

ESCAPE DYNAMICS: A CONTINUOUS—TIME APPROXIMATION

Dmitri Kolyuzhnov
Anna Bogomolova
Sergey Slobodyan

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

Working Paper Series
(ISSN 1211-3298)

285

**Escape Dynamics:
A Continuous—Time Approximation**

Dmitri Kolyuzhnov
Anna Bogomolova
Sergey Slobodyan

CERGE-EI
Prague, January 2006

ISBN 80-7343-080-0 (Univerzita Karlova v Praze, CERGE)
ISBN 80-7344-069-5 (Národohospodářský ústav AV ČR, Praha)

Escape Dynamics: A Continuous–Time Approximation

Dmitri Kolyuzhnov, Anna Bogomolova, and Sergey Slobodyan*

CERGE-EI[†]
Politických vězňů 7, 111 21 Praha 1,
Czech Republic

Abstract

We extend a continuous–time approach to the analysis of escape dynamics in economic models with adaptive learning with constant gain. This approach is based on applying results of continuous–time version of large deviations theory to the diffusion approximation of the original discrete–time dynamics under learning. We characterize escape dynamics by analytically deriving the most probable escape point and mean escape time. The continuous–time approach is tested on the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent (1999) and Cho, Williams and Sargent (2002) (henceforth, CWS). We compare the results with simulations and the results obtained by CWS. We express reservations regarding applicability of escape dynamics theory to characterization of mean escape time for economically plausible values of constant gain in the model of CWS. We show that for these values of the gain simple considerations and formulae generate much better mean escape time results than the large deviations theory. We explain it by insufficient averaging near the point of self–confirming equilibrium for relatively large gains and suggest two changes which might help the approaches based on large deviation theory to work better in this gain interval.

Abstrakt

Rozšiřujeme časově spojitý přístup v analýze únikové dynamiky v ekonomických modelech s adaptivním učením a konstantním přínosem. Tento přístup je založen na aplikaci výsledků časově spojitě verze teorie velkých odchylek při aproximaci limitního časově spojitého procesu (difuze) u originálního diskretního modelu ve verzi s učením. Únikovou dynamiku charakterizujeme analytickým odvozením nejpravděpodobnějšího bodu úniku a průměrné doby úniku. Časově spojitý přístup je testován na Phelpsově problému vlády kontrolující inflaci při jejím adaptivním učením přiblížení se Phillipsově křivce, dříve zkoumaném Sargentem (1999) a Cho, Williams a Sargentem (2002) (dále jen CWS). Porovnáваме výsledky se simulacemi a výsledky, ke kterým dospěli CWS. Vyjadřujeme výhrady ohledně aplikovatelnosti teorie

*{Dmitri.Kolyuzhnov, Anna.Bogomolova, Sergey.Slobodyan}@cerge-ei.cz. The second author acknowledges support provided by the World Bank through 2nd year student fellowship. We thank Radim Bohacek, Jeong Byeongju, Michal Kejak, Andreas Ortmann, and especially Viatcheslav Vinogradov for comments and suggestions. All remaining errors are ours.

[†]CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

únikové dynamiky při charakterizaci průměrné doby úniku u ekonomicky přijatelných hodnot konstantního přínosu v modelu CWS. My ukazujeme, že pro tyto hodnoty přínosu poskytují jednoduché úvahy a rovnice o mnoho lepší výsledky ohledně průměrné doby úniku než teorie velkých odchylek. Vysvětlujeme to nedostatečným průměrováním v blízkosti bodu „samopotvrzujícího“ bodu rovnováhy pro relativně velké přínosy a navrhuje dvě změny, které by mohly pomoci přístupům založeným na teorii velkých odchylek fungovat lépe tomto intervalu.

JEL Classification: C62, C65, D83, E10, E17

Keywords: constant gain adaptive learning, E-stability, recursive least squares, large deviations theory

1 Introduction

The aim of this paper is to extend the continuous-time approach to the analysis of escape dynamics in economic models with adaptive learning and to test it on the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent (1999) and Cho et al. (2002). The idea to extend the continuous-time approach is motivated by the restricted applicability and computational intensity of the approach used to derive theoretical characteristics of escape dynamics in the recent economic literature. Theoretical analysis of escape dynamics in economic models with adaptive learning allows to theoretically characterize diverse economic phenomena such as currency crises, inflation episodes, endogenous collusion in oligopoly, cycles of economic activity, see Cho and Kasa (2003); Williams (2001, 2002, 2003, 2004); Bullard and Cho (2005); Cho et al. (2002); and Kasa (2004). Escape dynamics also was used to study large mutations in evolutionary games, see Kandori, Mailath and Rob (1993); and Binmore and Samuelson (1997).

In this literature, these phenomena are modeled as a result of escape dynamics in economic models with boundedly rational economic agents who use adaptive learning in a form recently summarized in Evans and Honkapohja (2001) to update their beliefs about economic models. Among the literature devoted to this form of adaptive learning are Bray (1982); Bray and Savin (1986); Fourgeaud, Gourieroux and Pradel (1986); Marcet and Sargent (1989); Evans and Honkapohja (1994a); Evans and Honkapohja (1994b); Evans and Honkapohja (1995); Marimon (1997); and many others. In this literature, agents are considered as econometricians who estimate forecasting models using standard statistical procedures, such as recursive least squares, stochastic gradient, or Bayesian learning, and form beliefs about an economic model. The beliefs thus formed are then used to generate agents' actions, and thus influence the realized values of economic variables which are taken as a new data point by the agents. In the next period, the agents update their beliefs with the new data. New beliefs then affect actions and economic variables, and this process repeats period after period.

