Vanderbilt University Department of Economics Working Paper Series
Nashville, TN 37235-1819

Vanderbilt University Department of Economics Working Papers 13-00003

The Equilibrium Dynamics of Economic Epidemiology

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. 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.

Citation: David Aadland and David Finnoff and Kevin x.d. Huang, (2013) "The Equilibrium Dynamics of Economic Epidemiology", *Vanderbilt University Department of Economics Working Papers*, VUECON-13-00003.
Contact: David Aadland - Aadland@uwyo.edu, David Finnoff - Finnoff@uwyo.edu, Kevin x.d. Huang - kevin.huang@vanderbilt.edu.
Submitted: February 14, 2013. **Published:** March 25, 2013.
URL:http://www.accessecon.com/Pubs/VUECON/VUECON-13-00003.pdf

The Equilibrium Dynamics of Economic Epidemiology

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

July 2012

Abstract

In this paper we investigate the nature of rational expectations equilibria for economic epidemiological models. 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, equilibria, 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.

1 Introduction

In this paper we investigate the dynamic properties of rational expectations 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 the stability properties of rational expectations EE equilibria and the relationship to public health policy.[1] Similar to the macroeconomic literature on the stability properties of monetary and fiscal policy (Guo and Lansing (1998); Clarida, Gali, and Gertler (2000); Fatás and Mihov (2003); Meng (2002); Meng and Yip (2004); Xiao (2008)), we show that well-intentioned public policy has the potential to contribute to aggregate instability and volatility. For instance, government policy designed to lower the transmission probability or raise the quality-of-life associated with infectious diseases can push the EE system from a stable equilibrium path to ones exhibiting instability or indeterminate equilibrium paths. The latter also have the potential of contributing to self-fulfilling "sunspot" equilibria, which can contribute to the volatility and unpredictability of the system. We also show how the incentives and choices of susceptible individuals may cause the system to gravitate toward a socially sub-optimal transition path when multiple equilibrium paths exist. This is a type of dynamic externality imposed on the economic system but one that is slightly different than the infection externality typically stressed in the literature (Kremer (1996); Gersovitz and Hammer (2004)). 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.

[1] The 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 Economic Epidemiological Model

Following work by Philipson and Posner (1993), we specify an integrated economic epidemiological model to describe communicable disease dynamics. The model is set in discrete time (Auld (2003)), where t indexes the decision interval.[2] There is a constant population of N individuals, which are all identical except for their state of the disease. A similar model is presented in Aadland, Finnoff, and Huang (2011) to specifically examine the dynamics and potential eradication of syphilis.

2.1 Epidemiology

The epidemiological portion of the model describes the evolution of three mutually exclusive disease categories: susceptible (s), infected (in), and recovered with immunity (r). Each category is measured as a proportion of the overall population so that $s + in + r = 1$. This is the classical 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 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 HIV/AIDS), diseases with recovery but no immunity (SIS diseases such as the common cold), and diseases with permanent immunity (SIR diseases such as measles and chicken pox).

The model is represented by three equations:

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

$$in_{t+1} = (1 - v - \mu)in_t + p_t s_t \tag{2}$$

$$r_{t+1} = (1 - \gamma - \mu)r_t + vin_t, \tag{3}$$

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

$$in_{t+1} = (1 - p_t - v - \mu)in_t + p_t(1 - r_t) \tag{4}$$

$$r_{t+1} = (1 - \gamma - \mu)r_t + vin_t. \tag{5}$$

[2] 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.

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 and are treated in Appendix A.

Assuming that individuals independently choose x_t contacts and engage in a fixed number of interactions (a) with each contact, the probability that susceptible individuals become infected is

$$p_t = \Pr(\text{infection}) = 1 - (1 - \lambda_p in_t)^{x_t}, \tag{6}$$

where $\lambda_p = 1 - (1 - \lambda_a)^a$ is the probability of contracting the disease from a single infected contact, and λ_a is the probability of contracting the disease from a single interaction with an infected contact (Kaplan (1990); Oster (2005)). The dependence on the chosen number of contacts distinguishes the analysis from standard mathematical epidemiology.

We now turn our attention to the economic analysis and the optimal choice of contacts.

2.2 Economics

Representative individual i maximizes expected lifetime utility by choosing the number of contacts, $x_{i,t}$. The objective function is

$$E_t \sum_{j=0}^{\infty} \beta^j [\ln(x_{i,t+j}) + h_{i,t+j}] \tag{7}$$

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, and $\bar{x}$ is the maximum number of contacts per period. The parameter $h_{i,t}$ captures the individual's health status with infected individuals experiencing lower values of h. The core tradeoff in the model is that additional contacts bring immediate satisfaction but also the risk of future infection. Infection in turn causes a deterioration of health.

In any period t, individual i is in one of three epidemiological states as measured 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. Because all individuals are identical other than disease state and health level, we drop the i subscript and consider a single representative individual in each disease category.

The value functions for each category – susceptible, infected, and recovered – are given by

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

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

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

where $h^S > h^{IN}$.

All individuals maximize (7) without concern for the welfare of the general population. Infected and immune individuals will therefore choose the maximum number of contacts, $\bar{x}$, because they face no risk of immediate infection (Geoffard and Philipson (1996)). Assuming an interior solution, susceptible individuals will choose the number of contacts to satisfy 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 number of contacts is $p_{x,t} = -\ln(1-p_t)(1-p_t)/x_t$.[3] The contact rate in mathematical epidemiological models is typically constant or varies deterministically (Korobeinikov (2006)). As equation (11) shows, the contact rate in EE models is instead based on behavioral responses to changes in disease risk**.** We look at two cases depending on individuals' ability to observe their own host immunity.

2.2.1 Unobservable Host Immunity

In this case, individuals with host immunity believe they are susceptible. Therefore, equation (10) is not relevant and equation (9) becomes

$$V_t^{IN} = \ln(\bar{x}) + h^{IN} + \beta E_t[vV_{t+1}^S + (1-v)V_{t+1}^{IN}]. \quad (12)$$

Substituting out the value functions (V_{t+1}^S and V_{t+1}^{IN}), equation (11) can be rewritten as

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

[3] The second-order sufficiency conditions require that the marginal cost curve with respect to contacts (right side of equation (11)) must slope up or if it slopes down, it must be locally flatter than the marginal benefit curve (left side of equation (11)).

where $h = h^S - h^{IN}$ is the health gap between being susceptible and infected. This equation states that rational individuals choose the number of contacts to balance the marginal benefits (left side) with the discounted, expected costs (right side).

2.2.2 Observable Host Immunity

When individuals observe their own immunity, they rationally choose the maximum number of contacts $\bar{x}$ and have health level h^S. Susceptible individuals, on the other hand, choose x_t to satisfy

$$x_t^{-1} = \beta p_{x,t} E_t \left[\ln(x_{t+1}/\bar{x}) + h + \frac{(1 - v - p_{t+1})}{x_{t+1} p_{x,t+1}} - \beta \Delta_{t+2} \right], \tag{14}$$

where

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

The Euler equations in (13) and (14) are identical except for Δ_{t+2}. This term captures the expected future "costs" of infection associated with observed acquired immunity. Because Δ_{t+2} enters the right side of (14) with a negative sign, the possibility of future immunity is a benefit of becoming infected. See Appendix B for a derivation of equation (14).

3 Equilibria

We focus on the nature of the transition dynamics around the endemic EE steady states.[4]

3.1 Steady States

The endemic steady states solve time-invariant versions of (4), (5), and the Euler equation. The Euler equation either takes the form of (13) when the indicator variable is set at $\phi = 0$ or the form of (14) when

[4] There 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 number of contacts or engage in preventative behavior. 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.

