze the structure of the model and its components. We have created a simplified model of Equation 3.1 by removing the age stratification and the funding portion in order to do some analysis.

Condensed Variables	
$P(t)$	Number of potential businesses at time t [businesses]
$A(t)$	Number of active businesses at time t [businesses]
$D(t)$	Number of dropped out businesses at time t [businesses]

Table 4: Condensed population breakdown for the SIR model

Condensed Parameters	
μ	new business development rate $[\frac{businesses}{time}]$
θ	business failure rate $[\frac{businesses}{businesses \cdot time}]$
δ	salesmen [salesmen]
λ	business acquisition rate $[\frac{1}{time \cdot salesmen}]$
γ	organic growth rate $[\frac{1}{businesses \cdot time}]$
σ	drop out rate $[\frac{businesses}{businesses \cdot time}]$
η	retrieval rate $[\frac{businesses}{businesses \cdot time}]$
ξ	refusal rate $[\frac{businesses}{businesses \cdot time}]$

Table 5: Condensed transition parameters for the SIR model

After condensing the system into its essential components listed in Tables 4 and 5, Equation 3.1 can be now re-written:

$$\begin{aligned}
\frac{dP}{dt} &= \mu - \lambda\delta P(t) - \gamma P(t)A(t) + \eta\sigma A(t) - \xi P(t) - \theta P(t), \\
\frac{dA}{dt} &= \lambda\delta P(t) + \gamma P(t)A(t) - \sigma A(t) - \theta A(t), \\
\frac{dD}{dt} &= (1-\eta)\sigma A(t) + \xi P(t) - \theta D(t).
\end{aligned} \tag{4.1}$$

All of the coefficients in Table 5 are treated as known constants that have been predetermined in order to simplify the analysis of the system.

4.1 Finding the Equilibrium

To begin to break down the system and start solving for the variables $P(t)$, $A(t)$, and $D(t)$, the first step is to add together the first three lines of Equation 4.1:

$$\begin{aligned}\frac{dP}{dt}+\frac{dA}{dt}+\frac{dD}{dt}=&\mu\lambda\delta P(t)-\gamma P(t)A(t)-\eta\sigma A(t)-\xi P(t)-\theta P(t)+\lambda\delta P(t)\\&+\gamma P(t)A(t)-\sigma A(t)-\theta A(t)+(1-\eta)\sigma A(t)+\xi P(t)-\theta D(t).\end{aligned}$$

Many of the terms are reduced to zero by addition/subtraction cancellations, simplifying to show that the total population changes as new businesses are added and failed businesses are removed from each compartment.

$$\frac{dP}{dt}+\frac{dA}{dt}+\frac{dD}{dt}=\mu-\theta P(t)-\theta A(t)-\theta D(t)$$

For the sake of simplicity $N(t)$, representing the entire population, is substituted for $P(t)+A(t)+D(t)$. This relation is seen below:

$$\frac{dN}{dt}=\mu-\theta N(t). \tag{4.2}$$

Equation 4.2 is a first-order linear ordinary differential equation which can be solved using an integration factor of $e^{\theta t}$. After applying the integration factor the result is:

$$N(t)=\frac{\mu}{\theta}+ce^{-\theta t}. \tag{4.3}$$

The initial conditions are used to solve for the integration constant, c. At time $t=0$ the exponential term in Equation 4.3 simplifies to 1. The dropping out of the exponential terms makes intuitive sense as it reduces the equation to show that the initial population is equal to the ratio of new to failed businesses plus a constant.

$$N(0)=N_0=\frac{\mu}{\theta}+c$$

After we have solved the prior equation for c, we can rewrite Equation 4.3 in terms of known coefficients and initial conditions, which are assumed to have been given.

$$N(t)=\frac{\mu}{\theta}\left(1-e^{-\theta t}\right)+N_0e^{-\theta t} \tag{4.4}$$

By evaluating the equation as t approaches infinity, the system's steady state behavior can be found. Finding this behavior serves as a self-check for the work done so far. As time approaches infinity, the total population of businesses tends towards the ratio of new to failed businesses. If businesses are failing quicker than they are starting up, the total population will decrease over time accordingly.

$$\lim_{t\to\infty} N(t) = \frac{\mu}{\theta}$$

At this point it is now okay to undo the substitution of $N(t)$ for $P(t)+A(t)+D(t)$ to begin considering the next steps in analyzing the system.

$$P(t) + A(t) + D(t) = (P_0 + A_0 + D_0)\, e^{-\theta t} + \frac{\mu}{\theta}\left(1 - e^{-\theta t}\right) \tag{4.5}$$

Equation 4.5 shows that the total population broken up by SIR compartments, potential, active, and dropped out, can be determined at each time step based on initial populations and known parameters. To determine the total population, we consider the number of initial businesses still surviving at time t and the number of non-initial businesses still surviving. The transient state of the equation is represented by $(P_0 + A_0 + D_0)\, e^{-\theta t}$ and the steady state by $\frac{\mu}{\theta}$. According to the structure of the model, all businesses eventually fail over time as in a real market where no one business is immune to a weak economy.

The next step in breaking down the system in Equation 4.1 is to consider what happens at the system's equilibrium. When the system is in equilibrium, the rates $\frac{dP}{dt}$, $\frac{dA}{dt}$, and $\frac{dD}{dt}$ are set to zero and the compartments are no longer considered time-dependent. Instead, they are at a steady state, denoted by P_∞, A_∞, and D_∞.

$$\begin{aligned}
0 &= \mu - \lambda\delta P_\infty - \gamma P_\infty A_\infty + \eta\sigma A_\infty - \xi P_\infty - \theta P_\infty \\
0 &= \lambda\delta P_\infty + \gamma P_\infty A_\infty - \sigma A_\infty - \theta A_\infty \\
0 &= (1-\eta)\sigma A_\infty + \xi P_\infty - \theta D_\infty
\end{aligned} \tag{4.6}$$

The dropped out compartment, D_∞, of the system in Equation 4.6 is only present in the third line of the equations. Because this term is only found in the one line, it can be solved for in terms of P_∞ and A_∞ with various constants:

$$D_\infty = \frac{1}{\theta}\left[(1-\eta)\,\sigma A_\infty + \xi P_\infty\right]. \tag{4.7}$$

Once P_∞ and A_∞ have been solved for, D_∞ is determined, so we set about solving for them in the first two lines of Equation 4.6. Without the third line, we can consider Equation 4.6 to be a system of two equations and two unknowns, P_∞ and A_∞. In this system we begin with the second line of Equation 4.6 where we solve for A_∞ in terms of P_∞.

$$A_\infty = \frac{\lambda\delta P_\infty}{\theta + \sigma - \gamma P_\infty}. \tag{4.8}$$

Having solved for A_∞ in terms of P_∞ and some given coefficients in Equation 4.8 we can now solve for P_∞. To do this we look at the first line of Equation 4.6 and substitute all A_∞ terms for Equation 4.8. Doing so gives us an equation for P_∞ in terms of only the given constants,

$$0 = \mu - (\lambda\delta + \theta + \xi) P_\infty - \gamma P_\infty \left(\frac{\lambda\delta P_\infty}{\theta + \sigma - \gamma P_\infty} \right) + \eta\sigma \left(\frac{\lambda\delta P_\infty}{\theta + \sigma - \gamma P_\infty} \right).$$

