ESSAYS IN HEALTH ECONOMICS

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Economics

By

Jason Albert, M.A.

Washington, DC
April 23, 2015

Essays in Health Economics

Jason Albert, M.A.

Dissertation Advisors: Luca Anderlini and Roger Lagunoff

Abstract

The first chapter studies a dynamic model of a fee-for-service healthcare system in which healthcare providers compete for patients by prescribing antibiotics. Using antibiotics limits antibiotic-treatable infections, but fosters the development of antibiotic-resistant infections. The chapter first demonstrates a 'Goldilocks' effect from competition. A perfectly competitive market for providers over-prescribes antibiotics because providers do not bear the cost of antibiotic-resistant infections. A patient monopolist under-prescribes antibiotics in order to increase the level of treatable infection. This is because while infection is a 'bad' for society, infection is a 'good' for a provider of antibiotics under a fee-for-service regime. Due to more moderate antibiotic use, oligopolistic competition can be the optimal market structure. The chapter then demonstrates how the model can be used for policy analysis by computing the optimal licensing regime, prescription quota, and tax on antibiotics.

The second chapter studies a model in which court delay limits a court's ability to implement a welfare enhancing liability scheme. The model describes how court delay affects a firm's decision to enter a market and invest in worker safety. Workers are unaware of the risk at a particular firm, but can demand a wage that compensates for the expected risk of employment in the industry. Delay causes firms to under-invest in safety. When firms are symmetric, firms fully bear the burden of the under-investment in safety through the market wage. As delay increases, profits can decrease to the point that efficient firms do not enter the market. When firms are asymmetric, the wage

over-compensates workers at the high-type firm and under-compensates workers at the low-type firm. As delay increases, this wage subsidy from a high-type to a low-type firm can increase and induce inefficient firms to enter the market. The model provides support for applying prejudgment interest to accident damages.

INDEX WORDS: Antibiotic resistance, Healthcare competition, Markov
                 equilibria, Court delay, Worker safety

This dissertation is dedicated to my family, for helping me.

CHAPTER 1

STRATEGIC DYNAMICS OF ANTIBIOTIC USE AND THE EVOLUTION OF
ANTIBIOTIC-RESISTANT INFECTIONS

1.1 INTRODUCTION

Antibiotics have long been a boon to human well-being by fighting infection. However, using antibiotics also fosters the development of antibiotic-resistant infections (Levy 1992; Seppälä et al 1995). While antibiotic use has tempered the level of treatable infection, antibiotic-resistant infections have grown to become a costly public health problem (Klevens et al 2007; Roberts et al 2009). The Center for Disease Control and Prevention goes so far as to say "antimicrobial resistance is one of our most serious health threats" (CDC 2013). Understanding how economic incentives affect antibiotic use and the infection/resistance balance is a problem of first-order importance.

This chapter develops a model of a fee-for-service healthcare system in which healthcare providers compete for patients and must choose whether to prescribe antibiotics. This a natural component of the provider-patient relationship to model, as numerous studies have found that patients frequently request antibiotics from their providers and that providers comply to keep patients satisfied, even when use is inappropriate (Bauchner et al 1999; Stivers 2005). A recent study found that providers in the United States prescribe antibiotics to patients with a sore throat about seventy percent of the time, even though antibiotics are only required about ten percent of the

time (Barnett and Linder 2014). Providers are thus an important source of antibiotic misuse.

The model is used to study two questions. First, what is the optimal market structure of healthcare providers? There is significant variation in the market concentration of providers across counties in the United States (Schneider et al 2008). More competitive markets for providers are associated with more frequent antibiotic use than less competitive markets (Bennet et al 2014; Fogelberg 2013). However, given the competing, dynamic effects of antibiotic use on treatable and resistant infection, the optimal market structure of providers is an open question in the economics literature on antibiotic resistance. Second, what role can public policy play in correcting the misuse of antibiotics by providers?

The chapter embeds an economic model of the provider-patient matching process into an epidemiological model of infection and antibiotic resistance. The economic model describes how provider and patient behavior determine antibiotic use and the epidemiological model describes how infection evolves in response to antibiotic use.

The epidemiological model describes an infection that has two strains: an antibiotic-treatable strain and an antibiotic-resistant strain. The two strains compete for resources (healthy bodies) in the ecosystem. When a patient infected with treatable bacteria takes the antibiotic, the patient raises his probability of recovery. The patient is then less likely to infect other healthy agents. However, by removing a source of competition for the resistant bacteria, the patient's use of the antibiotics allows the resistant bacteria to grow and spread at a faster rate in the future.

Figures 1.1 and 1.2 show the evolution of Penicillin-treatable and Penicillin-resistant invasive *Streptoccocus Pneumoniae* infections in Baltimore City, Maryland and the Nashville-Davidson region of Tennessee from 2003-2009.[1]

Baltimore City, featured on the top, displays patterns typical from antibiotic use: a decrease in the rate of treatable infections and an increase in the rate of resistant infections. Nashville-Davidson, on the other hand, displays the opposite pattern: an increase in the rate of treatable infections and a decrease in the rate of resistant infections. Interestingly, in 2001, Nashville-Davidson started a campaign to reduce antibiotic use.[2] The graphs highlight competition between the resistant and treatable infections for resources: one strain of the infection grows at the expense of the other. It is this competition between the resistant and treatable infections, the winner of which is determined by antibiotic use, that the epidemiological model captures.

The economic model describes a matching process between providers and patients. Providers choose how frequently to prescribe antibiotics for their patients and patients choose whether to see a provider and which provider to see. Patients gain utility from a provider's antibiotic prescription behavior, the effectiveness of the antibiotic, and an idiosyncratic component related to the quality of their match with a provider. Patients choose whether or not to pay a fixed fee to see their most preferred provider, ignoring the external effects of their antibiotic use on infection and resistance. Providers choose a prescription rate to maximize profits, given the behavior of other providers and patients, internalizing their own effect on the future levels of infection and resistance. Provider and patient behavior together determine the aggregate amount of antibiotic use, which is then embedded into the epidemiological model.

---

[1] I thank Dr. David Blythe and the Maryland Department of Health and Mental Hygiene for access to the Baltimore City data. The Tennessee data is available online at: http://health.tn.gov/Ceds/WebAim/

[2] http://health.state.tn.us/ceds/Antibiotics/index.htm

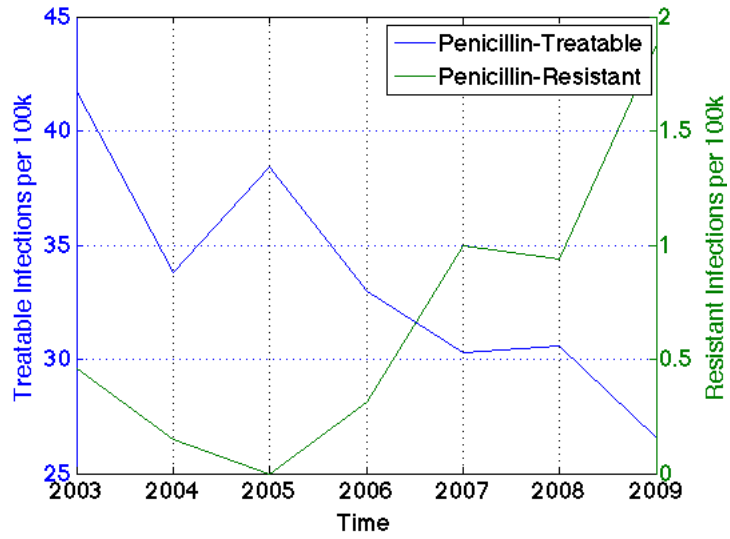Invasive *Streptococcus Pneumoniae* Dynamics
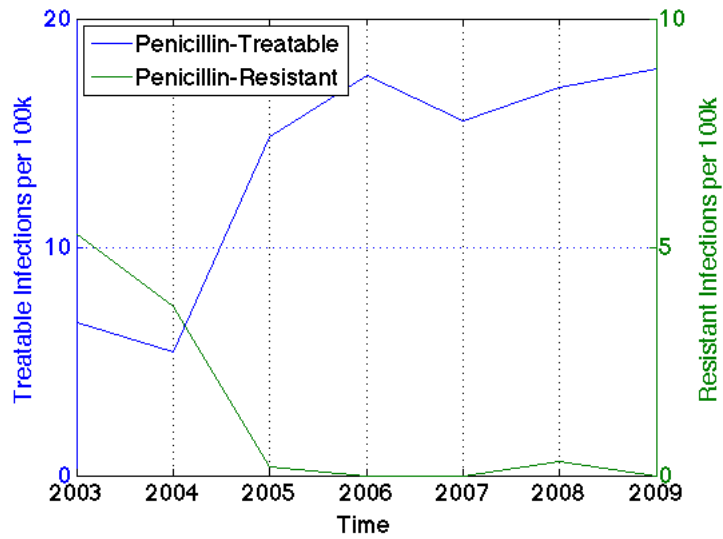

Figure 1.1: Baltimore City

Figure 1.2: Nashville-Davidson

The fully fleshed model is a variant of the resource extraction "fish war" model, where treatable infection is a scarce resource that gets extracted as providers treat patients with antibiotics. In the classic extraction models a la Levhari and Mirman (1980), the social objective is to extract the resource so as to maximize the lifetime utility from consumption. In the set-up here, the social goal is to use antibiotics to extract the treatable disease so as to minimize the lifetime disutility from total infection, taking into account the effect that extracting the treatable infection accelerates the growth of the resistant infection.

Due to the idiosyncratic utility shocks, providers sell differentiated products to patients. This allows for the model to study antibiotic provision under any number of providers.

I first study the market allocation under perfect competition. When the market is competitive, providers have a negligible effect on the evolution of the system and therefore no incentive to conserve the treatable infection. This incentivizes providers to perpetually *over-prescribe* antibiotics in order to attract as many patients in the current period as possible. While this is initially a boon for society as the level of treatable infection quickly falls, the resistant infection grows at a quicker rate than is socially optimal.

At the other end of the competitive spectrum, a patient monopolist *under-prescribes* antibiotics relative to the planner. This increases profit along two margins for the monopolist. First, the monopolist maintains a higher antibiotic effectiveness, which increases the value of the monopolist's product and, in the long run, induces more patients to pay the fee. Second, the monopolist can maintain a perpetually higher total level of infection than the social planner desires, which results in a larger pool of potential fee-payers. The paper finds conditions for the monopolist to maintain a steady-state characterized by an undesirably high level of infection.

Increasing the level of competition from monopoly to duopoly generates a steady-state with a lower level of infection. Competition for patients induces the duopolists to prescribe antibiotics more frequently than the monopolist, leading to a steady-state that yields strictly lower treatable infection and higher welfare than the monopolist's steady-state. However, as the market structure becomes increasingly competitive, antibiotics eventually become over-prescribed. Using numerical analysis, I show that social welfare as a function of the number of providers can take on an inverted U-shape (illustrated in Section 4.4). Social welfare can initially increase as the market becomes more competitive from monopoly, peak at some level of oligopolistic competition, and then decrease as the market moves towards perfect competition. Oligopolistic competition can thus be the optimal market structure.

This "Goldilocks" effect arises because fee-for-service rewards patient volume. At low levels of competition, this aspect of fee-for-service incentivizes providers to use their market power to manipulate the epidemiological states to increase the number of patients willing to pay the fee. This is accomplished by under-prescribing antibiotics to prevent the growth of the resistant infection and encourage the growth of the treatable infection.

At high levels of competition, the reward for patient volume incentives providers who lack any market power to try to attract as many patients as possible in the current period. This leads to over-prescription of antibiotics that, while it quickly limits the treatable infection, encourages faster long-run growth of the resistant infection than is desirable.

An intermediate level of competition creates a better incentive structure for providers to trade off limiting the growth of the treatable infection in the current and limiting the growth of the resistant infection in the future. This is because an intermediate level of competition prevents providers from letting treatable infection

grow as a monopolist does, but gives enough of a stake in the future that providers do not drive down the treatable infection as quickly as a perfectly competitive market does, thus slowing the growth of the resistant infection.

I then demonstrate how the model can evaluate the welfare effects of different public policies. Using numerical analysis, I compute the optimal provider licensing regime, quota on providers' prescription rates and tax on antibiotics and compare the welfare associated with each policy.

The chapter makes two primary contributions to the economics literature on antibiotic resistance (which is discussed in detail in Section 2). First, the chapter provides a novel model of imperfectly competitive provision of antibiotics. The chapter demonstrates new results about the optimality of oligopolistic competition and generalizes results previously discussed in the literature. Second, the chapter develops a framework for policy analysis.

The rest of the chapter will proceed as follows: Section 2 discusses the related literature. Section 3 presents the epidemiological model of infection and antibiotic resistance and the economic model of provider competition. Section 4 studies the market allocation of antibiotics. Section 5 studies the effect of public policy. Section 6 concludes. For ease of exposition, all proofs can be found in the appendix.

1.2 Related Literature

This section reviews the economics literature on infection and antibiotic resistance. Layton and Brown (1996) first formalize the competing externalities from antibiotic use. Laxminarayan and Brown (2001) and Laxminarayan and Weitzman (2002) study the optimal use of antibiotics. Recently, attention has turned to how markets allocate antibiotics. Elbasha (2003) estimates a static model of market provision of antibiotics

in which use only causes a negative externality. Tisdell (1982) is an early contribution of a stylized two period model in which a competitive market over-uses an antibiotic-like good relative to the planner. Herrmann and Gaudet (2009) extend the analysis to an infinite-horizon model of infection and resistance. Mechoulan (2007) uses numerical simulations to show that a monopolist can settle into a steady-state with a positive level of infection whereas a planner may prefer to use antibiotics to achieve full eradication of the disease. Finally, Herrmann (2010) studies the pricing behavior of a pharmaceutical company that produces an antibiotic that is temporarily protected by a patent after which there is open-access to the antibiotic.

Mechoulan (2007) and Herrmann and Gaudet (2009) are the closest to the present chapter. These papers study an environment with market demand for antibiotics given by a demand curve that gets supplied by a monopolist and competitive pharmaceutical companies respectively. The primary contribution of this chapter is a model that allows for the study of antibiotic provision under imperfect market competition. This is useful because, as the chapter shows, oligopolistic competition can be the optimal market structure.

Further, a theory of imperfectly competitive antibiotics provision in a dynamic setting is necessary to study the effects of public policy. While the literature has started examining policy responses to the problem of antibiotic resistance, most of the work has focused on policy in simplified settings. Elbasha (2003), for example, describes the optimal tax in a static model of market provision of antibiotics. This chapter develops a broad framework for policy analysis in a dynamic model of infection and resistance.

The chapter also generalizes results discussed in the aforementioned literature. For example, this chapter finds conditions for which a monopolist has an undesirably high

steady-state level of infection, a broad extension of the numerical results in Mechoulan (2007).

1.3 THE MODEL

The model is infinite-horizon and in discrete time. There are two components: an epidemiological model describing how infection evolves in response to antibiotic use and an economic model describing how provider and patient behavior determine antibiotic use. Each are considered in turn.

1.3.1 EPIDEMIOLOGICAL MODEL

The model of infection and antibiotic resistance is based on the *Susceptible - Infected - Susceptible* model of disease transmission (Kermack and McKendrick 1927). Related models of antibiotic resistance have appeared in the epidemiological literature such as Bonhoeffer et al (1997), Débarre et al (2009), and Porco et al (2012) as well as in the economics literature such as Herrmann and Gaudet (2009) and Laxminarayan and Brown (2001).

Resistant and non-resistant bacteria compete for resources (healthy bodies) in the ecosystem. When a patient infected with non-resistant bacteria takes the antibiotic, the patient raises his probability of recovery. The patient is then less likely to infect other healthy people. However, by removing a source of competition for the resistant bacteria, the patient's use of the antibiotics allows the resistant bacteria to grow and spread at a faster rate in the future. This is referred to as the 'natural selection' effect in the literature.

There is a society of measure one. Agents in the society cycle between being susceptible to an infection (S), being infected with a treatable infection (I^T), or being

infected with an antibiotic-resistant strain (I^R). Critically, agents can be infected with the treatable or the resistant strain, but not both. This captures the notion of 'bacterial interference' (Reid et al. 2001), a feature that has been observed in the data.[3]

Agents know if they are infected, but do not know the strain of the infection. Antibiotic effectiveness, E, is defined as the ratio of treatable disease to total disease in society:

$$E = \frac{I^T}{I^T + I^R} \tag{1.1}$$

E is the probability that, conditional on being infected, an agent is infected with the treatable strain of the infection.

Both strains of the infection are equally contagious. Let β reflect the contagiousness of the disease. Infected agents recover naturally from the infection with some probability. Let p^T denote the probability that an agent infected with the treatable disease recovers. Let p^R denote the probability that an agent infected with the resistant disease recovers. Let $\Delta p \equiv p^R - p^T$. I assume that the difference is strictly positive. A positive Δp describes what is known in the epidemiological literature as the 'fitness cost of resistance'. The intuition is that resistance comes at a cost to the bacteria that allows the treatable bacteria to outcompete the resistant bacteria "if the selective pressure from antibiotics is reduced" (Anderson and Hughes 2010).

If an agent is infected with the treatable disease and takes the antibiotic, then they have an increase in their recovery probability of p^A, for a total recovery probability of $p^T + p^A$. I assume that $p^T < p^R < p^T + p^A$. That is, agents who are infected with

---

[3]In a more elaborate epidemiological framework, patients could be colonized with both the treatable and resistant bacteria. The externalities from antibiotic use would operate through a similar albeit more complicated mechanism than the one studied here.

the treatable disease and are treated with antibiotics recover quickest while agents infected with the treatable disease and are not treated recover slowest. Agents infected with the resistant disease recover at an intermediate rate. This ordinal ranking of recovery probabilities makes antibiotic effectiveness a renewable resource in the model.

