



Vanderbilt University Department of Economics Working Papers 16-00004

Behavioral Origins of Epidemiological Bifurcations

David Aadland
University of Wyoming

David Finnoff
University of Wyoming

Kevin X. D. Huang
Vanderbilt University

Abstract

In this paper, we investigate the nature of rational expectations equilibria for economic epidemiological models, with a particular focus on the behavioral origins of epidemiological bifurcations. Unlike mathematical epidemiological models, economic epidemiological models can produce regions of indeterminacy or instability around the endemic steady states. We consider SI, SIS, SIR and SIRS versions of economic compartmental models and show how well-intentioned public policy may contribute to disease instability and uncertainty.

Behavioral Origins of Epidemiological Bifurcations

David Aadland, David Finnoff and Kevin X.D. Huang*

January 2016

Abstract

In this paper, we investigate the nature of rational expectations equilibria for economic epidemiological models, with a particular focus on the behavioral origins of epidemiological bifurcations. Unlike mathematical epidemiological models, economic epidemiological models can produce regions of indeterminacy or instability around the endemic steady states. We consider SI, SIS, SIR and SIRS versions of economic compartmental models and show how well-intentioned public policy may contribute to disease instability and uncertainty.

JEL Codes: D1, I1.

Keywords: economic epidemiology, bifurcation, dynamics, disease, indeterminacy, rational expectations

*Aadland and Finnoff: Department of Economics and Finance, University of Wyoming, 1000 E. University Avenue, Laramie, WY, 82071. Huang: Department of Economics, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235. This publication was made possible in part by grant number 1R01GM100471-01 from the National Institute of General Medical Sciences (NIGMS) at the National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS.

1 Introduction

In this paper we investigate the behavioral origins of epidemiological bifurcations.¹ Human responses in the face of endogenous disease risk lead to bifurcations in epidemiological models with a variety of dynamic implications. The dynamic implications lead to the potential for unintended consequences from public health policies. As seen in the macroeconomic literature on monetary and fiscal policy (Clarida, Gali, and Gertler (2000)), we show how well-intentioned public health policy may contribute to multiple equilibria and aggregate instability. For instance, government policy designed to lower the transmission probability or raise the quality-of-life associated with infectious diseases may push these equilibria from being stable to exhibiting instability or indeterminacy. The latter also have the potential of contributing to self-fulfilling “sunspot” equilibria, which can contribute to the volatility and unpredictability of the system (Grandmont (1986); Smith (1989); Woodford (1986)). We show how public health interventions that reduce the consequences of infection can lead to a bifurcation with excess variability in prevalence for diseases such as the common cold and tuberculosis. We also show how the observability of a disease status leads to a type of dynamic externality imposed on the economic system that differs from the infection externality typically stressed in the literature (Kremer (1996); Gersovitz and Hammer (2004)).

The results indicate a richer set of implications for how human behavior impacts the spread of infectious diseases than typically found in economic epidemiological (EE) models. The economic epidemiology field integrates traditional mathematical epidemiology and rational economic decision making. Economic research in this area began in response to the AIDS epidemic and has led to an improved understanding of how decision making by individuals and policymakers influences infectious disease dynamics. For example, policymakers may have limited ability to eradicate infectious diseases if rational individuals respond to lower prevalence by reducing protection (Geoffard and Philipson (1996)) or may increase disease prevalence and induce fatalistic behavior with the introduction of imperfect vaccines (Kremer (1996)). These examples highlight the need to understand how economic incentives can alter policy prescriptions in the presence of infectious diseases.

Our focus is on how economic and epidemiological characteristics influence the dynamic properties of EE equilibria and the relationship to public health policy.² To the best of our knowledge, these are new findings in the EE literature and an additional reason for policymakers to consider the predictions of integrated economic and epidemiological models.

¹By epidemiological bifurcations, we are referring to the general idea that “...the state of the [epidemiological] system depends on some parameter usually with a concomitant change in stability” Shivamoggi (2014).

²The stability properties of continuous-time epidemiological systems have been studied in detail (see e.g., Korobeinikov and Wake (2002)). In general, the endemic equilibrium has been found to be globally stable.

2 An Economic Epidemiological Model

Following work by Philipson and Posner (1993) and Lightwood and Goldman (1995), we specify an integrated economic epidemiological model to describe individual behavior in the face of endogenous risk of infection from a communicable disease. In any period t , individual i from a constant population of N individuals is in one of three epidemiological states as indicated by the binary variables: susceptible ($s_{i,t}$), infected ($in_{i,t}$), or recovered and immune ($r_{i,t}$). The proportions of susceptible, infected and recovered individuals in the entire population are given by averaging over all i .³ A similar model is presented in Aadland, Finnoff, and Huang (2013) to examine the dynamics and potential eradication of syphilis.

The risk of infection is endogenous because individuals choose to engage in activities which affect their exposure to the disease. Additional exposure brings immediate satisfaction but also the risk of future infection, leading to a deterioration of the individual's health. We assume current utility of an individual depends on their choice of exposure in a period, $x_{i,t}$, and their health status $h_{i,t}$. As exposure is a direct choice and health status an indirect consequence of past choices, we assume utility is diminishing in exposure and linear in health status. Individuals maximize expected lifetime utility by choosing exposure in each period

$$\max_{x_{i,t}} E_t \sum_{j=0}^{\infty} \beta^j [\ln(x_{i,t+j}) + h_{i,t+j}] \quad (1)$$

where $0 < \beta < 1$ is the discount factor, E_t represents an individual's rational expectation of future outcomes conditional on all information dated t and earlier, \bar{x} is the maximum feasible amount of exposure per period. For tractability, we assume the health status of individuals in each epidemiological state is constant and the health status of susceptible (or recovered/immune) individuals exceeds that of infected individuals, $h^S > h^I$. All individuals maximize (1) without concern for the welfare of the general population, subject to the epidemiological dynamics.

2.1 Epidemiology

The epidemiological portion of the model describes the evolution of the three mutually exclusive disease categories measured as proportions of the overall population. We specify the most general SIRS model (Anderson and May (1991)), where individuals transition from being susceptible to infected to recovered (and immune) and then back to susceptible. The SIRS model has previously been used to model infectious

³The model is set in discrete time (Auld (2003), Lightwood and Goldman (1995)), where t indexes the decision interval. Allen (1994) finds that endemic equilibria from discrete-time epidemiological models have the potential to be stable, exhibit periodicity or be chaotic. Instability tends to be driven by high contact rates and high birth/death rates per time interval.

diseases such as syphilis and whooping cough (Grassly, Fraser, and Garnett (2005); Rohani, Zhong, and King (2010)). The SIRS model is sufficiently general to handle cases with permanent infection (SI diseases such as HIV/AIDS), diseases with recovery but no immunity (SIS diseases such as the common cold and active tuberculosis), and diseases with permanent immunity (SIR diseases such as measles and chicken pox).

The disease dynamics are represented by the system

$$s_{t+1} = \mu + (1 - p_t - \mu)s_t + \gamma r_t \quad (2)$$

$$in_{t+1} = (1 - v - \mu)in_t + p_t s_t \quad (3)$$

$$r_{t+1} = (1 - \gamma - \mu)r_t + v in_t, \quad (4)$$

where μ is the common birth/death rate, $1/\gamma$ is the average duration of immunity, v is the recovery rate, and p_t is the probability of infection for a susceptible individual, which depends on their exposure choices, $x_{i,t}$. If the disease is transmitted person-to-person such as an STD, then $x_{i,t}$ can be interpreted as the chosen number of partners. If the disease is vector-borne such as the common cold, measles or malaria, then $x_{i,t}$ can be interpreted as exposure time to the disease. Assuming independence across episodes of exposure, the probability that (identical) susceptible individuals become infected is

$$p_t = \Pr(\text{infection}) = 1 - (1 - \lambda in_t)^{x_{i,t}}, \quad (5)$$

where λ is the per unit exposure (or per partner) transmission rate (Kaplan (1990); Oster (2005)). The fact that p_t depends on the chosen degree of exposure distinguishes the analysis from standard mathematical epidemiology (ME).

Imposing the condition that all three categories sum to one, the model simplifies to a two-variable system in in and r :

$$in_{t+1} = (1 - p_t - v - \mu)in_t + p_t(1 - r_t) \quad (6)$$

$$r_{t+1} = (1 - \gamma - \mu)r_t + v in_t. \quad (7)$$

The SIR model sets $\gamma = 0$ so that individuals are permanently recovered and immune to the disease. The SIS and SI models omit the immunity category.