Combined dynamics of parameters describing agents' beliefs and of observed economic variables form a stochastic recursive algorithm (SRA). Under some regularity conditions, the SRA corresponding to a particular adaptive learning process converges to the rational expectations equilibrium (REE) of the model,¹ and thus limit dynamics

¹or one of the REE in multiple equilibria models

under adaptive learning is the same as that under rational expectations. Stability under adaptive learning which guaranteed such convergence has been considered a very important characteristic of the REE in recent monetary policy literature, c.f., Evans and Honkapohja (2003) or Bullard and Mitra (2002).

Beyond using adaptive learning as a *de facto* equilibrium selection mechanism or a tool for designing policy rules, one could concentrate on the dynamics of the model under adaptive learning as such, in particular, in a case of adaptive learning with a constant gain.² In this case, convergence of the learning process to REE is only in distribution: there are persistent fluctuations around the REE caused by such learning, and thus rare events — large distance movements called “escapes” — may occur with nonzero probability. During an escape, agents’ beliefs about the model move away from nearly rational expectations. As a rule, their actions and the values of realized economic variables also deviate from those observed in the REE.

The analysis of such escape dynamics caused by the adaptive learning process is possible using the theory of large deviations by Freidlin and Wentzell (1998) (FW henceforth); Dupuis and Kushner (1989); and others. Depending on what version of the large deviations theory — continuous-time by FW or discrete-time by Dupuis and Kushner (1989) — one wants to utilize, there are two possible approaches to the theoretical analysis of escape dynamics: the discrete-time approach and the continuous-time approach. The discrete-time approach, which has received wider attention in the literature, is based on the analytical derivation of escape dynamics for the original discrete-time SRA used to describe a learning process. In the continuous-time approach, a continuous-time diffusion approximation of the discrete-time SRA is derived, and then escape dynamics is studied for this approximation.³

The first approach was used in the majority of the papers cited above, in particular in Cho et al. (2002) (henceforth CWS). These papers work directly with discrete-time SRA processes and use the recent results of Williams (2001), who derived numerically the action functional for a linear-quadratic case when the state variable process is autoregressive with Gaussian noise.

There are three basic problems associated with the above approach. First, if the state variable process is subject to unbounded (for example, Gaussian) shocks, the

²Constant gain learning discounts the past by assigning more weight to more recent data.

³The idea of approximating discrete-time learning with continuous-time process in order to apply FW large deviations theory results is due to Kasa (2004) who considered a simple one-dimensional model.

discrete time version of large deviations theory does not contain theoretical results allowing for a full description of escape dynamics. In particular, the most probable point of escape from the neighborhood of convergence point (as stated above, this point is usually a REE) and the expected time until escape, are unavailable, see CWS, Theorem 5.3. Second, characterizing escape dynamics for the discrete-time process in a way proposed by Williams (2001) implies numerical calculation of a functional in a calculus-of-variation problem that leads to a system of non-linear differential equations with numerically derived right hand side functions. For complicated problems (many lags, high dimensionality) this approach can become numerically intractable. Finally, analytical solution for escape dynamics of a discrete-time process can be derived only for a restrictive form of learning processes, such as recursive least squares with a constant gain or stochastic gradient learning.

The continuous time approach developed here resolves these problems. Since a diffusion is a natural approximation for a difference equation with Gaussian noise and since FW have developed the theory of large deviations for diffusions, the problem of insufficient theoretical results is removed. The second and the third problems are partially alleviated because the diffusion, derived by approximation around REE — the stationary point of the SRA, is linear. In the large deviations theory, all escape dynamics characteristics — expected time until the beliefs escape any given neighborhood D , the point through which this escape is most likely, and probability of leaving D within a given amount of time — are obtained by minimizing a so-called action functional on the boundary of the neighborhood, ∂D . Given our choice of a linear approximating diffusion, this is a standard linear control theory problem, and the problem of minimizing the action functional is reduced to the trivial task of finding a minimum of a quadratic form on ∂D .

In order to compare the performance of the two approaches of deriving escape-dynamics characteristics, the continuous-time approach is tested on the model where the escape dynamics characteristics were already derived using the discrete-time approach. This is the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent (1999) and CWS.⁴

⁴CWS show that under a constant-gain recursive least squares algorithm, the self-confirming equilibrium (SCE) — a unique set of beliefs corresponding to a time-consistent Nash equilibrium of the RE version of the model — is weakly stable. In this equilibrium, the government believes in strong inflation-unemployment tradeoff. Attempts to exploit this tradeoff, combined with the private sector's rational expectations, lead to high average inflation. However, inflation periodically performs large

The rest of the paper is organized as follows. We briefly describe the dynamic and static versions of the model of CWS in Section 2. We develop the continuous-time approach in Section 3. In Section 4, we present the results of testing the continuous-time approach developed in Section 3 on the model of CWS and compare the approach prediction results with the results of simulations. In Section 5, we discuss the results presented in Section 4 and compare them with the results of CWS, and Section 6 concludes.

2 The model

2.1 Setup: Two Versions of The Model

The economy consists of the government and the private sector. The government uses the monetary policy instrument x_n to control inflation rate π_n and attempts to minimize losses from inflation and unemployment U_n . It believes (in general, incorrectly) that an exploitable tradeoff between π_n and U_n (the Phillips curve) exists. The true Phillips curve is subject to random shifts and contains this tradeoff only for unexpected inflation shocks. The private sector possesses rational expectations $\hat{x}_n = x_n$ about the inflation rate, and thus unexpected inflation shocks come only from monetary policy errors.