I denote the fraction of the infected population that is treated with antibiotics at time t by F_t. The law of motion describing the evolution of treatable disease is given by:

$$I_{t+1}^T = I_t^T + I_t^T[\beta S_t - p^T - p^A F_t] \tag{1.2}$$

The law of motion describing the evolution of resistant disease is similar, except that there is no benefit of taking antibiotics. This is given by:

$$I_{t+1}^R = I_t^R + I_t^R[\beta S_t - p^R] \tag{1.3}$$

For simplicity, I impose a "no-death" condition.[4] Because the population measure stays constant, the change in the susceptible population is simply the inverse of the change in the infected population:

$$S_{t+1} = S_t - (I_{t+1}^T - I_t^T) - (I_{t+1}^R - I_t^R) \tag{1.4}$$

The mechanism through which antibiotic use affects the evolution of the system can be understood through the following sequence of diagrams:

---

[4]The paper studies infinitely-lived agents. However, the model can be extended to one with finitely lived agents that die due to the infection.
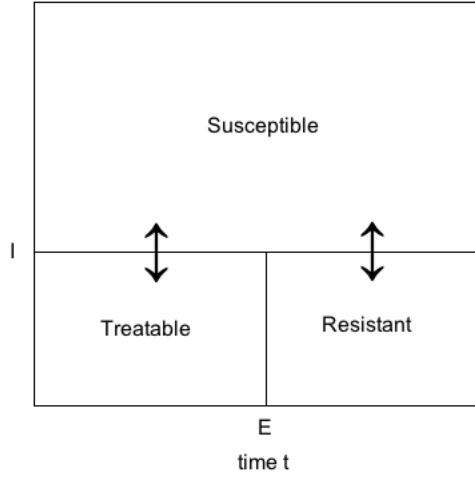
Figure 1.3: Initial state

The relative size of each segment in the box represents the proportion of society in that state. There is some natural flow from the susceptible population into the infected states and vice versa. Antibiotic use at time t increases the flow of agents moving from being infected with the treatable strain to being susceptible in the following period (captured by an increase S_{t+1}).
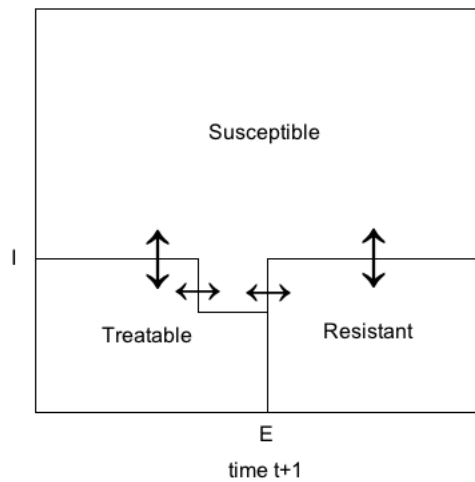
Figure 1.4: Antibiotic use clears the treatable infection
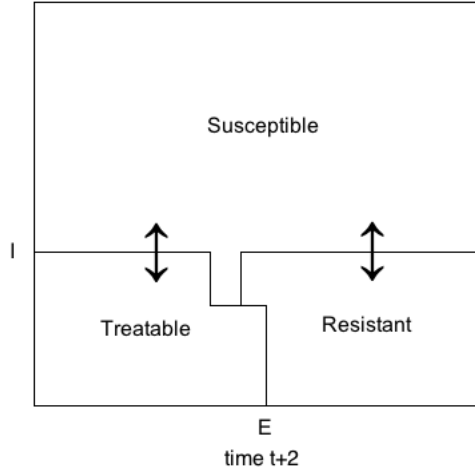
Figure 1.5: The resistant infection has more room to grow

However, because there is now a larger susceptible population at $t+1$, the resistant infection will grow at a faster rate than if there was no antibiotic treatment. This is captured by an increase in I_{t+2}^R.[5]

The dynamic externalities can be understood as follows: patients who are treatable and take the antibiotic recover quicker, which means they infect less of the susceptible population with the treatable strain (the positive externality). However, the patients that recovered are now susceptible to reinfection by the resistant strain, which accelerates its growth (the negative externality). This notion that patients who take antibiotics to clear a treatable infection can quickly be reinfected with resistant bacteria is observed in Kuster et al 2014, who find that patients who recently received a course of antibiotics were more likely to be infected with a resistant strain

---

[5]Epidemiological models frequently include some period of immunity to becoming infected again after recovery. In the set-up here, agents are immediately susceptible to the disease again after recovery. Were the immune response to be included, the mechanism through which antibiotics affects the system would fundamentally remain the same, only with a time lag.

of *Streptoccocus Pneumoniae* than patients who had not recently received a course of antibiotics.

Combining the definition of antibiotic effectiveness with the laws of motion governing susceptibility and infection, the dynamics of the system can succinctly be written as:

$$I_{t+1} = I_t + I_t[\beta(1 - I_t) - p^R + E_t(\Delta p - p^A F_t)] \tag{1.5}$$

and

$$E_{t+1} = E_t + \frac{E_t(1 - E_t)(\Delta p - p^A F_t)}{1 + \beta(1 - I_t) - p^R + E_t(\Delta p - p^A F_t)} \tag{1.6}$$

Equation (5) describes the evolution of total infection. Equation (6) describes the evolution of antibiotic effectiveness - the proportion of treatable to total infection. Notice that for $F_t > \frac{\Delta p}{p^A}$, antibiotic effectiveness is decreasing. Effectiveness decreases at high levels of antibiotic use because sufficiently high antibiotic use drives out the treatable infection at a quicker rate than the resistant infection (the natural selection effect). For $F_t < \frac{\Delta p}{p^A}$, effectiveness is increasing. When there is sufficiently low antibiotic use, patients with the resistant strain recover quicker on average than patients with the treatable strain (the fitness cost effect), leading to an increase in antibiotic effectiveness. At $F_t = \frac{\Delta p}{p^A}$, the two effects exactly offset and effectiveness stays constant.

A steady-state is a fixed point of Equations (5) and (6) - the laws of motion governing infection and antibiotic effectiveness. For a constant treatment rate F, the dynamic system tends to one of three possible steady-states described below.

(1) For $F > \frac{\Delta p}{p^A}$, the treatable strain clears the system at a faster rate than the resistant strain due to the natural selection effect. The system tends to a corner

14

steady-state where the treatable strain goes extinct and only the resistant strain remains. This steady-state is reached asymptotically:

$$(\bar{I}, \bar{E}) = (\frac{\beta - p^R}{\beta}, 0) \tag{1.7}$$

(2) For $F < \frac{\Delta p}{p^A}$, the resistant strain clears the system faster than the treatable strain due to the fitness cost effect. The system tends to a corner steady-state where the resistant strain goes extinct and only the treatable strain remains. This steady-state is also reached asymptotically. Note that this steady-state has a higher level of infection than the other two steady-states:

$$(\bar{I}, \bar{E}) = (\frac{\beta - p^T - p^A F}{\beta}, 1) \tag{1.8}$$

(3) For $F = \frac{\Delta p}{p^A}$, the fitness cost effect and the natural selection effect exactly offset. The system maintains an interior steady-state level of antibiotic effectiveness:

$$(\bar{I}, \bar{E}) = (\frac{\beta - p^R}{\beta}, E) \text{ for } E \in (0, 1) \tag{1.9}$$

When applicable, the chapter characterizes the steady-states of economic actors. However, the time horizon of the chapter is the *human time-scale* rather than the *evolutionary time-scale*, and so, particularly in the case of asymptotic steady-states, the analysis of the steady-state is intended largely as a point of reference. The 'action' in the chapter occurs along the transition path.

1.3.2 ECONOMIC COMPONENT

I now introduce an economy with a provider-patient matching process and a fee-for-service healthcare system. Infected patients search for healthcare providers, who are the gatekeepers of antibiotics. Patients gain utility from a provider's prescription

rate, the current effectiveness of the antibiotic (higher antibiotic quality increases the utility that agents get from providers), and some idiosyncratic component related to the quality of their match with a provider. Patients choose whether or not to pay a fee to see their most preferred medical provider, who in turn chooses whether or not to prescribe antibiotics. Providers prescribe antibiotics to maximize their profits given the behavior of the other providers, patients, and the laws of motion governing the system.

Whether or not providers maximize profits is a debated issue in the literature. While financial incentives affect providers' treatment behavior (Clemens and Gottlieb 2014), many models of the provider-patient interaction study environments in which providers maximize a convex combination of profits and patient welfare (see Chonè and Ma 2011 or Jacobson et al 2013 for two recent examples). It would be possible to build this feature or an ethical constraint (e.g. the Hippocratic oath) into the model, but in order to both make the model more tractable and not impose structure that could assume the problem away, I refrain from doing so.

Patients are risk neutral. Each patient knows if he is infected, but does not know if he is infected with the treatable or resistant version of the disease. Patients and providers have common knowledge of the states I and E. Utility is comprised of a component related to a patient's health status and a component related to the medical care they may purchase. The health status portion of utility is given by:

$$
\begin{cases}
U & \text{if healthy} \\
[E_t p^T + (1 - E_t) p^R]U & \text{if infected}
\end{cases}
\tag{1.10}
$$

Infected patients can choose to be seen by a provider who may possibly prescribe antibiotics. I assume that providers cannot diagnose whether patients have the treatable or resistant version of the disease and that their prescription rates are public

16

knowledge. Patients pay the provider a fixed price whether or not an antibiotic is prescribed. The price is taken to be exogenous, but can be imagined to be set by an insurance company or government. The exogoneity of the price reflects the fact that providers often have limited price-setting capacity and instead compete based on the quantity of services offered. I normalize the price to 1. Providers incur a constant marginal cost of c in administering antibiotics.

Providers pick a prescription rate, which from a patient's perspective is the probability that the provider will prescribe him antibiotics. Patients randomly draw idiosyncratic utility for each provider at time 0. This can be thought of as the degree to which a healthcare provider's non-medical features, e.g. bed-side manner, distance, etc. agree with a patient. I introduce this match-specific utility component so that providers sell differentiated products. Patient m's idiosyncratic utility for provider i is given by:

$$\varepsilon_m^i \sim U[0,1] \tag{1.11}$$

If f^i is provider i's prescription rate, then the net utility that patient m receives from seeing provider i is:

$$U_m^i(f^i, \varepsilon_m^i, E) = f^i E p^A U + \varepsilon_m^i - 1 \tag{1.12}$$

The first term captures the utility gained from healthcare, the second the idiosyncratic utility from provider i, and the third the cost of the visit. I assume three additional conditions. First, that $p^A U \leq 1$. This assumption ensures that demand for provider care is elastic over the entire range of admissible values in the model. Second, that $U > (p^T + p^A)U - c + 1$. Third, that $U > p^R U + 1$. These latter two assumptions ensure that from a social welfare perspective, being healthy is always preferred to being infected and receiving care from a provider.

1.3.3 Social Planner's Problem

To provide a benchmark for future analysis, I now describe the social planner's problem. The social planner treats a fraction of the infected population with antibiotics to solve the problem:

$$\max_{\{F_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \delta^t \left[(1 - I_t)U + I_t \left([E_t p^T + (1 - E_t)p^R + F_t E_t p^A]U + 1 \right) - cF_t I_t \right] \quad (1.13)$$

such that (5), (6), I_0, E_0, $F_t \in [0, 1]$

The discount rate is $\delta < 1$. The first term is the utility of healthy agents, the second is the utility of infected patients, and the third is the cost of treating patients with antibiotics. Note that the utility of infected patients is based on the expected recovery probability given the planner's prescription rate and the upper bound on the support of idiosyncratic utility.[6] Equations (5) and (6) are the laws of motion describing the evolution of infection and antibiotic effectiveness. I_0 and E_0 are the initial levels of infection and effectiveness.

The planner's objective is to minimize the disutility from total infection, taking into account the cost of antibiotic use and the effect of treatment on the growth resistant infection. The planner's solution can be characterized recursively through the Bellman equation:

$$V^{SP}(I, E) = \max_{F} \left\{ (1-I)U + I\left([Ep^T + (1-E)p^R + FEp^A]U + 1 \right) - cFI + \delta V^{SP}(I', E') \right\}$$

$$(1.14)$$

such that (5), (6), $F \in [0, 1]$

For a positive F, the planner has the first-order condition:

---

[6]I assume that the planner is able to give the best possible idiosyncratic provider service to each infected patient.

$$Ep^A U I - cI + \delta \left[\frac{\partial V^{SP}(I', E')}{\partial I'} \frac{\partial I'}{\partial F} |_{F=F^*} + \frac{\partial V^{SP}(I', E')}{\partial E'} \frac{\partial E'}{\partial F} |_{F=F^*} \right] \geq 0 \qquad (1.15)$$

The first-order condition reconciles the benefit of treating patients with antibiotics today and the benefit from the future decrease in infection due to treatment today with the cost of treating agents with antibiotics today and the cost of lower effectiveness in the future. Let $F^{SP}(I, E)$ refer to the planner's prescription rate when the states are are (I, E).

I use value function iteration to numerically investigate the planner's problem. I use the parameter values $\delta = .9, U = 2, \beta = .4, p^R = .25, p^T = .1, p^A = .5$, and $c = .9$ for the numerical analysis of the paper.

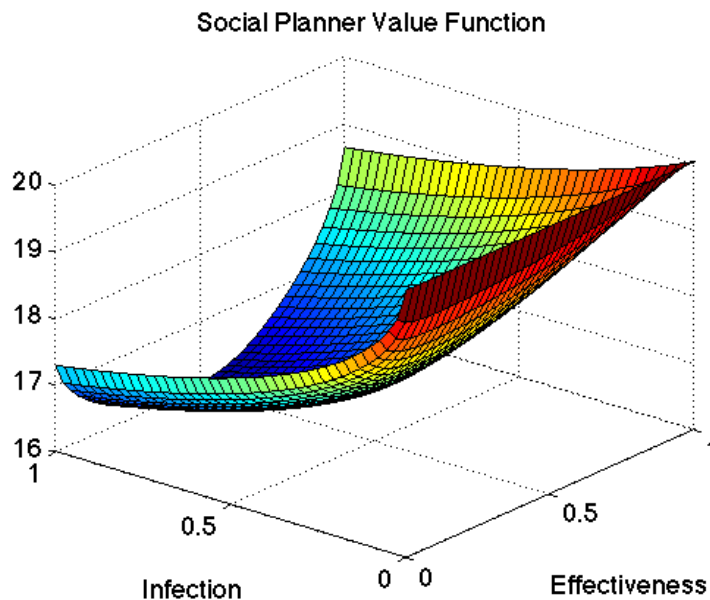The planner's value function is displayed in Figure 1.6.

Figure 1.6: Planner's value function

The value function is convex in both infection and effectiveness. Convexity in infection is due to the non-linear spread of infection. Recall that new infections are given

19

by $I\beta(1-I)$. Note that the rate of increase is decreasing in I. Intuitively, increasing infection generates a larger negative externality when there is a large susceptible population than when there is a small susceptible population. Hence, while the planner's value function is decreasing in infection, it decreases at a less-than-linear rate.

The non-monotonicity in effectiveness stems from the cost of using antibiotics. As an extreme example, consider the case in which effectiveness is zero, i.e. every infection is resistant. Infected patient recovers in each period with probability p^R. Now suppose that effectiveness E is in the interior. For a low enough effectiveness, using antibiotics is not cost-justified. Initially, the average recovery probability is $(1-E)p^R + Ep^T < p^R$. Therefore, initially, patients are infected for on average a longer time than when effectiveness is zero. Due to this effect, the planner's value function initially decreases in effectiveness. However, as effectiveness increases, prescribing antibiotics eventually becomes worthwhile. Welfare begins to increase as the average recovery time decreases and the spread of infection is lessened.

The value function can be used to compute the optimal prescription rate, which is displayed in Figure 1.7.

The optimal prescription rate tends to be either zero or one. Note that at low levels of antibiotic effectiveness, the optimal prescription rate is one when infection is low but zero when infection is high. When infection is lower, the planner faces a smaller marginal cost of antibiotic treatment and a higher marginal future benefit from the decrease in infection than when infection is higher. The planner is therefore willing to use antibiotics even when they are ineffective if the level of infection is low, but not when the level of infection is high.

The policy function can be used to compute the dynamic path of infection and effectiveness in Figure 1.8. For initial values, I use $I_0 = .6$ and $E_0 = .92$.
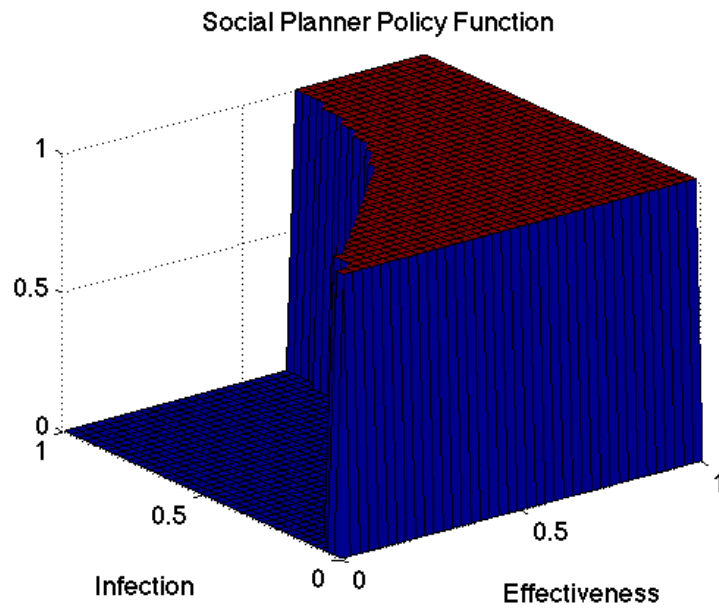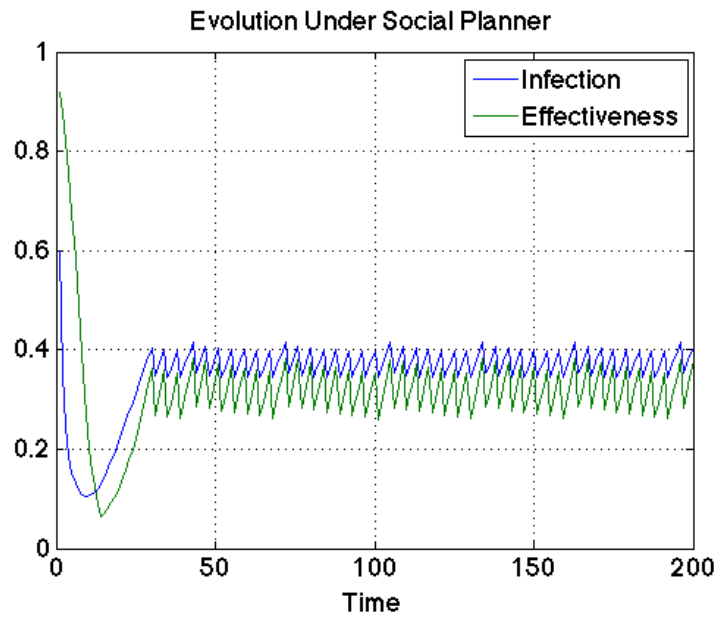
Figure 1.7: Optimal prescription rate

Figure 1.8: Planner's dynamic path

21

Note that the planner does not drive out the treatable strain entirely (because of the crowding out effect it has on the resistant strain), nor does the planner drive the resistant strain out entirely (because doing so would generate perpetually higher level of infection). Rather than settle into a steady-state, the planner cycles antibiotic use. This stems from the convexity of the planner's value function and the renewability of antibiotic effectiveness. Given the parameter values, the level of infection in an interior steady-state is $\frac{\beta - p^R}{\beta} = .375$. By cycling antibiotic use rather than prescribing at the constant steady-state fraction $\frac{\Delta p}{p^A}$, the planner is able to induce a level of infection that is on average lower than .375 at a cost that is on average lower than the cost of maintaining the steady-state.