After reduction and manipulation, we can reduce this equation into the following form:

$$(\lambda\delta + \theta + \xi) P_\infty - \frac{(\eta\sigma - \gamma P_\infty) \lambda\delta P_\infty}{\theta + \sigma - \gamma P_\infty} = \mu. \tag{4.9}$$

Since Equation 4.9 is a quadratic equation, it can have at most two positive roots where the left-hand side is equal to the right. The roots indicate where the equilibrium for the system can be found. To find these equilibrium points, it helps to rewrite Equation 4.9 in a more familiar form of a quadratic:

$$-\gamma (\theta + \xi) P_\infty^2 + [(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu] P_\infty - (\theta + \sigma) \mu = 0. \tag{4.10}$$

To find the roots of Equation 4.10, the quadratic formula is used:

$$\begin{aligned} P_\infty =& \frac{1}{2\gamma (\theta + \xi)} [(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu] \\ &\mp \frac{\sqrt{[(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu]^2 - 4 [-\gamma (\theta + \xi)] [-(\theta + \sigma) \mu]}}{2\gamma (\theta + \xi)}. \end{aligned} \tag{4.11}$$

Equation 4.11 shows that there are two roots of P_∞. We must establish whether these roots are positive; only positive values of P_∞ may be associated with valid equilibria. Because all of the parameters in Table 5 are positive, the leading fractional coefficient and the denominator under the term with the radical are positive. The first term can be shown to be positive after distributing the $(\theta + \sigma)$. After this we compare the two products containing σ to ensure the negative term is lesser in magnitude than the sum of the other terms. Since η is a rate in the range $0 < \eta < 1$, the following inequality is true:

$$0 < \sigma\lambda\delta (1 - \eta) = \sigma\lambda\delta - \eta\sigma\lambda\delta. \tag{4.12}$$

Hence, all the parameters being added in the first half of Equation 4.11 are positive. The next step is to verify that the discriminant is positive. We have just shown that the squared term under the radical in Equation 4.11 is positive. From this positive squared term we are subtracting another positive term. For the roots of P_∞ to be real, the subtracted term must be less than the squared term. We can check that the discriminant, Δ, is positive in the following steps:

$$\begin{aligned} \Delta &= [(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu]^2 - 4 [\gamma (\theta + \xi)] [(\theta + \sigma) \mu] \\ &> [(\theta + \xi)(\theta + \sigma) + \gamma\mu]^2 - 4\gamma\mu (\theta + \xi)(\theta + \sigma) \\ &= [(\theta + \xi)(\theta + \sigma) - \gamma\mu]^2 . \end{aligned} \tag{4.13}$$

The series of steps in Equation 4.13 show that the discriminant is greater than a simplification in which

we remove the subtracted term, $\eta\sigma\lambda\delta$. After multiplying out the square, we find that the subtracted term simplifies the equation in such a way that it can be factored into a squared term which is always positive. Hence, the discriminant must be positive.

$$0 < \sqrt{\left[(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu\right]^2 - 4\left[\gamma(\theta + \xi)\right]\left[(\theta + \sigma)\mu\right]}$$

We have shown the discriminant to be positive so the roots must at least be real. The term being squared in the discriminant is the same one seen outside of the radical in Equation 4.11. By squaring this value, subtracting from it a positive term, and taking the square root of it, we are reducing its original value. This means the first bracketed line in Equation 4.11 is greater in magnitude than the second line, the radical term. Because the first term is always greater and the discriminant is positive, when subtracting the second half of Equation 4.11 from the first we get a positive root. Since both halves of Equation 4.11 are positive, their sum is also positive and we find that there is a second positive root.

$$\begin{aligned} P_\infty^- =& \frac{1}{2\gamma(\theta + \xi)}\left[(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu\right] \\ &- \frac{\sqrt{\left[(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu\right]^2 - 4\left[\gamma(\theta + \xi)\right]\left[(\theta + \sigma)\mu\right]}}{2\gamma(\theta + \xi)} \end{aligned} \tag{4.14}$$

$$\begin{aligned} P_\infty^+ =& \frac{1}{2\gamma(\theta + \xi)}\left[(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu\right] \\ &+ \frac{\sqrt{\left[(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu\right]^2 - 4\left[\gamma(\theta + \xi)\right]\left[(\theta + \sigma)\mu\right]}}{2\gamma(\theta + \xi)} \end{aligned} \tag{4.15}$$

Now that we have two positive roots for P_∞, we want to see if both equilibrium points are possible in the overall system detailed in Equation 4.6. To do this, we substitute each of the roots, Equations 4.14 and 4.15 for P_∞ in Equation 4.8 and check if the corresponding roots of A_∞ are also positive. A look at Equation 4.8 shows that we are only concerned with what happens when the roots for P_∞ cause the denominator of A_∞ to become negative, thus creating a negative root. Again, negative roots are not considered in the domain of this SIR Model because $P(t)$, $A(t)$, and $D(t)$ are populations and you cannot have a negative population value. To find possible roots for A_∞, we want to maintain the following inequality derived from Equation 4.8:

$$P_\infty < \frac{\theta + \sigma}{\gamma}. \tag{4.16}$$

Returning to Equation 4.9, we can geometrically determine approximately where the two roots P_∞^- and P_∞^+ lie. Consider a Cartesian plane in which the x-axis is P_∞ and the y-axis is a combination of the parameters that define P_∞, $f(P_\infty) = (\lambda\delta + \sigma + xi)P_\infty - \frac{(\eta\sigma - \gamma P_\infty)\lambda\delta P_\infty}{\theta + \sigma - \gamma P_\infty} = \mu$. Then define the left-hand side of the equality in Equation 4.9 as $f(P_\infty)$, to find the equilibrium points we must find

where $f(P_\infty)$ crosses the line $y = \mu$. For very small values of P_∞, P_∞^2 is even smaller so the dominant term is $(\lambda\delta + \theta + \xi) P_\infty$ and $f(P_\infty)$ increases linearly. To ensure that P_∞ increases monotonically, we take the derivative of Equation 4.10 with respect to P_∞ and solve for P_∞:

$$P_\infty = \frac{(\lambda\delta + \theta + \xi)(\theta + \sigma) - \eta\sigma\lambda\delta + \gamma\mu}{2\gamma(\theta + \xi)}. \tag{4.17}$$

The numerator of Equation 4.17 is strictly positive as we discussed when showing the discriminant of Equation 4.11 was positive. We note that the denominator is strictly positive as gamma, theta and eta are all strictly positive, therefore $f(P_\infty)$ is an increasing function. Since $f(P_\infty)$ increases monotonically, as P_∞ increases, the function approaches the vertical asymptote $P_\infty = \frac{\theta+\sigma}{\gamma}$. On the right-hand side of $P_\infty = \frac{\theta+\sigma}{\gamma}$, $f(P_\infty)$ ascends from the asymptote and continues to increase monotonically as P_∞ increases. A sketch of this behavior can be seen in Figure 2:

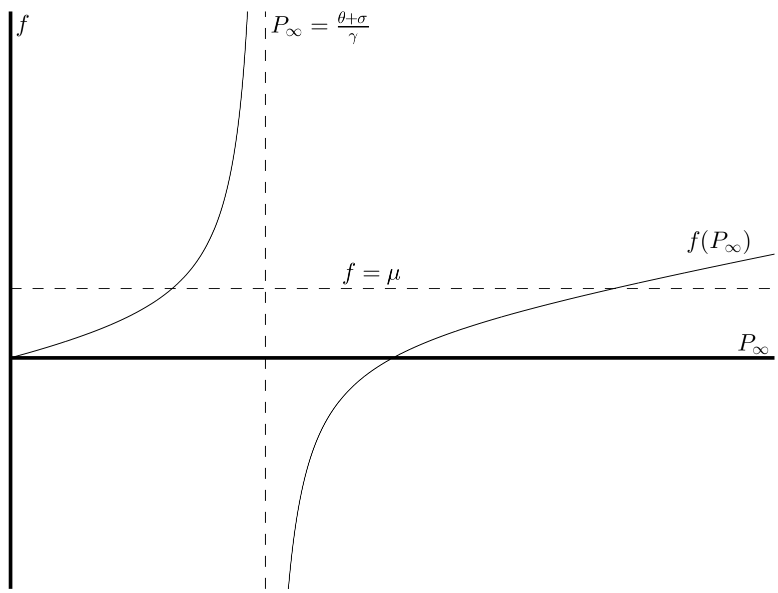

Figure 2: Sketch of P_∞

Because the function $f(P_\infty)$ is always monotonically increasing, proven by Equation 4.17, its graph crosses the line $y = \mu$ in exactly two places, once before the vertical asymptote, and once after. This coincides with our finding of two roots in Equation 4.14 and Equation 4.15. According to Equation 4.16 though, the lesser of the two roots, 4.14, results in a positive A_∞ and 4.15 results in a negative value for A_∞. This means only the lesser of the two roots is a valid equilibrium point since a negative A_∞ is not realistic. The value of P_∞^- from Equation 4.14, will then be used to determine A_∞ and D_∞ using Equations 4.8 and 4.7 respectively. Therefore the system has a unique equilibrium when $P = P_\infty^-$, $A = A_\infty$, and $D = D_\infty$.

4.2 Determining Stability

After finding the equilibrium values for the system defined by Equation 4.1, it is important to see how stable the point is. We do this by looking at global and local stability using phase plane analysis to investigate the behavior at and around the system's equilibrium point.

4.2.1 Global Stability

Starting with our initial system from Equation 4.1 and based on our results from finding a unique equilibrium, we begin our phase plane analysis in the first quadrant of a Cartesian Plane where P is the x-axis and A is the y-axis. We know that there exists a single equilibrium point in the first quadrant, which is the only realistic solution to the system that we will consider. Beginning on the y-axis, we set $P = 0$ in the first line of Equation 4.1:

$$\frac{dP}{dt} = \mu + \eta\sigma A(t). \tag{4.18}$$

Because all of the parameters in Table 5 are positive and $A(t)$ is always non-negative, Equation 4.18 is always positive. This means that along the y-axis of our phase plane, the vectors that originate there all point to the right, in the positive direction. Also, as we increase $A(t)$, the vectors representing the derivatives defined by Equation 4.18 become longer. The vectors become longer because their lengths represent the magnitude of Equation 4.18 which increases as we increase $A(t)$. Next, we check the x-axis by setting $A = 0$ in the second line of Equation 4.1:

$$\frac{dA}{dt} = \lambda\delta P(t). \tag{4.19}$$

In analyzing Equation 4.19 we proceed as we did in analyzing Equation 4.18. Along the x-axis $P(t)$ is positive according to our problem domain since it is a population implying that $\frac{dA}{dt}$ is also positive. A positive $\frac{dA}{dt}$ causes vectors originating on the x-axis to point into the first quadrant. Because any vector on either the x- or y-axis points towards the first quadrant where the equilibrium point is, the system is positively invariant. This outlines a trapping region forcing a solution that begins in the first quadrant to remain there.

Now that we know Equation 4.1 is positively invariant, we take a closer look at the axes to find the slope of the vectors. Returning to the y-axis, we divide line two of Equation 4.1 by line one. Then set $P = 0$ for all the values in the equation, the result is as follows:

$$\frac{dP}{dA} = \frac{\mu + \eta\sigma A(t)}{-(\theta + \sigma)A(t)}. \tag{4.20}$$

This shows that the slopes along the y-axis point down and to the right, seemingly approaching the equilibrium point. From the x-axis we do a similar calculation but invert the division and set $A = 0$:

$$\frac{dA}{dP} = \frac{\lambda\delta P(t)}{\mu - (\lambda\delta + \xi + \theta) P(t)}. \tag{4.21}$$

Along the x-axis there is a vertical asymptote when Equation 4.21 is undefined:

$$P(t) = \frac{\mu}{\lambda\delta + \xi + \sigma}.$$

To the left of this asymptote and along the x-axis, the vectors point up and to the right, while $(\lambda\delta + \xi + \sigma) P(t) < \mu$. When the opposite inequality is true, the vectors point up and to the left, both sides going towards the asymptote. This shows that when $A = 0$ the equilibrium is stable in one dimension.

4.2.2 Local Stability

We would like to check the local stability of the system by focusing near the equilibrium point. To do so in the Cartesian plane we have set up, we conduct local linearization about the point (P_∞, A_∞), where P_∞ and A_∞ are defined in Equations 4.14 and 4.8, respectively. A simple way of doing this is by assuming there exists a Taylor expansion of the first two lines of Equation 4.1 and taking the Jacobian of this expansion where $P(t)$ and $A(t)$ have been replaced by $P - P_\infty$ and $A - A_\infty$. After this substitution, all the appropriate steps are taken to reconcile the additions of P_∞ and A_∞ resulting in the following:

$$\begin{aligned}
\frac{d(P - P_\infty)}{dt} =& \mu - \lambda\delta(P - P_\infty) - \gamma(P - P_\infty)(A - A_\infty) + \eta\sigma(A - A_\infty) - \xi(P - P_\infty) \\
& - \theta(P - P_\infty) - \lambda\delta P_\infty - \gamma P_\infty A - \gamma P A_\infty + \gamma P_\infty A_\infty + \eta\sigma A_\infty - \xi P_\infty - \theta P_\infty, \quad (4.22) \\
\frac{d(A - A_\infty)}{dt} =& \lambda\delta(P - P_\infty) + \gamma(P - P_\infty)(A - A_\infty) - \sigma(A - A_\infty) - \theta(A - A_\infty) \\
& + \lambda\delta P_\infty + \gamma P_\infty A + \gamma P A_\infty - \gamma P_\infty A_\infty - \sigma A_\infty - \theta A_\infty. \quad (4.23)
\end{aligned}$$