$\phi = 1$. The steady-state system can therefore be rewritten as three equations:

$$in = p(1 - in - r)/(v + \mu) \tag{15}$$

$$r = vin/(\gamma + \mu) \tag{16}$$

$$x^{-1} = \beta[p_x(\ln(x/\bar{x}) + h - \phi\beta\Delta) + (1 - v - p)/x] \tag{17}$$

in three unknown variables (in, r, x) where

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

As is well-known, economic epidemiological models are capable of producing multiple endemic steady states (Goldman and Lightwood (2002)). The possibility of multiple steady states is shown in panel (a) of Figure 1 for the economic SIS model. The marginal benefit (MB) curve is given by the left side of equation (17) and is everywhere downward sloping; the marginal cost (MC) curve is given by the right side of equation (17). At low levels of h, the marginal costs of an additional contact are low and there is no endemic steady-state contact choice, x. As h increases, the MB and MC curves eventually reach a tangency point such that there is a unique endemic steady state contact choice. As h increases further, the MB and MC curves intersect twice and produce two endemic steady states – one with a low contact choice and one with a high contact choice. This bifurcation is shown in panel (b) of Figure 1. As the health gap h increases, the system progresses from no endemic steady state to a single steady state to a pair of endemic steady states. At low levels of h, both steady states are stable. As h increases, the high-contact steady state becomes unstable, while the low-contact steady state remains stable. We now turn to an analysis of the transition dynamics near the steady states.

3.2 Transition Dynamics

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

$$\widehat{in}_{t+1} = (1 - v - \mu - p)\widehat{in}_t + (1 - in - r)\hat{p}_t - p\hat{r}_t \tag{18}$$

$$\hat{r}_{t+1} = (1 - \gamma - \mu)\hat{r}_t + v\widehat{in}_t, \tag{19}$$

where hats (ˆ) over the variables indicate deviation from one of the steady states. The linearized Euler equation is:

$$\begin{aligned} p_x\hat{x}_t + x\hat{p}_{x,t} &= \beta p_x(1-v-p-xp_x)E_t\hat{x}_{t+1} + \beta x(1-v-p)E_t\hat{p}_{x,t+1} + \beta x p_x E_t\hat{p}_{t+1} \\ &+\phi\beta^2 E_t \left\{ \begin{array}{c} p_x[v\gamma + (1-v-\gamma)(1-p-xp_x)]\hat{x}_{t+2} + x[v\gamma + (1-v-\gamma)(1-p)]\hat{p}_{x,t+2} + \\ [(1-v-\gamma)xp_x]\hat{p}_{t+2} - [(1-\gamma)p_x/\beta]\hat{x}_{t+1} - [(1-\gamma)x/\beta]\hat{p}_{x,t+1} \end{array} \right\} \end{aligned} \tag{20}$$

where

$$\hat{p}_t = p_{in}\widehat{in}_t + p_x\hat{x}_t \tag{21}$$

$$\hat{p}_{x,t} = [(1+\ln[1-p])/x]\hat{p}_t - (p_x/x)\hat{x}_t \tag{22}$$

and

$$p_{in} = x\lambda_p(1-\lambda_p in)^{x-1} \tag{23}$$

$$p_x = -\ln(1-p)(1-p)/x. \tag{24}$$

In matrix form, the EE system can be written as

$$\hat{z}_t = J\,\hat{z}_{t+1}, \tag{25}$$

where $\hat{z}_t = (\hat{x}_t, \widehat{in}_t, \hat{r}_t)'$ when $\phi = 0$ or $\hat{z}_t = (\hat{x}_t, \widehat{in}_t, \hat{r}_t, \hat{x}_{t+1}, \widehat{in}_{t+1})'$ when $\phi = 1$. See the Appendix C for the derivation of (25) and the transition matrix J.

We use the method of Blanchard and Kahn (1980) to analyze the nature of the rational expectation EE equilibrium. The three-variable system (25) contains one jump ($\hat{x}_t$) and two predetermined ($\widehat{in}_t$ 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. The five-variable system contains three jump ($\hat{x}_t$, $\hat{x}_{t+1}$ and $\widehat{in}_{t+1}$) and two predetermined ($\widehat{in}_t$ and $\hat{r}_t$) variables. The fifth equation in (25) is an identity for $\widehat{in}_{t+1}$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.

3.3 Parameter Values and Mathematical Program

The parameter values in Table 1 are fixed and not calibrated to a particular disease.

Table 1. Fixed Parameter Values

Parameters	β	μ	v	γ	$\bar{x}$
Value	0.96	0.05	1	0.2	100

The value of β implies a 4% annual discount rate, μ gives a 5% birth and death rate for the population, v implies a 100% recovery rate within a year of infection, γ gives an expected 5-year immunity duration, and the maximum number of feasible annual contacts is 100.

4 Results and Policy Implications

Figures 2-4 show the types of dynamic paths for the EE model under a range of values for the health gap (h) and infection rate (λ_p). These two parameters represent possible public health policy targets. The health gap parameter ($h = h^S - h^{IN}$) can be lowered through the discovery and introduction of drug treatments, while the infection rate (λ_p) can be lowered through the introduction of vaccines or new protection technologies.

Figure 2 shows the map of path types for the SI and SIS models around the low-contact endemic steady state (left panels) and high-contact endemic steady state (right panels). For low values of the health gap h, there is no endemic steady state (see Figure 1 and the discussion above). In this case, the dynamics of the system are evaluated around the maximum number of partners, $\bar{x}$, and shown in both the left and right panels. 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 contact choice (x_0) that puts the system on a convergent equilibrium path to the endemic steady state. All other initial contact levels lead to divergent paths that violate non-negativity or non-explosion conditions. Because both the low-contact and high-contact steady states exhibit local saddle-path stability, the system may gravitate toward either the low-contact or high-contact steady state. The welfare contours show that, in the long run, society is better off at the low-contact, low-prevalence steady state.[5]

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

The bottom panels of Figure 2 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. Appendix D discusses individual and social welfare along the nonlinear model's transition paths and shows that of the two paths, the one into the low-contact steady state is both privately and socially optimal.[6] The parameter region for the low-contact steady steady state is primarily a saddle-path equilibrium, but there is also an explosive region for the low-contact steady state near the bifurcation band. Therefore, public health 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 prevalence as individuals rationally take more risk.

Figure 2 shows the types of dynamic paths and contact choices for economic SI and SIS models. Mathematical epidemiological models, by contrast, do not vary the number of contacts in response to changes in disease prevalence. For reasonable contact rates, the mathematical SI and SIS models are characterized by stable dynamic paths with a fixed number of contacts and no dependence on the health gap, h.

To gain intuition for the types of dynamic paths in the economic and mathematical SIS models, consider a simple heuristic, din_{t+1}/din_t, relating changes in future prevalence to a change in current prevalence. Nearby the endemic steady state, this metric is given by

$$\frac{d\widehat{in}_{t+1}}{d\widehat{in}_t} = (1 - v - \mu) + (1 - in)\left(p_{in} + \kappa p_x x/in\right), \tag{26}$$

where κ is the contact elasticity with respect to prevalence. Prevalence elasticity measures the percentage change in contacts 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 fewer contacts.[7] For example, whether the economic SIS system depicted in the lower left panel of Figure 2 is saddle-path stable or explosive depends on the magnitude of κ. The critical prevalence elasticity along the stable-explosive boundary can be found by setting (26) equal to -1 and solving for κ:

$$\kappa_c = -\frac{in}{p_x x}\frac{[1 + (1 - v - \mu) + p_{in}(1 - in)]}{1 - in}.$$

For the parameter values in Table 1, the critical prevalence elasticity is approximately $\kappa_c = -1.3$. Parameter combinations in the SIS saddle-path region are associated with prevalence elasticities that have a smaller

[6] The simulated nonlinear transition paths in GAMS confirm the types of equilibrium transition paths from the analysis of the linearized system.