2.2 Decision Making

The value functions for each state – susceptible (S), infected (IN), and recovered (R) – are given by

$$V_t^S = \ln(x_t) + h^S + \beta E_t[p_t V_{t+1}^{IN} + (1 - p_t)V_{t+1}^S] \quad (8)$$

$$V_t^{IN} = \ln(\bar{x}) + h^{IN} + \beta E_t[v V_{t+1}^R + (1 - v)V_{t+1}^{IN}] \quad (9)$$

$$V_t^R = \ln(\bar{x}) + h^S + \beta E_t[\gamma V_{t+1}^S + (1 - \gamma)V_{t+1}^R], \quad (10)$$

where infected and immune individuals choose the maximum feasible amount of exposure, \bar{x} , because they face no risk of immediate infection (Geoffard and Philipson (1996)). Similar to Gersovitz and Hammer (2004), this behavior imposes an externality on susceptible individuals. We discuss this in detail later in the paper. In contrast, susceptible individuals will choose the degree of exposure that balances the current marginal benefits of exposure with the discounted expected future marginal costs of exposure as given by the Euler equation:

$$x_t^{-1} = \beta p_{x,t} E_t[V_{t+1}^S - V_{t+1}^{IN}], \quad (11)$$

where the partial derivative of p_t with respect to the degree of exposure is $p_{x,t} = -\ln(1-p_t)(1-p_t)/x_{it}$.⁴ The left side of (11) gives the marginal benefit of exposure, while the right side of (11) gives the expected future marginal costs of exposure. The implication from (11) is that individuals will choose to expose themselves until the marginal benefits from additional exposure cease to outweigh the marginal costs of exposure. However, unlike ME models where the infection rate is constant or varies deterministically (Korobeinikov (2006)), the term $p_{x,t}$ highlights the fact that the transition probability from susceptible to infected is endogenous in EE models. Aadland, Finnoff, and Huang (2013) demonstrate the implications of the decision rule in (11). While the marginal benefits of exposure diminish with additional exposure, the marginal costs of exposure rise (fall) when exposure is low (high). This curious shape follows from exposure increasing the probability of infection, putting an upward pressure on marginal costs, but as the probability rises its rate of change decreases, pushing marginal costs back down. At low levels of exposure the former effect overwhelms the latter effect, while the reverse occurs at high levels of exposure. This type of endogenous risk creates a non-convexity.

⁴The standard sufficiency conditions for this epidemiology problem do not hold (Goenka and Liu (2012); Gersovitz and Hammer (2004)). Therefore, we use numerical simulation to check the optimality of the transition path given by equation (11).

2.2.1 Endogenous Risk, Non-Convexities and Multiple Equilibria

The implication from the non-convexity is the potential for multiple equilibria, which is well-known in economic epidemiological models (Goldman and Lightwood (2002)). What is not so well known are the dynamic consequences of the non-convexities. They cause both bifurcations in equilibria and have the potential to alter the stability properties of equilibria. Bifurcations can occur from changes in any parameter that alter the marginal benefit or cost of exposure. We focus on two parameters that are likely levers of public policy, the per unit transmission rate, λ , and the health status of the infected state, h^{IN} .

Consider an SIS variant of the model, with parameter values as shown in Table 1 (based on those in Aadland, Finnoff, and Huang (2013)). These values are used to create Figure 1a, which plots the steady-state marginal benefit and marginal cost curves across a variety of values of λ and h^{IN} .

Table 1. SIS Parameter Values

Parameters	β	μ	v	\bar{x}	λ	h^{IN}	h^S
Value	0.96	0.05	1	100	0.5	20	50

The value of β implies a 4% annual discount rate, μ gives a 5% birth and death rate for the population (Garnett et al. (1997)), and the maximum annual number of exposures is set at 100. For each parameter combination there are up to two equilibria, one with a low degree of exposure (equilibrium points 1, 3 and 5 in Figure 1) and one with a high degree of exposure (equilibrium points 2 and 4 in Figure 1). Changes in economic and epidemiological characteristics that shift either curve can lead to bifurcations and complex dynamics.

SIS Equilibria For example, in the right Panel (a) of Figure 1, increases in the health status of the infected state (i.e., a reduction in the consequences of being infected, an increase in h^{IN} from 20 to 40) scale the marginal cost curve downward. The marginal costs associated with all levels of exposure decline, yet the exposure required to attain a certain level of marginal costs increases for low exposures, and falls for high exposures. The result is a small increase in the low-exposure equilibrium (such as the movement from equilibrium 1 to 3) and a larger decrease in the high-exposure equilibrium (movement from equilibrium 2 to 4). Further increases in h^{IN} – beyond where the curves are tangent – cause the marginal benefits of exposure to be everywhere greater than the marginal costs (as seen with $h^{IN} = 43$). Here, a bifurcation occurs and the two equilibrium points merge into one, leading susceptible individuals to choose the maximum level of exposure.

Public policies that reduce the per unit transmission rate λ provide an interesting contrast. While the

increase in h^{IN} scaled marginal costs downward, a policy that reduces λ leads to a flattening and shifting of the marginal cost curve, as shown in the left Panel (a) of Figure 1. Lower λ 's lead to a reduction in marginal costs for low exposures and an increase in marginal costs for high exposures. The result is that lower λ 's cause increases in exposure for both the low-exposure and high-exposure equilibria (i.e., the movements from equilibrium 1 to 3 to 5 and the movement from equilibrium 2 to 4).

SIS Dynamics and Stability Analysis Figure 1b demonstrates the bifurcations that occur for variations in λ or h^{IN} (holding the other constant at the benchmark values of $\lambda = 0.5$ or $h^{IN} = 20$). The figure plots both equilibria and notes their local stability properties using the method of Blanchard and Kahn (1980). The equilibria can exhibit local saddle-path stability, indeterminate paths (infinite number of paths), or unstable explosive paths. In the presence of saddle-path stability, public intervention can play a meaningful, predictable role. In the presence of indeterminacy, there are multiple equilibrium paths and the possibility of “sunspot” equilibria (Benhabib and Farmer (1999)). Sunspot equilibria are often associated with self-fulfilling prophecies and additional aggregate volatility.

Figure 1b illustrates that variations in the key policy parameters lead to significantly different equilibrium exposures and a wide variety of approach paths. Reductions in λ below the benchmark lead to increased equilibrium exposures and differing approach paths. The approach paths at the benchmark levels are both saddle (unique) paths, but as λ is lowered, the path to the upper equilibrium switches to being indeterminate. In contrast, increasing h^{IN} above the benchmark pulls the two equilibria together and the saddle paths switch to unstable spirals. As h^{IN} approaches the health status of a susceptible individual (h^S), the equilibria merge to become a single unstable equilibrium. Thus, public policy aimed at lowering λ or raising h^{IN} may induce indeterminacy or aggregate instability, in either case lowering the predictability of policy interventions and leading to increased fluctuations in disease prevalence.

2.2.2 SIRS Immunity Externality

Human beliefs over diseases, or lack thereof, lead to significant changes in behavior and dynamics of the system. This is especially clear for SIRS diseases such as pertussis or syphilis. Although these diseases have an interval of immunity that is typically known by the scientific or medical community, it is not often well known in the bulk of the population. Whether or not individuals know (or believe) they will experience an interval of immunity leads to a shift in the marginal cost of exposure. Specifically, when recovered individuals know or can observe their own immunity, they rationally choose the maximum feasible degree

of exposure \bar{x} and have health level h^S . Susceptible individuals under the same set of beliefs, on the other hand, choose a specific degree of exposure x_t to satisfy

$$x_t^{-1} = \beta p_{x,t} E_t \left[\ln(x_{t+1}/\bar{x}) + h + \frac{(1-v-p_{t+1})}{x_{t+1} p_{x,t+1}} - \beta \Delta_{t+2} \right], \quad (12)$$

where

$$\Delta_{t+2} = \frac{v\gamma}{x_{t+2} p_{x,t+2}} + (1-v-\gamma) \left[\ln\left(\frac{x_{t+2}}{\bar{x}}\right) + \frac{1-p_{t+2}}{x_{t+2} p_{x,t+2}} \right] + (1-\gamma) \left[h - \frac{1}{\beta x_{t+1} p_{x,t+1}} \right]$$

and $h = h^S - h^{IN}$ is the health gap between being susceptible and infected. See Appendix A for a derivation of equation (12). The term Δ_{t+2} captures the expected future personal “costs” of infection associated with acquired immunity. Because Δ_{t+2} enters the right side of (12) with a negative sign, the possibility of future immunity is a personal benefit of becoming infected. In the decentralized equilibrium considered here, immunity (if known) thus imposes an additional externality on society. The externality is dynamic and different than the infection externalities discussed in Gersovitz and Hammer (2004). Here, forward-looking susceptible individuals expose themselves to more risk when the possibility of immunity is known, because future infection also carries along the benefit of future immunity.