1.3.4 MARKET PAYOFFS

If there are n total providers, then an infected agent m is willing to pay the fee to see provider i if:

$$U_m^i(f^i, \varepsilon_m^i, E) \geq U_m^j(f^j, \varepsilon_m^j, E) \quad j = 0, ..., i-1, i+1, ..., n \qquad (1.16)$$

where $j = 0$ denotes the outside alternative of not seeing a provider and $j > 0$ denotes provider i's competitors. This condition means that m is willing to see provider i if provider i is preferred to all other providers and provider i is preferred to not seeking treatment.

I denote provider i's market share at time t by $\Omega(f_t^i, f_t^{-i}, E_t)$, where f_t^i is provider i's prescription rate at time t and f_t^{-i} is a vector of the other providers' prescription rates at time t. Market share can be written as:

$$\Omega(f_t^i, f_t^{-i}, E_t) = Pr\left(U_m^i(f_t^i, \varepsilon_m^i, E_t) \geq U_m^j(f_t^j, \varepsilon_m^j, E_t) \ \ j = 0, ..., i-1, i+1, ..., n\right)$$

$$(1.17)$$

Provider i's per period payoff is:

$$[1 - cf_t^i]\Omega(f_t^i, f_t^{-i}, E_t)I_t \tag{1.18}$$

Provider and patient behavior determines an aggregate prescription rate:

$$F_t = \sum_{i=1}^{n} f_t^i \Omega(f_t^i, f_t^{-i}, E_t) \tag{1.19}$$

which can be inserted into Equations (5) and (6) to determine how infection and antibiotic effectiveness evolve. Critically, both provider prescription behavior and patient matching behavior determine the aggregate prescription rate.

Given a sequence of prescription rates, the lifetime profits of provider i are:

$$\sum_{t=0}^{\infty} \delta^t \left[1 - cf_t^i\right]\Omega(f_t^i, f_t^{-i}, E_t)I_t \tag{1.20}$$

such that $(5), (6), F_t = \sum_{i=1}^{n} f_t^i \Omega(f_t^i, f_t^{-i}, E_t), I_0, E_0$

The chapter studies symmetric Markov Perfect Equilibrium. Focusing on symmetric equilibrium makes the analysis more tractable and yields intuitive insights. In equilibrium, for example, patients will either match with the provider for whom they have drawn the highest idiosyncratic utility for or they will not pay the fee to any provider. Patient choice of provider in the model is therefore sticky, a feature that has been empirically observed in previous studies of patient-provider choice (Mold et al 2004).

A Markov strategy for provider i is a mapping:

23

$$\sigma^i : [0,1]^2 \rightarrow [0,1] \tag{1.21}$$

where $\sigma^i(I, E) = f^i$. The strategy function takes the level of infection and effectiveness as inputs and gives a prescription rate as an output. Given a strategy function, lifetime profits for provider i can be written recursively using the value function:

$$V^i(I, E; \sigma) = [1 - c\sigma^i(I, E)]\Omega\Big(\sigma^i(I, E), \sigma^{-i}(I, E), E\Big)I + \delta V^i(I', E'; \sigma) \tag{1.22}$$

such that $(5), (6), F = \sum_{i=1}^n \sigma^i(I, E)\Omega(\sigma^i(I, E), \sigma^{-i}(I, E), E)$

A Markov strategy profile $\sigma = (\sigma^1, \sigma^2, ..., \sigma^n)$ constitutes a Markov Perfect Equilibrium if σ constitutes a sub-game perfect equilibrium in Markov Strategies. That is, σ is a Markov Perfect Equilibrium if for all $i = 1, ..., n$:

$$V^i(I, E; \sigma) \geq V^i(I, E; \hat{\sigma}^i, \sigma^{-i}), \forall(I, E), \forall \hat{\sigma}^i \text{ Markov strategies} \tag{1.23}$$

The equilibrium is symmetric if every provider plays the same strategy function. The Markov Perfect equilibrium can be characterized as the solution to the system of Bellman equations. Provider i's Bellman equation is:

$$V^i(I, E; \sigma) = \max_f \left\{ [1 - cf]\Omega(f^i, f^{-i}, E)I + \delta V^i(I', E'; \sigma) \right\} \tag{1.24}$$

such that $(5), (6), \quad F = \sum_{j=1}^n f^j \Omega(f^j, f^{-j}, E), \quad f \in [0,1]$

The first-order condition for a positive f^* in the symmetric equilibrium is:

$$[1 - cf^*]\frac{\partial\Omega(f^i, f^{-i*}, E)}{\partial f^i}\Big|_{f^i=f^*}I - c\Omega(f^*, f^{-i*}, E)I$$
$$+ \delta\frac{\partial V^i(I', E'; \sigma)}{\partial I'}\frac{\partial I'}{\partial f^i}\Big|_{f^i=f^*} + \delta\frac{\partial V^i(I', E'; \sigma)}{\partial E'}\frac{\partial E'}{\partial f^i}\Big|_{f^i=f^*} \geq 0 \tag{1.25}$$

24

The first-order condition reconciles the marginal benefit from increasing the prescription rate and attracting more patients today with the cost of using more antibiotics and the future effects from decreasing infection and effectiveness. The equilibrium first-order condition is derived in Appendix B.

1.4 MARKET PROVISION

This section studies the market allocation of antibiotics. I first study the limit case as the number of providers, n, tends towards infinity, a perfectly competitive market. I then study the case in which $n = 1$, monopoly, before analyzing the model under arbitrary n, oligopoly.

1.4.1 PERFECT COMPETITION

The equilibrium under perfect competition can be computed analytically. Let $F^{PC}(I, E)$ denote the aggregate prescription rate under perfect competition when the states are (I, E).

Theorem 1

1. The symmetric equilibrium outcome under perfect competition is for every provider to set $\sigma^i(I, E) = 1$ for all (I, E).

2. $F^{PC}(I, E) \geq F^{SP}(I, E)$ for all (I, E) and strictly greater for some $(\tilde{I}, \tilde{E})$.

3. The steady-state level of antibiotic effectiveness is $\bar{E}_{PC} = 0$.

4. The steady-state level of infection is $\bar{I}_{PC} = \frac{\beta - p^R}{\beta}$.

Competition with other providers for patients induces providers to prescribe antibiotics at rate one in all periods, i.e. every patient that the provider sees is prescribed antibiotics. Intuitively, when other providers prescribe at rate one, provider

i has no incentive to deviate because he will not gain any patients in the current period and, due to the behavior of others, his own behavior has a negligible effect on the evolution of the system. Under perfect competition, providers have no incentive to preserve antibiotic effectiveness for future use, and so they extract the treatable infection as quickly as the market will allow.

This generates perpetual over-use of antibiotics by a competitive market relative to the planner and leads to a steady-state in which the treatable strain is driven out in the limit. Contributing to this strong over-use result is the assumption that the upper bound on the support of idiosyncratic utility is the same as the provider fee (both equal to 1). This assumption means that the idiosyncratic utility patients can gain from providers is sufficiently high that patients are willing to purchase healthcare from their most preferred provider in a competitive setting along the entire dynamic path as effectiveness tends to zero. While it simplifies the remaining analysis, the key takeaways from the paper do not explicitly depend on this assumption.

I display the dynamic path of infection and effectiveness under perfect competition in Figure 1.9. The competitive market extracts the treatable infection as quickly as possible. This causes total infection to initially fall, but then rebound towards the steady-state level as the resistant infection quickly grows to fill the void left from antibiotic use.
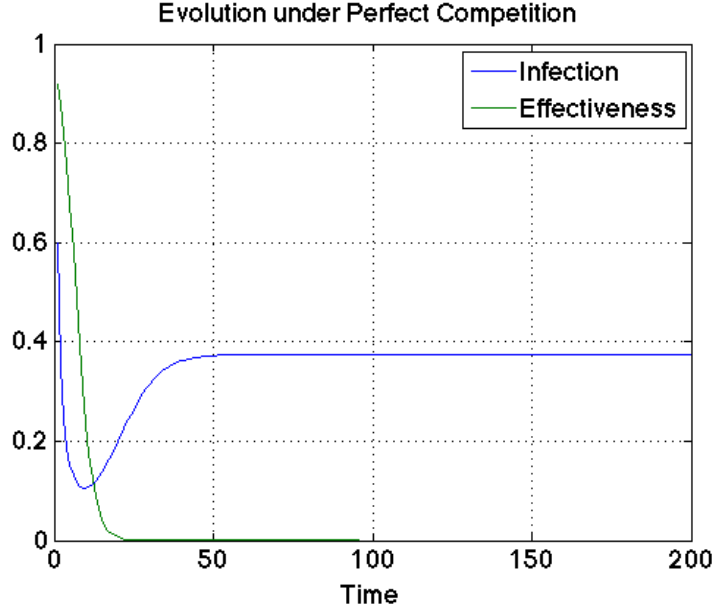
Figure 1.9: Perfect competition's dynamic path

1.4.2 Monopolist

In contrast to provision under perfect competition, under monopoly, the provider completely controls the flow of antibiotics to patients. The monopolist therefore has a large effect on the evolution of the system which he fully internalizes as he chooses his prescription rate. Before characterizing the monopolist's steady-state, I introduce the following two assumptions.

First, that $\frac{\Delta p}{p^A} < p^A U$. If this assumption did not hold, then regardless of the monopolist's prescription behavior, the resistant infection would clear the system at a faster rate than the treatable infection and effectiveness would tend towards one.[7] I use this assumption to ensure that, if such behavior arises, then it is due in part to economic considerations by the monopolist.

---

[7]The monopolist's market share is $Pr\big(fEp^A U + \varepsilon - 1 \geq 0\big) = fEp^A U$ and so the aggregate prescription rate is $F = f^2 Ep^A U$. If $\frac{\Delta p}{p^A} > p^A U$, then the aggregate prescription rate is always $F < \frac{\Delta p}{p^A}$ and effectiveness tends to one.

27

Second, that:

$$\left[1 - 2c\left(\frac{\Delta p}{p^{A2}U}\right)^{\frac{1}{2}}\right] - \frac{\left[1 - c\left(\frac{\Delta p}{p^{A2}U}\right)^{\frac{1}{2}}\right]}{\beta - p^R} 2\Delta p < 0 \tag{A1}$$

I refer to this assumption in later text as Assumption $A1$.

Theorem 2

1. The monopolist's steady-state level of effectiveness $\bar{E}_M \to 1$ as $\delta \to 1$.

2. A sufficiently patient monopolist has a steady-state level of infection $\bar{I}_M > \frac{\beta - p^R}{\beta}$.

The patient monopolist has a steady-state with an inefficiently high level of infection: any steady-state with infection I such that $\bar{I}_M > I > \frac{\beta - p^R}{\beta}$ generates strictly higher welfare than the monopolist's steady-state.

Three reasons compel a monopolist provider to under-prescribe antibiotics. These effects can be understood through the monopolist's first-order condition:

$$\underbrace{[1 - 2f_M^* c]Ep^A UI}_{\text{positive marginal payoff requirement}} + \delta\left[\overbrace{\frac{\partial V^M(I', E'; \sigma)}{\partial I'}}^{\text{infection a good}} \frac{\partial I'}{\partial f}\Big|_{f=f_M^*} + \underbrace{\frac{\partial V^M(I', E', \sigma)}{\partial E'}}_{\text{demand-inducement effect}} \frac{\partial E'}{\partial f}\Big|_{f=f_M^*}\right] \geq 0 \tag{1.26}$$

First, the monopolist has a positive current period marginal payoff requirement. When the monopolist increases his prescription rate, he has a marginal benefit from attracting new patients, but bears a loss from increasing costly treatment on patients already purchasing his services.[8] As elaborated on below, the latter two terms in the monopolist's first-order condition are negative. This means that in order to satisfy

---

[8]This effect is analogous to the standard case in which a price-setting monopolist lowers his price to attract more consumers but takes a loss on the consumers already purchasing his product. As in the case of the standard monopolist, this effect can cause under-provision.

28

the first-order condition, the monopolist requires that his steady-state current period marginal payoff is positive. This implies that:

$$\frac{1}{2c} \geq f_M^* \tag{1.27}$$

This positive current period marginal payoff requirement puts an upper bound on the monopolist's prescription rate. If the upper bound is low enough, then the aggregate steady-state prescription rate will be less than the critical threshold $\frac{\Delta p}{p^A}$, which implies a steady-state antibiotic effectiveness of one (note this also implies a steady-state with a strictly higher level of infection than is desirable).

Second, the monopolist's steady-state payoff is strictly increasing in antibiotic effectiveness. This is because of a demand-inducement effect: at higher levels of antibiotic effectiveness, more patients are willing to pay the fee to see the monopolist. Increasing his prescription rate to attract more patients today causes a future loss to the monopolist by lowering the willingness of future patients to purchase his services. As the monopolist becomes patient, his steady-state level of antibiotic effectiveness tends towards one because of this effect. Interestingly, although more patients are willing to pay the monopolist when the steady-state level of antibiotic efficacy is one, patients are not, on net, better off because of the higher antibiotic efficacy. This is because the monopolist's under-use of antibiotics causes patients to be sick in expectation for a longer time.

Third, under fee-for-service, infection is a good for a healthcare provider. Increasing his prescription rate to attract more patients today causes a future loss to the monopolist by lowering the number of infected patients and hence the size of the monopolist's future market. A monopolist can under-prescribe antibiotics so that there is a higher level of infection - resulting in a greater amount of potential fee-payers - than is socially optimal.
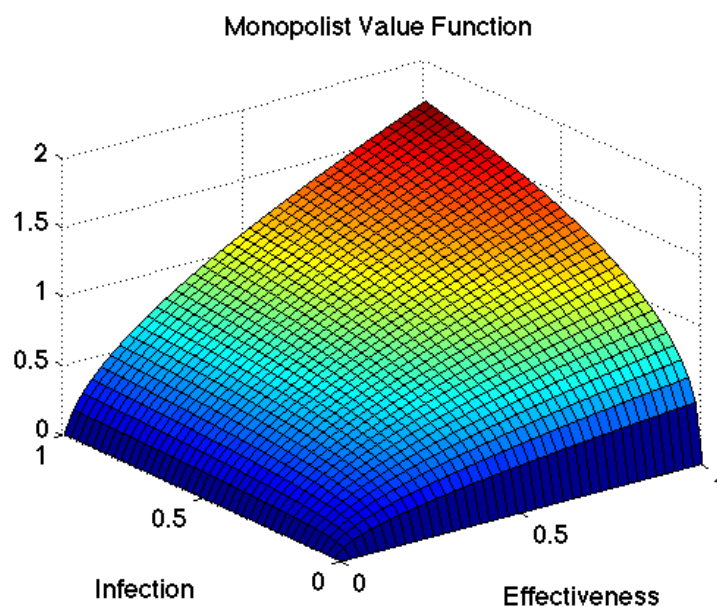
Figure 1.10: Monopolist's value function

These distortions arise because of the linear payment structure of fee-for-service. By tying revenue to the volume of patients, fee-for-service incentivizes the monopolist to use his market power to manipulate the epidemiological states so as to increase the total number of patients willing to pay the fee.

I use value function iteration to study the behavior of the monopolist outside of the steady-state. The monopolist's value function is displayed in Figure 1.10.

Note that the value function is strictly increasing in both infection and antibiotic effectiveness. This reflects the fact that fee-for-service makes infection and antibiotic effectiveness 'goods' for the monopolist. The value function can be used to derive the monopolist's optimal policy rule, which is displayed in Figure 1.11.