Using Equation 4.6, $\mu - \lambda\delta P_\infty + \eta\sigma A_\infty - \xi P_\infty - \theta P_\infty$ in Equation 4.22 can be replaced with $\gamma P_\infty A_\infty$, and $\lambda\delta P_\infty - \sigma A_\infty - \theta A_\infty$ in Equation 4.23 with $\gamma P_\infty A_\infty$ also. After reorganization of the two equations (second order terms are reduced to zero), the above can be expressed as a linear system:

$$\begin{bmatrix} -(\lambda\delta + \gamma A_\infty + \xi + \theta) & \eta\sigma - \gamma P_\infty \\ \lambda\delta + \gamma A_\infty & \gamma P_\infty - \sigma - \theta \end{bmatrix} \begin{bmatrix} P - P_\infty \\ A - A_\infty \end{bmatrix} = \begin{bmatrix} \frac{d(P-P_\infty)}{dt} \\ \frac{d(A-A_\infty)}{dt} \end{bmatrix}. \tag{4.24}$$

We now wish to focus on the first matrix of Equation 4.24 containing the coefficients as the next step in checking the local stability of Equation 4.1:

$$\begin{bmatrix} -(\lambda\delta + \gamma A_\infty + \xi + \theta) & \eta\sigma - \gamma P_\infty \\ \lambda\delta + \gamma A_\infty & \gamma P_\infty - \sigma - \theta \end{bmatrix}.$$

Remember, as defined in Table 5 and Equations 4.8 and 4.14, all the parameters of the entries in the above matrix are non-negative. This means that in the first column the first element is non-positive and the second is non-negative. For the time being we will assume none of the parameters are zero to simplify the language, hence column one has a negative entry and a positive entry. By Equation 4.16, the bottom element of the second column is negative since $\gamma P_\infty < \theta + \sigma$. The only element of the matrix that is not definitely positive or negative is the upper element of the second column, this is because depending on the values chosen for η, σ, and γ, the value can be either. Based on this information about the entries of the matrix, we know the trace is negative. Next we look at the determinant of the matrix:

$$-\left(\frac{\theta+\sigma}{\theta+\sigma-\gamma P_\infty}+\lambda\delta+\xi+\theta\right)(\gamma P_\infty-\sigma-\theta)-\left(\frac{\theta+\sigma}{\theta+\sigma-\gamma P_\infty}\lambda\delta\right)(\eta\sigma-\gamma P_\infty)\,.$$

In the determinant, the first product is positive and the negative component of the second product is small enough in magnitude to not cause the entire determinant to be negative. There is a similar term in the first product to ensure it is positive, the re-grouping below shows this:

$$(1-\lambda\delta\eta)\,\frac{\theta+\sigma}{\theta+\sigma-\gamma P_\infty}\sigma.$$

The term $1-\lambda\delta\eta$ is positive so the determinant is also positive. Since the trace of the matrix is negative and the determinant is positive, the real parts of the eigenvalues are negative, hence the system is locally stable about the equilibrium point. Knowing we have local stability, we can say the equilibrium is a sink and wish to check if it is a direct sink or a spiral sink. The two eigenvalues must at least have negative real components, if they also have imaginary components then the equilibrium is a spiral sink. To check the type of sink at the equilibrium we tested different parameter values to find real and complex eigenvalue pairs. The most likely cases with realistic parameter estimates had real eigenvalues but more obscure estimates resulted in complex eigenvalues. These cases are shown as examples.

The parameters resulting in real eigenvalues for a direct sink are seen below along with the location of the equilibrium, (P_∞, A_∞).

Parameter Values	
δ	4
λ	0.0034
γ	0.000011
σ	6.08
η	0.25
ξ	0.25
μ	22
θ	0.00035
P_∞	84.44
A_∞	0.1889

Table 6: Parameter values for a direct sink

The phase plane plot for the values in Table 6 is seen in the Figure 3.

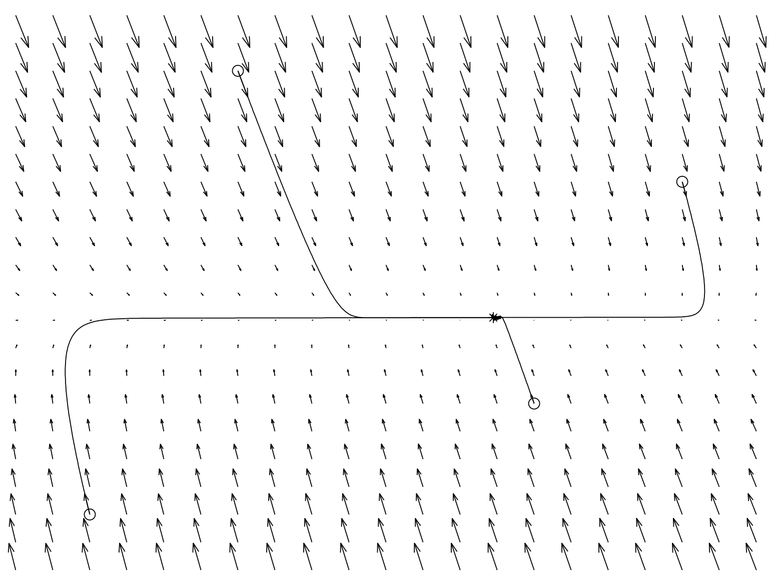

Figure 3: Direct sink phase plane plot

Notice the vectors in the surrounding area point directly inwards, towards the equilibrium point. In the plot are four solution curves, originating at the circles and ending all at the same point, the asterisk.

For an example of a spiral sink, the following parameter values are used:

Parameter Values	
δ	0.9578
λ	1.0543
γ	1.2943
σ	1.2382
η	0.2581
ξ	0.2231
μ	2.2642
θ	0.9665
P_∞	0.8024
A_∞	0.6948

Table 7: Parameter values for a spiral sink

The phase plane plot for the spiral sink example values in Table 7 is seen in the following figure:

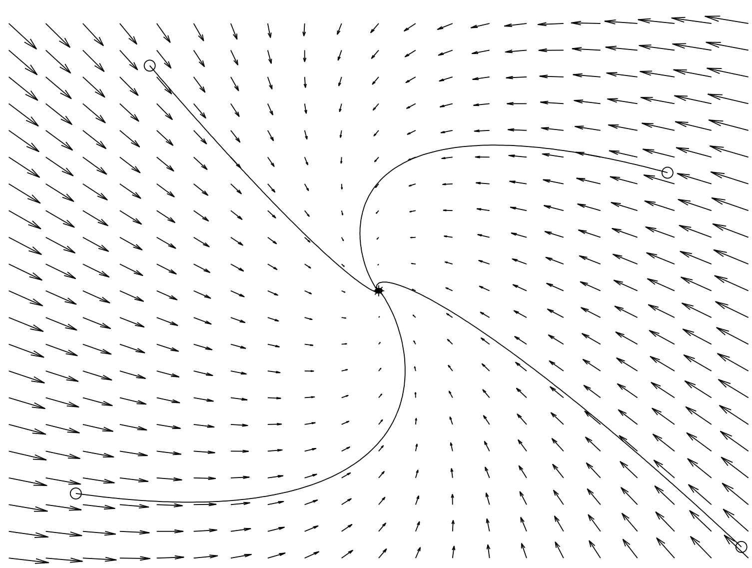

Figure 4: Spiral sink phase plane plot

Here we see that the trajectories follow a more spiral pattern towards the equilibrium point noted by the asterix instead of a direct route like in Figure 4.