$$U_n = u - \theta(\pi_n - \hat{x}_n) + \sigma_1 W_{1n}, \quad u > 0, \theta > 0, \quad (1a)$$

$$\pi_n = x_n + \sigma_2 W_{2n}, \quad (1b)$$

$$\hat{x}_n = x_n, \quad (1c)$$

$$U_n = \gamma_1 \pi_n + \gamma_{-1}^T X_{n-1} + \eta_n. \quad (1d)$$

Vector $\gamma = (\gamma_1, \gamma_{-1}^T)^T$ represents government's beliefs about the Phillips curve. W_{1n} and W_{2n} are two uncorrelated Gaussian shocks with zero mean and unit variance. η_n is the Phillips curve shock as perceived by the government, believed to be a white noise uncorrelated with regressors π_n and X_{n-1} . Following CWS, we consider two versions of the model: "dynamic" and "static" ones. In the "dynamic" model, vector X_{n-1} contains two lags of inflation and unemployment rates and a constant,

$$X_{n-1} = (U_{n-1}, U_{n-2}, \pi_{n-1}, \pi_{n-2}, 1)^T, \quad (2)$$

deviations, or "escapes", from a neighborhood of the RE Nash equilibrium toward the low inflation time-inconsistent Ramsey outcome of the RE version of the model. This happens when a sequence of stochastic shocks makes the government learn that there is very little tradeoff between unemployment and inflation. These beliefs about the Phillips curve force the government to set inflation low and thus approach the Ramsey outcome.

while only the constant is present in X_{n-1} in the “static” version. In other words, the only difference between the two versions of the model lies in the structure of the government’s beliefs (1d), which are more “sophisticated” in the dynamic model. In the sequel, we concentrate on the dynamic model and consider the static one only in Section 6.

Given beliefs γ , the government solves

$$\min_{\{x_n\}_{n=0}^{\infty}} E \sum_{n=0}^{\infty} \beta^n (U_n^2 + \pi_n^2), \quad (3)$$

subject to (1b) and (1d). This Linear–Quadratic problem produces a linear monetary policy rule

$$x_n = h(\gamma)^T X_{n-1}. \quad (4)$$

2.2 Nash, Ramsey, and Self–Confirming Equilibria

CWS identify three beliefs consistent with the model. Belief 1, $\gamma = (-\theta, 0, 0, 0, 0, u(1 + \theta^2))^T$, generates policy function $x_n = \theta u$. In a model where the government knows the true Phillips curve (1a), this is the Nash, or discretionary equilibrium of Sargent (1999) and Barro and Gordon (1983). Beliefs 2 of the form $\gamma = (0, 0, 0, 0, 0, u^*)^T$ lead to $x_n = 0$ and zero average inflation for any u^* : Ramsey, or the optimal time–inconsistent equilibrium of Kydland and Prescott (1977). Finally, Beliefs 3 where $\gamma_1 + \gamma_4 + \gamma_5 = 0$ (sum of coefficients on current and lagged inflation is zero) asymptotically lead to $x_n = 0$: this is an “induction hypothesis” belief, see Sargent (1999).

In the model with learning, equilibrium is defined as a vector of beliefs at which the government’s assumptions about orthogonality of η_n to the space of regressors are consistent with observations:

$$E \left[\eta_n \cdot (\pi_n, X_{n-1})^T \right] = 0. \quad (5)$$

CWS call this point a *self–confirming equilibrium*, or SCE: despite the fact that government believes in an incorrect Phillips curve (1d), a particular assumption about it, exemplified by (5), turns out to be true. Williams (2001) shows that the only SCE in the model is Belief 1.

2.3 Adaptive Learning and SRA

In a period n , the government uses its current vector of beliefs γ_n to solve (3), assuming the beliefs will never change. The generated monetary policy action x_n is correctly

anticipated by the private sector and produces U_n according to (1a). Then the government adjusts its beliefs about the Phillips curve coefficients γ_n and their second moments matrix R_n in an adaptive learning step. Define $\xi_n = [W_{1n} \ W_{2n} \ X_{n-1}^T]^T$, $g(\gamma_n, \xi_n) = \eta_n \cdot (\pi_n, X_{n-1}^T)^T$, and $M_n(\gamma_n, \xi_n) = (\pi_n, X_{n-1}^T)^T \cdot (\pi_n, X_{n-1}^T)$. Next period's beliefs γ_{n+1} and R_{n+1} are given by

$$\gamma_{n+1} = \gamma_n + \epsilon_n R_n^{-1} g(\gamma_n, \xi_n), \quad (6a)$$

$$R_{n+1} = R_n + \epsilon_n (M_n(\gamma_n, \xi_n) - R_n). \quad (6b)$$

Equations (6) represent a specific form of a recursive learning algorithm. When the gain sequence ϵ_n is given by $1/n$, an appropriate choice of γ_0 and R_0 generates OLS in a recursive form. When $\epsilon_n = \text{const}$, this is a *constant gain* learning or tracking algorithm.⁵

As $U_n = u - \theta\sigma_2 W_{2n} + \sigma_1 W_{1n}$ and $\pi_n = h(\gamma_n)^T X_{n-1} + \sigma_2 W_{2n}$, the evolution of the state vector ξ_n can be written as

$$\xi_{n+1} = A(\gamma_n)\xi_n + BW_{n+1}, \quad (7)$$

where $W_{n+1} = [W_{1n+1} \ W_{2n+1}]^T$, for some matrices $A(\gamma_n)$ and B . Finally, stack lower-triangular elements of the symmetric matrix R_n into a vector, $\text{vech}(R_n)$, and form the parameter vector

$$\theta_n^\epsilon = [\gamma_n^T, \text{vech}^T(R_n)]^T \quad (8)$$

and the right-hand side vector

$$H(\theta_n^\epsilon, \xi_n) = \left[(R_n^{-1}g(\gamma_n, \xi_n))^T, \text{vech}^T(M_n(\gamma_n, \xi_n) - R_n) \right]^T. \quad (9)$$

Then the dynamics of the model under constant-gain learning can be written as

$$\theta_n^\epsilon = \theta_n^\epsilon + \epsilon H(\theta_n^\epsilon, \xi_n), \quad (10a)$$

$$\xi_{n+1} = A(\gamma_n)\xi_n + BW_{n+1}, \quad (10b)$$

which is the standard SRA form.⁶

⁵Constant gain algorithm's assigning more weight to recent data makes sense when agents suspect the world around them to be non-stationary. Presence of sudden breaks in data generating processes, for example as a result of an unpredicted change in the the government policy, also calls for tracking algorithms such as constant gain learning. See Evans and Honkapohja (2001) for an extensive discussion of constant gain learning and its relation to decreasing gain learning such as OLS.