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

magnitude than κ_c while parameter combinations in the SIS explosive band are associated with prevalence elasticities larger in magnitude than κ_c. If a 10% increase in disease prevalence triggers susceptible individuals to reduce the number of contacts by more than 13%, then prevalence at $t+1$ falls more than the initial rise in prevalence at t. As a result, the system oscillates in an explosive manner. Conversely, prevalence elasticity is zero in the mathematical SIS model because susceptible individuals do not alter their behavior in response to changes in disease prevalence. This implies that

$$\frac{d\widehat{in}_{t+1}}{d\widehat{in}_t} = (1 - v - \mu) + (1 - in)p_{in},$$

which is positive and less than one for all values of the infection parameter λ_p. Increases in prevalence near the endemic steady state cause the mathematical SIS system to convergence monotonically back to the endemic steady state.

Figure 3 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_p and h, there is a single endemic steady state. For most of the parameter region, susceptible individuals choose the maximum number of contacts 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 high levels of the health gap 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 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. Thus, public health policy aimed at improving the quality of life for individuals infected with diseases that have known temporary immunity may induce aggregate instability and indeterminacy. Furthermore, although the transition paths to the low-contact steady state are socially optimal (see Appendix D), susceptible individuals will find it privately optimal to be on the transition path to the high-contact steady state. The higher transition path is privately optimal because susceptible individuals are aware that risky behavior has the benefit of a known period of host immunity. This is a dynamic disease externality associated with the existence of multiple equilibria paths. Public health policy that encourages less risky behavior and internalizes the infection externality (Gersovitz and Hammer (2004)) has

the additional benefit of placing society on the optimal equilibrium transition path.

Figure 4 depicts the SIRS counterpart to Figure 3 but with unobservable host immunity. Unobservable host immunity causes two primary changes. First, individuals in the economic SIR system now choose fewer contacts for any parameter combination. Knowledge of perfectly effective treatment and permanent immunity, as depicted in Figure 3, greatly reduces the future cost of current risky behavior. Second, the indeterminacy region for the SIRS system now covers a much smaller range of health gap parameters.

5 Conclusion

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 influences the nature of communicable disease dynamics as the system transitions toward the endemic long-run equilibrium. In this paper, we explore the nature of the short-run equilibrium dynamics for rational expectation economic epidemiological systems. The analysis digs beneath a comparison of fixed parameter values and demonstrates the behavioral origin for changes in the dynamic properties of the system. 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 applying and calibrating the methods outlined in this paper to specific diseases.

References

Aadland, D., D. Finnoff, and K. Huang (2011): "Syphilis Cycles," unpublished manuscript.

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.

Fatás, A., and I. Mihov (2003): "The case for restricting fiscal policy discretion," *Quarterly journal of economics*, 118(4), 1419–1447.

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.

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.

Grassly, N., C. Fraser, and G. Garnett (2005): "Host immunity and synchronized epidemics of syphilis across the United States," *Nature*, 433, 417–421.

Guo, J., and K. Lansing (1998): "Indeterminacy and stabilization policy," *Journal of economic theory*, 82(2), 481–490.

Kaplan, E. (1990): "Modeling HIV infectivity: must sex acts be counted?," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 3(1), 55.

Korobeinikov, 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.

Korobeinikov, 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.

Meng, Q. (2002): "Monetary policy and multiple equilibria in a cash-in-advance economy," *Economics Letters*, 74(2), 165–170.

Meng, Q., and C. Yip (2004): "Investment, interest rate rules, and equilibrium determinacy," *Economic Theory*, 23(4), 863–878.

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.

Xiao, W. (2008): "Increasing Returns and the design of interest rate rules," *Macroeconomic Dynamics*, 12(1), 22.

Appendix A. SIS Economic Epidemiological System

Here we describe the SIS economic epidemiological model. In the SIS model, infected individuals transition directly back to the susceptible category and do not experience a period of immunity. The SIS dynamic equations are

$$s_{t+1} = \mu + (1 - p_t - \mu)s_t + v in_t \tag{A.1}$$

$$in_{t+1} = (1 - v - \mu)in_t + p_t s_t, \tag{A.2}$$

while the steady-state values are

$$s = (v + \mu)/(p + v + \mu) \tag{A.3}$$

$$in = p/(p + v + \mu). \tag{A.4}$$

The SI model is defined by $v = 0$ so that infection is permanent. The value functions are

$$V_t^S = \ln(x_t) + h + \beta E_t[p_t V_{t+1}^{IN} + (1 - p_t)V_{t+1}^S] \tag{A.5}$$

$$V_t^{IN} = \ln(\bar{x}) + \beta E_t[v V_{t+1}^S + (1 - v)V_{t+1}^{IN}]. \tag{A.6}$$

The Euler equation for susceptible individuals is given by equation (11) in the main text. Using (A.5) and (A.6) to substitute out the value functions, the Euler equation can be rewritten as

$$x_t^{-1} = \beta p_{x,t} E_t \left[\ln(x_{t+1}/\bar{x}) + h + \frac{(1 - v - p_{t+1})}{x_{t+1} p_{x,t+1}}\right]. \tag{A.7}$$

The linearized EE system is

$$\widehat{in}_{t+1} = (1 - v - \mu)\widehat{in}_t + (1 - in)\hat{p}_t \tag{A.8}$$

$$p_x \hat{x}_t + x\hat{p}_{x,t} = \beta p_x (1 - v - p - x p_x) E_t \hat{x}_{t+1} + \beta x (1 - v - p) E_t \hat{p}_{x,t+1} + \beta x p_x E_t \hat{p}_{t+1} \tag{A.9}$$

along with equations (21) and (22). Assuming perfect foresight, equations (A.8) and (A.9) in matrix form are

$$\underbrace{\begin{bmatrix} 0 & 1-v-\mu-p \\ p_x & 0 \end{bmatrix}}_{A} \begin{bmatrix} \hat{x}_t \\ \widehat{in}_t \end{bmatrix} + \underbrace{\begin{bmatrix} 1-in & 0 \\ 0 & x \end{bmatrix}}_{B} \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0 & 1 \\ \beta p_x(1-v-p-xp_x) & 0 \end{bmatrix}}_{C} \begin{bmatrix} \hat{x}_{t+1} \\ \widehat{in}_{t+1} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 \\ \beta x p_x & \beta x(1-v-p) \end{bmatrix}}_{D} \begin{bmatrix} \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix}. \tag{A.10}$$

Along with (A.19), the SIS economic epidemiological model can then be written in the form of equation (25) where the coefficient matrices A, B, C and D are redefined and $\hat{z}_t = (\hat{x}_t, \widehat{in}_t)$.