In contrast, if individuals do not know or believe they will have an interval of immunity (i.e., unobserved immunity), individuals with immunity (i.e., the R state) believe they are susceptible, and susceptibles do not account for being immune in their decision making. In this case the term Δ_{t+2} is equal to zero and equation (12) can be rewritten as

$$x_t^{-1} = \beta p_{x,t} E_t \left[\ln(x_{t+1}/\bar{x}) + h + \frac{(1-v-p_{t+1})}{x_{t+1} p_{x,t+1}} \right]. \quad (13)$$

Figure 2 demonstrates the implications of the externality. If the externality is internalized by the individual (i.e., they observe immunity) it lowers the marginal costs of exposure (similar to the effect of increasing h^{IN} in the SIS model, Figure 1a). For the sketch given, this results in the low-exposure choice increasing, and the high-exposure choice decreasing. While there remains two alternative equilibrium choices, the magnitudes are altered in different directions and to different degrees.

2.3 Complete SIRS Bifurcation and Stability Analysis

To perform a complete bifurcation and stability analysis⁵, we extend the analysis to cover several variants of the models across all combinations of the policy parameters, λ and h^{IN} . Similar to Goenka, Liu, and Nguyen (2012), we note the existence of an eradication equilibrium but focus on endemic EE equilibria.⁶ The analysis uses the parameter values in Table 1 and a few additional parameters for SIR and SIRS models. We consider treatable SIR and SIRS diseases with a 100% recovery rate within a year of infection (i.e., $v = 1$). Following Garnett et al. (1997), the SIRS variant is calibrated so individuals experience a five-year period of protective immunity (i.e., $\gamma = 0.2$).

Figures 3-5 summarize the stability and bifurcation of the exposure equilibria, across ranges of the per unit transmission rate, λ , and infected health, h^{IN} , for all of the SI, SIS, SIR and SIRS models. Infected health h^{IN} can be lowered by public policy that results in the discovery and introduction of drug treatments, while the per unit infection rate λ can be lowered through the introduction of vaccines or new prevention technologies. Our main findings are:

- EE models typically produce two endemic steady states. Aggregate welfare is always higher along the transition path to the low-exposure endemic steady state. For the SIRS model, the dynamic disease externality leads to increased risk-taking by susceptible individuals (along the transition path to the low-exposure steady state). Rational, forward-looking individuals who know that future infection will also carry a period of immunity will choose to be riskier than if they were unaware of the immunity.
- Increasing infected health status h^{IN} or reducing the per unit infection rate λ increases aggregate welfare for most model variants, yet can move the system from a unique stable equilibrium to one that exhibits aggregate instability and indeterminacy.

We now turn to more detailed analyses for each variant of the EE model.

⁵See Appendix B for the complete economic SIRS system used in the bifurcation and stability analysis.

⁶There is also an eradication steady state where $in = r = 0$, $s = 1$, and $x = \bar{x}$. For the range of parameter values we consider, the economic eradication steady state is locally unstable because susceptible individuals have no incentive to reduce the degree of exposure. More specifically, the local stability condition requires that the basic reproduction number, R_0 , is less than one (Anderson and May (1991)). The basic reproduction number is defined as the number of secondary infections generated by a single infected individual in an otherwise susceptible population. For the SIRS model described above, we have $R_0 = p/(in(v + \mu))$. Using L'Hôpital's rule, this reduces to $R_0 = \lambda_p \bar{x}/(v + \mu)$, which is greater than one for all the parameter combinations considered below.

2.3.1 Economic SI and SIS Models

Figure 3 shows the map of path types for the SI and SIS models around the low-exposure endemic steady state (left panels) and high-exposure endemic steady state (right panels). For high values of infected health h^{IN} , there is a single endemic steady state. In this case, the dynamics of the system are evaluated around the maximum feasible level of exposure, \bar{x} , because the marginal benefits of exposure are everywhere greater than the marginal costs. The top panels show the type of localized dynamic paths for the economic SI model with no available treatment, $v = 0$. The majority of the parameter space for the economic SI model is defined by saddle-path equilibria. For a given initial prevalence level (in_0), there is a unique initial exposure choice (x_0) that puts the system on a convergent path to the endemic steady state. All other initial exposure levels lead to divergent paths that violate non-negativity or non-explosion conditions. Because both the low-exposure and high-exposure steady states exhibit local saddle-path stability, the system has the potential to gravitate toward either steady state. The welfare contours show that, in the long run, society is better off at the low-exposure, low-prevalence steady state.⁷

The bottom panels of Figure 3 show the dynamic path types for the SIS model where infected individuals have access to perfectly effective treatment ($v = 1$) and return to the susceptible pool after treatment. The parameter region for the low-exposure steady state is primarily a saddle-path equilibrium, but there is also an explosive region for the low-exposure steady state at high levels of h^{IN} . Public policy aimed at improving the health of infected individuals could inadvertently move the system from a stable saddle-path region to an explosive system with higher waves of prevalence as individuals rationally take more risk.⁸

To gain intuition for the types of dynamic paths in the economic SIS models, consider the exposure elasticity with respect to prevalence near the endemic steady state, $\kappa = \frac{\partial x_{ss}}{\partial in_{ss}} \frac{in_{ss}}{x_{ss}}$. Prevalence elasticity measures the percentage change in exposure for a one-percent change in prevalence. This elasticity is generally negative, indicating that susceptible individuals respond to the increased risk of infection by choosing less exposure.⁹ For example, the prevalence elasticities for the economic SIS system depicted in the lower left panel of Figure 3 are negative and inelastic in the saddle-path region. In the explosive band (unstable spiral) the magnitude of κ (for the interior solutions) flips to being positive (an increase in prevalence leads to an increase in exposure) with elastic and inelastic magnitudes.

⁷Total welfare is calculated at steady state using a weighted average of the value functions for the three disease types: $s \cdot V^S + in \cdot V^{IN} + r \cdot V^R$.

⁸Mathematical epidemiological models, by contrast, do not vary the number of partners in response to changes in disease prevalence. For reasonable partner rates, the mathematical SI and SIS models are characterized by stable dynamic paths with a fixed number of partners.

⁹Kremer (1996) discusses the possibility of a positive prevalence elasticity and fatalistic behavior.

2.3.2 Economic SIR and SIRS Models with Observable Immunity

Figure 4 shows a similar map for the dynamic paths in SIR and SIRS models where immunity is observable. The top panels show the dynamic equilibrium types for the economic SIR model with permanent immunity, $\gamma = 0$. For all combinations of λ and h^{IN} , there is a single endemic steady state. For most of the parameter region, susceptible individuals choose the maximum level of exposure and the system is locally stable. This makes intuitive sense because susceptible individuals know that if they become infected, they can receive immediate treatment and enjoy a lifetime of disease immunity. There is a small range of indeterminacy for low levels of infected health and low levels of the infection parameter.

The lower panels show the path types for the economic SIRS model with $\gamma = 0.2$ (i.e., average immunity duration of five years). Unlike the economic SIR model, the economic SIRS model produces large regions of indeterminacy where there are multiple approach paths and the possibility of “sunspot” equilibria (Benhabib and Farmer (1999)). Considering that welfare is higher in the lower left panel, public policy aimed at improving the quality of life for individuals infected with diseases that have known temporary immunity may induce aggregate instability and indeterminacy.

2.3.3 Economic SIR and SIRS Models with Unobservable Immunity

Figure 5 depicts the SIRS counterpart to Figure 4 but with unobservable immunity. Unobservable immunity causes two primary changes. First, individuals in the economic SIR system now choose less exposure for any parameter combination. Knowledge of perfectly effective treatment and permanent immunity, as depicted in Figure 4, greatly reduces the future cost of current risky behavior. Second, the indeterminacy region for the SIRS system now covers a much smaller range of the health parameter. While well-intentioned public policy that improves the quality of life for infected individuals leads to equilibria with higher welfare, indeterminate paths that approach the equilibrium with spiral trajectories imply aggregate volatility and potentially wide swings in prevalence.