⁶Note that vector θ is 27-dimensional, with 6 components representing the government's beliefs γ , and the remaining 21 its beliefs about the second moments matrix of γ .

3 Continuous–Time Approach

3.1 Convergence of SRA and Diffusion Approximation

Define the approximating ordinary differential equations corresponding to our SRA as

$$\dot{\gamma} = R^{-1}\bar{g}(\gamma) = R^{-1}E[g(\gamma, \xi_n)], \quad (11a)$$

$$\dot{R} = \bar{M}(\gamma) - R = E[M_n(\gamma, \xi_n)] - R. \quad (11b)$$

Vector $\bar{\gamma}$ that forms Belief 1 and corresponding 2^{nd} moments matrix \bar{R} are the only equilibrium of the above ODE. This equilibrium is stable. CWS show that under some assumptions, the continuous–time process θ_t^ε defined as $\theta_t^\varepsilon = \theta_n^\varepsilon$ for $t \in [n\varepsilon, (n+1)\varepsilon)$ converges weakly (in distribution) to $\theta(t, a) = [\gamma^T, vech^T(R)]^T$, solution of the ODE (11), where $a = \theta(0)$ is the initial condition for the ODE (11), and starting point of the process θ_t^ε . This solution is also called the “mean dynamics trajectory” of the SRA (10), with the right–hand side of (10) being the “mean dynamics”.

Because of the constant gain learning, the convergence of θ_n^ε to the mean dynamics trajectory $\theta(t)$ is only weak (in distribution). This implies persistent fluctuations around the trajectory $\theta(t, a)$ and its stationary point $\bar{\theta}$. Large deviations theory studies the probability of rare events, during which these fluctuations force the stochastic process θ_n^ε out of any given region around the converging trajectory $\theta(t, a)$. Freidlin and Wentzell (1998, p.6) state that the probabilities of these rare events “have asymptotics of the form $\exp\{-C\varepsilon^{-2}\}$ as $\varepsilon \rightarrow 0$ (rough asymptotics, i.e., not up to equivalence but logarithmic equivalence)”.

The theoretical results of FW on escape dynamics characteristics in continuous time can be applied to the continuous–time approximation of the original discrete–time SRA. This is the essence of the proposed continuous–time approach.⁷ Evans and Honkapohja (2001, Prop. 7.8) show that as $\varepsilon \rightarrow 0$ the process $U_t^\varepsilon = \frac{\theta_t^\varepsilon - \theta(t, a)}{\sqrt{\varepsilon}}$ converges (weakly) to a following diffusion:

$$dU_t^\varepsilon = D_\theta p(\theta(t, a))U_t^\varepsilon dt + \Sigma^{1/2}(\theta(t, a)) dW_t, \quad (12)$$

where W_t is a multi–dimensional Brownian process with dimensionality equal to that of θ . $p(\theta)$ is the mean dynamics vector, and Σ the matrix whose elements are covariances of different components of the mean dynamics vector, both with respect to the unique

⁷The advantages of the continuous–time approach are discussed in the introduction.

invariant probability distribution $\Gamma_\theta(dy)$ of the state vector X :⁸

$$p(\theta) = \int H(\theta, y) \Gamma_\theta(dy), \quad (13)$$

$$\Sigma_{ij} = \sum_{k=-\infty}^{\infty} \text{Cov} [H_i(\theta, X_k(\theta)), H_j(\theta, X_0(\theta))]. \quad (14)$$

This result is used to get continuous-time approximation of SRA, given any initial condition:

$$d\theta_t^\varepsilon = D_\theta p(\theta(t, a)) [\theta_t^\varepsilon - \theta(t, a)] dt + \sqrt{\varepsilon} \Sigma^{1/2}(\theta(t, a)) dW_t. \quad (15)$$

Williams (2001, Theorem 3.2) shows that the above results can be used to derive a local continuous-time approximation of the SRA around the limit point $\bar{\theta}$ (stable point of the associated ODE (11), SCE):

$$d\varphi_t = D_\theta p(\bar{\theta}) \varphi_t dt + \sqrt{\varepsilon} \Sigma^{1/2}(\bar{\theta}) dW_t, \quad (16)$$

where $\varphi_t = \theta_t - \bar{\theta}$ are deviations from the SCE. The 6×6 upper left corner of $\Sigma(\bar{\theta})$ equals the fourth moments matrix Q of CWS evaluated at $\bar{\theta}$. Matrices $D_\theta p(\bar{\theta})$ and $\Sigma(\bar{\theta})$ need to be evaluated only at the SCE. This could be performed analytically (the technical appendix with these derivations is available from the authors upon request).

Diffusion (16), used in this paper, approximates a highly nonlinear multidimensional SRA only at the stationary point of the mean dynamics. Dembo and Zeitouni (1998, p. 223) argue that “the rationale here is that any excursion off the stable point has an overwhelmingly high probability of being pulled back there, and it is not the time spent near any part of ∂D that matters but the *a priori* chance for a direct, fast exit due to a rare segment in the Brownian motion’s path.”