Since the discriminant can be either positive or negative, the eigenvalues of the coefficient matrix in Equation 4.24 can be real or complex. For this reason we conclude that depending on the parameter values chosen, the stable sink can be either direct or spiral.

5 Data and Parameter Collection

In order to validate the model we have developed to visualize the behavior of the eBay Local marketplace, reasonable parameter estimates must be made. Based on these estimates we can project the potential, active, and dropped out business populations in the SIS epidemiology model Equation 4.1 and predict the success of the program. The marketplace has been running for a few months and initial estimates can be derived from the data collected thus far. After removing all of the personal identification information from the database, the data was sent to us by eBay employee Ted Dziuba[1] to analyze and use in the model.

By using the program to simulate approximately six months of the marketplace running we can estimate the building of the marketplace and attempt to predict its long-term success. The first step when working with the data was to build a correlation matrix to see if any businesses would have a higher predisposition to joining the marketplace. We found that the database ID values for merchants were highly correlated to the ones for accounts. This correlation shows that many merchants can belong to a single account. The merchants are the ones who post the coupons but these coupons can sometimes be used at multiple locations if the account is also associated with other merchants. For the marketplace, the test city was Los Angeles, California but because of the multiple locations aspect of larger companies, some coupons were posted for other cities. We also believe that a few of the merchants were online distributors without a physical storefront that would ship goods to residents of Los Angeles and other cities.

Apart from the location of the merchants, there are no other correlations in the data. In the data, if a merchant was rated on the site Yelp, it was recorded. Yelp is a social website where individuals can post ratings and reviews for local businesses. There was no correlation between Yelp ratings, or lack thereof, and the coupons posted, including price or likelihood of posting. It also showed that while most of the active merchants were classified as active, beauty, restaurants, and shopping, there were merchants from each of the 20 different categories participating. Merchants posted deals valued from five dollars to over 2000 with expiration dates ranging from three months to up to two years. The age of merchants was not tracked by eBay in the marketplace.

In order to estimate our parameters for use in the model we start with the economic demographics of Los Angeles, California. According to the United States Census Bureau, in 2011 there were approximately 495,000 businesses in California with fewer than 20 employees. Twenty is an appropriate cap on the number of employees that defines a small business for the purposes of this thesis. Of these 495,000 businesses, approximately 87,000 were less than two years old and were classified as newly established, 101,000 were moderately established, and the majority, 307,000 were well established. Therefore we estimated the population breakdown to be $18 : 20 : 62$ of newly, moderately, and well established businesses. Also from the Census Bureau, nearly 76,500 businesses with fewer than 20 employees were classified as retail, arts, entertainment, recreation, accommodation, or food services. We estimated that approximately 80% of the businesses in that Los Angeles sold goods or services in these categories and would be interested in the marketplace. For simplicity, the total potential population before the start of

[1] Data and information exchanged with Ted Dziuba were sent as personal communications

eBay Local is 62,500 with a breakdown of 20 : 20 : 60. Because the program has not yet started, there are no active or dropped out businesses at time $t = 0$. Hence the initial conditions for the model are: $P_2(0) = P_1(0) = 12500$, $P_0(0) = 37500$, and $A_2(0) = A_1(0) = A_0(0) = D_2(0) = D_1(0) = D_0(0) = 0$.

On average half of the newly established potential businesses have been in operation for less than one year. However, the younger the business, the higher its likelihood to fail, so we estimate that of the newly established businesses closer to 8,000 are in their first year of business, hence $\mu = 8000$. If we divide the population of newly established businesses into half, then every year each half ages another year which causes the older half to transition into the moderately established compartment. For a similar reason, the oldest third of the moderately established businesses mature into well established businesses. These parameters are represented by $\beta = \frac{1}{2}$ and $\kappa = \frac{1}{3}$. The values for θ_2, θ_1, and θ_0 are estimated using μ, β, and κ to ensure a stable population for each age compartment. Based on the above values we define the business failure rates as $\theta_2 = 0.14$, $\theta_1 = 0.1667$, and $\theta_0 = 0.1111$.

For the initial building of the marketplace, eBay set a budget of $\$175,000$ for paying sales staff. The sales staff initially consists of four salesmen who were paid an average \$450 commission for each new business brought into the marketplace. Using this information we can estimate the following initial condition and parameters: $S(0) = 175000$, $\delta = 4$, and $\alpha = 450$.

Because eBay did not take into account the ages of businesses, the transition parameters for merchant involvement are the same for each age compartment. From the data, 440 merchants posted coupons in the marketplace. Of the total 440, 365 were a direct result of a salesman calls and 75 were from organic growth. We estimate λ and γ by setting the rest of the merchant involvement transition parameters to zero. Through the process of checking the output of the model and comparing the results to the actual data, we found that $\lambda = 0.0034$ and $\gamma = 0.00011$. After eBay Local had been running for six months, it was found that all active businesses dropped out of the marketplace. In fact, the number of merchants that posted a new coupon unprompted was on the order of round-off error. With an infectious period of 60 days and a daily time step, $\sigma = \frac{365}{60}$. After only having run the marketplace a few months, there was not enough data to base estimates for η and ξ on. Instead we used the rationale that since Groupons are profitable and effective for 67% of businesses [4], those businesses are likely to try the marketplace again. This is why we estimated $\eta = 0.67$. Without any data or literature to support an ξ value, we removed it from the model by setting it to zero for the time being. An organized list of these parameter values can be found in Table 8.

Initial Values			
$P_2(0)$	12500	μ	8000
$P_1(0)$	12500	β	$\frac{1}{2}$
$P_0(0)$	37500	κ	$\frac{1}{3}$
$A_2(0)$	0	θ_2	0.14
$A_1(0)$	0	θ_1	0.1667
$A_0(0)$	0	θ_0	0.1111
$D_2(0)$	0	δ	4
$D_1(0)$	0	α	450
$D_0(0)$	0	λ	0.0034
$S(0)$	175000	γ	0.00011
		σ	$\frac{365}{60}$
		η	0.67
		ξ	0

Table 8: Variable and transition parameter estimates based on initial data

Using the estimated values from Table 8 in Equation 3.1, we were able to gain some insight into the success of the eBay Local marketplace.

6 Results

After building an SIR model for eBay Local in Equation 3.1 we would like to run simulations of the marketplace to predict the success of the program. Using a fourth-order Runge-Kutta method written in MATLAB to integrate the system, we can create visual output resembling Figure 1 for merchant participation. Our forecast based on initial data can be a powerful tool for eBay when it considers marketing strategies.

6.1 The Original Model

The prediction plots we show throughout this section show the steady population two years prior to the start of eBay Local and 18 years after the creation. We have labeled the time at which the the salesmen begin recruiting merchants as time 0 on the x-axis. With the variable and transition parameter estimates from Table 8, we can plot our first graph of the marketplace.