3 Nonlinear Dynamic Analysis

To illustrate the full dynamic implications of the public policies, here we simulate a nonlinear economic SIS model. As the epidemiology choice problem is non-convex (Goenka and Liu (2012)), we use GAMS to solve for the nonlinear transition paths as a dynamic mixed complementary problem. Overall, the nonlinear dynamic analysis confirms the linearized, local analysis performed earlier. More specifically, the analysis

confirms that public policy which reduces λ or increases h^{IN} , while successful at improving welfare, has the potential to increase the volatility in disease prevalence and make the transition path less predictable.

Figure 6 plots the approach paths for prevalence (in_t) and exposure (x_t) that correspond from changes in λ and h^{IN} . In each panel of Figure 6, the equilibrium points of Figure 1 are shown along with the transition paths (where arrows indicate the direction of change). The parameter changes in λ and h^{IN} cause the low-exposure equilibrium to move from equilibrium (1) to (3) and the high-exposure equilibrium to move from equilibrium (2) to (4). The 45 degree lines represent time-invariant levels of each variable.

Start with the policies that reduce λ from the benchmark value of 0.5 to 0.15, as shown in panel (a). Figures 1 and 3 classify the low-exposure equilibrium in both cases as being saddle-path stable (points 1 and 3 in Figure 1b), while the high-exposure equilibrium switches from a saddle path to indeterminate (points 2 and 4 in Figure 1b). The nonlinear dynamics for the low-exposure equilibrium are simple – there is a small increase in equilibrium prevalence and exposure with a quick convergence from point 1 to 3. The transition path for the high-exposure equilibrium demonstrates the issues of indeterminacy - the reduced per unit transmission leads to a large increase in exposure (point 2 to 4), and wide fluctuations in exposure and prevalence as the system converges. As shown by Figure 7, the reduction in λ leads to an increase in welfare for both equilibria with the low-exposure equilibrium path preferred.

The implications from policies to increase the quality of life for those infected, h^{IN} , are shown in panels (b) and (c) of Figure 6. For an increase in h^{IN} from the benchmark to a level of 40 in panel (b), both equilibria remain on the saddle paths, converging to the new equilibria. At the low-exposure equilibrium, prevalence and exposure increase; more individuals are infected but their quality of life has significantly improved and aggregate welfare increase (as shown in Figure 7). At the high-exposure equilibrium, prevalence and exposure decrease with a similar increase in welfare. Similar to above, the low-exposure equilibrium path is preferred. Increasing h^{IN} further to a level high enough that the two equilibria merge into one (i.e., $h^{IN} = 0.43$) as shown in panel (c), causes individuals to choose the maximum level of exposure and the equilibrium to be unstable and volatile.

4 Conclusions

Economic epidemiology has made significant advances in educating health officials about the behavioral implications of public policies. However, one area that has received little attention is how policy and human

responses influences the nature of communicable disease dynamics. In this paper, we explore the nature of the dynamics for rational expectation economic epidemiological systems. The analysis digs beneath a comparison of fixed parameter values and demonstrates the behavioral origin for epidemiological bifurcations. Indeed, we show that well-intentioned policy has the potential to create instability and indeterminacy when individuals behave rationally and in a self-interested manner. Future research should focus on providing precise policy recommendations by considering the costs of policy and applying the methods outlined in this paper to specific diseases.

References

- AADLAND, D., D. FINNOFF, AND K. HUANG (2013): “Syphilis Cycles,” *The B.E. Journal of Economic Analysis and Policy*, 14, 297–348.
- ALLEN, L. (1994): “Some discrete-time SI, SIR, and SIS epidemic models,” *Mathematical biosciences*, 124(1), 83–105.
- ANDERSON, R., AND R. MAY (1991): *Infectious Diseases of Humans, Dynamics and Control*. Oxford University Press.
- AULD, M. (2003): “Choices, beliefs, and infections disease dynamics,” *Journal of Health Economics*, 22, 361–377.
- BENHABIB, J., AND R. FARMER (1999): “Indeterminacy and sunspots in macroeconomics,” *Handbook of macroeconomics*, 1, 387–448.
- BLANCHARD, O. J., AND C. M. KAHN (1980): “The solution of linear difference models under rational expectations,” *Econometrica*, 48(5), 1305–1311.
- CLARIDA, R., J. GALI, AND M. GERTLER (2000): “Monetary policy rules and macroeconomic stability: evidence and some theory,” *Quarterly journal of economics*, 115(1), 147–180.
- GARNETT, G., ET AL. (1997): “The Natural History of Syphilis: Implications for the Transition Dynamics and Control of Infection,” *Sexually Transmitted Diseases*, 24(4), 185–200.
- GEOFFARD, P., AND T. PHILIPSON (1996): “Rational epidemics and their public control,” *International Economic Review*, 37(3), 603–624.

- GERSOVITZ, M., AND J. HAMMER (2004): “The Economical Control of Infectious Diseases,” *The Economic Journal*, 114, 1–27.
- GOENKA, A., AND L. LIU (2012): “Infectious diseases and endogenous fluctuations,” *Economic Theory*, 50(1), 125–149.
- GOENKA, A., L. LIU, AND M.-H. NGUYEN (2012): “Infectious Diseases and Endogenous Growth,” Discussion paper, Mimeo: National University of Singapore.
- GOLDMAN, S., AND J. LIGHTWOOD (2002): “Cost Optimization in the SIS Model of Infectious Disease with Treatment,” *Topics in Economic Analysis and Policy*, 2(1), 1–22.
- GRANDMONT, J.-M. (1986): *On endogenous competitive business cycles*. Springer.
- GRASSLY, N., C. FRASER, AND G. GARNETT (2005): “Host immunity and synchronized epidemics of syphilis across the United States,” *Nature*, 433, 417–421.
- KAPLAN, E. (1990): “Modeling HIV infectivity: must sex acts be counted?,” *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 3(1), 55.
- KOROBENIKOV, A. (2006): “Lyapunov functions and global stability for SIR and SIRS epidemiological models with non-linear transmission,” *Bulletin of Mathematical Biology*, 68(3), 615–626.
- KOROBENIKOV, A., AND G. WAKE (2002): “Lyapunov functions and global stability for SIR, SIRS, and SIS epidemiological models,” *Applied Mathematics Letters*, 15(8), 955–960.
- KREMER, M. (1996): “Integrating behavioral choice into epidemiological models of AIDS,” *Quarterly Journal of Economics*, 111(2), 549–573.
- LIGHTWOOD, J., AND S. GOLDMAN (1995): “The SIS Model of Infectious Disease with Treatment,” Unpublished Manuscript.
- OSTER, E. (2005): “Sexually transmitted infections, sexual behavior, and the HIV/AIDS epidemic,” *Quarterly Journal of Economics*, 120(2), 467–515.
- PHILIPSON, T., AND R. POSNER (1993): *Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective*. Harvard University Press.
- ROHANI, P., X. ZHONG, AND A. KING (2010): “Contact network structure explains the changing epidemiology of pertussis,” *Science*, 330(6006), 982.

SHIVAMOGGI, B. K. (2014): *Nonlinear dynamics and chaotic phenomena: an introduction*, vol. 103. Springer.

SMITH, B. D. (1989): “Legal restrictions, sunspots, and cycles,” *Journal of Economic Theory*, 47(2), 369–392.

WOODFORD, M. (1986): “Stationary sunspot equilibria in a finance constrained economy,” *Journal of Economic Theory*, 40(1), 128–137.