3.2 Action Functional and Escapes

Suppose that we have a stochastic process, for example some diffusion. The basic idea of the theory of large deviations for paths of stochastic processes is that the probability of a stochastic process’s deviating from a given path along a specific trajectory can be determined by the value of a certain functional (called *action functional*) on this trajectory. Action functional $I_{0T}(\varphi)$ represents the costs associated with moving along some trajectory φ for a period of time $[0, T]$. Cost function $I(T, x, y) = \min_{\varphi_0=x, \varphi_T=y} I_{0T}(\varphi)$

⁸State vector ξ_n has a unique invariant probability distribution: it contains stationary Gaussian random variables W_{1n} and W_{2n} , a constant, and a stable 4-dimensional AR(1) variable. This distribution can be calculated explicitly.

is the minimal cost required for transition from x to y in time T . *Quasipotential* $I(x, y) = \inf_{T>0} I(T, x, y)$ is the minimal cost necessary to move from x to y given arbitrary (potentially infinite) time. The idea here is that the system moves in the direction along which it incurs the least cost.

Suppose that such a functional exists. We are given some neighborhood D of a stationary point of the diffusion's drift, O . Under certain assumptions one can derive the probability that a stochastic process belongs to D from the minimum value of the quasipotential $I(O, y)$ on the boundary of D , $\{y : y \in \partial D\}$. The most probable point at which the stochastic process leaves (escapes) D , is the point where $I(O, y)$ has a minimum. The minimum of $I(O, y)$ also allows one to derive asymptotic behavior of the mean escape time, *i.e.*, the expected time needed for the stochastic process to cross the boundary of D for the first time.

The exact results on the mean exit time and the dominant escape point are given in Dembo and Zeitouni (1998, Theorem 5.7.11). In particular, the limiting behavior of the mean escape time, $E_x(\tau^\varepsilon)$, is given by

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \ln E_x(\tau^\varepsilon) = \bar{I}, \quad (17)$$

where \bar{I} is a minimum value of the quasipotential on ∂D . The most probable escape point is an extremal of the quasipotential on ∂D , see Appendix A for exact definitions.

3.3 Minimizing the Action Functional

For a diffusion $d\varphi_t = A\varphi_t dt + \sqrt{\varepsilon} B dW_t$, Dembo and Zeitouni (1998, p. 214) provide the following expression for the action functional:

$$I_{0T}(\varphi) = \inf \frac{1}{2} \int_0^T \left| \dot{g}_t \right|^2 dt, \quad (18a)$$

$$\text{s.t. } \dot{\varphi}_t = A\varphi_t + B\dot{g}_t, \quad (18b)$$

$$\varphi_0 = 0, \quad (18c)$$

where a stationary point of the drift O is assumed to be the origin. Minimization is performed over all possible trajectories of $\dot{g}_t = u_t$ which take the system from the origin to φ_T in exactly T time units. In the approximating diffusion (16) the matrix A equals $D_{\theta p}(\bar{\theta})$, and $B = \Sigma^{1/2}(\bar{\theta})$.

The only complication with this formulation stems from the fact that matrix $B = \Sigma^{1/2}(\bar{\theta})$ can be singular.⁹ As a result, there might be points in the state space that

⁹It turns out that it is singular in the model of CWS. Singularity comes from collinearity of

could not be reached in any time using any control trajectory $\{u_t\}_{t=0}^{\infty}$: the system (A, B) is not necessarily reachable. The way to proceed with the control problem for an unreachable system is to transform the state space so that first k new coordinates (z_1) form the basis of the *reachable subspace*, where the remaining $n-k$ (z_2) coordinates all equal zero. In these coordinates, the system's evolution on the reachable subspace is governed by

$$\dot{z}_1 = \bar{A}_1 z_1 + \bar{B}_1 u, \quad (19)$$

where $z_1 = T_1^T \varphi$, T_1 is the basis of the reachable subspace, and the system (\bar{A}_1, \bar{B}_1) is by construction reachable, see Dahleh, Dahleh and Verghese (2004, Ch. 22) for the construction. The action functional (18) is then rewritten as

$$I_{0T}(z_1) = \inf \frac{1}{2} \int_0^T |\dot{g}_t|^2 dt, \quad (20a)$$

$$\text{s.t. } \dot{z}_1 = \bar{A}_1 z_1 + \bar{B}_1 \dot{g}_t, \quad (20b)$$

$$z_1(0) = 0. \quad (20c)$$

To find \bar{I} , one has to minimize $I_{0T}(z_1)$ over the time to escape T and all points $z_{1,D}$ such that $T_1 z_{1,D} \in \partial D$. In other words, the problem of finding minimum value of the action functional over all trajectories starting at the origin and terminating on ∂D in an arbitrary time is split into two separate problems: first, find the minimum norm control path, \dot{g}_t , which takes the linear control system from the origin to $z_{1,D}$ in arbitrary time, and then minimize over all possible terminal points $z_{1,D}$.

The first problem is a standard control problem with the following solution:

$$I(z_{1,D}) = \frac{1}{2} z_{1,D}^T \cdot \bar{G}^{-1} \cdot z_{1,D}, \quad (21)$$

where \bar{G} is Gramian of the reachable subsystem. See Appendix B for details and definition of matrices T_1 and \bar{G} . The problem of finding the minimum value of the action functional then becomes a trivial one: minimize the quadratic function of $z_{1,D}$ on $\{z_{1,D} : T_1 z_{1,D} \in \partial D\}$. By solving this problem, we find the most probable point of escape, $T_1 z_{1,D}$, and the rate of convergence, \bar{I} , that characterizes the limiting behavior of the mean escape time by the limit expression (17).

regressors at the SCE: Inflation rate and unemployment rate equal a constant plus *i.i.d.* noise. As a result, 14 out of 21 entries in \bar{R} are constants which do not depend on the noise magnitude, and the rank of matrix Σ equals $27 - 14 = 13$.