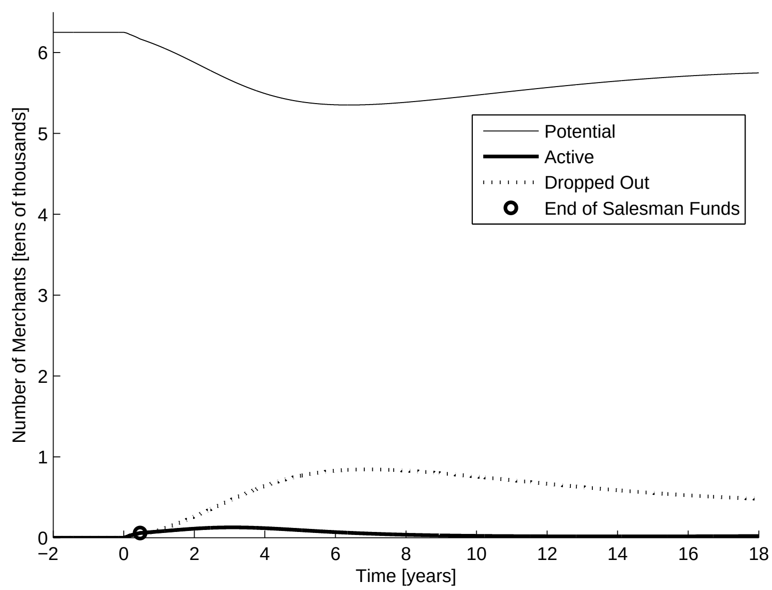

Figure 5: Prediction plot based on original data estimates

Figure 5 shows the important role the salesmen play. Six months after starting the marketplace, funds for the salesmen run out, indicated by the circle. When this happens, it is up to the merchants to grow the marketplace via organic growth. At its peak, the marketplace has over 1,000 merchants actively posting coupons for distributors to bid on. This sounds like a successful marketplace but this peak doesn't occur until three years after the start of eBay Local. In addition to the late peak of the marketplace, we observe that without salesmen the population of active merchants eventually drops to 200 merchants.

6.2 Salesman Related Model Variations

Knowing the important role that salesmen play in marketplace we adjusted the model in MATLAB. We assume the only parameters we are able to adjust is the number of salesmen, when they work, and how they are paid. Instead of counting down the funds available for salesmen compensation, we adjusted the model to keep a running total of the amount spent by eBay on the marketplace. In addition to changing the funding scheme, we factored in day of week and time of year for the salesmen. Since the time step is each day, we gave a five day work week to the salesmen and around the holidays we doubled the number of working salesmen from four to eight. We have plotted our predictions for the marketplace based on the above changes in the next figure.

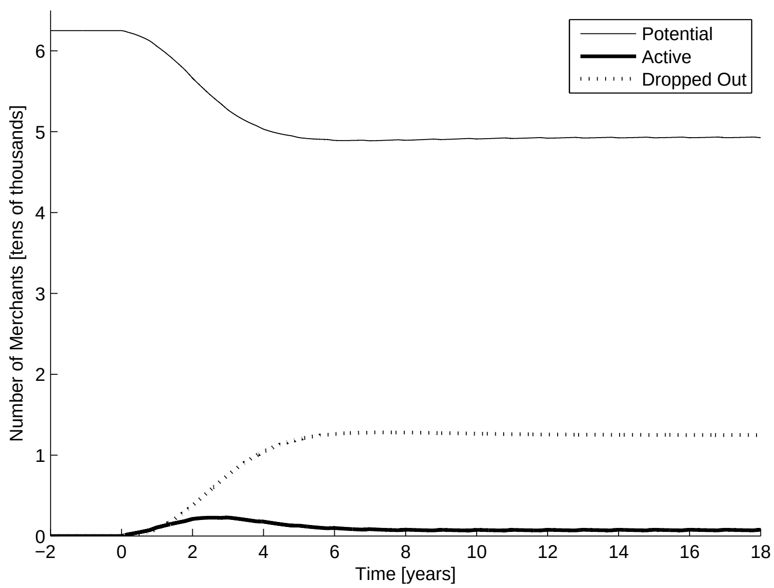

Figure 6: Prediction plot with continuous salesmen and date considerations

Here in Figure 6 we see a plot that looks similar to the plot in Figure 5 with a few important differences. Again it take approximately three years for the marketplace to hit a maximum enrollment but this time there are over 2,200 active merchants instead of 1,000. We also see that because of a continuous salesman influence on the marketplace after 18 years there are still 700 merchants posting coupons, an improvement on the 100 from Figure 5. The downside of this model variation is the price. Eighteen years after starting the marketplace, the cost to eBay would be more than $4,800,000 for a single city, much more than the $175,000 spent during the trial period.

If a merchant had a pleasant experience in the marketplace then dropped out, there is a greater likelihood that it can be persuaded into re-joining. Having a list of previously active dropped out merchants means less work for the salesmen and a lower cost of recruitment to eBay. Initially we were given an estimated cost per acquisition of $450 but this number really only applies to new merchants. For returning merchants, this acquisition cost is actually closer to $200. In order to account for this variation in spent funding, we adjusted the system, changing it from an SIR model, to an SIRS model.

The SIRS model is simply an extension of the SIR model. Equation 3.1, is technically an SIRS model because after being infected, merchants can become susceptible again. Now what we would like to do is differentiate those who are susceptible again as a separate class. In addition to the above compartments of P, A, and D, we added R for merchants Returning to the potential pool as those merchants that are susceptible after infection. Again, R is broken down by age and has the same subscripts of 2, 1, and 0 where merchants age by rates of β and κ, and fail by corresponding rates of θ. Merchants join R at a rate of $\eta\sigma A$, entering the Returning pool instead of the Potential pool. Returning merchants can then become active by salesman influence or interacting with currently active merchants. To distinguish the rates at which previously active merchants can be brought in we use λ_P, γ_P, α_P, λ_R, γ_R, and α_R where the subscript P is for the coefficients related to potential merchants and R for returning potential merchants. We adjusted Equation 3.1 to reflect these changes.