Note that this is not the aggregate prescription rate, but rather the rate that the monopolist prescribes antibiotics to patients who pay the fee. The policy rule can

Figure 1.11: Monopolist's prescription rate

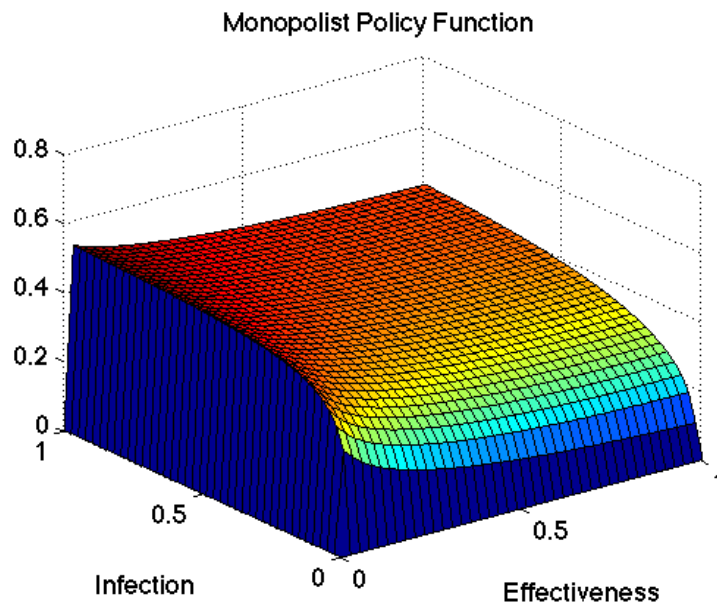be used to simulate the evolution of infection and antibiotic effectiveness under the monopolist, which is displayed in Figure 1.12. The monopolist uses his market power to let infection and effectiveness grow towards a corner steady-state with a higher level of infection than is socially optimal.
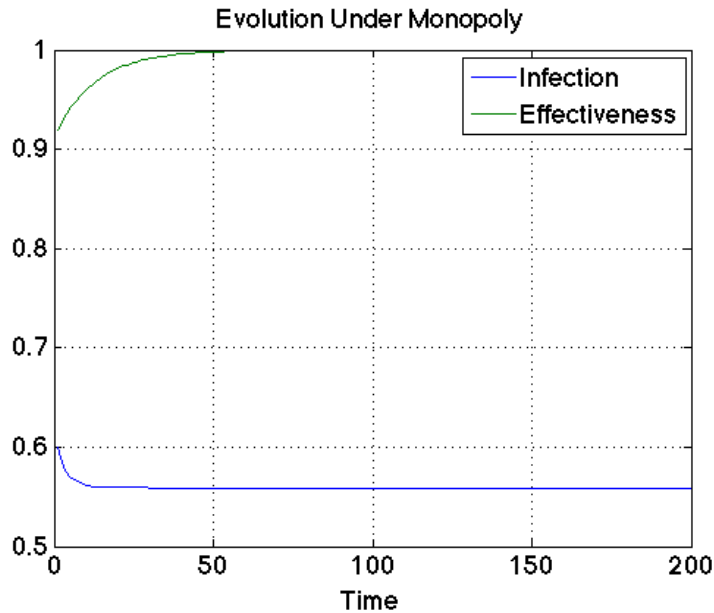
Figure 1.12: Monopolist's dynamic path

1.4.3 OLIGOPOLY

I now turn to imperfectly competitive provision of antibiotics. Critically, unlike in the case of monopoly or perfectly competitive provision, providers now act strategically. The strategic interaction affects providers in several ways. First, whereas a patient would pay the fee to the monopolist if his individual rationality constraint was met, a patient will only pay the fee to the oligopolist if his individual rationality constraint is met and the oligopolist is preferred to all other providers. The oligopolist's prescription behavior thus affects his competitors' market share (and similarly, the oligopolist's market share is affected by his competitors' prescription behavior). Second, whereas the monopolist fully internalizes the effects of his prescription behavior, the oligopolist only partially internalizes the effects of his prescription behavior. The oligopolist's prescription behavior affects his competitors by changing the levels of infection and effectiveness in the future (and similarly, the oligopolist is

affected by his competitors in this fashion). The chapter's next result contrasts the steady-state conditions of the monopolist and duopolist.

Theorem 3

A sufficiently patient duopolist has a strictly lower steady-state level of infection than the equally patient monopolist, i.e. $\bar{I}_D < \bar{I}_M$.

When there is competitive pressure in the market, the higher levels of infection that a monopolist can maintain in the steady-state are unsustainable. The presence of a competitor lowers an individual provider's benefit from withholding antibiotic treatment to generate more favorable states in the future. This forces more weight on attracting patients in the current period, which leads to higher antibiotic use and a welfare gain for society.

I use numerical analysis to compute the Markov Perfect equilibrium. The numerical algorithm used is described below.

Computational Algorithm

Step 1. Start with an initial value function $v^0(I, E) = 0$.

Step 2. Plug the proposed valued function into the oligopolist's first-order condition. Solve for the optimal symmetric prescription rate f^*.

Step 3. Use f^* to update the value function. Specifically, $v^{k+1}(I, E) = [1 - cf^*]\Omega(f^*, f^{-i*}, E) + \delta v^k(I', E')$.

Step 4. Iterate until the difference between $v^{k+1}(I, E)$ and $v^k(I, E)$ becomes small.

The algorithm computes the Markov Perfect equilibrium of a finite horizon game and takes the limit as the time horizon tends to infinity. I display the duopolist's symmetric equilibrium value function in Figure 1.13.

33
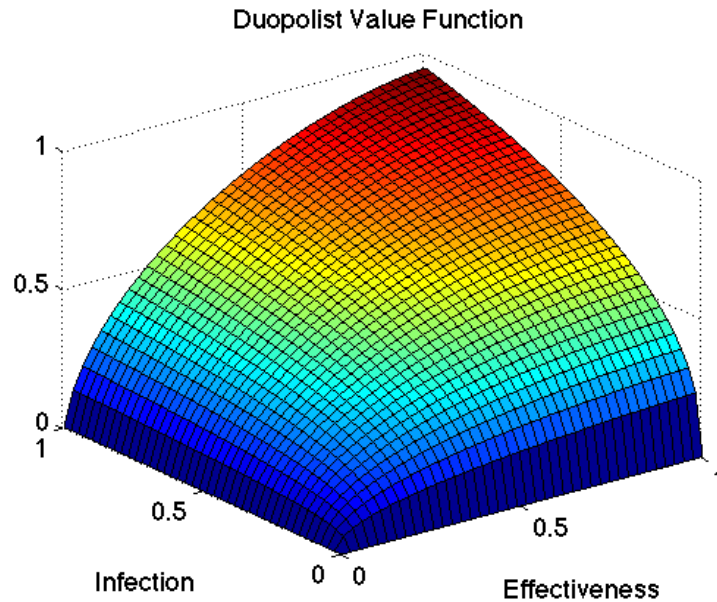
Figure 1.13: Duopolist's value function

The value function can be used to derive the policy function, displayed in Figure 1.14. In the simulation, the duopolist prescribes at a strictly higher rate than the monopolist for all values of infection and antibiotic effectiveness. The presence of another provider cuts into a provider's market share, as well as diminishes the marginal effect of one's own prescription behavior on the future levels of infection and resistance relative to monopoly provision. These factors encourage greater antibiotic use under duopoly than monopoly. The policy rule can be used to simulate the dynamic path of infection and effectiveness under the duopolist, which is shown in Figure 1.15

Notice that the duopolist has a strictly lower level of infection and effectiveness than the monopolist along the entire dynamic path. While fee-for-service still incentives infection as a good for the duopolist, competition for patients with the other provider prevents a provider from letting infection grow as a monopolist would prefer.
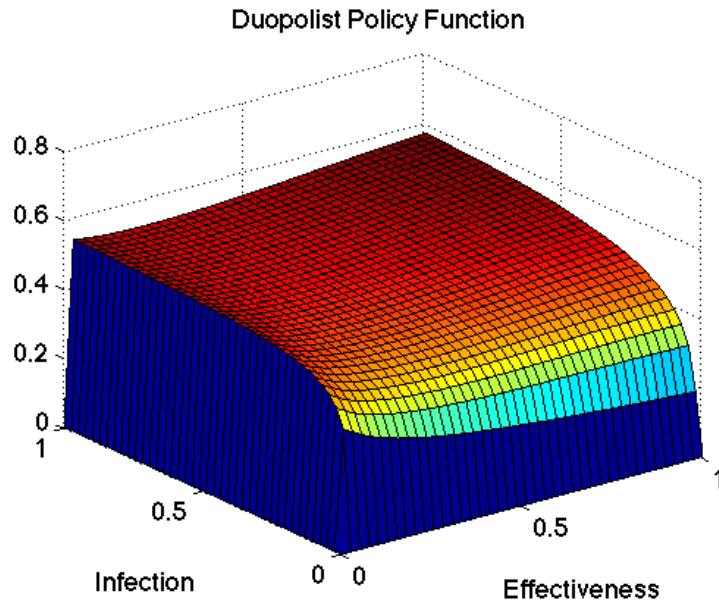
Figure 1.14: Duopolist's prescription rate
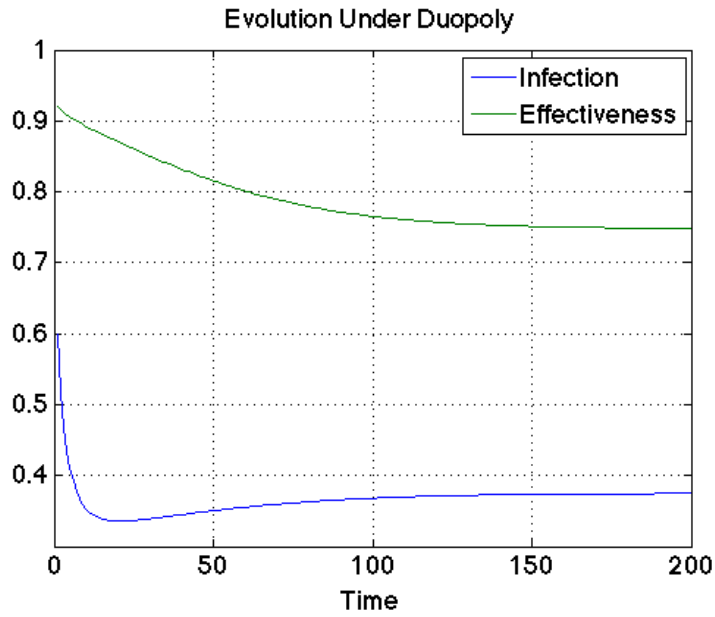
Figure 1.15: Duopolist's dynamic path
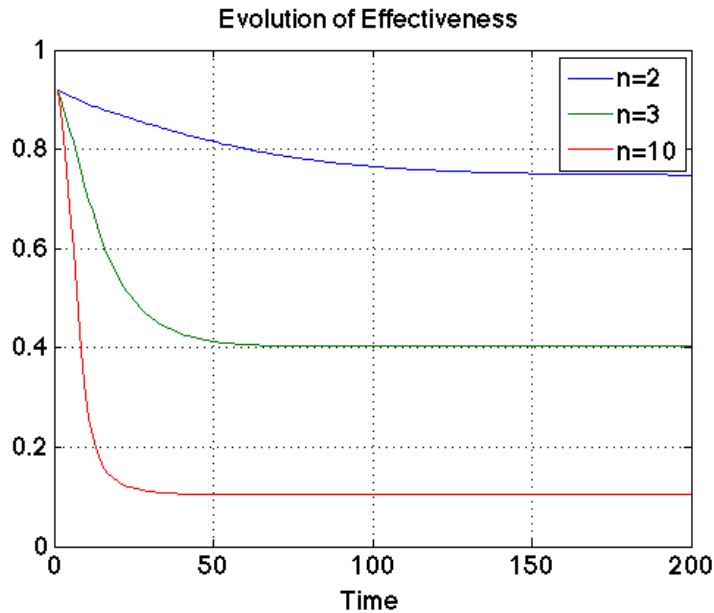
Figure 1.16: Effectiveness under different market structures

Intuitively, increasing the level of competition increases antibiotic use as providers focus less on maintaining favorable states in the future and more on attracting patients in the present. Figure 1.16 shows the evolution of antibiotic effectiveness under increasing levels of competition.

When there is more competition, antibiotic effectiveness decreases at a quicker rate and reaches a lower steady-state level than when there is less competition. This is because higher antibiotic use in more competitive markets results in quicker and more prolonged extraction of the treatable infection. The effect of competition on total infection is shown in Figure 1.17.

Due to the increased antibiotic use, more competitive market structures generate a quicker initial decrease in total infection than less competitive markets as the treatable strain is cleared from the system. However, this also allows the resistant strain to grow

36
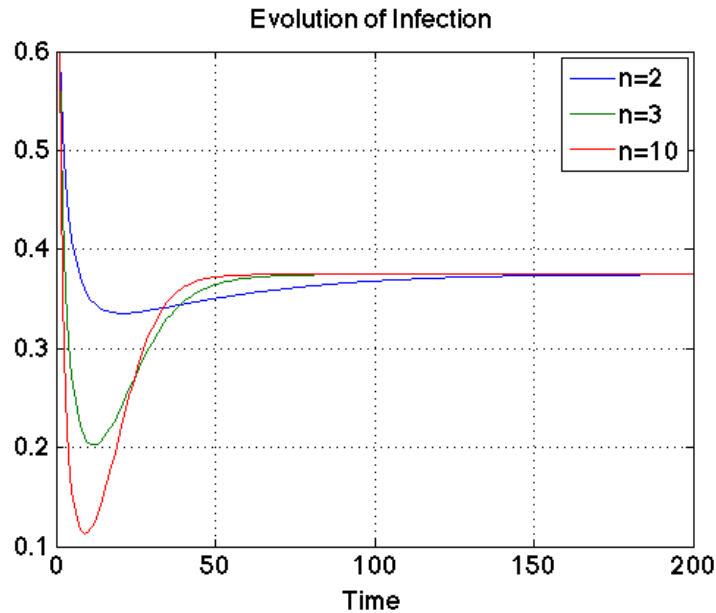
Figure 1.17: Infection under different market structures

at a faster rate. In fact, the resistant strain can grow sufficiently faster that the total level of infection in more competitive markets can eventually exceed the total level of infection in less competitive markets. This effect can be seen on the graph above by following the dynamic path of infection for $n = 10$.

Note that all the market structures illustrated converge to the same steady-state level of infection, i.e. they all implement the same aggregate prescription rate in the steady-state. To see how this can occur, suppose that the market is in the steady-state and the level of competition is increased. Increased competition induces individual providers to prescribe at a higher prescription rate. However, this drives down the level of antibiotic effectiveness, which makes patients less willing to see a provider. Hence, even though individual providers are prescribing at a higher rate, enough

37

patients become unwilling to see a provider so that the aggregate prescription rates eventually become the same.

1.4.4 Optimal Number of Providers

One of the central questions of the chapter asks about the welfare effects of provider competition. Figure 1.18 gives an answer.
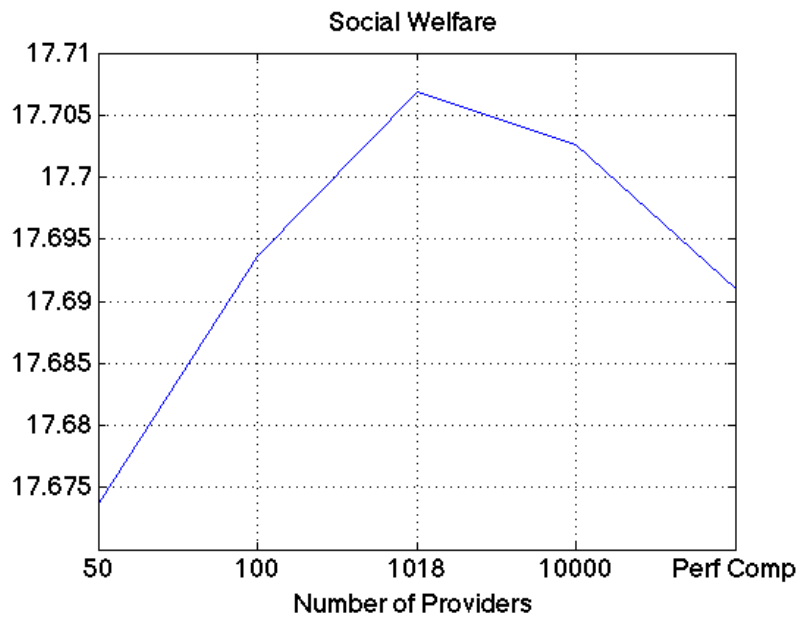
Figure 1.18: Oligopolistic competition is the optimal market structure

This graph plots social welfare as a function of the number of providers in a 200 period simulation. One of the key takeaways from this chapter is that the graph can take on an inverted U-shape. That is, social welfare can initially increase as the market structure becomes more competitive from monopoly, peak at some level of oligopolistic competition, and then decrease as the market structure becomes perfectly competitive.

Intuitively, welfare initially increases as competition increases from monopoly because this generates lower steady-state level of infection. The reason that welfare

can decrease at high levels of competition is more subtle. As the market becomes more competitive, treatable infection decreases at a quicker rate, which initially generates higher welfare for society since the level of total infection becomes lower. Further, as competition increases, so too does the idiosyncratic utility patients gain from their providers (since they are more likely to draw a higher quality match). However, the future loss from the increase in resistant infections eventually becomes sufficiently high so as to outweigh those benefits.

The growth of the resistant infection causes two sources of welfare loss. First, it makes it less likely that the antibiotic will work for future users. Second, it increases the level of total infection. As illustrated in the previous section, this effect can cause the total level of infection to become higher in a more competitive market than a less competitive market.

The model thus features a 'Goldilocks' effect: to a rough approximation, low levels of competition use too few antibiotics and high levels of competition use too much antibiotics. The optimal number of providers in the simulation is $n^* = 1018$. An interpretation is that a provider should serve about $\frac{1}{1000}$ of their market. While n^* does not implement the planner's solution, n^* better treads the balance between limiting the current growth of the treatable infection and the future growth of the resistant infection given the economic constraints in the model than other market structures.

1.5 Public Policy

The model can be used to study the effects of public policies aimed at mitigating the misuse of antibiotics by providers. I consider three types of policies: restricting

the numbers of providers (licensing), capping how frequently a provider can prescribe antibiotics (a prescription quota) and charging a fee to use antibiotics (a tax).

Up to this point, this chapter has studied the normative effects of market structure without explicitly considering the number of providers as a policy variable. However, a possible public policy is restricting the number of providers through licensing. The optimal licensing regime implements the optimal decentralized market structure. The model provides a simple way to compute this: compute the equilibrium and welfare of different market structures and choose the maximum. Given the parameter values used for the simulation, the optimal number of providers is 1018 and this generates total social welfare of 17.707.

Another possible public policy takes the market structure as given and restricts antibiotic prescriptions instead. In the context of the model, a prescription quota sets a ceiling on a provider's prescription rate. The process of computing equilibrium is essentially unchanged, except now rather than being bounded by one, a provider's prescription rate is bounded by the prescription quota. Again, the model provides a simple tool to determine the optimal quota: compute and the equilibrium and welfare induced by different quotas and choose the maximum. I set the number of providers $n = 10,000$. The optimal quota caps the prescription rate at $\hat{f} = .93$ and generates social welfare of 17.703.

The final policy considered here is a tax on antibiotics. I impose the tax, τ, on providers. The tax increases the provider's marginal cost of prescribing antibiotics from c to $c + \tau$. Given $n = 10,000$, the optimal tax on antibiotics sets $\tau = 0.17$ and generates social welfare of 17.722. Given the parameter values, the model predicts that a tax on antibiotics is a better policy than licensing or a prescription quota.

1.6 Conclusion

This chapter studies a model of provider competition embedded within a dynamic epidemiological model of infection and antibiotic resistance. One of the main findings is that oligopolistic competition can be the optimal decentralized market structure. This is because a perfectly competitive market over-uses antibiotics because providers do not bear the cost of antibiotic resistance and a monopolist under-uses antibiotics to increase infection and antibiotic efficacy. An interior level of competition has more moderate antibiotic use which generates higher welfare. The chapter also demonstrates how the model can be used for public policy analysis.