4 Testing the Approach on the Phelps Problem

4.1 Simulations and Reduced Dimensionality of the Model

In this section, we present the results of applying the continuous-time approach developed in the previous section to the Phelps problem studied previously by CWS who exploited the discrete-time approach. All the model parameters are taken to be the same as in CWS in order to guarantee the possibility of comparison between the results.¹⁰ Before applying the theory, we first run simulations of the model to have an initial picture of dynamics in the system.

Following CWS, we plot simulation runs in two-dimensional plane, where the abscissa is set to be the “inflation slope coefficient” (the sum of beliefs coefficients before inflation and lagged inflation, $\gamma_1 + \gamma_4 + \gamma_5$) in order to see how the system moves towards the “induction hypothesis” plane, $\gamma_1 + \gamma_4 + \gamma_5 = 0$. In contrast to CWS, we use as an ordinate the intercept coefficient summed with the “lagged unemployment slope coefficient” multiplied by the average unemployment rate, $\gamma_6 + u \cdot (\gamma_2 + \gamma_3)$, rather than the intercept coefficient, γ_6 , alone. The exact algebraic form of the coordinates used is explained by the following consideration.

Suppose that the government’s beliefs are fixed for a number of periods at γ_n , so that the state dynamics become stationary, with unconditional expectation of U_n being u and that of π_n being some $\tilde{\pi}$. What is the expected value of η_n ? $\eta_n = U_n - \gamma_1\pi_n - \gamma_{-1}^T X_{n-1}$, and so

$$E[\eta_n] = u - (\gamma_1 + \gamma_4 + \gamma_5)\tilde{\pi} - (\gamma_2 + \gamma_3)u - \gamma_6. \quad (22)$$

This back-of-the-envelope calculation suggests that from the government’s point of view, linear combinations $\gamma_1 + \gamma_4 + \gamma_5$ and $(\gamma_2 + \gamma_3)u + \gamma_6$ rather than the whole vector γ matter. As it is exactly a perceived error η_n which matters for the adjustment of θ in (6), one presumes that coordinates

$$(\tilde{\gamma}_1, \tilde{\gamma}_2) = (\gamma_1 + \gamma_4 + \gamma_5, u \cdot (\gamma_2 + \gamma_3) + \gamma_6) \quad (23)$$

are useful in thinking about the model.

The above coordinates are used to plot a typical simulation run started at SCE with $\epsilon = 0.001$, including an escape towards an “induction hypothesis” belief (Belief 3 of CWS) and very low inflation. In the $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ plane, all simulation points are very close

¹⁰Throughout the paper, we use the same parameter values as in CWS: $\sigma_1 = \sigma_2 = 0.3$, $u = 5$, $\theta = 1$, $\beta = 0.98$. All the figures are for simulations with $\epsilon = 0.001$, unless otherwise noted.

to a one-dimensional curve (a straight line). Very similar graphs are obtained in all 1000 runs, which strongly suggests that we could use coordinates $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ to effectively reduce the dimensionality of the system.

The reason for such almost “one-dimensionality” of the dynamics can be found by looking at the parameterical structure of the SRA. Consider equation (6a). At the SCE,¹¹ the largest eigenvalue of \bar{R}^{-1} is $\lambda_1 = 3083.8$ and the next largest $\lambda_2 = 29.1$, less than 1% of λ_1 .¹² Therefore, if one writes $g(\gamma_n, \xi_n)$ as a linear combination of eigenvectors of \bar{R}^{-1} , then the projection of $g(\gamma_n, \xi_n)$ onto v_1 , the eigenvector corresponding to λ_1 , is magnified 100 times as strongly as the projection onto v_2 . The dynamics described by $\bar{R}^{-1}g(\gamma_n, \xi_n)$ is, thus, almost 1-dimensional. In coordinates $(\tilde{\gamma}_1, \tilde{\gamma}_2)$, v_1 is approximately proportional to $(1, -5)'$. Fig. 1 shows that indeed, all the simulation run points are aligned along this vector.

One cannot help noticing the striking similarity between our Fig. 1 and Fig. 6 of CWS, which uses coordinates $(\hat{\gamma}_1, \hat{\gamma}_2) = (\gamma_1 + \gamma_4 + \gamma_5, \gamma_6)$. In $(\hat{\gamma}_1, \hat{\gamma}_2)$ space v_1 is proportional to $(1, -7.86)'$, which is again sufficiently close to the line drawn by simulation run points of CWS. Therefore, straight lines drawn in different coordinates by CWS and by us in this paper are nothing but projections of the “largest” eigenvector of \bar{R}^{-1} , v_1 , onto different hyperplanes.¹³

It is possible to neglect dynamics of R in considering escapes because the covariance matrix $\Sigma_{\bar{\theta}} = cov[H(\bar{\theta}), X(\bar{\theta})]$ contains \bar{R}^{-1} in its upper-left corner, and its largest eigenvalue’s eigenvector \tilde{v}_1 is proportional to $(1, -5)'$ in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ coordinates. The parameterical structure of $\Sigma_{\bar{\theta}}$ is close to block-diagonal, so there is very little interaction between RHS terms of (11), $R_n^{-1}g(\gamma_n, \xi_n)$ and $M_n(\gamma_n, \xi_n) - R_n$, which influence the components of γ and of R , at least for the “largest” eigenvectors which determine the dynamics of the model.¹⁴

After reducing the dimensionality of the model, it is possible to analyze the behavior

¹¹SCE is a starting point of the process for all simulations and theoretical derivations.

¹²This dramatic difference is due to the fact that u is so large. Entries of \bar{R} are of order u^2 , u , and 1. Large u leads to significantly different entries of \bar{R} and thus to a large λ_1/λ_2 ratio. For $u = 1$, the ratio λ_1/λ_2 drops to 5.17.