Appendix A. Derivation of the Economic SIRS Euler Equation with Observable Immunity

Here we derive the Euler equation for the economic SIRS model with observable immunity. To begin, note that equations (9) and (10) imply

$$V_t^R - V_t^{IN} = h + \beta E_t [\gamma(V_{t+1}^S - V_{t+1}^R) + (1 - v)(V_{t+1}^R - V_{t+1}^{IN})], \quad (\text{A.1})$$

while equations (8) and (9) imply

$$V_t^S - V_t^{IN} = \ln(x_t/\bar{x}) + h + \beta E_t [(1 - p_t)(V_{t+1}^S - V_{t+1}^{IN}) - v(V_{t+1}^R - V_{t+1}^{IN})]. \quad (\text{A.2})$$

Using equation (11), we have

$$E_t(V_{t+1}^S - V_{t+1}^{IN}) = (\beta x_t p_{x,t})^{-1}, \quad (\text{A.3})$$

for all t . Next, rearrange (A.2) as

$$V_{t+1}^R - V_{t+1}^{IN} = \frac{1}{\beta v} [\ln(x_t/\bar{x}) + h] + \frac{1}{v} (1 - p_t) E_t (V_{t+1}^S - V_{t+1}^{IN}) - \frac{1}{\beta v} (V_t^S - V_t^{IN}). \quad (\text{A.4})$$

Take E_{t-1} on both sides of (A.4) and substitute (A.3) to get

$$E_{t-1}(V_{t+1}^R - V_{t+1}^{IN}) = \frac{1}{\beta v} E_{t-1} [\ln(x_t/\bar{x}) + h] + \frac{1}{\beta v} E_{t-1} \left(\frac{1 - p_t}{x_t p_{x,t}} \right) - \frac{1}{\beta^2 v} \left(\frac{1}{x_{t-1} p_{x,t-1}} \right). \quad (\text{A.5})$$

Now rewrite equation (A.1) as

$$V_t^R - V_t^{IN} = h + \beta E_t [\gamma(V_{t+1}^S - V_{t+1}^{IN}) + (1 - v - \gamma)(V_{t+1}^R - V_{t+1}^{IN})]. \quad (\text{A.6})$$

Move (A.6) ahead one period, take E_{t-1} of both sides, and set equal to (A.5) to get

$$\begin{aligned} & \frac{1}{\beta v} E_{t-1} [\ln(x_t/\bar{x}) + h] + \frac{1}{\beta v} E_{t-1} \left(\frac{1 - p_t}{x_t p_{x,t}} \right) - \frac{1}{\beta^2 v} \left(\frac{1}{x_{t-1} p_{x,t-1}} \right) \\ &= h + \beta E_{t-1} \left\{ \gamma (\beta x_t p_{x,t})^{-1} + (1 - v - \gamma) \left(\frac{1}{\beta v} [\ln(x_t/\bar{x}) + h] + \frac{1}{\beta v} \left(\frac{1 - p_t}{x_t p_{x,t}} \right) - \frac{1}{\beta^2 v} \left(\frac{1}{x_{t-1} p_{x,t-1}} \right) \right) \right\}. \end{aligned} \quad (\text{A.7})$$

Impose perfect foresight, move ahead one period, and rearrange to get

$$x_t^{-1} = \beta p_{x,t} \left[\ln(x_{t+1}/\bar{x}) + h + \frac{(1 - v - p_{t+1})}{x_{t+1} p_{x,t+1}} - \beta \Delta_{t+2} \right],$$

where

$$\Delta_{t+2} = \frac{v\gamma}{x_{t+2}p_{x,t+2}} + (1-v-\gamma) \left[\ln\left(\frac{x_{t+2}}{\bar{x}}\right) + \frac{1-p_{t+2}}{x_{t+2}p_{x,t+2}} \right] + (1-\gamma) \left[h - \frac{1}{\beta x_{t+1}p_{x,t+1}} \right].$$

Appendix B. SIRS Economic Epidemiological (EE) Steady-State and Matrix Systems

Here, we describe the steady state EE system and the linearized EE matrix system used in the bifurcation and stability analyses. The endemic steady states solve time-invariant versions of (6), (7), and the Euler equation. The Euler equation either takes the form of (12) when an indicator variable set at $\phi = 1$ (observable immunity) or the form of (13) when $\phi = 0$ (unobservable immunity). The steady-state system can therefore be rewritten as three equations:

$$in = p(1 - in - r)/(v + \mu) \tag{B.1}$$

$$r = vin/(\gamma + \mu) \tag{B.2}$$

$$x^{-1} = \beta[p_x(\ln(x/\bar{x}) + h - \phi\beta\Delta) + (1 - v - p)/x] \tag{B.3}$$

in three unknown variables (in, r, x) , where the immunity externality is given by

$$\Delta = \frac{1}{p_x x} [v\gamma + (1 - v - \gamma)(1 - p) - (1 - \gamma)/\beta] + (1 - v - \gamma) \ln(x/\bar{x}) + (1 - \gamma)h.$$

Similar to Goenka, Liu, and Nguyen (2012), we also note the existence of an eradication steady and focus on the local stability properties around the endemic steady states.

To analyze these transition dynamics, we linearize around the endemic steady states:

$$\widehat{in}_{t+1} = (1 - v - \mu - p)\widehat{in}_t + (1 - in - r)\widehat{p}_t - p\widehat{r}_t \tag{B.4}$$

$$\widehat{r}_{t+1} = (1 - \gamma - \mu)\widehat{r}_t + v\widehat{in}_t, \tag{B.5}$$

where hats ($\hat{\cdot}$) over the variables indicate deviation from one of the steady states. The linearized Euler equation is:

$$p_x \hat{x}_t + x \hat{p}_{x,t} = \beta p_x (1 - v - p - xp_x) E_t \hat{x}_{t+1} + \beta x (1 - v - p) E_t \hat{p}_{x,t+1} + \beta x p_x E_t \hat{p}_{t+1} \quad (\text{B.6})$$

$$+ \phi \beta^2 E_t \left\{ \begin{array}{l} p_x [v\gamma + (1 - v - \gamma)(1 - p - xp_x)] \hat{x}_{t+2} + x [v\gamma + (1 - v - \gamma)(1 - p)] \hat{p}_{x,t+2} + \\ [(1 - v - \gamma)xp_x] \hat{p}_{t+2} - [(1 - \gamma)p_x/\beta] \hat{x}_{t+1} - [(1 - \gamma)x/\beta] \hat{p}_{x,t+1} \end{array} \right\}$$

where

$$\hat{p}_t = p_{in} \hat{in}_t + p_x \hat{x}_t \quad (\text{B.7})$$

$$\hat{p}_{x,t} = [(1 + \ln[1 - p])/x] \hat{p}_t - (p_x/x) \hat{x}_t \quad (\text{B.8})$$

and

$$p_{in} = x\lambda(1 - \lambda in)^{x-1} \quad (\text{B.9})$$

$$p_x = -\ln(1 - p)(1 - p)/x. \quad (\text{B.10})$$

In matrix form, the EE system can be written as

$$\hat{z}_t = J \hat{z}_{t+1}, \quad (\text{B.11})$$

where $\hat{z}_t = (\hat{x}_t, \hat{in}_t, \hat{r}_t)'$ when $\phi = 0$ or $\hat{z}_t = (\hat{x}_t, \hat{in}_t, \hat{r}_t, \hat{x}_{t+1}, \hat{in}_{t+1})'$ when $\phi = 1$, and J is the transition matrix.

Specifically, if we impose perfect foresight, the $\phi = 0$ linearized EE matrix system can be written as:

$$\begin{aligned} & \underbrace{\begin{bmatrix} 0 & 1 - v - \mu - p & -p \\ 0 & v & 1 - \gamma - \mu \\ p_x & 0 & 0 \end{bmatrix}}_A \begin{bmatrix} \hat{x}_t \\ \hat{in}_t \\ \hat{r}_t \end{bmatrix} + \underbrace{\begin{bmatrix} 1 - in - r & 0 \\ 0 & 0 \\ 0 & x \end{bmatrix}}_B \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \end{bmatrix} \\ = & \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \beta p_x(1 - v - p - xp_x) & 0 & 0 \end{bmatrix}}_C \begin{bmatrix} \hat{x}_{t+1} \\ \hat{in}_{t+1} \\ \hat{r}_{t+1} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \beta x p_x & \beta x(1 - v - p) \end{bmatrix}}_D \begin{bmatrix} \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix} \quad (\text{B.12}) \end{aligned}$$

and

$$\underbrace{\begin{bmatrix} -1 & 0 \\ -(1 + \ln(1-p))/x & 1 \end{bmatrix}}_F \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \end{bmatrix} = - \underbrace{\begin{bmatrix} p_x & p_{in} & 0 \\ p_x/x & 0 & 0 \end{bmatrix}}_G \begin{bmatrix} \hat{x}_t \\ \hat{in}_t \\ \hat{r}_t \end{bmatrix}. \quad (\text{B.13})$$