Appendix B. 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

$$V_t^R - V_t^{IN} = h + \beta E_t \left[\gamma(V_{t+1}^S - V_{t+1}^R) + (1-v)(V_{t+1}^R - V_{t+1}^{IN})\right], \tag{A.11}$$

while equations (8) and (9) imply

$$V_t^S - V_t^{IN} = \ln(x_t/\bar{x}) + h + \beta E_t \left[(1-p_t)(V_{t+1}^S - V_{t+1}^{IN}) - v(V_{t+1}^R - V_{t+1}^{IN})\right]. \tag{A.12}$$

Using equation (11), we have

$$E_t(V_{t+1}^S - V_{t+1}^{IN}) = (\beta x_t p_{x,t})^{-1}, \tag{A.13}$$

for all t. Next, rearrange (A.12) as

$$V_{t+1}^R - V_{t+1}^{IN} = \frac{1}{\beta v}[\ln(x_t/\bar{x}) + h] + \frac{1}{v}(1-p_t)E_t(V_{t+1}^S - V_{t+1}^{IN}) - \frac{1}{\beta v}\left(V_t^S - V_t^{IN}\right). \tag{A.14}$$

Take E_{t-1} on both sides of (A.14) and substitute (A.13) to get

$$E_{t-1}(V_{t+1}^R - V_{t+1}^{IN}) = \frac{1}{\beta v}E_{t-1}[\ln(x_t/\bar{x}) + h] + \frac{1}{\beta v}E_{t-1}\left(\frac{1-p_t}{x_t p_{x,t}}\right) - \frac{1}{\beta^2 v}\left(\frac{1}{x_{t-1}p_{x,t-1}}\right). \tag{A.15}$$

Now rewrite equation (A.11) as

$$V_t^R - V_t^{IN} = h + \beta E_t \left[\gamma(V_{t+1}^S - V_{t+1}^{IN}) + (1-v-\gamma)(V_{t+1}^R - V_{t+1}^{IN})\right]. \tag{A.16}$$

Move (A.16) ahead one period, take E_{t-1} of both sides, and set equal to (A.15) to get

$$\begin{aligned} &\frac{1}{\beta v}E_{t-1}\left[\ln(x_t/\bar{x}) + h\right] + \frac{1}{\beta v}E_{t-1}\left(\frac{1-p_t}{x_t p_{x,t}}\right) - \frac{1}{\beta^2 v}\left(\frac{1}{x_{t-1}p_{x,t-1}}\right) \\ = \;& h + \beta E_{t-1}\left\{\gamma(\beta x_t p_{x,t})^{-1} + (1-v-\gamma)\left(\frac{1}{\beta v}\left[\ln(x_t/\bar{x}) + h\right] + \frac{1}{\beta v}\left(\frac{1-p_t}{x_t p_{x,t}}\right) - \frac{1}{\beta^2 v}\left(\frac{1}{x_{t-1}p_{x,t-1}}\right)\right)\right\}. \end{aligned} \tag{A.17}$$

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 C. SIRS Economic Epidemiological Matrix System

Imposing perfect foresight, the $\phi = 0$ linearized EE matrix system can be written as:

$$\underbrace{\begin{bmatrix} 0 & 1-v-\mu-p & -p \\ 0 & v & 1-\gamma-\mu \\ p_x & 0 & 0 \end{bmatrix}}_{A} \begin{bmatrix} \hat{x}_t \\ \widehat{in}_t \\ \hat{r}_t \end{bmatrix} + \underbrace{\begin{bmatrix} 1-in-r & 0 \\ 0 & 0 \\ 0 & x \end{bmatrix}}_{B} \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \beta p_x(1-v-p-xp_x) & 0 & 0 \end{bmatrix}}_{C} \begin{bmatrix} \hat{x}_{t+1} \\ \widehat{in}_{t+1} \\ \hat{r}_{t+1} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \beta x p_x & \beta x(1-v-p) \end{bmatrix}}_{D} \begin{bmatrix} \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix} \tag{A.18}$$

and

$$\underbrace{\begin{bmatrix} -1 & 0 \\ -(1+\ln(1-p))/x & 1 \end{bmatrix}}_{F} \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \end{bmatrix} = -\underbrace{\begin{bmatrix} p_x & p_{in} & 0 \\ p_x/x & 0 & 0 \end{bmatrix}}_{G} \begin{bmatrix} \hat{x}_t \\ \widehat{in}_t \\ \hat{r}_t \end{bmatrix}. \tag{A.19}$$

When $\phi = 1$, we have

$$\underbrace{\begin{bmatrix} 0 & 1-v-\mu-p & -p & 0 & 0 \\ 0 & v & 1-\mu-\gamma & 0 & 0 \\ p_x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} \hat{x}_t \\ \widehat{in}_t \\ \hat{r}_t \\ \hat{x}_{t+1} \\ \widehat{in}_{t+1} \end{bmatrix} + \underbrace{\begin{bmatrix} s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{B} \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \\ \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \beta p_x(1-v-p-xp_x) - \beta^2[(1-\gamma)p_x/\beta] & 0 & 0 & \beta^2 p_x[v\gamma + (1-v-\gamma)(1-p-xp_x)] & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}}_{C} \begin{bmatrix} \hat{x}_{t+1} \\ \widehat{in}_{t+1} \\ \hat{r}_{t+1} \\ \hat{x}_{t+2} \\ \widehat{in}_{t+2} \end{bmatrix} +$$

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta x p_x & \beta x(1-v-p) - \beta^2[(1-\gamma)x/\beta] & \beta^2(1-v-\gamma)xp_x & \beta^2 x[v\gamma + (1-v-\gamma)(1-p)] \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{D} \begin{bmatrix} \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \\ \hat{p}_{t+2} \\ \hat{p}_{x,t+2} \end{bmatrix}$$

and

$$\underbrace{\begin{bmatrix} -1 & 0 & 0 & 0 \\ -(1+\ln(1-p))/x & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -(1+\ln(1-p))/x & 1 \end{bmatrix}}_{F} \begin{bmatrix} \hat{p}_t \\ \hat{p}_{x,t} \\ \hat{p}_{t+1} \\ \hat{p}_{x,t+1} \end{bmatrix} = -\underbrace{\begin{bmatrix} p_x & p_{in} & 0 & 0 & 0 \\ p_x/x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_x & p_{in} \\ 0 & 0 & 0 & p_x/x & 0 \end{bmatrix}}_{G} \begin{bmatrix} \hat{x}_t \\ \widehat{in}_t \\ \hat{r}_t \\ \hat{x}_{t+1} \\ \widehat{in}_{t+1} \end{bmatrix}. \tag{A.20}$$

where

$$J = (A - BF^{-1}G)^{-1}(C - DF^{-1}G). \tag{A.21}$$

Appendix D. Transition Paths and Welfare

In this section we conduct a welfare analysis across nonlinear transition paths to the multiple steady states. For brevity we consider two representative cases: the SIS model and the SIRS model with observable immunity.

The analysis proceeds by simulating the nonlinear transition paths for each model and steady state. Conducted in GAMS, the paths for each steady state were found by solving the Euler conditions and difference equations for each model (the SIS system as given by equations A.2 and A.7 with A.1 substituted in; the SIRS system given by equations 4, 5, and 14). The method solves the system as a dynamic mixed complementary problem where the initial conditions for the states are specified and terminal condition for the control given by the respective steady-state value.