$$
\begin{aligned}
\frac{dP_2}{dt} &= \mu - \beta P_2(t) - \lambda_P \delta P_2(t) - \gamma_P P_2(t) \sum_{j=0}^{2} A_j(t) - \xi P_2(t) - \theta_2 P_2(t), \\
\frac{dP_1}{dt} &= \beta P_2(t) - \kappa P_1(t) - \lambda_P \delta P_1(t) - \gamma_P P_1(t) \sum_{j=0}^{2} A_j(t) - \xi P_1(t) - \theta_1 P_1(t), \\
\frac{dP_0}{dt} &= \kappa P_1(t) - \lambda_P \delta P_0(t) - \gamma_P P_0(t) \sum_{j=0}^{2} A_j(t) - \xi P_0(t) - \theta_0 P_0(t), \\
\frac{dA_2}{dt} &= -\beta A_2(t) + \delta \left[\lambda_P P_2(t) + \lambda_R R_2(t)\right] + \left[\gamma_P P_2(t) + \gamma_R R_2(t)\right] \sum_{j=0}^{2} A_j(t) - \sigma A_2(t) - \theta_2 A_2(t), \\
\frac{dA_1}{dt} &= \beta A_2(t) - \kappa A_1(t) + \delta \left[\lambda_P P_1(t) + \lambda_R R_1(t)\right] + \left[\gamma_P P_1(t) + \gamma_R R_1(t)\right] \sum_{j=0}^{2} A_j(t) - \sigma A_1(t) - \theta_1 A_1(t), \\
\frac{dA_0}{dt} &= \kappa A_1 + \delta \left[\lambda_P P_0(t) + \lambda_R R_0(t)\right] + \left[\gamma_P P_0(t) + \gamma_R R_0(t)\right] \sum_{j=0}^{2} A_j(t) - \sigma A_0(t) - \theta_0 A_0(t), \\
\frac{dD_2}{dt} &= -\beta D_2(t) + (1-\eta)\sigma A_2(t) + \xi P_2(t) - \theta_2 D_2(t), \\
\frac{dD_1}{dt} &= \beta D_2(t) - \kappa D_1(t) + (1-\eta)\sigma A_1(t) + \xi P_1(t) - \theta_1 D_1(t), \\
\frac{dD_0}{dt} &= \kappa D_1(t) + (1-\eta)\sigma A_0(t) + \xi P_0(t) - \theta_0 D_0(t), \\
\frac{dR_2}{dt} &= -\beta R_2(t) + \eta\sigma A_2(t) - \lambda_R \delta R_2(t) - \gamma_R R_2(t) \sum_{j=0}^{2} A_j(t) - \theta_2 R_2(t), \\
\frac{dR_1}{dt} &= \beta R_2(t) - \kappa R_1(t) + \eta\sigma A_1(t) - \lambda_R \delta R_1(t) - \gamma_R R_1(t) \sum_{j=0}^{2} A_j(t) - \theta_1 R_1(t), \\
\frac{dR_0}{dt} &= \kappa R_1(t) + \eta\sigma A_0(t) - \lambda_R \delta R_0(t) - \gamma_R R_0(t) \sum_{j=0}^{2} A_j(t) - \theta_0 R_0(t), \\
\frac{dS}{dt} &= \alpha_P \lambda_P \delta \sum_{j=0}^{2} P_j(t) + \alpha_R \lambda_R \delta \sum_{j=0}^{2} R_j(t). \qquad (6.1)
\end{aligned}
$$

For the time being, we will assume $\lambda_P = \lambda_R$ and $\gamma_P = \gamma_R$ since we don't have enough data to determine new estimates. Since the reason for this new model, Equation 6.1 was based on funds we will use the estimates of $\alpha_P = 450$ and $\alpha_R = 200$. An updated version of Table 8 can be found in the next table.

Initial Values			
$P_2(0)$	12500	μ	8000
$P_1(0)$	12500	β	$\frac{1}{2}$
$P_0(0)$	37500	κ	$\frac{1}{3}$
$A_2(0)$	0	θ_2	0.14
$A_1(0)$	0	θ_1	0.1667
$A_0(0)$	0	θ_0	0.1111
$D_2(0)$	0	δ	4
$D_1(0)$	0	α_P	450
$D_0(0)$	0	α_R	200
$R_2(0)$	0	λ_P	0.0034
$R_1(0)$	0	λ_R	0.0034
$R_0(0)$	0	γ_P	0.00011
$S(0)$	0	γ_R	0.00011
		σ	$\frac{365}{60}$
		η	0.67
		ξ	0

Table 9: Variable and transition parameter estimates for the SIRS model

The MATLAB generated prediction plot based on Equation 6.1 and values in Table 9 is shown as Figure 7.

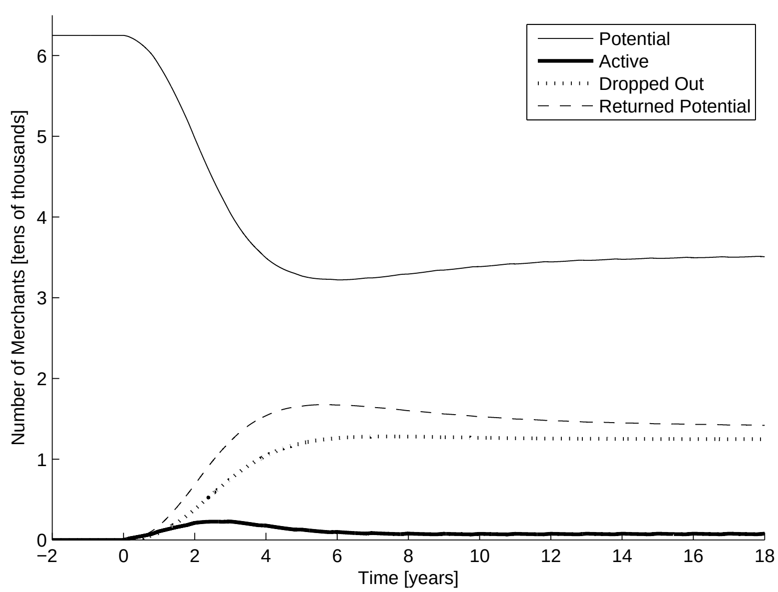

Figure 7: SIRS model prediction plot

As expected, this figure looks very similar to Figure 6 with the exception that potential merchants are split into the original potential pool and the returning pool. Again, the maximum participation occurs just shy of three years with 2,277 active merchants but for $4,100,000. While a savings of $700,000 is worth noting, we would like to save eBay more money if possible. In order to minimize the maintenance cost of eBay Local, we decided to adjust some of the other variables. By still keeping our estimates reasonable we can advise eBay on the best course to take with the marketplace.

By simply adjusting our allocation of sales staff we can greatly reduce the expense of the marketplace. First we focused all four salesmen, 8 during the holidays, on the potential population for a year. After that first year we organized their time so that each salesman spent 25% of his time finding new merchants and 75% re-recruiting returning merchants. To represent this in the model we divided δ into δ_P and δ_R similar to what we did to λ, γ, and α. In the first year we used $\delta_P = 4$ and $\delta_R = 0$, $\delta_P = 8$ and $\delta_R = 0$ at the end of the year, to show the collection of businesses from the potential compartment. Once the first year had passed we changed the values to $\delta_P = 1$ and $\delta_R = 3$, holiday values: $\delta_P = 2$ and $\delta_R = 6$. Using this organization of the salesmen we saw a graph similar to Figure 7. The peak enrollment of active members was again just before the third year of the marketplace with 1,750 merchants. At the end of 18 years there were still 500 active merchants but instead of costing eBay $4,100,000, this marketplace only cost $1,360,000. Over the course of 18 years this averages to just over $75,000 annually to maintain the marketplace in a single city, Los Angeles.

The main problem eBay Local seems to face is the 100% drop out rate. Without any merchants rejoining the marketplace on their own, the burden of building eBay Local falls on the salesmen at the expense of eBay. If the merchants who had successful coupons were to rejoin on their own, eBay could see a maximum of 31,000 active merchants just after a year and a half and 6,600 merchants at 18 for a mere $600,000. By changing the drop out rate from $\sigma = \frac{365}{60}$ to $\sigma = \frac{365}{120}$, our prediction plots changed drastically.