¹³The fact that simulation run points plot almost an ideal straight line suggests that the matrix R does not change much along the typical escape path, preserving the ratio of the largest to the second largest eigenvalue and the direction of the “largest” eigenvector. This conjecture turns out to be correct for relatively large values of ϵ .

¹⁴Matrix $\Sigma_{\bar{\theta}}$ is close to being block-diagonal, similarly to $R(\bar{\theta})$ in Evans and Honkapohja (2001, Eq. 14.6), in the following sense: If one takes one “largest” eigenvector of 6×6 upper left corner of $\Sigma_{\bar{\theta}}$ and two “largest” eigenvectors of 21×21 lower right corner and pads them with zeros appropriately, the resulting vectors are almost undistinguishable from the three largest eigenvectors of the whole $\Sigma_{\bar{\theta}}$ (for a block-diagonal matrix, they should be exactly equal).

of the mean and the stochastic part of dynamics using the simulations results. Several trajectories of the mean dynamics of the model, given by (11), are presented in Fig. 2 in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ coordinates. The region around the SCE where the mean dynamics points back towards it is very small; if initial deviation from the SCE is relatively large, the mean dynamics trajectory treks towards Belief 3, or the “induction hypothesis” plane, where $\tilde{\gamma}_1=0$. After spending some time in the neighborhood of $\tilde{\gamma}_1=0$, the trajectory slowly returns back to the SCE. The right panel of Figure 2 tracks several mean dynamics trajectories as they travel to the “induction hypothesis” plane. The paths are almost indistinguishable at this scale: Away from the immediate neighborhood of the SCE, mean dynamics trajectories are rapidly converging to the line connecting the SCE with (0,5).

To understand the relation between the mean and the stochastic parts of the dynamics of (6), consider Fig. 3. It plots a ratio of the relative magnitude of the stochastic dynamics, given by $\|R^{-1}\{g(\gamma, \xi_n) - E[g(\gamma, \xi_n)]\}\|$ averaged over 4000 realizations of ξ_n , and of the mean dynamics $\|R^{-1}E[g(\gamma, \xi_n)]\|$. The ratio is evaluated at different points θ along the eigenvector \tilde{v}_1 . For large deviations from the SCE, the mean dynamics dominates the stochastic part. In the small region around SCE where the mean dynamics points back towards it, stochastic dynamics is on average hundreds and thousands times larger than the mean dynamics.

Using the theoretical derivations and the knowledge of the system obtained through simulations, we think about escapes in the following way. Consider a small neighborhood D of the SCE. After the trajectory crosses the boundary ∂D , we assume that the stochastic dynamics does not play any role, and the model’s behavior is determined exclusively by its mean dynamics (11). Arguments presented in the previous paragraph allow us to claim that this is a very good approximation far from the SCE. Moreover, as all mean dynamics trajectories are very close in this region, one does not need to know the exact escape point to predict the most likely behavior of the system during travel to the low inflation outcome. A process of excursion towards $\tilde{\gamma}_1 = 0$ is, therefore, split into two parts: first, stochastic “escape” from D , and second, almost deterministic movement to $\tilde{\gamma}_1 = 0$ and back to the SCE. If our selection of ∂D is such that after crossing it the mean dynamics points towards $\tilde{\gamma}_1 = 0$, there is no additional contribution to the quasipotential as the system does not need any additional energy to move away from the SCE. Therefore, we concentrate on the first part, the stochastic “escape” from the set D , such that after the escape the mean dynamics points away

from the SCE.

4.2 Analytical Results vs. Simulations: Point of Escape

We select the set D and its boundary ∂D in several ways. The first way is to use a cylinder: a sphere in a six-dimensional γ space, and no binding restrictions in 21-dimensional space of components of R . This approach is similar to the road taken by CWS. The theoretical results of the problem of minimizing action functional on the cylinder are presented in Appendix B. This simple calculation does not replicate the behavior observed in simulations for gain value $\epsilon = 0.001$:¹⁵ no matter what the cylinder’s radius is, the escape from the SCE is predicted to occur in the approximate direction $(1, -6.5)'$ in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ coordinates. When one selects the cylinder’s radius so that the cylinder crosses \tilde{v}_1 at $\tilde{\gamma}_1 = -0.985$ (more on this choice below), the distance between the mean of observed escape points and the theoretically derived escape point, $d1$, equals 100% of the distance from the SCE to the theoretical escape point. Fig. 5 presents a histogram of observed escape directions.¹⁶ As is easy to see, the theoretical prediction of approximately -6.5 is way off the mode of the empirical distribution, which is -5. Fewer than 1% of simulation runs result in escape in the direction with tangent less than -6. Therefore, we conclude that the cylinder is not a good choice for the escape region D , at least at this value of ϵ . Additionally, simulation runs show that many of the “escapes” generated in this way violate our basic assumption: mean dynamics paths, initiated at the escape point, do not deviate towards Belief 3, see Fig. 6, where such points are marked by green crosses.

Our second way of selecting the boundary of the set D is the curve in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space, defined numerically as a path “separating” trajectories coming back to the SCE from those which first travel to the “induction hypothesis” plane under the mean dynamics.¹⁷ We derive this curve as follows. Find two points on the eigenvector \tilde{v}_1 , such that one of them generates a trajectory of (11) immediately coming back to the SCE, while another starts an excursion toward the induction hypothesis plane. Using a sequence

¹⁵Gain value $\epsilon = 0.001$ is chosen for a comparison of simulations with analytical results as the lower boundary for economically plausible values of ϵ . Large deviation theory has to work for 0.001 if it works for economically plausible values of $\epsilon \in [0.001, 0.01]$ at all.