When $\phi = 1$, we have

$$\begin{aligned} & \underbrace{\begin{bmatrix} 0 & 1-v-\mu-p & -p & 0 & 0 \\ 0 & v & 1-\mu-\gamma & 0 & 0 \\ p_x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_A \begin{bmatrix} \hat{x}_t \\ \hat{in}_t \\ \hat{r}_t \\ \hat{x}_{t+1} \\ \hat{in}_{t+1} \end{bmatrix} + \underbrace{\begin{bmatrix} s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_B \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \\ \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix} \\ = & \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \beta p_x(1-v-p-xp_x) - \beta^2[(1-\gamma)p_x/\beta] & 0 & 0 & \beta^2 p_x[v\gamma + (1-v-\gamma)(1-p-xp_x)] & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}}_C \begin{bmatrix} \hat{x}_{t+1} \\ \hat{in}_{t+1} \\ \hat{r}_{t+1} \\ \hat{x}_{t+2} \\ \hat{in}_{t+2} \end{bmatrix} + \\ & \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta x p_x & \beta x(1-v-p) - \beta^2[(1-\gamma)x/\beta] & \beta^2(1-v-\gamma)x p_x & \beta^2 x[v\gamma + (1-v-\gamma)(1-p)] \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_D \begin{bmatrix} \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \\ \hat{p}_{t+2} \\ \hat{p}_{x,t+2} \end{bmatrix} \end{aligned}$$

and

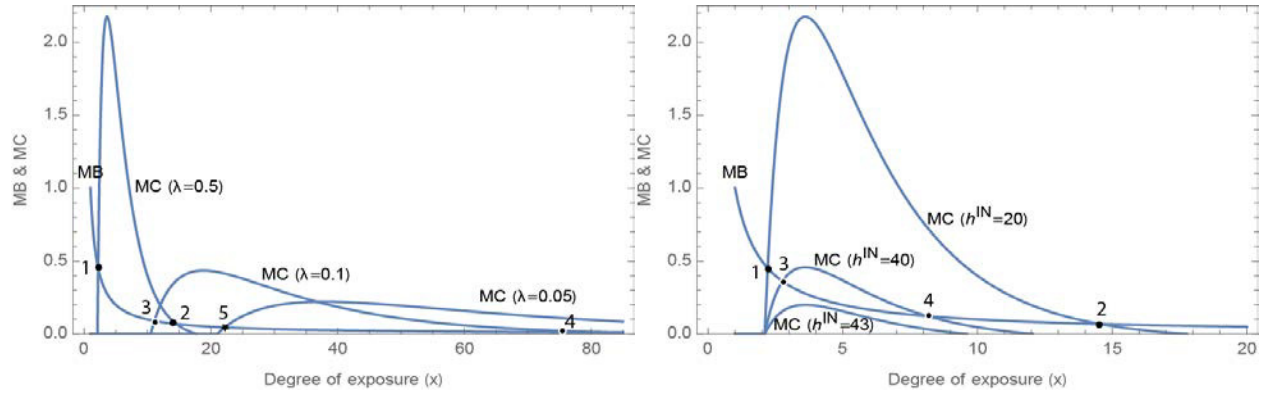
$$\underbrace{\begin{bmatrix} -1 & 0 & 0 & 0 \\ -(1 + \ln(1-p))/x & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -(1 + \ln(1-p))/x & 1 \end{bmatrix}}_F \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \\ \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix} = - \underbrace{\begin{bmatrix} p_x & p_{in} & 0 & 0 & 0 \\ p_x/x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_x & p_{in} \\ 0 & 0 & 0 & p_x/x & 0 \end{bmatrix}}_G \begin{bmatrix} \hat{x}_t \\ \hat{in}_t \\ \hat{r}_t \\ \hat{x}_{t+1} \\ \hat{in}_{t+1} \end{bmatrix}. \quad (\text{B.14})$$

where

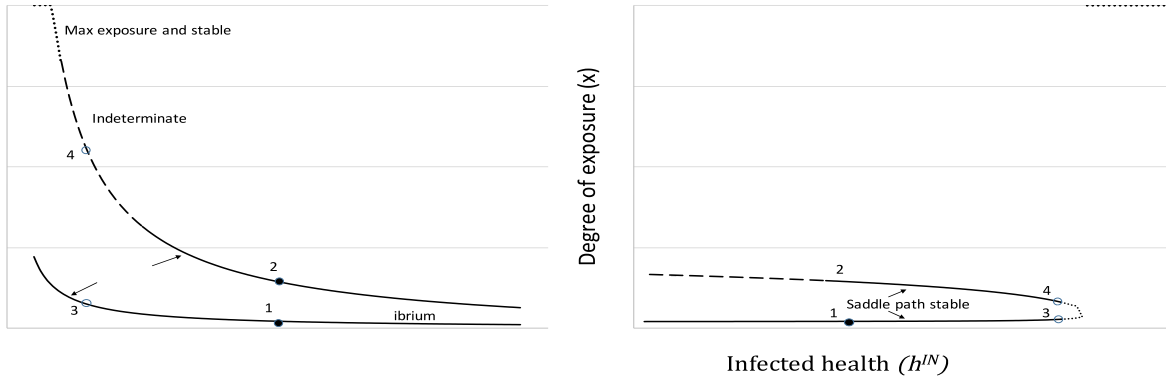
$$J = (A - BF^{-1}G)^{-1}(C - DF^{-1}G). \quad (\text{B.15})$$

We use the method of Blanchard and Kahn (1980) to analyze the nature of the rational expectation EE equilibrium. When $\phi = 0$ the three-variable system contains one jump (\hat{x}_t) and two predetermined (\hat{in}_t and \hat{r}_t) variables. The system will exhibit saddle-path stability if there are two eigenvalues of J outside the unit circle, indeterminate multiple stable paths if there are no forward stable eigenvalues, and explosive paths if there is more than one forward-stable eigenvalue. When $\phi = 1$ the five-variable system contains three jump (\hat{x}_t , \hat{x}_{t+1} and \hat{in}_{t+1}) and two predetermined (\hat{in}_t and \hat{r}_t) variables. The fifth equation is an identity for \hat{in}_{t+1} with a zero eigenvalue. Considering the other four eigenvalues, the system will exhibit saddle-path stability if exactly two of the eigenvalues are outside the unit circle, indeterminate multiple stable paths if there are three or more eigenvalues outside the unit circle, and explosive paths if there is less than two eigenvalues outside the unit circle.