A precise determination of the optimal market structure or public policy depends on the parameter values of the model. Estimating the model is an important empirical exercise for future work. Other extensions of the model include incorporating innovation of new antibiotics by pharmaceutical companies and extending the analysis to a global setting.

1.7 Appendix A

Throughout, I use the notation f^* to denote a provider's optimal prescription rate and $\bar{f}$ to denote the steady-state prescription rate.

Proof of Theorem 1

Suppose that all providers prescribe at rate one for all states. Without loss of generality, suppose that provider 1 defects and prescribes at a rate less then one. Let N denote the set of providers and n the number of providers. Note that:

$$lim_{n \to \infty} Pr\left(f^1 Ep^A U + \varepsilon^1 \geq Ep^A U + \max_{j \in N \setminus i} \varepsilon^j\right) = Pr\left(f^1 Ep^A U + \varepsilon^1 \geq Ep^A U + 1\right) = 0 \tag{1.28}$$

Under perfect competition, patients can always find a provider who is preferred to provider 1. Hence, by deviating, provider 1 attracts no patients and has a current period payoff of zero. Since provider 1 attracts a non-measurable set of patients whether he adheres to the equilibrium strategy or not, deviating has no effect on his continuation payoff. Therefore there is no incentive to defect from prescribing at rate one.

Since the price to see a provider is 1 and the upper bound on the support of idiosyncratic utility is also 1, given the prescription behavior of providers, the utility a patient gains from their most preferred provider is $Ep^A U > 0$. Thus all patients are seen by a provider and prescribed antibiotics. This generates an aggregate prescription rate of $F^{PC}(I, E) = 1$ for all (I, E). The prescription behavior generates the asymptotic steady-state in which $\bar{E}_{PC} = 0$ and $\bar{I}_{PC} = \frac{\beta - p^R}{\beta}$.

To show perpetual overuse by the competitive market, it suffices to show that the the social planner does not prescribe at rate one for all state variables. The planner's first-order condition for a positive F^{SP*} is:

$$Ep^A UI - cI + \delta \left[\frac{\partial V^{SP}(I', E')}{\partial I'} \frac{\partial I'}{\partial F}|_{F=F^{SP*}} + \frac{\partial V^{SP}(I', E')}{\partial E'} \frac{\partial E'}{\partial F}|_{F=F^{SP*}} \right] \geq 0 \quad (1.29)$$

By (5) and (6):

$$\frac{\partial I'}{\partial F} = -EI p^A \quad (1.30)$$

and

$$\frac{\partial E'}{\partial F} = -\frac{E(1-E)p^A[1 + \beta(1-I) - p^R]}{\left[1 + \beta(1-I) - p^R + E[\Delta p - p^A F]\right]^2} \quad (1.31)$$

42

Assuming differentiability of the value function, the envelope conditions are:

$$\frac{\partial V^{SP}(I,E)}{\partial I} = [-1 + Ep^T + (1-E)p^R + F^{SP*}Ep^A]U + 1 - cF^{SP*} +$$

$$\delta \frac{\partial V^{SP}(I',E')}{\partial I'}\frac{\partial I'}{\partial I} + \delta \frac{\partial V^{SP}(I',E')}{\partial E'}\frac{\partial E'}{\partial I} \quad (1.32)$$

and

$$\frac{\partial V^{SP}(I,E)}{\partial E} = I[p^T - p^R + F^{SP*}p^A]U + \delta\frac{\partial V^{SP}(I',E')}{\partial E'}\frac{\partial E'}{\partial E} + \delta\frac{\partial V^{SP}(I',E')}{\partial I'}\frac{\partial I'}{\partial E} \quad (1.33)$$

Suppose by contradiction that the planner prescribes at rate one for all state variables. Then, $I \to \frac{\beta - p^R}{\beta}$ and $E \to 0$. Notice that $\frac{\partial I'}{\partial F} \to 0$ and $\frac{\partial E'}{\partial F} \to 0$ as well, and that the derivatives of the value function are finite in the limit. The planner's first-order condition converges to:

$$-c\frac{\beta - p^R}{\beta} < 0 \quad (1.34)$$

which is incompatible with a positive F^*, a contradiction.

□

Proof of Theorem 2

The monopolist's market share is:

$$Pr\left(fEp^AU + \varepsilon - 1 \geq 0\right) = fEp^AU \quad (1.35)$$

Hence the Bellman equation is:

$$V^M(I,E;\sigma) = \max_f \left\{[1 - cf]fEp^AUI + \delta V^M(I',E';\sigma)\right\} \quad (1.36)$$

such that (5), (6), $f \in [0,1], F = f^2Ep^AU$

The first-order condition for a positive f_M^* is:

$$[1-2f_M^*c]Ep^AUI+\delta\left[\frac{\partial V^M(I',E';\sigma)}{\partial I'}\frac{\partial I'}{\partial f}|_{f=f_M^*}+\frac{\partial V^M(I',E',\sigma)}{\partial E'}\frac{\partial E'}{\partial f}|_{f=f_M^*}\right]\geq 0 \quad (1.37)$$

The effect of the monopolist's prescription rate on the states is:

$$\frac{\partial I'}{\partial f}=-EIp^A(2fEp^AU) \quad (1.38)$$

and

$$\frac{\partial E'}{\partial f}=-\frac{E(1-E)p^A[1+\beta(1-I)-p^R]}{\left[1+\beta(1-I)-p^R+E[\Delta p-p^Af(fEp^AU)]\right]^2}(2fEp^AU) \quad (1.39)$$

Assuming differentiability of the value function, the envelope conditions are:

$$\frac{\partial V^M(I,E;\sigma)}{\partial I}=[1-cf_M^*]f_M^*Ep^AU+\delta\frac{\partial V^M(I',E';\sigma)}{\partial I'}\frac{\partial I'}{\partial I}+\delta\frac{\partial V^M(I',E';\sigma)}{\partial E'}\frac{\partial E'}{\partial I}$$
$$(1.40)$$

and

$$\frac{\partial V^M(I,E;\sigma)}{\partial E}=[1-cf_M^*]f_M^*p^AUI+\delta\frac{\partial V^M(I',E';\sigma)}{\partial E'}\frac{\partial E'}{\partial E}+\delta\frac{\partial V^M(I',E';\sigma)}{\partial I'}\frac{\partial I'}{\partial E}$$
$$(1.41)$$

In the steady-state, $I=I'$ and $E=E'$. The monopolist cannot have a steady-state in which $E=0$ because $F=f^2Ep^AU$ which for low enough E is below the threshold $\frac{\Delta p}{p^A}$. Enough patients stop going to see the monopolist that he cannot maintain an aggregate prescription rate high enough to drive effectiveness to 0. There are two

possible steady-state configurations: one with an interior level of effectiveness and one in which effectiveness equals one.

In an interior steady-state, the envelope conditions are:

$$\frac{\partial V^M(\bar{I}_M, \bar{E}_M; \sigma)}{\partial I} = \frac{[1 - c\bar{f}_M]\bar{f}_M \bar{E}_M p^A U}{1 - \delta[1 - (\beta - p^r)]} \tag{1.42}$$

and

$$\frac{\partial V^M(\bar{I}_M, \bar{E}_M)}{\partial E} = \frac{[1 - c\bar{f}_M]\frac{\beta - p^r}{\beta}\bar{f}_M p^A U}{1 - \delta} \tag{1.43}$$

Using the fact that in an interior steady-state, $\bar{f}_M^2 \bar{E}_M p^A U = \frac{\Delta p}{p^A}$, the interior steady-state first-order condition can written as:

$$\frac{\Delta p}{p^A \bar{f}_M^2} \frac{\beta - p^r}{\beta}$$
$$\left([1 - 2c\bar{f}_M] - \frac{\delta[1 - c\bar{f}_M]}{1 - \delta[1 - (\beta - p^r)]} 2\left(\frac{\Delta p^2}{p^{A2} U \bar{f}_M^2}\right) - \frac{\delta[1 - c\bar{f}_M]}{1 - \delta} 2\Delta p\left(1 - \left(\frac{\Delta p}{p^{A2} U \bar{f}_M^2}\right)\right)\right) \geq 0 \tag{1.44}$$

such that $\bar{f}_M \in (\frac{\Delta p}{p^{A2} U}]^{\frac{1}{2}}, 1]$

In the steady-state in which effectiveness equals one, the steady-state envelope condition governing infection is:

$$\frac{\partial V^M(\bar{I}_M, \bar{E}_M; \sigma)}{\partial I} = \frac{[1 - c\bar{f}_M]\bar{f}_M p^A U}{1 - \delta(1 - (\beta - p^T - \bar{f}_M^2 p^{A2} U))} \tag{1.45}$$

and since $\frac{\partial E}{\partial f} = 0$ in this steady-state, the monopolist's the steady-state first-order condition when effectiveness is one can be written as:

$$(\bar{I}_M p^A U)\left([1 - 2c\bar{f}_M] - \delta \frac{[1 - c\bar{f}_M]2p^{A2} U \bar{f}_M^2}{1 - \delta(1 - (\beta - p^T - \bar{f}_M^2 p^{A2} U))}\right) = 0 \tag{1.46}$$

45

such that $\bar{f}_M \in [0, \frac{\Delta p}{p^{A2}U}]^{\frac{1}{2}}$, $\bar{I}_M = \frac{\beta - p^T - \bar{f}_M^2 p^{A2}U}{\beta}$

Notice that either Equation 1.44 or 1.46 admits a solution.

I now argue that the monopolist's steady-state level of effectiveness $\bar{E}_M \to 1$ as $\delta \to 1$. Suppose by contradiction that $\bar{E}_M$ is bounded away from 1 for all δ. This means that $\bar{f}_M$ is bounded above $[\frac{\Delta p}{p^{A2}U}]^{\frac{1}{2}}$ for all δ.

Notice that for all values of δ, the first two terms in the monopolist's interior-steady state first-order condition (1.44) are finite. However, because $\bar{f}_M$ is bounded above $[\frac{\Delta p}{p^{A2}U}]^{\frac{1}{2}}$, the third term of the monopolist's first-order condition becomes arbitrarily small as $\delta \to 1$. Hence for high enough δ, the monopolist's first-order condition becomes negative, a contradiction. Therefore, the monopolist's steady-state level of antibiotic effectiveness tends towards one as the monopolist becomes patient.

I now argue that for sufficiently high δ, the monopolist's steady-state level of infection is $\bar{I}_M > \frac{\beta - p^r}{\beta}$. Suppose this was not the case. Since $\bar{E}_M \to 1$ and $\bar{I}_M = \frac{\beta - p^r}{\beta}$, as $\delta \to 1$, the monopolist's first-order condition tends to:

$$\left[1 - 2c\left(\frac{\Delta p}{p^{A2}U}\right)^{\frac{1}{2}}\right] - \frac{\left[1 - c\left(\frac{\Delta p}{p^{A2}U}\right)^{\frac{1}{2}}\right]}{\beta - p^R} 2\Delta p < 0 \tag{1.47}$$

which is negative by assumption A1. Hence, a sufficiently patient monopolist has an inefficiently high steady-state level of infection.

□

Lemma 1

I now state and prove a lemma that will be used in proving Theorem 3.

Lemma 1: Suppose that 1.46 is negative for some $f \in [0, 1]$. Then, the the equation is negative for all $\hat{f} \geq f$.

Proof:

Re-write (1.46) as:

$$[1 - 2cf] - \delta \frac{[1 - cf][2f^2 p^{A2} U]}{1 - \delta(1 - \beta I(f))} \tag{1.48}$$

where $I(f) = \frac{\beta - p^T - f^2 p^{A2} U}{\beta}$. Suppose that the expression is negative at f. Differentiating with respect to f gives:

$$-2c - \frac{4fp^{A2}U[1 - \frac{6}{4}cf]\Big(1 - \delta\big(1 - \beta I(f)\big)\Big) - (1 - cf)2fp^{A2}U\delta\beta\frac{\partial I}{\partial f}}{\Big(1 - \delta\big(1 - \beta I(f)\big)\Big)^2} \tag{1.49}$$

which is negative unless $[1 - \frac{6}{4}cf]\Big(1 - \delta\big(1 - \beta I(f)\big)\Big)$ is sufficiently negative. However, if $[1 - \frac{6}{4}cf] < 0$, then $[1 - 2cf] < 0$, which means that 1.46 is negative anyway. Hence, if 1.46 evaluated at f is negative, then 1.46 is negative for all $\hat{f} \geq f$.

□

Proof of Theorem 3

Let δ be sufficiently high that the monopolist's steady-state level of infection is $\bar{I}_M > \frac{\beta - p^r}{\beta}$. By the analysis in Appendix B, the duopolist's first-order condition is:

$$[1 - 2cf_D^*]Ep^A UI + \frac{1}{2}c(Ep^A Uf_D^*)^2 I + \delta\frac{\partial V^D(I', E'; \sigma)}{\partial I'}(-EIp^A)(2Ep^A Uf_D^* - \frac{3}{2}(Ep^A Uf_D^*)^2) +$$

$$\delta\frac{\partial V^D(I', E'; \sigma)}{\partial E'}\frac{-E(1 - E)p^A[1 + \beta(1 - I) - p^R](2Ep^A Uf_D^* - \frac{3}{2}(Ep^A Uf_D^*)^2)}{\Big[1 + \beta(1 - I) - p^R + E\Big(\Delta p - p^A\big(f_D^*(1 - (1 - Ep^A Uf_D^*)^2)\big)\Big)\Big]^2}$$

$$\geq 0 \quad (1.50)$$

The envelope condition governing infection is:

$$\frac{\partial V^D(I, E; \sigma)}{\partial I} = [1 - cf_D^*]\frac{1}{2}[1 - (1 - Ep^A Uf_D^*)^2] + \delta\frac{\partial V^D(I', E'; \sigma)}{\partial I'}\frac{\partial I'}{\partial I} + \delta\frac{\partial V^D(I', E'; \sigma)}{\partial E'}\frac{\partial E'}{\partial I}$$

$$(1.51)$$

47

The envelope condition governing effectiveness is:

$$\frac{\partial V(I,E;\sigma)}{\partial E} = [1-cf_D^*]I[1-Ep^AUf_D^*]p^AUf_D^* + \delta\frac{\partial V^D(I',E';\sigma)}{\partial I'}\frac{\partial I'}{\partial E} + \delta\frac{\partial V^D(I',E';\sigma)}{\partial E'}\frac{\partial E'}{\partial E}$$

(1.52)

If the duopolist has an interior steady-state level of effectiveness, then the monopolist automatically has a higher level of infection. Suppose the duopolist has a steady-state in which effectiveness is one. For notational convenience, let $a = p^AU$. The steady-state first-order condition in which antibiotic effectiveness is one can be written as:

$$[1-2c\bar{f}_D]a\bar{I}_D + \frac{1}{2}c(a\bar{f}_D)^2\bar{I}_D + \delta\frac{[1-c\bar{f}_D][a\bar{f}_D - \frac{(a\bar{f}_D)^2}{2}]}{1-\delta(1-\beta\bar{I}_D)}(-I_D^*p^A)(2a\bar{f}_D - \frac{3}{2}(a\bar{f}_D)^2) = 0$$

(1.53)

such that $\left(\bar{f}_D(2a\bar{f}_D - (a\bar{f}_D)^2)\right) < \frac{\Delta p}{p^A}, \bar{f}_D \in [0,\tilde{f}]$

where $\tilde{f}$ is defined implicitly by the equality:

$$2a\tilde{f}^2 - a^2\tilde{f}^3 = \frac{\Delta p}{p^A}$$

(1.54)

I will show that if the monopolist were to implement the duopolist's steady-state level of infection, then his first-order condition would be negative. By Lemma 1, this means that the monopolist must pick a higher level of infection, proving the claim.

The aggregate prescription rate in the duopolist's steady-state is:

$$F = \bar{f}_D(2a\bar{f}_D - (a\bar{f}_D)^2)$$

(1.55)

For the monopolist to implement the duopolist's steady-state level of infection, the aggregate prescription rate must be the same, so the monopolist must set:

48

$$f_M^2 a = \bar{f}_D(2a\bar{f}_D - (a\bar{f}_D)^2) \tag{1.56}$$

or $f^M = \bar{f}_D(2 - a\bar{f}_D)^{1/2}$.

The monopolist's steady-state first-order condition evaluated at the duopolist's steady-state level of infection is:

$$[1 - 2c\bar{f}_D(2 - a\bar{f}_D)^{1/2}]a\bar{I}_D - \delta\frac{(1 - c\bar{f}_D(2 - a\bar{f}_D)^{1/2})2a^2\bar{f}_D^2(2 - a\bar{f}_D)\bar{I}_D p^A}{1 - \delta(1 - \beta\bar{I}_D)} \tag{1.57}$$

The objective is to show that this expression is negative. Recall the positive current period marginal payoff requirement. In order for this expression to be non-negative, it is required that:

$$[1 - 2c\bar{f}_D(2 - a\bar{f}_D)^{1/2}] > 0 \tag{1.58}$$

Because $[1 - 2c\bar{f}_D]a\bar{I}_D + \frac{1}{2}c(a\bar{f}_D)^2\bar{I}_D > [1 - 2c\bar{f}_D(2 - a\bar{f}_D)^{1/2}]a\bar{I}_D$, Equation 1.57 is less than:

$$\delta\frac{[1 - c\bar{f}_D][a\bar{f}_D - \frac{(a\bar{f}_D)^2}{2}]}{1 - \delta[1 - \beta\bar{I}_D]}(\bar{I}_D p^A)(2a\bar{f}_D - \frac{3}{2}(a\bar{f}_D)^2) - \delta\frac{(1 - c\bar{f}_D(2 - a\bar{f}_D)^{1/2})2a^2\bar{f}_D^2(2 - a\bar{f}_D)\bar{I}_D p^A}{1 - \delta(1 - \beta\bar{I}_D)} \tag{1.59}$$

It suffices to show that 1.59 is either negative (which would imply that 1.57 is negative) or that the monopolist's current period positive marginal payoff requirement is violated (which would also imply that 1.57 is negative). 1.59 is negative if:

$$(1 - c\bar{f}_D(2 - a\bar{f}_D)^{1/2})2a^2\bar{f}_D^2(2 - a\bar{f}_D) > [1 - c\bar{f}_D][a\bar{f}_D - \frac{(a\bar{f}_D)^2}{2}](2a\bar{f}_D - \frac{3}{2}(a\bar{f}_D)^2) \tag{1.60}$$

Dividing both sides by $2a^2\bar{f}_d^2(1 - \frac{a\bar{f}_d}{2})$ and subtracting the right-hand side from the left-hand side yields:

49

$$1 - 2c\bar{f}_d(2 - a\bar{f}_d)^{\frac{1}{2}} + c\bar{f}_d + \frac{3}{4}a\bar{f}_d(1 - c) \tag{1.61}$$

This expression is positive unless the monopolist's positive current period marginal payoff requirement is violated. Either way, the implication is that Equation 1.57, the monopolist's steady-state first-order condition evaluated at the duopolist's solution, is negative. By Lemma 1, the monopolist's first-order condition is negative for all higher prescription rates, and so to reconcile the first-order condition, the monopolist must have a strictly higher steady-state level of infection than the duopolist.