For each path and steady state, aggregate and individual measures of welfare were calculated. Aggregate welfare along each path were specified as the discounted, weighted sum of individual utility in each health state, for the low contact (L) and high contact (H) steady states, $k = \{L, H\}$

$$W_k = \sum\nolimits_{j=0}^{\infty} \beta^j \left\{ \left(\ln(x_t) + h^S \right) s_t + \left(\ln(\bar{x}) + h^{IN} \right) in_t + \left(\ln(\bar{x}) + h^S \right) r_t \right\} \tag{A.22}$$

Individual measures of welfare are given directly by the optimized value functions, equations A.5 and A.6 for the SIS model and equations 8, 9 and 10 for the SIRS model. To compare the optimality of the transition paths, the key value functions are those of a susceptible individual in the first period, $V^S_{L,0}$, for the lower steady state and $V^S_{H,0}$ for the high steady state. These are critical in analyzing the individual choices that will start society along one path versus the other.

SIS Transition Paths and Welfare

Figure 2 from the main text illustrates the stability properties for the SIS model. There are two obvious combinations of interest. Across the majority of values of the infection parameter and health gap, the lower steady state has saddle-path stability. The higher steady state can be explosive (at high values of the health gap) or saddle-path stable (at lower values of the health gap). We restrict our comparisons to cases where both steady states are saddle-path stable as explosive paths do not converge.

For a parameter combination where both steady states are saddle-path stable ($h = 9, \lambda = 0.5$), we simulated transition paths into both steady states across a wide range of initial conditions. To depict the paths, we construct Figure A1 in the spirit of a continuous time phase diagram. The horizontal axis is the

infected proportion (state variable) at time t, in_t; the vertical axis is the contact choice (control variable) of the representative susceptible individual at time t, x_t. The discrete-time isoclines for the endogenous variables are drawn with the steady states given by the intersections of the isocline for disease prevalence with the two isoclines for contacts.

For a low initial condition, in_0, the path to the high-contact steady state is given by the upper dashed line, the path to the low-contact steady state given by the lower dashed line. From the initial condition, both paths transition to their respective steady states within a few periods, exhibiting dampened oscillations.

Aggregate welfare for the path to the lower steady state exceeds that for the path to the high steady state, $W_L > W_H$. This is true for the case depicted in the figure and for a wide range of initial conditions simulated between $in_0 = 0.1$ and $in_0 = 0.7$. In terms of individual welfare, the initial period value functions for susceptible individuals of the lower transition path are always larger than those of the high transition path, i.e., $V^S_{L,0} > V^S_{H,0}$. For the SIS model individual incentives and aggregate welfare are in tune for rational susceptible individuals.

SIRS Transition Paths and Welfare

The SIRS model with observable immunity introduces the potential for individual incentives and aggregate welfare to diverge. Figure 3 in the main text demonstrates the diverse possibilities for the system. The low-contact steady state can exhibit saddle-path and indeterminate stability, while the high-contact steady state is predominantly indeterminate with a narrow band of saddle-path stability.

Following the same methodology as with the SIS model, we simulated transition paths for both steady states for each of the following stability combinations:

Steady State (L, H)	(h, λ)
L-saddle, H-indeterminate	$(h = 40, \lambda = 0.5)$
L-indeterminate, H-indeterminate	$(h = 28, \lambda = 0.5)$
L-indeterminate, H-saddle	$(h = 23, \lambda = 0.5)$

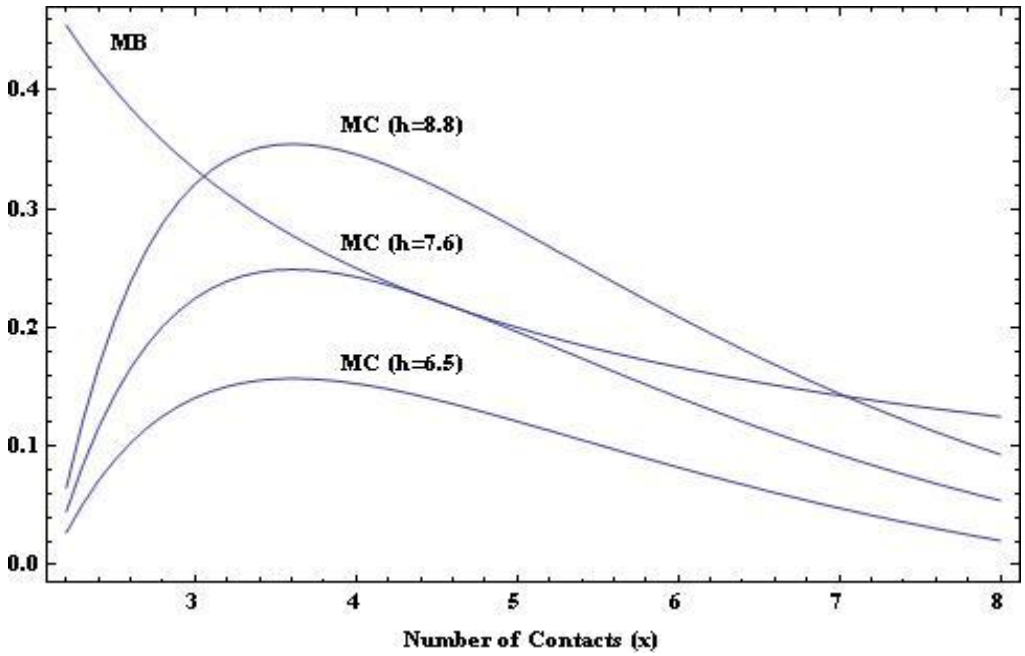
Figure A2 shows a pair of SIRS paths to the steady states for a low in_0 for the {L-saddle, H-indeterminate} combination, suppressing the r_t dimension in the plot. The saddle path to the low-contact steady state quickly converges. The indeterminate path to the high-contact steady state takes a few large oscillations that decline as the path converges to the steady state.

The interesting feature for this case is that aggregate and individual welfare measures diverge along the transition path. Aggregate welfare for the low-contact steady state path exceeds that of the high-contact steady state for each stability combination. However, the initial period value functions for susceptible individuals of the low-contact steady-state path are always less than those around the high-contact steady state as summarized below:

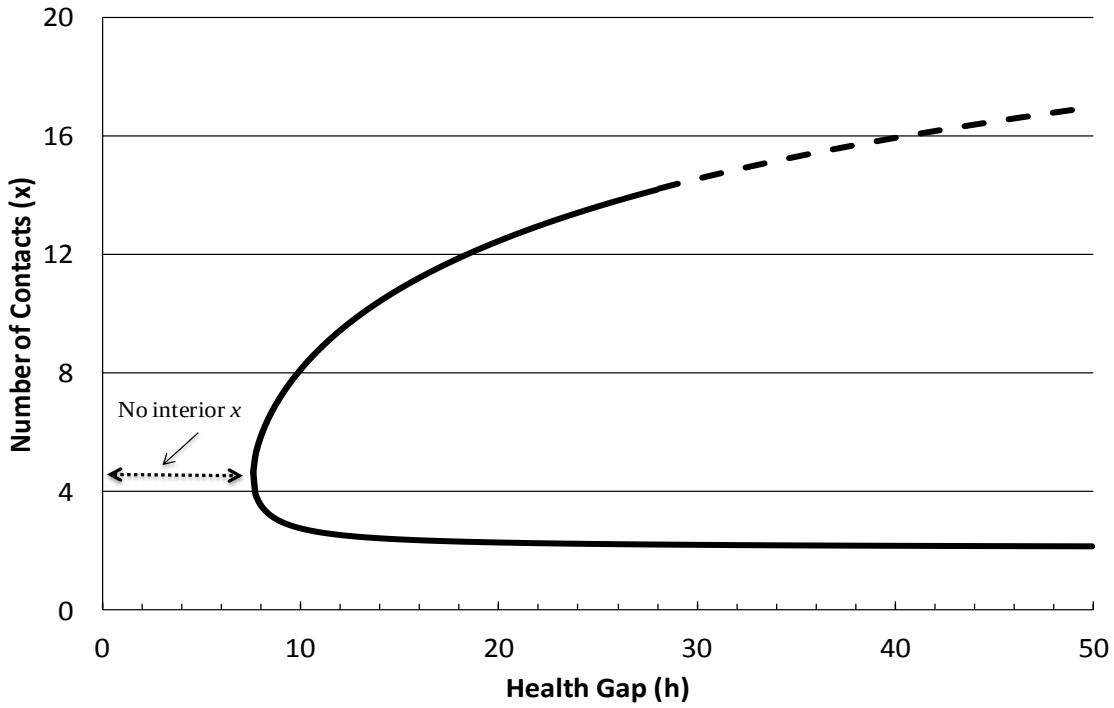
Aggregate	Individual	Equilibria Type
$W_L > W_H$	$V^S_{L,0} < V^S_{H,0}$	L-saddle, H-indeterminate
$W_L > W_H$	$V^S_{L,0} < V^S_{H,0}$	L-indeterminate, H-saddle
$W_L > W_H$	$V^S_{L,0} < V^S_{H,0}$	L-indeterminate, H-indeterminate