Figure 1. Optimal Choices and Bifurcation for the Economic SIS Model



(a) Optimal Choices



(b) Bifurcations

Figure 2. Immunity Externality

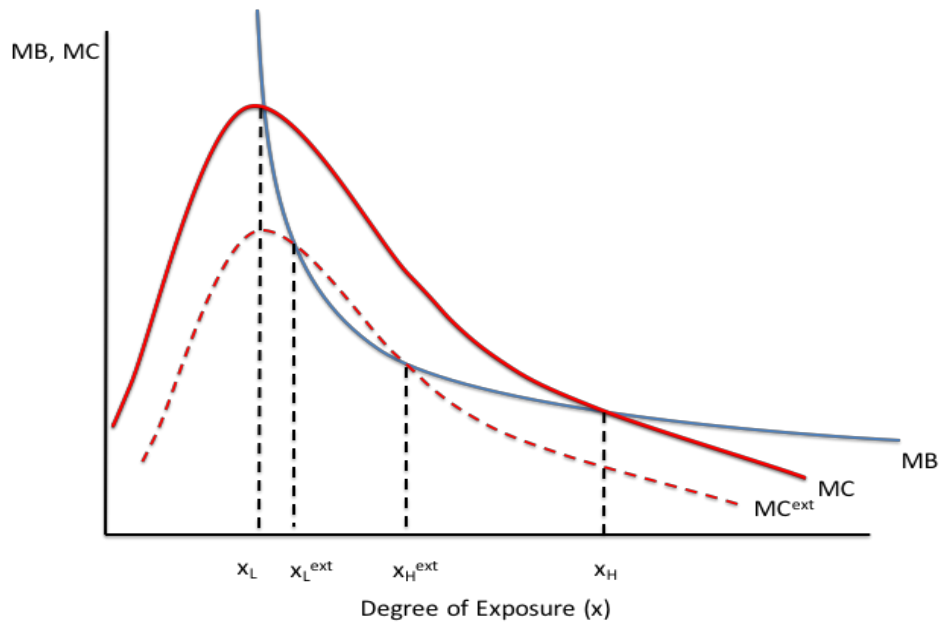
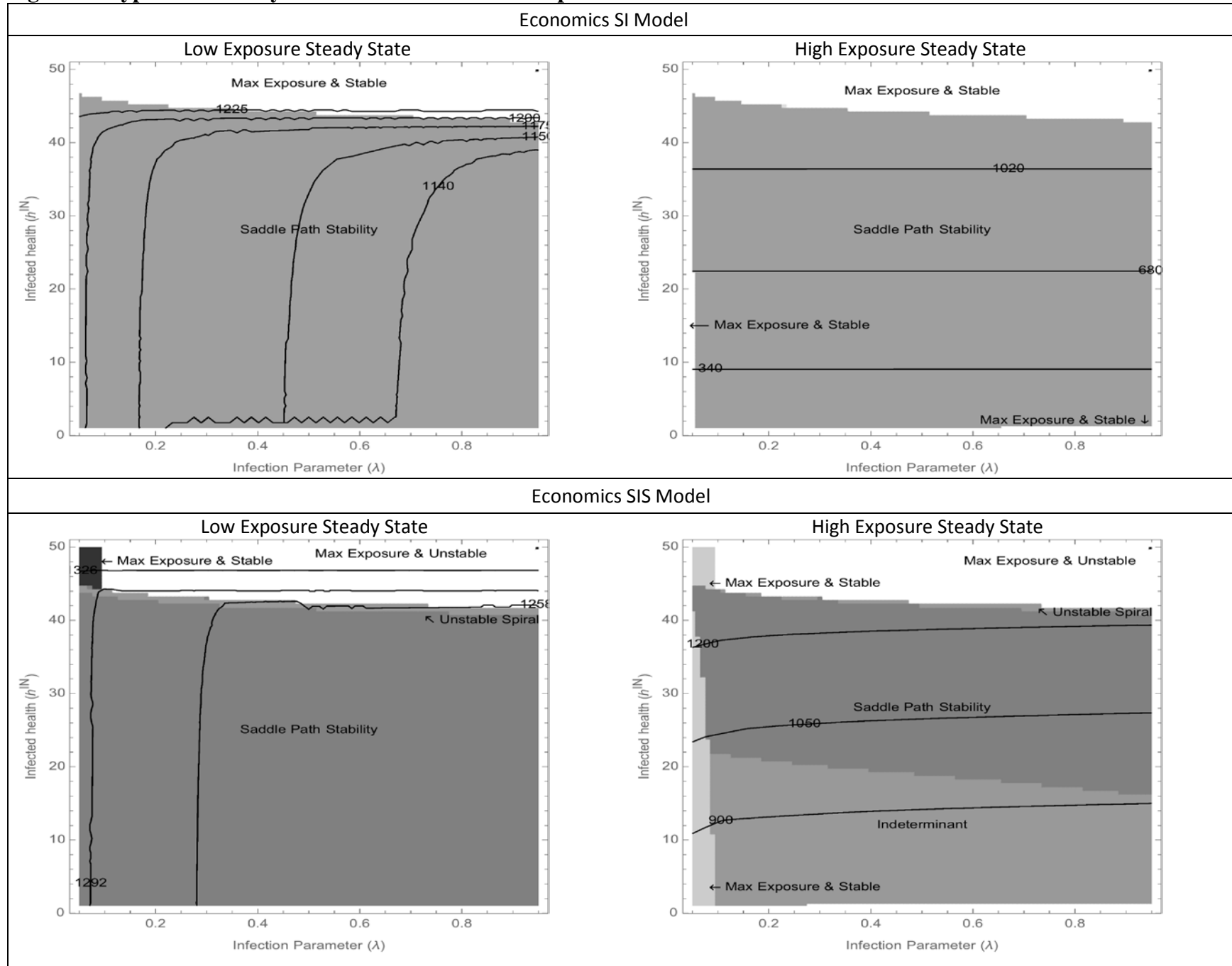
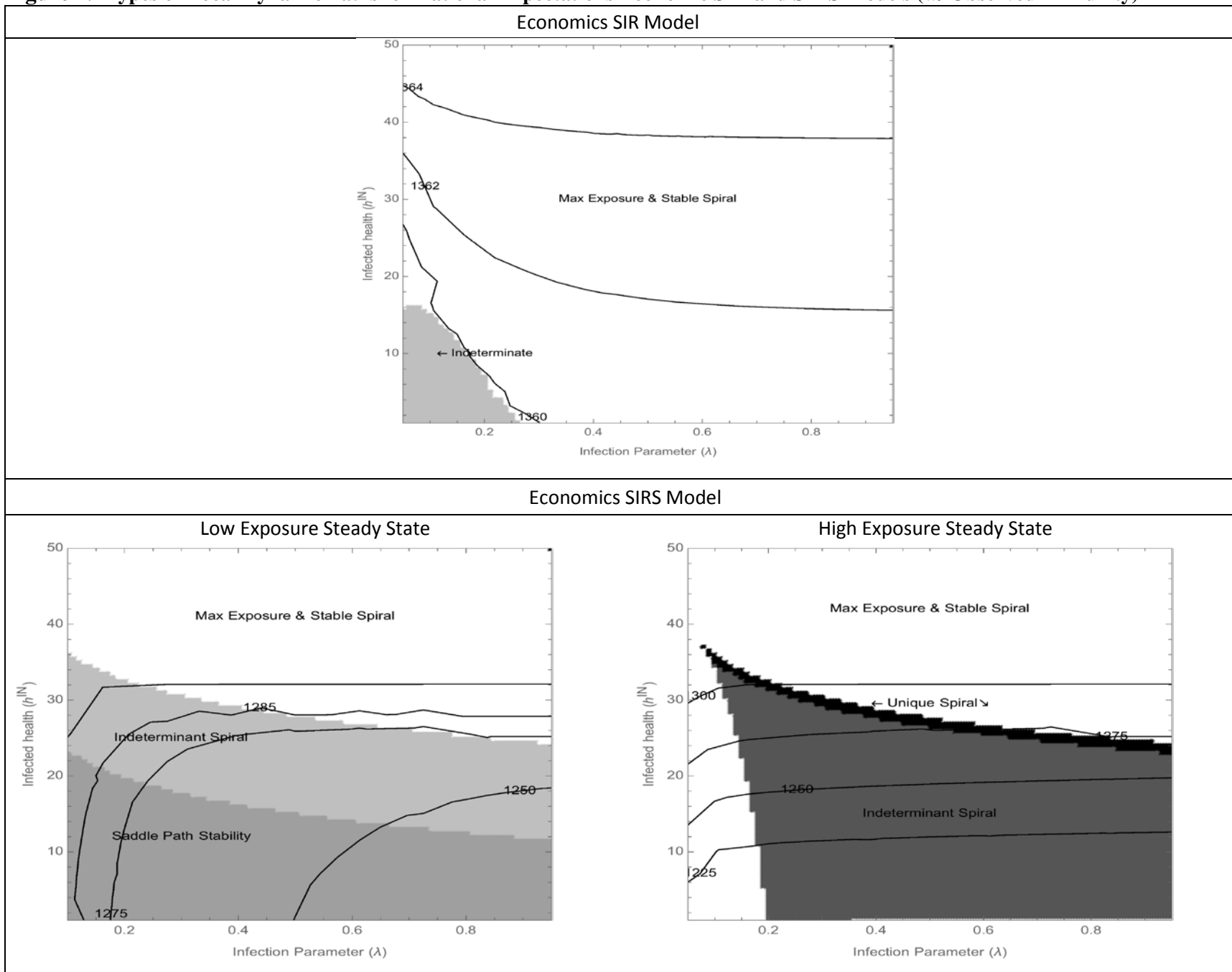


Figure 3. Types of Local Dynamic Paths for Rational Expectations Economic SI and SIS Models