$\square$

1.8 APPENDIX B

In this appendix, I derive the oligopolist's equilibrium first-order condition. The oligopolist's Bellman equation when there are n total providers is:

$$V(I, E; \sigma) = \max_{f^i} \left\{ [1 - cf^i]\Omega(f^i, f^{-i}, E)I + \delta V(I', E'; \sigma) \right\} \tag{1.62}$$

such that (5), (6), $F = \sum_{i=1}^{n} f^i \Omega(f^i, f^{-i}, E), f^i \in [0, 1]$

The first-order condition for a positive f^{i*} is:

$$[1 - cf^{i*}]\frac{\partial \Omega(f^i, f^{-i*}, E)}{\partial f^i}\Big|_{f^i = f^{i*}} I - c\Omega(f^{i*}, f^{-i*}, E)I +$$

$$\delta\frac{\partial V(I', E'; \sigma)}{\partial I'}\frac{\partial I'}{\partial F}\frac{\partial F}{\partial f^i}\Big|_{f^i = f^{i*}} + \delta\frac{\partial V(I', E'; \sigma)}{\partial E'}\frac{\partial E'}{\partial F}\frac{\partial F}{\partial f^i}\Big|_{f^i = f^{i*}} \geq 0 \quad (1.63)$$

To calculate the first-order condition, I need the functions that govern the marginal effect of the oligopolist increasing his prescription rate on his market share, the equilibrium market share, the aggregate prescription rate, and the marginal effect of the oligopolist increasing his prescription rate on the aggregate prescription rate.

Let N denote the set of doctors and suppose there are n total doctors. Consider doctor i who prescribes antibiotics at rate f^i. Let the remaining $n-1$ doctors prescribe antibiotics at the symmetric rate f. Then the probability that a consumer is willing to pay the fee to see provider i is:

$$\Omega(f^i, f^{-i}, E) = Pr\left(U^i(f^i, \varepsilon^i, E) \geq U^j(f, \varepsilon^j, E) \text{ for } j = 0, ..., i-1, i+1, ..., n\right) \quad (1.64)$$

This is the probability that a patient prefers provider i to all other providers and is preferred to the outside option. Since the other $n-1$ providers prescribe at the symmetric rate f, this can be written as:

$$\Omega(f^i, f^{-i}, E) = Pr\left(f^i Ep^A U + \varepsilon^i - 1 \geq 0 \bigcap f^i Ep^A U + \varepsilon^i \geq f Ep^A U + \max_{j \in N \setminus i} \varepsilon^j\right) \quad (1.65)$$

To ease notation, let $a = Ep^A U$. Using the Law of Conditional Probability,

$$\Omega(f^i, f^{-i}, E) = Pr(\varepsilon^i > 1 - f^i a) Pr\left(\varepsilon^i - \max_{j \in N \setminus i} \varepsilon^j > (f - f^i)a \middle| \varepsilon^i > 1 - f^i a\right) \quad (1.66)$$

since $\varepsilon^i \sim U[0, 1]$ Equation 1.66 can be re-written as:

$$\Omega(f^i, f^{-i}, E) = f^i a Pr\left(\varepsilon^i - \max_{j \in N \setminus i} \varepsilon^j > (f - f^i)a \middle| \varepsilon^i > 1 - f^i a\right) \quad (1.67)$$

I now focus on the latter term of Equation (81). This is the probability that provider i is preferred to all other providers conditional on i being preferred to the

51
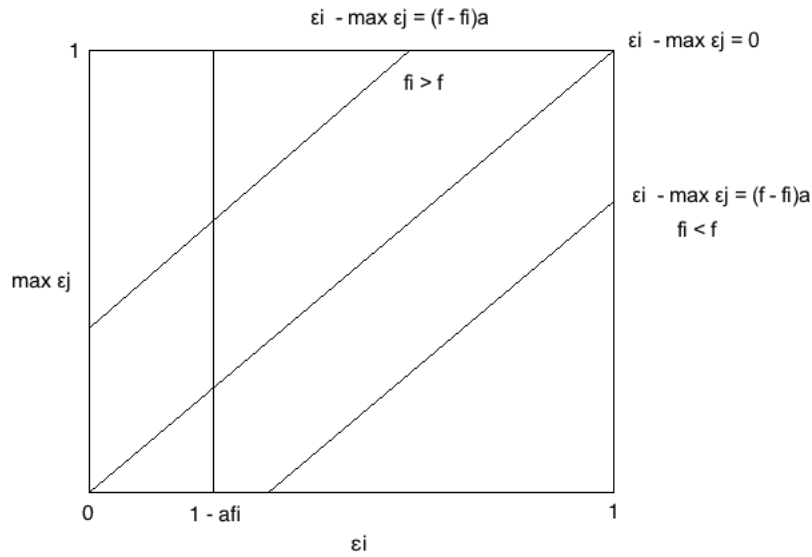
Figure 1.19: Illustration of parameter values for which i is preferred

outside option. Note that since there are n total providers including i, the distribution governing the maximum idiosyncratic utility a patient receives from the other $(n-1)$ providers is

$$\max_{j \in N \setminus i} \varepsilon^j \sim \beta(n - 1, 1) \tag{1.68}$$

which has the probability density function $f(x) = (n-1)x^{n-2}$.

I illustrate graphically the probability that provider i is preferred to all other providers conditional on i being preferred to the outside option in Figure 1.19.

ε^i runs along the x-axis and $\max_{j \in N \setminus i} \varepsilon^j$ runs along the y-axis. Conditioning on the fact that $\varepsilon^i > 1 - f^i a$, restricts the range of admissible values for ε^i. The iso-difference line $\varepsilon^i - \max_{j \in N \setminus i} \varepsilon^j = (f - f^i)a$ gives the value of ε^i that makes the patient indifferent between

seeing provider i and the next best provider, given prescription rates and a value for $\max_{j \in N \setminus i} \varepsilon^j$. The region below (above) the iso-difference line are the parameter values for which provider i is preferred (not preferred) to the best of the other providers.

The probability that provider i is preferred to all other providers conditional on being preferred to the outside option is the weighted region of admissible values under the iso-difference line divided by the weighted region of all admissible values. The region of admissible values has weight af^i.

The probability - and hence Equation 1.67 - is a piecewise function. I derive this expression for $f^i \geq f$. In equilibrium, all providers will prescribe antibiotics at the same rate, but this more general derivation is necessary to describe the marginal effect of the oligopolist increasing his prescription rate on his market share. It can easily be verified by performing the equivalent exercise for $f^i \leq f$ that the derivative at $f^i = f$ exists.

For simplicity, I calculate the weighted region of values for which i is not preferred, divide by the weight of admissible values af^i, and then subtract from one. The region for which i is not preferred is the triangle demarcated by the vertical line at $\varepsilon^i = 1 - af^i$, the corresponding the iso-difference line, and the top of the graph.

$$Pr\left(\varepsilon^i - \max_{j \in N \setminus i} \varepsilon^j \geq a(f - f^i) \,\middle|\, \varepsilon^i \geq 1 - af^i\right) = 1 - \frac{\int\limits_{1-af^i}^{1+a(f-f^i)} \int\limits_{x-a(f-f^i)}^{1} (n-1)y^{(n-2)} dy\, dx}{af^i}$$

(1.69)

$$= 1 - \frac{f}{f^i} - \frac{(1-af)^n - 1}{naf^i}$$

For $f^i \geq f$ provider i's market share is:

53

$$Pr\left(\varepsilon^i - \max_{j \in N \setminus i} \varepsilon^j \geq a(f - f^i) \bigcap \varepsilon^i \geq 1 - af^i\right) = af^i\left(1 - \frac{f}{f^i} - \frac{(1 - af)^n - 1}{naf^i}\right) \quad (1.70)$$

To calculate the marginal effect of the oligopolist increasing his prescription rate on his market share, differentiate 1.70 with respect to f^i to get:

$$\frac{\partial \Omega(f^i, f^{-i}, E)}{\partial f^i}\bigg|_{f^i = f} = a \quad (1.71)$$

To calculate i's equilibrium market share, set $f^i = f$ to get

$$\Omega(f, f^{-i}, E) = \frac{1}{n}\left(1 - (1 - af)^n\right) \quad (1.72)$$

The equilibrium aggregate prescription rate is:

$$F = fPr(\max_{j \in N} \varepsilon^j \geq 1 - af) = f\left(1 - (1 - af)^n\right) \quad (1.73)$$

To calculate an individual oligopolist's equilibrium marginal effect on the aggregate prescription rate, consider the more general formulation,

$$F = f^i \Omega(f^i, f^{-i}, E) + fPr\left(\max_{j \in N \setminus i} \varepsilon^j - \varepsilon^i \geq a(f^i - f) \bigcap \max_{j \in N \setminus i} \varepsilon^j \geq 1 - af\right) \quad (1.74)$$

This expression is to be differentiated with respect to f^i and evaluated at $f^i = f$. The first part of this expression is derived using Equations 1.71 and 1.72. Focusing on the second part of the expression and using the Law of Conditional Probability,

$$fPr\left(\max_{j \in N \setminus i} \varepsilon^j - \varepsilon^i \geq a(f^i - f) \bigcap \max_{j \in N \setminus i} \varepsilon^j \geq 1 - af\right) =$$
$$fPr\left(\max_{j \in N \setminus i} \varepsilon^j \geq 1 - af\right)Pr\left(\max_{j \in N \setminus i} \varepsilon^j - \varepsilon^i \geq a(f^i - f)\bigg|\max_{j \in N \setminus i} \varepsilon^j \geq 1 - af\right) \quad (1.75)$$

Note that

$$Pr\left(\max_{j\in N\backslash i}\varepsilon^j \geq 1-af\right) = \int\limits_{1-af}^{1}(n-1)x^{n-2}dx = 1-[1-af]^{n-1} \qquad (1.76)$$

The last component then is a closed-form expression for

$$Pr\left(\max_{j\in N\backslash i}\varepsilon^j - \varepsilon^i \geq a(f^i-f)\,\Big|\,\max_{j\in N\backslash i}\varepsilon^j \geq 1-af\right) \qquad (1.77)$$

I derive this for $f \geq f^i$ using a similar process as before, which can then be differentiated and evaluated at the symmetric equilibrium in which $f = f^i$.

For $f \geq f^i$,

$$Pr\left(\max_{j\in N\backslash i}\varepsilon^j - \varepsilon^i \geq a(f^i-f)\,\Big|\,\max_{j\in N\backslash i}\varepsilon^j \geq 1-af\right) = 1 - \frac{\int\limits_{1-af}^{1+a(f^i-f)}\int\limits_{x-a(f^i-f)}^{1}(n-1)x^{(n-2)}dydx}{1-[1-af]^{n-1}}$$

$$(1.78)$$

$$= 1 - \frac{\frac{1}{n}[1+a(f^i-f)]^n + \frac{n-1}{n}[1-af]^n - [1-af]^{n-1}[1+a(f^i-f)]}{1-[1-af]^{n-1}}$$

For $f \geq f^i$, Equation 1.75 can be re-written as:

$$fPr\left(max_{j\in N\backslash i}\varepsilon^j - \varepsilon^i \geq a(f^i-f)\bigcap max_{j\in N\backslash i}\varepsilon^j \geq 1-af\right) =$$

$$f\left(\left(1-[1-af]^{n-1}\right)\left(1-\frac{\frac{1}{n}[1+a(f^i-f)]^n + \frac{n-1}{n}[1-af]^n - [1-af]^{n-1}[1+a(f^i-f)]}{1-[1-af]^{n-1}}\right)\right)$$

$$(1.79)$$

Using Equations 1.71, 1.72, and 1.79, I can calculate the marginal effect that increasing an oligopolist's prescription rate has on the aggregate prescription rate in the symmetric equilibrium.

$$\frac{\partial F}{\partial f^i}\bigg|_{f^i=f} = \Omega(f, f^{-i}, E) + fa - fa + af(1-af)^{n-1} = \frac{1}{n}[1-(1-af)^n] + af(1-af)^{n-1} \tag{1.80}$$

I can now write the first-order condition that governs the oligopolist's symmetric Markov Perfect equilibrium. Incorporating Equations 1.71, 1.72, and 1.80 into the first-order condition for a positive f^* gives:

$$[1-cf^*]aI - c\frac{1}{n}[1-(1-af^*)^n]I + \delta\frac{\partial V(I', E'; \sigma)}{\partial I'}(-EIp^A)\left(\frac{1}{n}[1-(1-af^*)^n] + af^*(1-af^*)^{n-1}\right) +$$

$$\delta\frac{\partial V(I', E'; \sigma)}{\partial E'}\frac{-E(1-E)p^A[1+\beta(1-I)-p^R]\left(\frac{1}{n}[1-(1-af^*)^n] + af^*(1-af^*)^{n-1}\right)}{\left[1+\beta(1-I)-p^R + E\left(\Delta p - p^A\left(f^*(1-(1-af^*)^n)\right)\right)\right]^2}$$

$$\geq 0 \tag{1.81}$$

1.9 References

[1] Barnett, ML and JA Linder; Antibiotic prescribing to adults with sore throat in the United States, 1997-2010 *JAMA Intern Med.* Vol. 174(1) January 2014 pp. 138-140

[2] Bauchner, H, S Pelton, and J Klein; Parents, Physicians, and Antibiotic Use, *Pediatrics* February, 1999 pp. 1– 6

[3] Bennet, D, CL Hung, TL Lauderdale; Health Care Competition and Antibiotic Use in Tawain, Working Paper, 2012

[4] Brown, G and D Layton; Resistance economics: social cost and the evolution of antibiotic resistance, *Environment and Development Economics* July, 1996 pp. 349 – 355

[5] Clemens J and J Gottlieb; Do Physicians Financial Incentives Affect Medical Treatment and Health? *The American Economic Review* 2014, 104(4) pp. 1320 – 1349

[6] Chonè, P and CA Ma; Optimal Health Care Contract under Physician Agency, *Annals of Economics and Statistics* No. 101/102, January/June 2011pp. 229-256

[7] Débarre F, Lenormand T, Gandon S; Evolutionary Epidemiology of Drug-Resistance in Space, *PLoS Computational Biology* 5(4) 2009

[8] Elbasha, E; Deadweight loss of Bacterial Resistance Due to Overtreatment, *Health Economics* Vol. 12 pp. 125-138 February, 2003

[9] Fogelberg, S; Effects of Competition Between Healthcare Providers on Prescription of Antibiotics, Working Paper, 2013

[10] Herrmann, M; Monopoly pricing of an antibiotic subject to bacterial resistance, *Journal of Health Economics* 2010 pp. 137 – 150

[11] Herrmann, M and G Gaudet, The economic dynamics of antibiotic efficacy under open access, *Journal of Environmental Economics and Management* May, 2009 pp. 334 – 350

[12] Herrmann, M, B Nkuiya, A Dussault; Innovation and Antibiotic Use within Antibiotic Classes: Market Incentives and Economic Instruments, Working Paper 2014

[13] Jacobson, M T Chang, J Newhouse, and C Earle; Physician Agency and Competition: Evidence from a Major Change to Medicare Chemotherapy Reimbursement Policy, *NBER Working Paper No. 19247* July 2013

[14] Kermack, W and A McKendrick; A Contribution to the Mathematical Theory of Epidemics, *Proceedings of the Royal Society A* 115, pp. 700 - 721, 1927

[15] Klevens, RM, MA Morisson, J Nadle, S Petit, K Gershman, S Ray, LH Harrison, R Lynfield, G Dumyati, JM Townes, AS Craig, ER Zell, GE Fosheim, LK McDougal, RB Carey, SK Fridkin; Invasive methicillin-resistant Staphylococcus aureus infections in the United States, *The Journal of the American Medical Association* October, 2007 pp 1763 – 1771

[16] Kuster, SP, W Rudnick, A Shigayeva, K Green, M Baqi, WL Gold, R Lovinsky, MP Muller, JE Powis, N Rau, AE Smior, SL Walmsley, DE Low, A McGeer, Toronto Invasive Bacterial Diseases Network; Previous antibiotic exposure and antimicrobial resistance in invasive pneumococcal diseaseL results from prospective surveillance, *Clinical Infectious Diseases*, October 2014

[17] Layton, D and G Brown; Resistance Economics: Social Cost and the Evolution of Antibiotic Resistance, *Environmental and Development Economics*, 1 July 1996

[18] Laxminarayan, R and G Brown; Economics of Antibiotic Resistance: A Theory of Optimal Use, *Journal of Environmental Economics and Management*

[19] Laxminarayan, R and M Weitzman; On the implications of endogenous resistance to medications, *Journal of Health Economics* 2002 pp. 709 – 718

[20] Levhari, D and L Mirman; The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution, *Bell Journal of Economics* Spring, 1980 pp. 322 – 334

[21] Levy, SB; *The Antibiotic Paradox*, New York and London: Plenum Press, 1992

[22] Mechoulan, S; Market Structure and Communicable Diseases, *Canadian Journal of Economics* May, 2007 pp. 468 – 492

[23] Mold, S, G Fryer, A Roberts; When Do Older Patients Change Primary Care Physicians? *Journal of the American Board of Family Medicine* Vol. 17 No. 6, pp. 453-460, Nov. 2004

[24] Porco TC, Gao D, Scott JC, Shim E, Enanoria WT, et al; When Does Overuse of Antibiotics Become a Tragedy of the Commons? *PLoS ONE* 7(12), 2012

[25] Roberts RR, B Hota, I Ahmad, RD Scott 2nd, SD Foster, F Abbasi, S Schabowski, LM Kampe, GG Ciavarella, M Supino, J Naples, R Cordell, SB Levy, RA Weinstein; Hospital and societal costs of antimicrobial-resistant infections in a Chicago teaching hospital: implications for antibiotic stewardship, *Clinical Infectious Diseases* October 2009, pp. 1175-1184

[26] Schneider, J, Li, P, Klepser, D, Peterson, N, Brown, T, and Scheffler, R; The effect of physician and health plan market concentration on prices in commercial health insurance markets, *International Journal of Health Care Finance and Economics* 8:13Ð26. 2008

[27] Seppälä, H, T Klaukka, R Lehtonen, E Nenonen, the Finnish Study Group for Antimicrobial Resistance, P Huovinen; Outpatient Use of Erythromycin: Link to Increased Erythromycin Resistance in Group A Streptococci, *Journal of Clinical Microbiology*, 21: pp. 1378-1385, 1994

[28] Stivers, T; Parent Resistance to Physicians' Treatment Recommendations: One Resource for Initiating a Negotiation of the Treatment Decision, *Health Communication*, 18(1), pp 41-47, 2005

Court Delay, Firm Entry, and Worker Safety

2.1 Introduction

Courts are frequently called upon to resolve disputes between contracting parties. However, court delay, a common feature in many jurisdictions, can undermine a court's effectiveness. This chapter studies an economic model in which delay limits the court's ability to implement a welfare enhancing liability scheme. The context is worker's safety, but the model applies to cases of product safety as well. The model describes how delay affects a firm's decision to enter a market and invest in worker safety. After analyzing the economic effects of delay, I argue that the model provides support for applying prejudgment interest to accident claims, a policy that states currently stand divided on.