The fact that the high-contact transition path is privately optimal for susceptible individuals but lowers aggregate welfare provides a rationale for public intervention. Unlike the SIS model, susceptible individuals get the benefit of future immunity to disease infection. The presence of immunity tempers the expected future costs of infection to the individual, making the high-contact rate more advantageous. This dynamic disease externality places an external cost on the aggregate population and raises the need for public intervention with SIRS diseases as susceptible individuals will choose a socially suboptimal amount of risk.

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



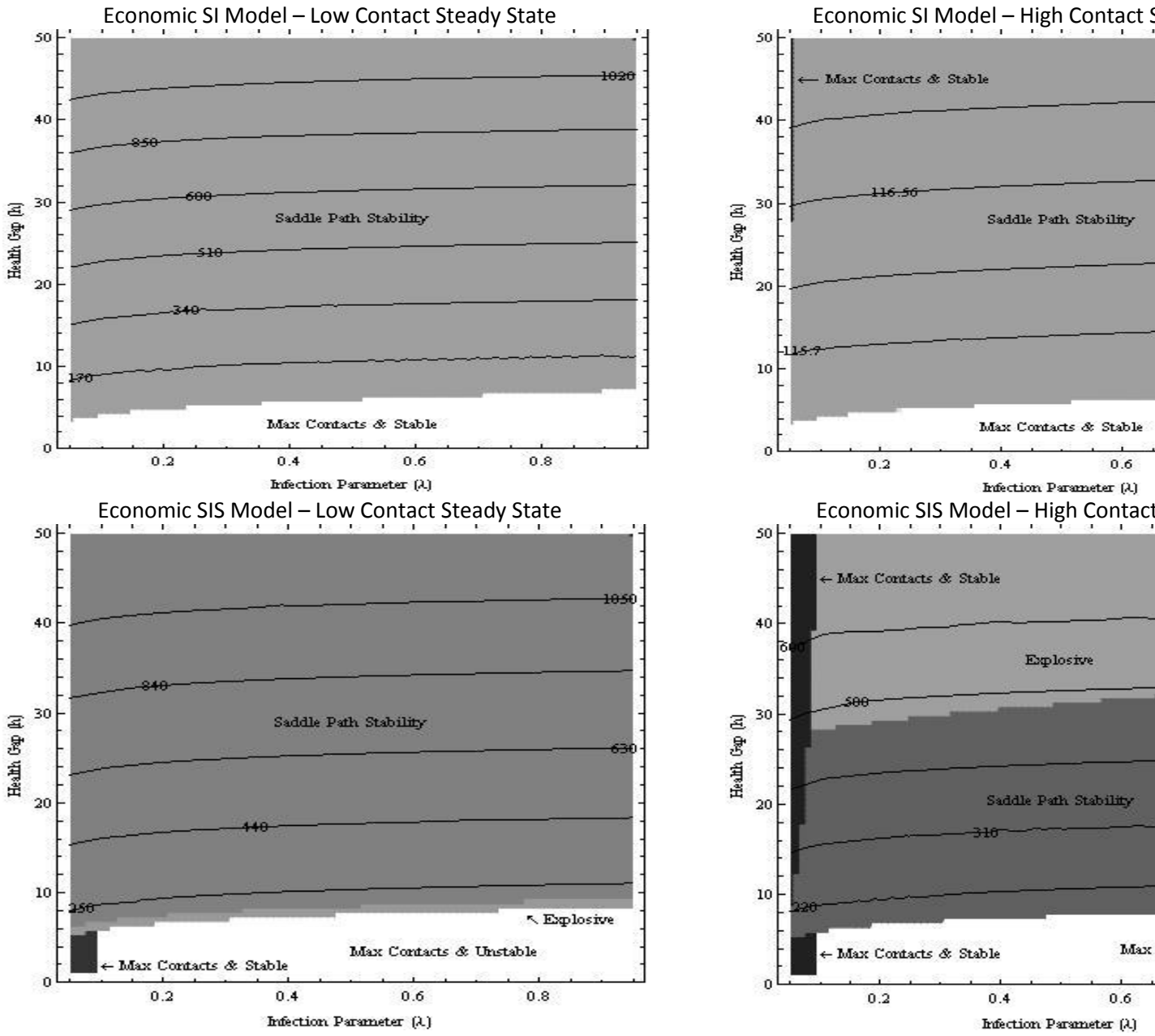
(a) Optimal Choices



(b) Bifurcation

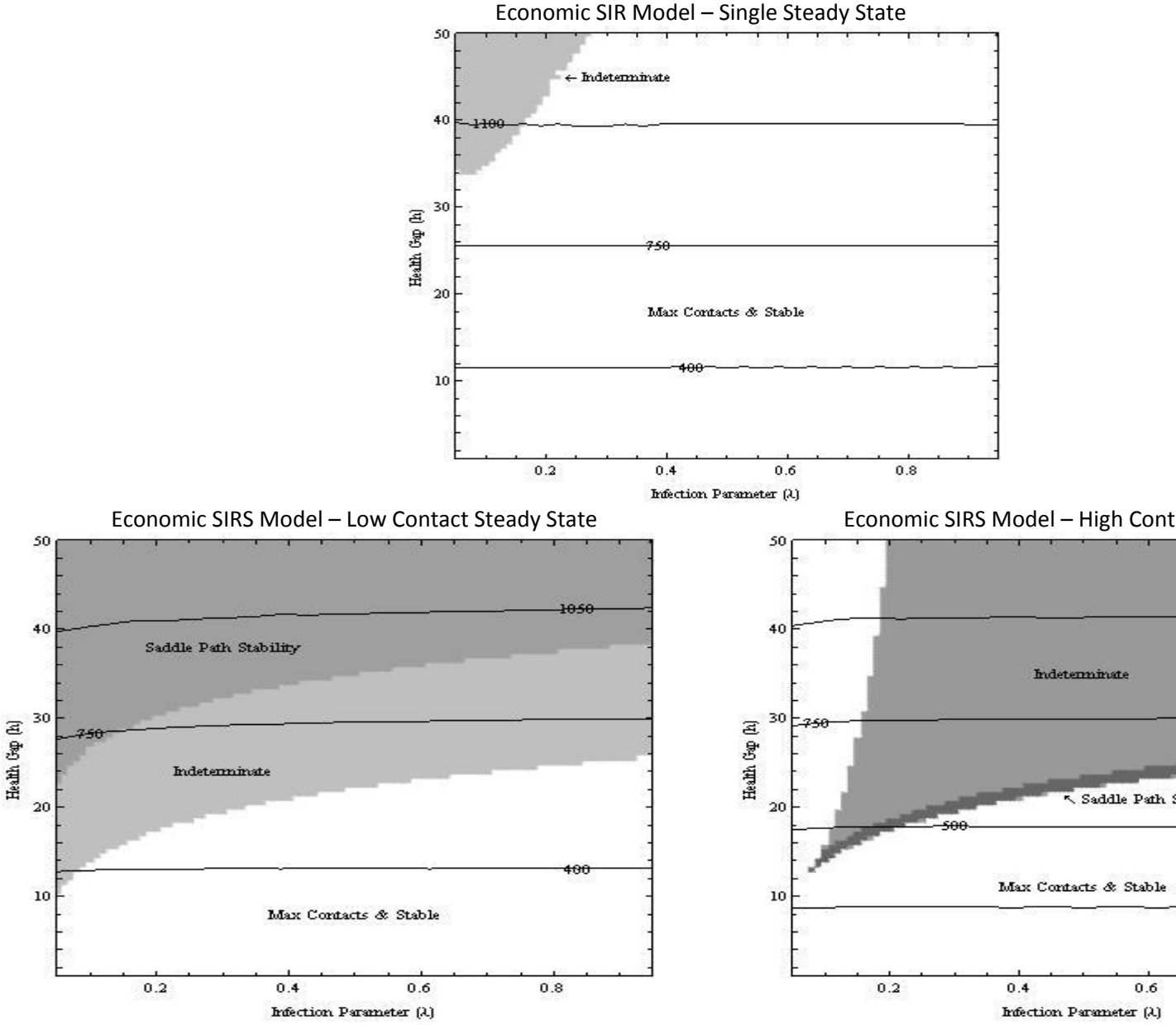