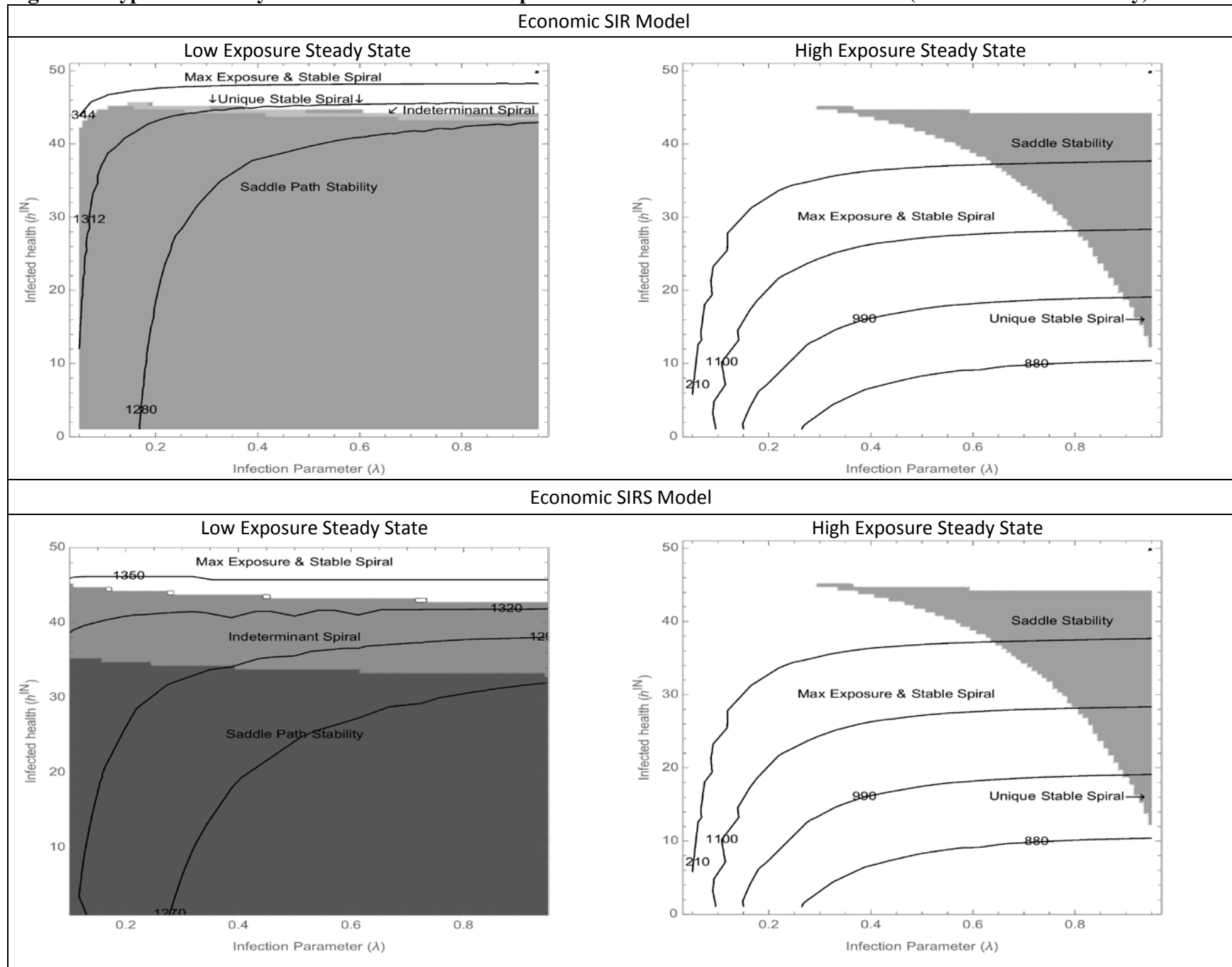
Notes: Contour lines indicate welfare at the steady state. Parameter values are given in Table 1.

Figure 4. Types of Local Dynamic Paths for Rational Expectations Economic SIR and SIRS Models (w/ Observed Immunity)



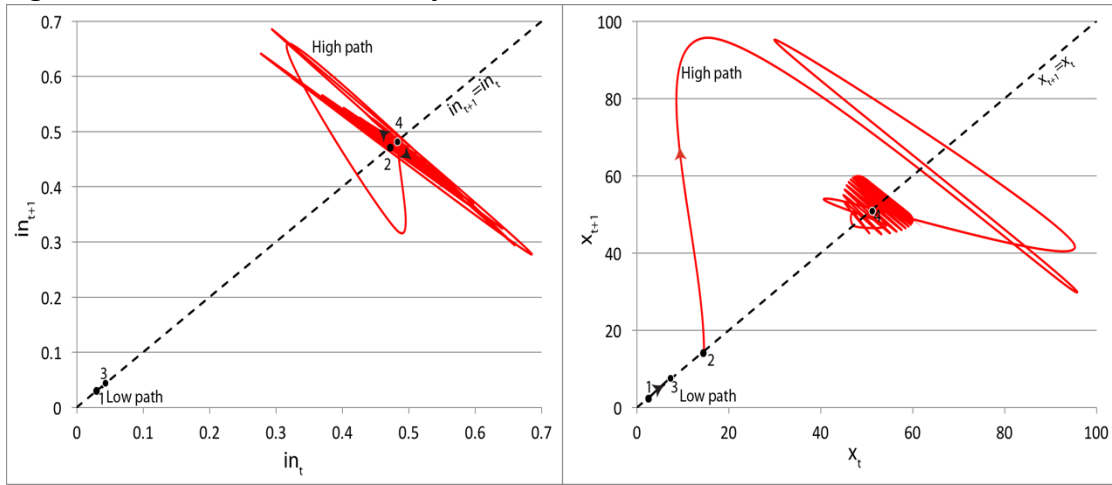
Notes: Contour lines indicate welfare at the steady state. Parameter values are given in Table 1.

Figure 5. Types of Local Dynamic Paths for Rational Expectations Economic SIR and SIRS Models (w/ Unobserved Immunity)

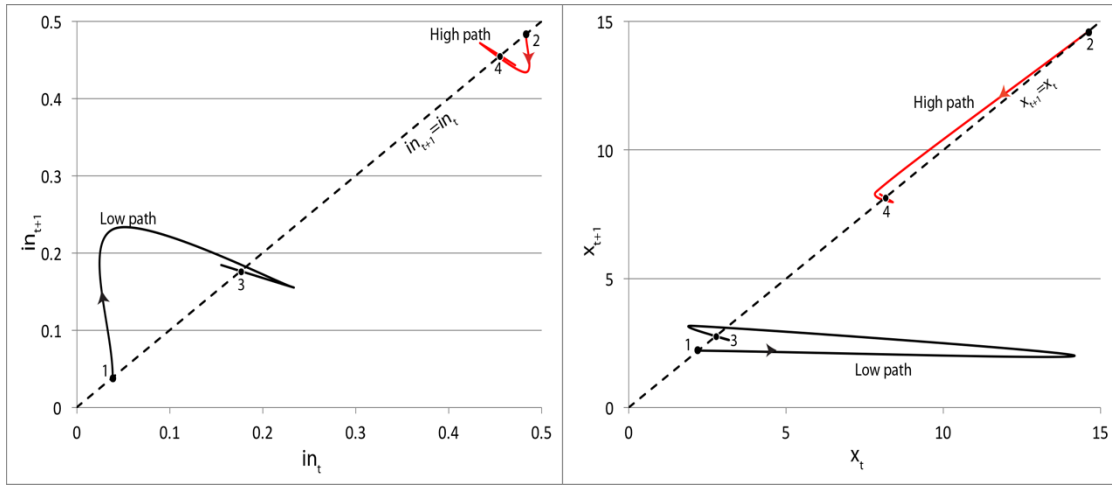


Notes: Contour lines indicate welfare at the steady state. Parameter values are given in Table 1.

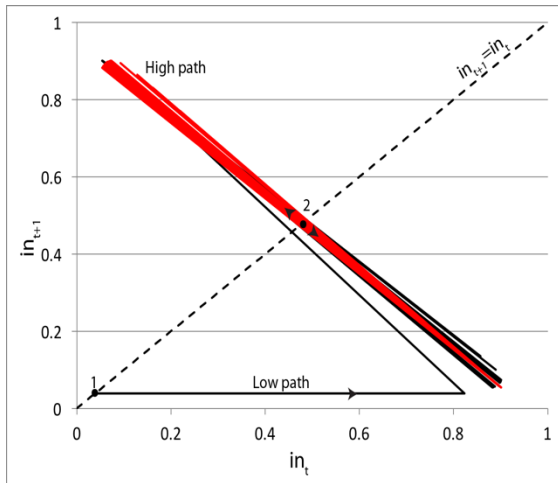
Figure 6. Nonlinear Transition Dynamics



(a) Paths from reduction to $\lambda = 0.15$



(b) Paths from increase $h^N = 40$



(c) Paths from increase to $h^N=43$ (exposure at upper bound)

Figure 7. Welfare