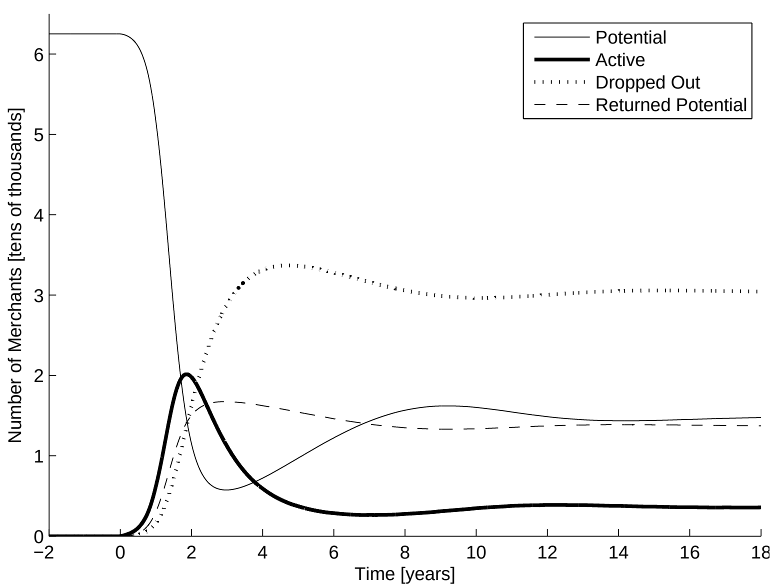

Figure 8: SIRS model prediction plot with improved drop out rate

Figure 8 is an excellent scenario for eBay Local. While the large peak does not continue beyond four years, there are still thousands of businesses remaining active at the end of the simulation. This marketplace costs eBay only \$33,000 to maintain. Such a large active population not only gives eBay a good reputation to encourage more merchants to join, it also offsets the cost of maintaining eBay Local by charging the distributors for access to the coupons. However; given the data we have seen thus far, this scenario seems highly unlikely so instead we examine what would happen if a mere 10% of merchants remained in the marketplace on their own.

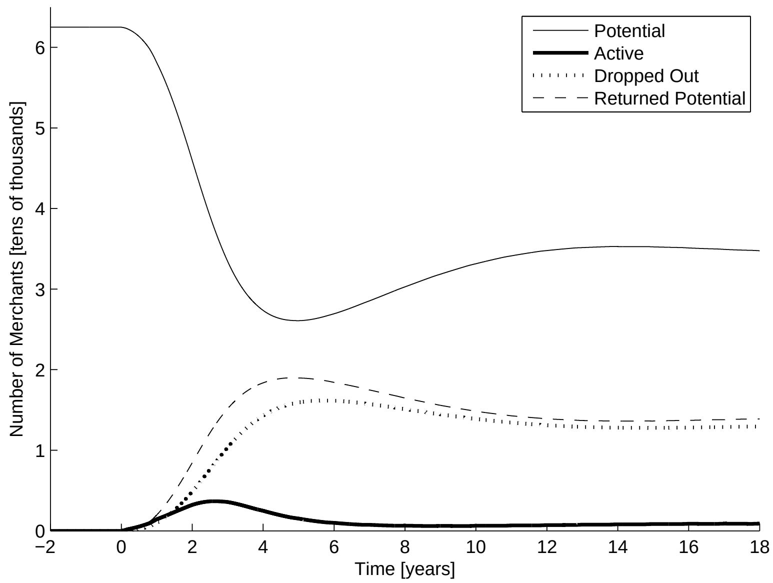

Figure 9: SIRS model prediction plot with reasonable drop out rate

Here we see a graph similar to Figure 7, peaking around two and a half years with 3,600 active merchants. The cost of building this marketplace to a steady active population of 850 merchants at 18 years was approximately \$1,225,000. While this price is higher than than the cost of the previous marketplace, it is still cheaper than one where 100% of the merchants drop out and seems reasonable enough to attain.

With more data, it would have been interesting to see what effects, if any, age had on the likelihood of a merchant joining the marketplace. It is possible that age would also play a role in organic growth. When forming an opinion on an product, individuals can experience many levels of influence from fellow individuals including age. Often a person will more seriously take the advice and suggestions of an older, more successful individual when it comes to the decision-making process [3]. Extending this idea to the merchants in our marketplace seems a natural step. If eBay were to collect data based on the age of the merchants recruited we would have more insight into the effectiveness of the age breakdown of the model. Currently we do not have enough data to accurately predict changes to the model based on age.

7 Conclusion and Future Work

In order to predict the growth and behavior of the eBay Local marketplace, we built a system of equations based on the SIR epidemiology model. In this model we found that there always exists a single possible stable equilibrium point for a given set of parameters. Depending on the parameter selection, the equilibrium is either a direct or a spiral sink. In terms of global stability, the system is positively invariant. Using a fourth-order Runge-Kutta method implemented in MATLAB, we generated numerical results for our model. To create output we used information about Los Angeles, California to estimate our merchant population sizes. From eBay we received data collected while building the marketplace that was used to estimate the transition parameters in the model. Based on these values we generated graphs to visualize the number of active merchants at any given point in time up to the first 18 years. By adjusting our parameter values and re-allocating the salesmen, we were able to create an inexpensive marketplace for eBay and come up with suggestions for improvement. The model is intended to be an all-encompassing system with room for adjustment if eBay were interested in collecting more merchant-related information such as age.

For future work we would recommend that eBay collect more data based on the age of recruited merchants and that they investigate drop out and refusal rates. With more data we could better customize the model for eBay and improve our prediction plots. We would also recommend studying the social structure of the merchants in each city. In viral marketing campaigns "social network structures have a significant impact on campaign performance"[1]. By studying these networks in each city, eBay could more effectively target and recruit merchants into the marketplace and greatly cut their costs. If they were to find an influential "hub" merchant that would encourage other merchants to post coupons via organic growth, they would see even lower operating costs.

References

[1] Mauro Bampo, Michael T. Ewing, Dineli R. Mather, David Stewart, and Mark Wallace. The effects of the social structure of digital networks on viral marketing performance. *Information Systems Research*, 19(3):273–290, September 2008.

[2] Lus M.A. Bettencourt, Ariel Cintrn-Arias, David I. Kaiser, and Carlos Castillo-Chvez. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications*, 364:513–536, May 2006.

[3] Arnaud De Bruyn and Gary L. Lilien. A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing*, 25(3):151–163, September 2008.

[4] U. Dholakia. How effective are groupon promotions for businesses? *Available at SSRN 1696327*, 2010.

[5] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A*, 115(772):700–721, August 1927.

[6] Amy E. Knaup. Survival and longevity in the business employment dynamics data. *Monthly Labor Review Online*, 128(5):50–56, May 2005.

[7] Jure Leskovec, Lada A. Adamic, and Bernardo A. Huberman. The dynamics of viral marketing. *ACM Trans. Web*, 1(1), May 2007.

[8] Daniel Trpevski, Wallace K. S. Tang, and Ljupco Kocarev. Model for rumor spreading over networks. *Physical Review E*, 81(5):056102, May 2010.