Notes. Parameters are given in Table 1 with $\lambda_p = 0.5$. Panel (a) depicts marginal benefit (MB) and marginal cost (MC) curves for three levels of the health gap (*h*). At a low *h* MB always lie above MC with no steady state *x*. As the health gap increases a single steady state *x* emerges at the tangency of MB and MC curves. Further increases in *h* lead to the emergence of two steady state *x* choices. Panel (b) demonstrates this bifurcation along a continuum of *h*. A solid (dashed) line depicts a locally stable (unstable) steady state.

Figure 2. Types of Local Dynamic Paths for Rational Expectations Economic SI and SIS Models



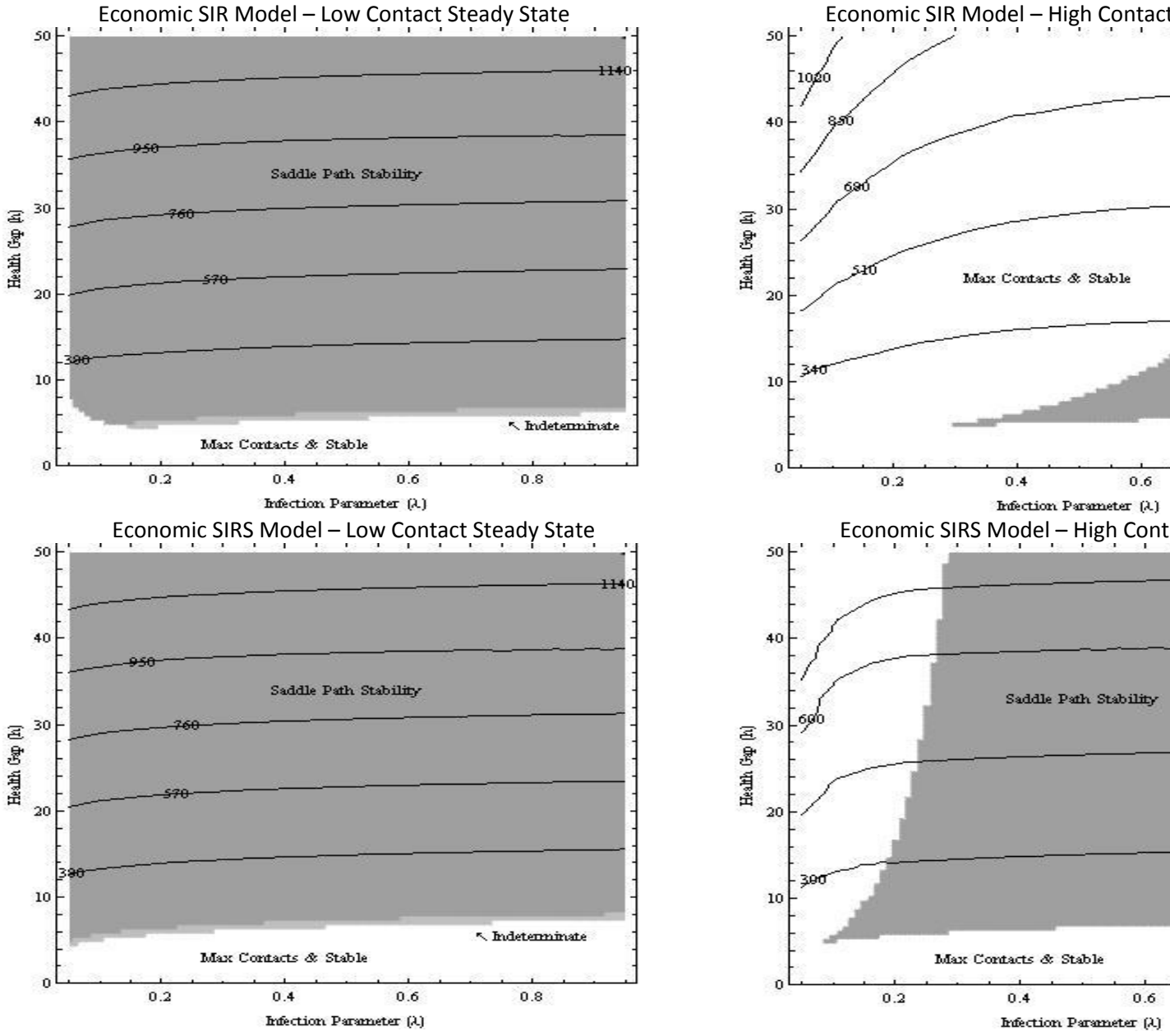
Notes: Contour lines indicate welfare at the steady state. Parameter values are given in Table 1.