A central feature of the model is an information asymmetry between firms and workers that creates scope for a liability scheme. Workers are unaware of the safety risk at a particular firm, but know the expectation of risk across the industry. In the absence of legal intervention, workers demand an ex-ante wage from the firm as compensation for risk. Due to the information asymmetry, a firm's investment in safety lowers the market wage by less than the social value of the investment, which incentivizes under-investing in safety. In this setting, a court can induce beneficial investment in safety by holding firms liable for damages.

In fact, if the court were perfect - which in the context of model studied here would be a court with no delay - then court intervention could induce the optimal safety investment. In this ideal setting, the court could award damages to injured workers that exactly compensates for the loss from an accident. The court acts as a commitment mechanism that binds the firm to paying the full social cost ex-post an accident, thus giving the firm proper incentive to invest in safety ex-ante. Efficient investment would occur despite the information asymmetry between firms and workers.

The model here describes a court that suffers delay in dispensing justice. I take the view of the legal proverb 'justice delayed is justice denied.' In the model, delay limits the court's ability to compensate injured workers after an accident. Workers, aware that they can only be partially compensated by the court, demand a wage that compensates for the remaining expected risk of employment. Workers are thus compensated for damages by a mix of the investment-inducing court and the investment-deterring wage. As delay increases, the reliance on the wage as the mechanism for compensation over the court becomes stronger.

The analysis of firm behavior is divided into two cases: one in which firms are symmetric with respect to their safety investment technology and one in which a high-type firm has superior investment technology to a low-type firm type.

In the symmetric case, firms choose the same accident probability which, due to court delay, is higher than is socially optimal. Workers cannot directly observe the firm's accident probability, but due to symmetry infer the accident probability and demand a wage that fully compensates for the loss from court delay. Symmetric firms therefore bear the full burden of the under-investment in safety. As court delay increases, firms invest less in safety, even though the equilibrium outcome is for profits to fall even further. This effect can become strong enough to prevent otherwise efficient

firms from entering market, where efficiency refers to a firm's ability to generate positive social welfare through production.

In the asymmetric case, as in the symmetric case, firms under-invest in safety. Under natural assumptions about the investment technology, the high-type firm has a lower accident probability than the low-type firm. Due to the information asymmetry, both firm types pay the same equilibrium wage. Workers at the high-type firm are therefore over-compensated for the expected damages from employment and workers at the low-type firm are under-compensated. This is effectively a wage subsidy from the high-type firm to the low-type firm.

As in the symmetric case, when court delay increases, firms invest less in safety. While this unambiguously drives down the profits of the high-type firms, the increase in delay can increase the profits of the low-type firms. This is because increasing delay increases reliance on the wage, which is subsidized for the low-type firm by the high-type firm. However, because increasing delay also increases the high-type firm's accident probability, increasing delay can eventually lower the value of this wage subsidy. Due to these effects, delay has more subtle effects on the entry behavior of asymmetric firms than symmetric firms. The paper focuses on the case in which the high-type firm is efficient and the low-type firm is inefficient.

Courts with sufficiently low delay generate a separating equilibrium in which efficient firms are active and inefficient firms are inactive. By forcing firms to bear most of the social cost of accidents, the low delay court is able to incentivize firms to make the socially correct entry decision. However courts with high delay may generate a pooling equilibrium in which both efficient and inefficient firms are active. This is because as court delay increases, the wage subsidy can become large enough to induce inefficient firms to enter the market.

The key takeaway from the model is that delay does not merely transfer welfare between plaintiffs and defendants: delay causes a welfare loss by encouraging firms to under-invest in safety and enter inefficiently.

A policy that can mitigate this effect is applying prejudgment interest, which is interest that begins to accrue from the time an injury, to damages. As discussed in Section 5, states differ widely in applying this multiplier. This chapter offers an economic argument in favor of applying prejudgment interest. Interest on damages can partially compensate an injured worker for delay in legal proceedings. By making firms more closely pay the true cost that workers bear for accidents, applying prejudgment interest encourages firms to make more beneficial investment and entry decisions.

The rest of the paper will proceed as follows: Section 2 discusses the related literature. Section 3 describes the model. Section 4 presents the main results. Section 5 concludes.

2.2 RELATED LITERATURE

Although observers have written about court delay since at least the ancient Greeks[1], the economics literature begins more recently with the work of Landes (1971) and Priest (1989). Their work focuses on how delay affects the decision to settle or proceed to trial in criminal and civil cases. Gravelle (1990) argues that some delay can be welfare enhancing because it prevents socially wasteful civil litigation from proceeding to trial. Chemin (2012) exploits a court reform in India that was enacted at different

---

[1]Lawyers and Litigants in Ancient Athens: The Genesis of the Legal Profession, Robert Johnson Bonner, pg 89

times by provinces and found that lowering court delay led to fewer contract breaches and increased pre-contractual investment, a result consistent with the present model.

Closer to the spirit of the present chapter are Vereeck and Mühl (2000) and Chappe (2012). Vereeck and Mühl fix a plaintiff and a defendant and study how delay affects their willingness to invest in safety. Chappe studies a model in which defendants choose to engage in actions that harm potential plaintiffs and delay affects a plaintiff's willingness to file suit.

This chapter adds to the literature by studying a model of court delay in a market setting with an endogenous participation decision by plaintiffs (workers) and defendants (firms). The chapter can thus shed insight on the broader economic effects of court scarcity rather than the effects on a particular conflict or set of contracting agents as typically studied in the literature. Specifically, the chapter studies how court delay endogenously affects workers' safety and wages, the market output, and the quality of active firms.

The chapter relates more generally to the extensive literature studying how courts can resolve inefficiencies stemming from information asymmetries. A small sample of the literature includes Ayres and Gertner (1989) and Bebchuck and Shavell (1991) which study how default rules can promote welfare enhancing information disclosure, and Anderlini et al (2011) which studies how a court that commits to voiding certain contracts can prevent inefficient pooling. This chapter also studies how a court can resolve inefficiencies resulting from information asymmetries. However in the model here, the court acts as an ex-post compensation mechanism that commits the firm to paying some of the cost of accidents. This decreases the reliance on the wage, an imperfect tool for encouraging investment due to the information asymmetry, as the compensation mechanism.

65

2.3 THE MODEL

There are three components to the model: courts, firms, and workers. Each are considered in turn.

2.3.1 COURTS

Courts are modeled as a black-box process that award a settlement on a case after an exogenous delay of time T. Delay can be thought of as arising either through congestion of the court system or simply because legal proceedings take time.[2]

I assume that all accidents are verifiable by the court and cause the worker damages d. However because of delay, injured workers only recover a fraction of the original damages from the firm in court. Let $f^T d$, where $f < 1$, denote the fraction of the original damages that injured employees expect to recover in court.

That delay causes a loss to injured workers that the court cannot account for is a critical assumption of the paper, and so it is worth elaborating on. As discussed in Section 5, some states do not award prejudgment interest on damages and so discounting causes a loss to the plaintiff. Even if the court does award prejudgement interest, there are other costs caused by delay. If workers can only access incomplete financial markets, then court delay can cause injuries to heal improperly because some workers may be unable to seek care until there is legal resolution. Further, evidence decays over time. Thus it seems reasonable to assume that court delay harms the plaintiff, the party who has been wronged in the framework described here.

When there is no delay, $f^0 = 1$ and the injured worker always recovers full damages. However, as T increases, the effect of decaying evidence is felt and the injured

---

[2]Delay can be made endogenous by modeling the court as a queueing process, but because this complicates the analysis without adding intuition, I assume delay is exogenous.

worker's expected payoff from the court decreases. The parameter f can be interpreted as the rate that the court payoff is discounted. As f decreases, the effects of delay are felt more strongly.

2.3.2 FIRMS

Risk-neutral firms observe T and decide whether or not to enter the market. Active firms make an unobservable, costly investment in safety. Production and injury risk happen discretely. Firm i chooses an accident probability α_i at cost $c_i(1 - \alpha_i)$. It is assumed that $c_i(.)$ is strictly increasing, strictly convex, that $c_i(0) = 0$, and that $\lim_{\alpha_i \to 0} c_i(1 - \alpha_i) = +\infty$. After the investment, the firm hires a sequence of single-period workers at the equilibrium wage w and production begins. Workers occasionally get injured during the production process. The injured worker takes the firm to court, where the firm has a loss of $f^T d$.

I assume that the value of the firm's per-period output is q, regardless of the number of active firms, and that the firm captures the entire value of its output. The analysis when there is consumer surplus or welfare effects from competition (through increased production) is essentially the same.

Conditional on entry, firm i's problem is to choose an accident probability α_i at time 0 to maximize profits:

$$\Pi_i = \sum_{t=0}^{\infty} \delta^t (q - w - \alpha_i f^T d) - c_i(1 - \alpha_i) \tag{2.1}$$

where $\delta < 1$ is the discount factor. Firm profits have four components: the value of goods created, the cost of wages, the cost of law suits, and the cost of the safety investment. Firm profits must be weakly positive in order for the firm to enter the market.

2.3.3 Workers

Risk-neutral workers do not observe the safety level at each firm, but know the distribution of safety across the industry. Since workers are partially compensated for the costs of an accident by the court, they demand a wage that compensates them for their expected loss not covered by the court. In each period there is equal probability of employment from any of the firms, and so the worker accepts a job offer if:

$$w \geq \frac{1}{n} \sum_{i=1}^{n} \alpha_i d[1 - f^T] \tag{2.2}$$

Given firm profit maximization, in equilibrium this must hold with equality.

Note that it is implicitly assumed that between the court and the wage the worker is fully compensated for risk, that the worker's outside option is valued at zero, and that there is no surplus sharing between the firm and worker.

2.4 Results

I briefly review the order in which events unfold in the model. The court is endowed with some delay T. Firms observe T and decide whether or not to enter the market. Firms observe the other firms who are active and then make their safety investment. Workers are hired and production begins. Due to the firm's two stage decision process, I focus on sub-game perfect equilibria. Before solving the model, it will be helpful to introduce a few definitions.

Definition: α_i^* is the socially optimal accident probability for firm i

α_i^* is the accident probability that minimizes the social costs of production at firm i, which are the cost of investing in safety and the cost of future accidents. α_i^* is the solution to the following problem:

$$\min_{\alpha_i} \sum_{t=0}^{\infty} \delta^t \alpha_i d + c_i(1 - \alpha_i) \tag{2.3}$$

The solution is characterized by the first order condition:

$$c_i'(1 - \alpha_i^*) = \frac{d}{1 - \delta}. \tag{2.4}$$

Definition: Firm i is *efficient* if $\frac{q - \alpha_i^* d}{1 - \delta} > c_i(1 - \alpha_i^*)$

A firm is efficient if the value of its output is greater than the cost of accidents and investment for some feasible safety investment.

Definition: Firm j is *inefficient* if $\frac{q - \alpha_j^* d}{1 - \delta} < c_j(1 - \alpha_j^*)$

A firm is inefficient if the value of its output is less than the cost of accidents and investment, even with the optimal safety investment.

I now analyze firm behavior in the model. For simplicity, I focus on the case in which there are $n = 2$ firms, although all of the results generalize to arbitrary n. I let $\alpha_k(T)$ and $\Pi_k(T)$ denote firm k's equilibrium accident probability and profits from entering respectively when there is court delay of length T.

2.4.1   SYMMETRIC FIRMS

I first look at the effect of court delay when firms are symmetric. Firms are symmetric if they have the same safety investment technology.

Definition: Firms i and j are *symmetric* if $c_i(1-\alpha) = c_j(1-\alpha)$ for all $\alpha \in [0,1]$

Throughout this section, I assume that firms are efficient.

Theorem 1

If q, the value of output, is sufficiently high, then for all T:

1. Firms enter with probability one.

2. Firms have the same accident probability, which is higher than the socially optimal accident probability, i.e. $\alpha_i(T) = \alpha_j(T) > \alpha^*$.

3. The accident probability is increasing in court delay, i.e. $\frac{\partial \alpha^k(T)}{\partial T} > 0$ for $k = i, j$.

4. Profits are decreasing in court delay, i.e. $\frac{\partial \Pi_k(T)}{\partial T} < 0$ for $k = i, j$.

Theorem 1 finds that court delay causes active symmetric firms to under-invest in safety. Due to court delay, workers are partially compensated for risk through the ex-ante wage. The social benefit of investing in safety is a decrease in the discounted future cost of accidents. Workers, however, are unaware of firm-level risk and so lowering the future cost of accidents at a particular firm only lowers the market wage by a fraction of the social benefit.

Firms are symmetric and have strictly convex cost functions, and so active firms pick the same accident probability. Workers can therefore infer the accident probability and demand a wage that exactly compensates for the risks from employment that the imperfect court cannot cover. Symmetric firms thus ultimately bear the full burden of the under-investment in safety.

As court delay increases, there is greater reliance on the wage as the compensation mechanism over the court. This lessens the return from investing in safety,

incentivizing firms to invest less in safety, even though equilibrium profits are driven down even further. In Theorem 1, the value of a firm's output is sufficiently high that firms earn positive profits regardless of court delay, ensuring that firms always enter. The next result looks at how delay can affect firm entry.

Theorem 2

If q is sufficiently low, then there exists a $\hat{T}$ such that:

1. For all $T \leq \hat{T}$, firms enter with probability one.

2. For all $T > \hat{T}$, there exists an equilibrium in which firms enter with probability $p(T) < 1$. This probability is decreasing in T.

Theorem 2 describes the effect of changes in court quality when the firms are less productive than in Theorem 1 (although still efficient). When court quality is high, firms behave similarly to the scenario described in Theorem 1. However, when the court quality is sufficiently low, equilibrium profits when both firms are active are driven below zero, in which case firms will no longer enter with probability one. There is a mixed strategy equilibrium in which both firms enter the market with some positive probability that is less than one. The longer is court delay, the lower is the probability of entry.

2.4.2 ASYMMETRIC FIRMS

I now turn to the effects of court delay when firms are asymmetric. I focus on a simple version of asymmetry in which a high-type firm has *superior* investment technology compared to a low-type firm.

Definition: Firm i has *superior* technology to firm j if $\forall \, \alpha \in [0, 1), \; c_j'(1 - \alpha) > c_i'(1 - \alpha)$.

That is, firm i is superior to firm j if the cost of i's investment increases at a slower rate than the cost of j's investment for any given accident probability. The first result about asymmetric firms is stated below.

Theorem 3

Suppose that firm i is superior to firm j. If q is sufficiently high, then for all T:

1. Both firm types enter with probability 1.

2. Both firm types under-invest in safety, i.e. $\alpha_k(T) > \alpha_k^*$ for $k = i, j$.

3. The superior firm has a lower accident probability, i.e. $\alpha_i(T) < \alpha_j(T)$.

4. The accident probability at each firm is increasing in delay, i.e. $\frac{\partial \alpha_k(T)}{\partial T} < 0$ for $k = i, j$.

5. The superior firm's profits decrease in court delay, i.e. $\frac{\partial \Pi_i(T)}{\partial T} > 0$.

6. The inferior firm's profits can increase or decrease in court delay, i.e. $\frac{\partial \Pi_j(T)}{\partial T} \gtrless 0$.

Theorem 3 finds that when active, both firm types under-invest in safety relative to the social planner. The intuition is the same as when firms are symmetric: investing in safety lowers the wage by less than the social value of decreased accidents. However, the high-type firm has a lower accident probability than the low-type firm.

As court delay increases, both firm types invest less in safety, again for the same intuition as with symmetric firms. Interestingly, while increasing delay negatively affects the superior firm's profits, increasing delay has ambiguous effects on the inferior firm's profits.

This is because because of a wage subsidy from the high-type firm to the low-type firm. The source of the wage subsidy is the information asymmetry between firms and workers, which causes both firm types pay the same market wage. Workers at the high-type firm are ex-post over-compensated by their firm since they were partially compensated for the possibility of becoming employed at the low-type firm. For the same reason, workers at the low-type firm are ex-post under-compensated by their firm. By lowering the cost of labor, the high-type firm is essentially giving a subsidy to the low-type firm.

As delay varies, the value of this subsidy to the low-type firm changes, which affects profits. Increasing delay increases the use of the wage, which increases the proportion of the low-type firm's costs that can be subsidized by the high-type firm. However, increasing delay causes the high-type firm to increase his accident probability, which can lower the value of the wage subsidy. This subsidy is a key driver of entry behavior when firms are asymmetric. I focus on the case in which the high-type firm is efficient and the low-type firm is inefficient.

Theorem 4

Suppose that firm i is efficient and that firm j is inefficient. Then, there exists a $\hat{T}$ such that for $T < \hat{T}$ the efficient firm enters the market and the inefficient firm is inactive.