Figure 3. Types of Local Dynamic Paths for Rational Expectations Economic SIR and SIRS Models (w/ C



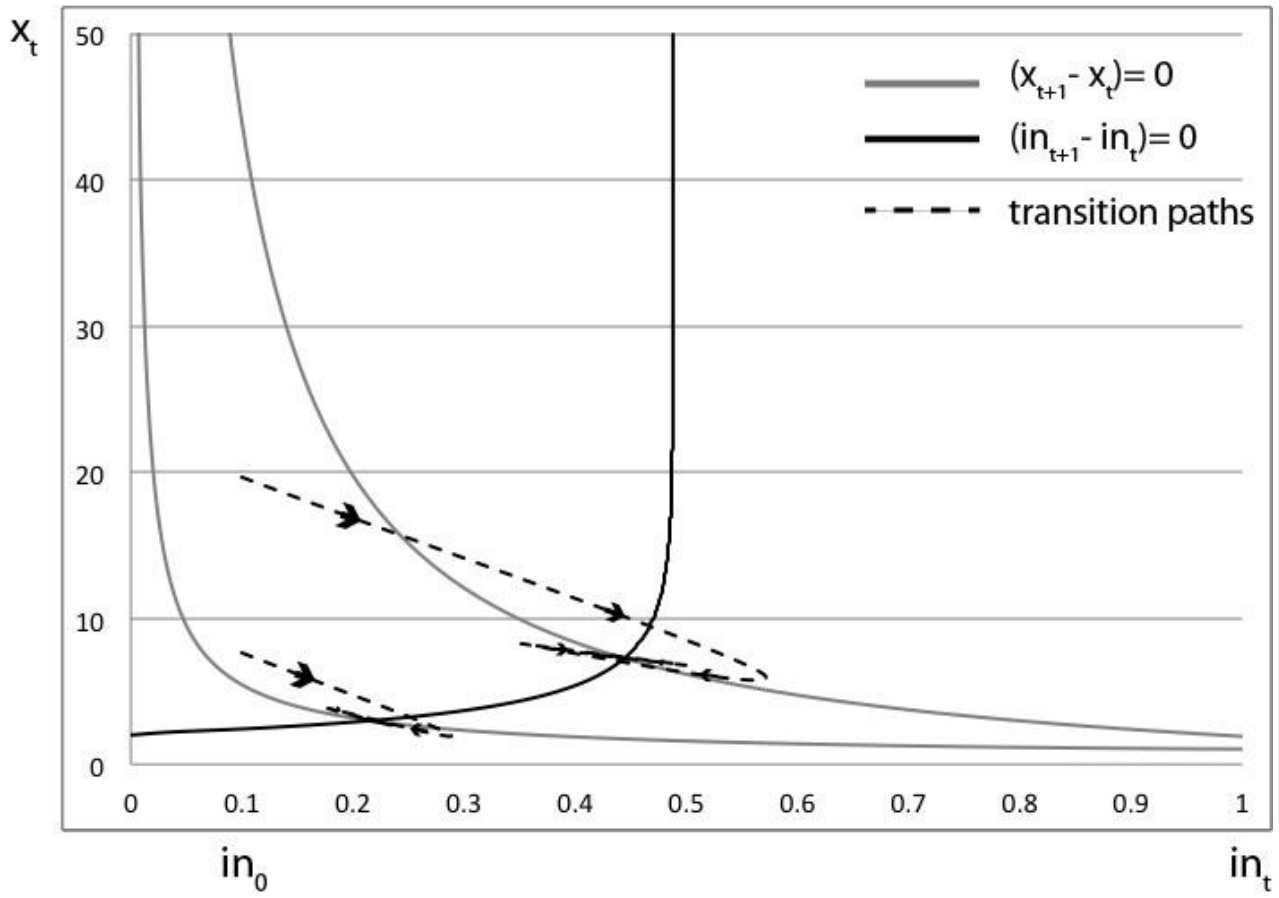
Notes: Contour lines indicate welfare at the steady state. Parameter values are given in Table 1.

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



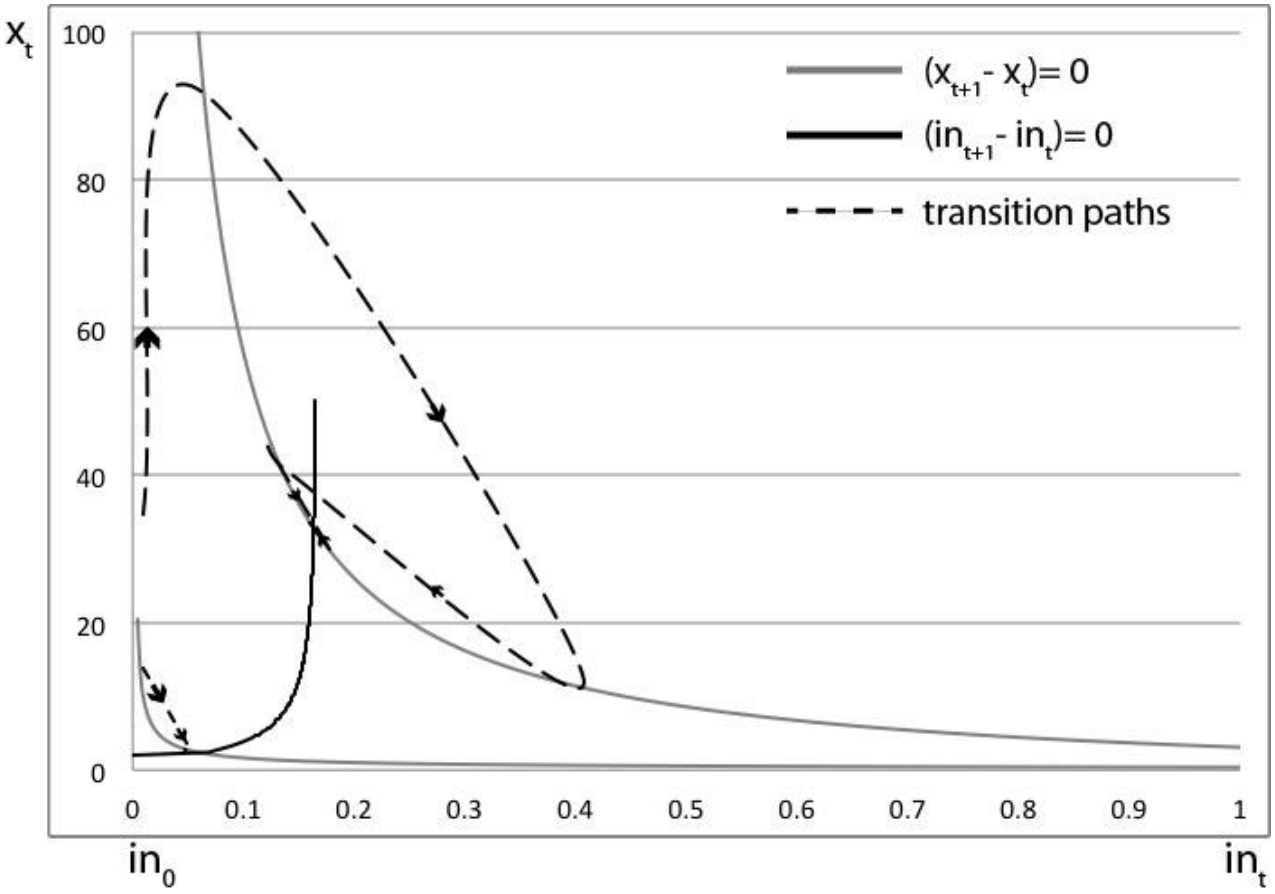
Notes: Contour lines indicate welfare at the steady state. Parameter values are given in Table 1.

Figure A1. Transition Paths for the Economic SIS Model



Notes. Parameters are given in Table 1. The solid black line is the prevalence isocline derived from equation (A.2). The grey lines are the contact isoclines derived from equation (A.7). The intersection of the prevalence and contact isoclines gives the two steady states. The dashed lines depict the transition paths to the respective steady states starting at an initial condition of in_0.

Figure A2. Transition Paths for the Economic SIRS Model with Observable Immunity



Notes. Parameters are given in Table 1. The solid black line is the prevalence isocline derived from equation (4) in the main text. The grey lines are the contact isoclines derived from equation (11) in the main text. The intersection of the prevalence and contact isoclines gives the two steady states. The dashed lines depict the transition paths to the respective steady states starting at an initial condition of in_0.