Theorem 4 finds that courts with sufficiently low delay generate a separating equilibrium in which only the efficient firm is active. Low delay courts reimburse workers for most of the cost of accidents, which limits the use of the wage as the mechanism for compensation and forces firms to bear most of the social cost of their investment decisions. Firms thus make the socially correct entry decision.
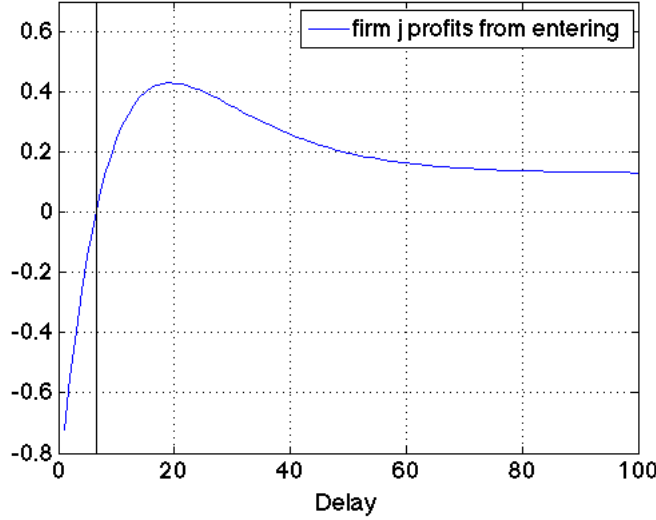
73

Figure 2.1: High levels of delay generate a pooling equilibrium

However, as delay increases, the court loses the ability to reimburse workers for the cost of accidents and the wage subsidy to the inefficient firm can become high enough to induce entry. I now present a numerical example in which high delay courts generate a pooling equilibrium in which both efficient and inefficient firms are active.

Consider the parameter values $\delta = .95$, $f = .93$, $q = 2.45$, $d = 5$, $c_i = \frac{3}{\alpha_i} - 3$, and $c_j = \frac{5}{\alpha_j} - 5$. The efficient firm i always has positive profits regardless of the inefficient firm j's behavior, and so always enters. Firm j's profits from entering as a function of court delay are plotted in Figure 2.1.

For values of court delay to the left of the bar, firm j has negative profits from entering, and so remains inactive. Low delay courts generate a separating equilibrium as discussed in Theorem 4. However firm j's profits are initially increasing in court delay - and for values of delay to the right of the bar, profits become positive. The wage subsidy from the efficient firm to the inefficient firm has become high enough
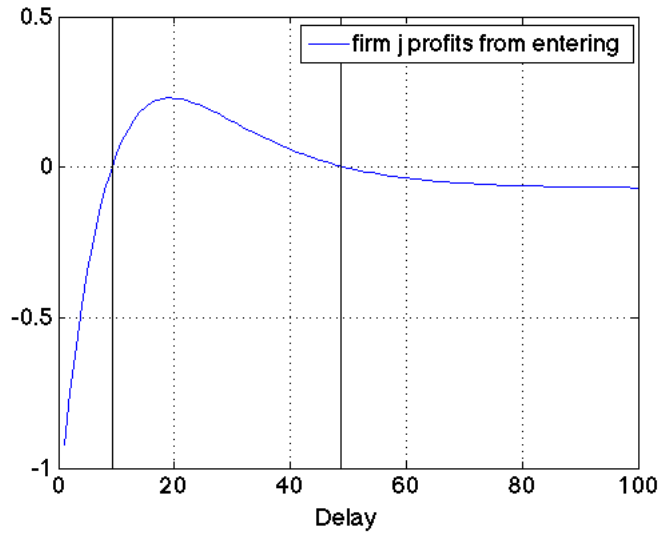
Figure 2.2: Intermediate levels of delay generate a pooling equilibrium

that the inefficient firm is induced to enter. High delay courts can thus generate an pooling equilibrium in which both firm types are active.

Increasing delay increases reliance on the wage as the tool for compensation, which increases the proportion of the inefficient firm's costs that can be subsidized by the efficient firm. However, increasing delay also increases the accident probability at the efficient firm, which lowers the value of the subsidy to the inefficient firm. These factors can cause the inefficient firm's profits as a function of delay to take on an inverse-U shape.

Because of this inverse-U shape, it is possible for low delay courts to generate a separating equilibrium, intermediate delay courts to generate a pooling equilibrium, and high delay courts to generate a separating equilibrium. Let the value of output be $q = 2.44$. Firm j's profits from entering as a function of court delay are plotted in Figure 2.2.

At low values of delay, the court generates a separating equilibrium by forcing firms to bear the social cost of their actions. As delay increases to an intermediate value, the wage subsidy to the inefficient firm can become high enough to induce entry. However, as delay increases, the efficient firm's accident probability rises. This eventually lowers the value of the wage subsidy to the inefficient firm. As Figure 2 demonstrates, this effect can become strong enough that under very high delay, inefficient firms do not gain a high enough subsidy to warrant entry. Hence, high delay courts can generate an efficient separating equilibrium.

2.5 CONCLUSION

This chapter presents an economic model of court delay in which delay limits the court's ability to implement a liability scheme. Delay generates several inefficiencies. First, delay decreases a firm's incentive to invest in worker safety. Second, delay can cause inefficient entry decisions. When firms are symmetric, delay can induce efficient firms to exit the market. When firms are asymmetric, delay can induce inefficient firms to enter the market.

A policy that can mitigate the effects of court delay is applying prejudgment interest to workers' compensation claims. Prejudgment interest is interest on a legal award applied from the time the injury occurs until a court awards damages.

State policy over awarding prejudgment interest on accident damages awards varies dramatically. In California, for example, the State Supreme Court ruled that, "interest under Civil Code section 3287 properly accompanies reinstatement and a backpay award in order to make the employee whole" (Currie v. Workers' Comp. Appeals Bd, 24 Cal 4th 1109, 2001). On the other hand, in Minnesota, Civil Procedure Chapter 549 Section 9 explicitly forbids awarding prejudgment interest on

workers' compensation awards. Still other states, such as Missouri, allow prejudgment interest only if certain conditions are met (McCormack v. Stewart Enterprises, 956 S.W. 2d 310, 1997).

The model presented in this chapter gives an argument in favor of applying prejudgment interest to workers' compensation claims. Recall that court compensation for an accident is given by $f^T d$. Applying prejudgment interest in the model is equivalent to increasing f, which increases the worker's compensation from the court following an accident. By more closely compensating the worker for the costs of an accident, prejudgment interest can encourage firms to make better investment and entry decisions.

2.6 APPENDIX

Throughout, I let $\alpha_k(T)$ denote firm k's equilibrium accident probability when there is delay T and α_i^* the socially optimal accident probability.

Proof of Theorem 1

Conditional on entry, firm i solves the problem:

$$\max_{\alpha_i} \; q - \sum_{t=0}^{\infty} \delta^t \left(\frac{1}{2}(\alpha_i + \alpha_j)[1 - f^T]d + \alpha_i f^T d \right) - c_i(1 - \alpha_i) \tag{2.5}$$

which has the equilibrium first order condition:

$$c_i'(1 - \alpha_i(T)) = \frac{\frac{1}{2}d(1 - f^T) + f^T d}{1 - \delta} \tag{2.6}$$

Note that if $T = 0$, then the firm chooses the socially optimal accident probability. However if $T > 0$:

$$\frac{\frac{1}{2}d(1 - f^T) + f^T d}{1 - \delta} < \frac{d}{1 - \delta} \Rightarrow c_i'(1 - \alpha_i(T)) < c_i'(1 - \alpha_i^*) \tag{2.7}$$

77

Because $c_i(.)$ is strictly convex, $\alpha_i(T) > \alpha^*$. Further, because of symmetry and strict convexity of the cost functions, firms i and j choose the same accident probability.

To show that the accident probability is increasing in court delay, re-write the first-order condition in a manner amenable to using the Implicit Function Theorem:

$$c_i'(1 - \alpha_i(T)) - \frac{\frac{1}{2}d(1 - f^T) + f^T d}{1 - \delta} = 0 \qquad (2.8)$$

By the Implicit Function Theorem:

$$\frac{\partial \alpha_i}{\partial T} = -\frac{\frac{1}{2}df^T log(f)}{c_i''(1 - \alpha_i)(1 - \delta)} > 0 \qquad (2.9)$$

This expression is positive because $log(f) < 0$ since $f < 1$ and $c_i''(1 - \alpha_i) > 0$ by the strict convexity of $c_i(.)$.

To show that profits are decreasing in court delay, apply the envelope theorem to equilibrium profits. Equilibrium profits are:

$$\Pi_i(T) = q - \frac{\frac{1}{2}[\alpha_i(T) + \alpha_j(T)][1 - f^T]d}{1 - \delta} - \frac{\alpha_i(T)f^T d}{1 - \delta} - c(1 - \alpha_i(T)) \qquad (2.10)$$

By the envelope theorem:

$$\frac{\partial \Pi_i}{\partial T} = -\frac{1}{2}\frac{\partial \alpha_j}{\partial T}[1 - f^T]d + \frac{\frac{1}{2}[\alpha_i(T) + \alpha_j(T)]\frac{\partial f^T}{\partial T}d}{1 - \delta} - \frac{\alpha_i \frac{\partial f^T}{\partial T}d}{1 - \delta} < 0 \qquad (2.11)$$

78

This expression is negative because $\frac{\partial \alpha_j}{\partial T} > 0$ and $\alpha_i(T) = \alpha_j(T)$.

Definition: α_i^{UB} is the accident probability firm i chooses when there is no court enforcement.

This is the accident probability the firm chooses in the limit as delay tends to infinity, i.e. there is no effective court enforcement. The only return the firm gets by investing in safety is a decrease in the compensatory wage. This is the upper bound on accident probabilities that a firm will choose in equilibrium. It is characterized by:

$$c_i'(1 - \alpha_i^{UB}) = \frac{1}{n}\frac{d}{1 - \delta} \tag{2.12}$$

As $T \to \infty$, the accident probability monotonically decreases and $\alpha_i(T) \to \alpha_i^{UB}$. Profits monotonically decrease and $\Pi_i(T) \to \frac{q - \alpha_i^{UB}d}{1 - \delta} - c_i(1 - \alpha_i^{UB})$, which is strictly positive if q is sufficiently high, ensuring that firms enter with probability 1.

$\square$

Proof of Lemma 1

I now state and prove a lemma that will be used in proving Theorems 2 and 4.

Lemma 1: If only one firm is active, then that firm invests optimally in safety.

Proof:

Suppose that only one firm is active. Conditional on entry, that firm solves the problem:

$$max_{\alpha_i} \ q - \sum_{t=0}^{\infty} \delta^t \left((\alpha_i)[1 - f^T]d + \alpha_i f^T d \right) - c_i(1 - \alpha_i) \tag{2.13}$$

which has the first-order condition:

$$c_i'(1 - \alpha_i(T)) = \frac{d(1 - f^T) + f^T d}{1 - \delta} = \frac{d}{1 - \delta} = c'(1 - \alpha_i^*) \tag{2.14}$$

When there is only one firm, there is no information asymmetry, and so the firm recoups the full social value of its investment through a decrease in the wage. The firm thus invests optimally in safety.

I denote firm i's profit when it is the only active form by Π_i^*. This value is positive when a firm is socially valuable and negative when a firm is socially wasteful.

$\square$

Proof of Theorem 2

Suppose that firms are socially valuable, but q is sufficiently low so that:

$$\frac{q - \alpha_i^{UB}}{1 - \delta} - c_i(1 - \alpha_i^{UB}) < 0 \tag{2.15}$$

Since firms invest optimally when $T = 0$, profits are:

$$\Pi_i(0) = \frac{q - \alpha_i^* d}{1 - \delta} - c(1 - \alpha_i^*) > 0 \tag{2.16}$$

which is positive by the assumption that firms are socially valuable. If both firms are always active, then profits $\Pi_i(T)$ tend monotonically from $\Pi_i(0)$ to $\frac{q - \alpha_i^{UB}}{1 - \delta} - c_i(1 - \alpha_i^{UB})$ as T increases.

Hence, there is a unique $\hat{T}$ below which profits if both enter are positive, in which case both firms enter in equilibrium with probability one, and above which profits if both enter are negative, in which case there are 3 equilibrium: a pure strategy equilibrium in which firm i enters with probability one and firm j does not enter, a pure strategy equilibrium in which firm j enters with probability one and firm i does

not enter, and a mixed strategy equilibrium in which firms i and j enter with some positive probability less than one.

The mixed strategy equilibrium makes a firm indifferent between entering and not entering. Let $p_j(T)$ denote the probability firm j enters when court delay is T. Due to symmetry of the problem, it is only necessary to find the value of $p_j(T)$ that makes firm i indifferent between entering and not entering; the same probability played by firm i will make firm j indifferent.

Not entering generates profits of 0. Entering generates profits of:

$$(1 - p_j(T))\Pi_i^* + p_j(T)\Pi_i(T) \tag{2.17}$$

which needs to equal zero in equilibrium. Hence the mixed strategy entry probability is:

$$p_j(T) = \frac{\Pi_i^*}{\Pi_i^* - \Pi_i(T)} \tag{2.18}$$

Because $\Pi_i(T)$ is decreasing in T, so is $p_j(T)$.

□

Proof of Theorem 3

Let firm i be superior to firm j. Conditional on entry of both firms, firm i has the equilibrium first order condition:

$$c_i'(1 - \alpha_i(T)) = \frac{\frac{1}{2}d(1 - f^T) + f^T d}{1 - \delta} \tag{2.19}$$

Conditional on entry of both firms, firm j has the equilibrium first order condition:

$$c_j'(1 - \alpha_j(T)) = \frac{\frac{1}{2}d(1 - f^T) + f^T d}{1 - \delta} \tag{2.20}$$

81

By strict convexity of the cost functions, each firm chooses a unique accident probability. Note that in order to satisfy the first-order conditions:

$$c_i'(1 - \alpha_i(T)) = c_j'(1 - \alpha_j(T)) \tag{2.21}$$

Because $c_j'(1 - x) > c_i'(1 - x)$ for all x (the definition of superiority), it must be the case that in equilibrium:

$$\alpha_j(T) > \alpha_i(T) \tag{2.22}$$

for all T in which both firms are active.

By the same logic in Theorem 1, when both firms are active, an increase in delay T causes an increase in both accident probabilities.

Applying the envelope theorem to the superior firm i's profits yields:

$$\frac{\partial \Pi_i}{\partial T} = -\frac{1}{2}\frac{\partial \alpha_j}{\partial T}[1 - f^T]d + \frac{\frac{1}{2}[\alpha_i(T) + \alpha_j(T)]\frac{\partial f^T}{\partial T}d}{1 - \delta} - \frac{\alpha_i \frac{\partial f^T}{\partial T}d}{1 - \delta} < 0 \tag{2.23}$$

This expression is negative because $\frac{\partial \alpha_j}{\partial T} > 0$, $\frac{\partial f^T}{\partial T} < 0$ and $\alpha_j(T) > \alpha_i(T)$.

On the other hand, applying the envelope to the inferior firm j's profits yields:

$$\frac{\partial \Pi_i}{\partial T} = -\frac{1}{2}\frac{\partial \alpha_i}{\partial T}[1 - f^T]d + \frac{\frac{1}{2}[\alpha_i(T) + \alpha_j(T)]\frac{\partial f^T}{\partial T}d}{1 - \delta} - \frac{\alpha_j \frac{\partial f^T}{\partial T}d}{1 - \delta} \tag{2.24}$$

Although the first term in the expression is negative, the net effect of the second and third term is positive because $\alpha_j(T) > \alpha_i(T)$. The sign of the above expression is ambiguous and depends on the parameter values of the model. Hence, the inferior firm's profits can increase or decrease with an increase in court delay.

If q is sufficiently high, then both the superior and inferior firm have positive profits for all T and enter with probability one.

□

Proof of Theorem 4

Because firm i is socially valuable, $\Pi_i^* = \Pi_i(0)$ is positive. Because firm j is socially wasteful, $\Pi_j^* = \Pi_j(0)$ is negative. By continuity of the profit function in T, there exists a $\hat{T}$ such that if $T < \hat{T}$, then $\Pi_j(T) < 0$, which means that firm j does not enter the market. Since firm i earns Π_i^*, i is active. Hence there is a separating equilibrium in which the efficient firm is active and the inefficient firm is inactive.

□

2.7 REFERENCES

[29] Anderlini, L, L Felli, and A Postelwaite; Should Courts Always Enforce What Contracting Parties Write? *Review of Law and Economics* 2011, Volume 7, Issue 1, pp 14-28

[30] Ayres I and R Gertner; Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules, *Yale Law Journal* 1989, Volume 99, Issue 1 pp 87-130

[31] Bebchuk, L and S Shavell; Information and the Scope of Liability for Breach of Contract: The Rule of Hadley V. Baxendale, *Journal of Law, Economics and Organization* 1991, Volume 7, Issue 2 pp 284-312

[32] Bonner, R; Lawyers and Litigants in Ancient Athens: The Genesis of the Legal Profession, W.S. Hein and Company, 1927

[33] Chappe, N; Demand for civil trials and court delay, *European Journal of Law and Economics* 2012, Volume 33, Issue 2, pp 342-357

[34] Chemin M; Does the Quality of the Judiciary Shape Economic Activity? Evidence from a Judicial Reform in India, *Journal of Law, Economics, and Organization* 2012 Volume 28, Issue 3 pp 460-485

[35] Coviello, D, A Ichino, and N Persico; Don't spread yourself too thin: The impact of task juggling on workers' speed of job completion, *NBER Working Paper* 2010

[36] Fishback, P and S Kantor; Compensation for Accidents before Workers' Compensation, *A Prelude to the Welfare State*, The University of Chicago Press, 2000

[37] Gravelle, H; Rationing Trials By Waiting: Welfare Implications, *International Review of Law and Economics* 1990, Volume 10, Issue 3 pp 255-270

[38] Landes, W; An Economic Analysis of the Court, *Journal of Law and Economics* 1971 Volume 14 Issue 1, pp 61-107

[39] Pace, N, R Reville, L Galway, A Geller, O Hayden, L Hill, C Mardesich, F Neuhauser, S Polich, J Yeom, and L Zakaras; Improving Dispute Resolution for California's Injured Workers, RAND Institute for Civil Justice, 2003

[40] Priest G; Private Litigants and the Court delay Problem, *Boston University Law Review* 1989, Volume 69, Issue 3, pp 527 - 559

[41] Vereeck, L and M Mühl' An Economic Theory of Court Delay, *European Journal of Law and Economics* 2000, Volume 10, Issue 3, pp 243-268