¹⁶These tangents are obtained as follows. For a given simulation run, determine the point of escape from the cylinder and project it into $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. Write the vector which starts at the SCE and points towards the escape point as $(1, Tg)'$. Then Tg is a simply tangent of the angle between $(1, 0)'$ and $(1, Tg)'$. Fig. 2 shows the histogram of Tg .

¹⁷This surface is not a separatrix in the strict sense of the word, as all trajectories eventually return back to the SCE and asymptotically converge to it. However, there is a sensitive dependence on the initial conditions in the neighborhood of this surface.

of binary bisections, shrink the interval between two such points to any desired small number. Starting from this point, solve (11) forward and backward in time. Project the resulting trajectory into $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space and call this curve ∂D . In the left panel of Figure 2, we plot two trajectories that turn back to the SCE (two leftmost lines), and two paths travelling to the “induction hypothesis” plane first and then coming back (two rightmost lines). The “separating” curve is plotted as a dashed line. At the point where the “separating” curve intersects with eigenvector \tilde{v}_1 , $\tilde{\gamma}_1$ approximately equals -0.985. Radius of the cylinder, described above, was selected in such a way that it intersects \tilde{v}_1 at the same point as the “separating” curve.

Deriving a “separating” surface has the advantage of taking into account some information about behavior of the mean dynamics away from the SCE. The linear approximating diffusion (16) discards this information by taking into account only $D_{\theta p}(\bar{\theta})$ rather than $D_{\theta p}(\theta)$. On the other hand, this procedure is very simplistic and heuristic: assuming that such a separating surface exists in the original 27-dimensional space, there is no particular reason to believe that its projection into two-dimensional $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space coincides with, or is concentrated around, the projection of one particular path. Nevertheless, Figure 6 shows that empirically the separating curve does a good job at least in some respects: consider projections of all 1000 simulated points of escape from the cylinder described above. We use these escape points as initial values for the trajectory of (11). If this trajectory comes back to the SCE, projection of the escape point is plotted using a ‘+’ symbol in Figure 6. If the path first travels to the “induction hypothesis” plane, the corresponding projection is marked as ‘o’. Escape points with projections to the left of the separating curve tend to start converging trajectories while those projected to the right of the curve initiate an excursion towards Belief 3. There are several points well to the left of the separating curve which nevertheless start excursions. These points represent very unlikely escapes, during which the structure of the matrix R changes a lot; they are actually very far from \tilde{v}_1 in the original 27D space.

Using a thus constructed boundary improves the match to simulations for gain value $\epsilon = 0.001$. The theoretical escape calculated in this way occurs through the point that lies in the approximate direction (1,-4) in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ coordinates. The distance between the average of empirical escape points and the theoretical escape point equals $d_2=31\%$ of the distance between the SCE and the theoretical point. Figure 7 shows that there is some accumulation of the escape tangents towards -4; however, the majority tends

to center on -5, exactly as in the case of escape from the cylinder.

According to point (b) of the Theorem A.1, as $\epsilon \rightarrow 0$, the probability of observing the escape in the neighborhood of the point minimizing action functional converges to one. This is not the behavior observed in our simulations: For both boundaries used above, simulated escapes tend to occur close to points in (1,-5) direction rather than the theoretically predicted (1,-6.5) or (1,-4).

The source of the discrepancy lies in the non-validity of continuous-time approximation for the values of ϵ used in the economic literature and for simulations in the comparison above. To see it, one could compare the mean dynamics given by (11) and the realizations of the stochastic process simulated in (6). As is mentioned above, for θ near the SCE such that the mean dynamics points towards $\bar{\theta}$, mean dynamics magnitude ($R^{-1}E[g(\gamma, \xi_n)]$) is dramatically lower than the typical realizations of $R^{-1}g(\gamma, \xi_n)$: the “noise-to-signal” ratio, $\frac{\|R^{-1}(g(\gamma, \xi_n) - E[g(\gamma, \xi_n)])\|}{\|R^{-1}E[g(\gamma, \xi_n)]\|}$, ranges from hundreds to several thousand, depending on θ . Figure 3 plots this ratio for points along the eigenvector \tilde{v}_1 , averaged over 4000 realizations of (W_{1n}, W_{2n}) for every point. The horizontal axis shows corresponding $\tilde{\gamma}_1$ values. Near the SCE, the noise-to-signal ratio tends to be extremely high. In this situation, one expects (16) to be a good approximation of (6) only if the system stays in the neighborhood of every point θ long enough to allow the average of $R^{-1}g(\gamma, \xi_n)$ to approach $R^{-1}E[g(\gamma, \xi_n)]$. With noise-to-signal ratios from 10 to 1000, this means hundreds and thousands of iterations near **every** point θ . However, for the values of ϵ used in CWS and commonly applied elsewhere in the literature ($\epsilon=0.001\div 0.01$), the expected escape time is measured in hundreds of iterations: this is the time spent by the system near **all** points around the SCE. Therefore, $\epsilon=0.001\div 0.01$ is not small enough for the approximation (16) to be valid.

Given such large noise-to-signal ratio, one could simply disregard the mean dynamics (set $D_{\theta p}(\theta) = 0$) and repeat the minimization of the action functional. This is our third way of deriving escape dynamics. The region D is the cylinder described above. The theoretical results of the problem of minimizing action functional on the cylinder in the case of diffusion without drift term are presented in Appendix B. Theoretical escape occurs in the direction of the largest eigenvalue of Σ , \tilde{v}_1 .¹⁸ This way provides a much better agreement between the theory and the simulations: for the same radius of the cylinder as in the first approach, distance $d\mathcal{B}$ between the mean of observed escape

¹⁸This escape direction could be derived more easily: take the largest eigenvector of R^{-1} and project it into $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. The result is the same as the result based on the formula in Appendix B up to the third decimal point.

points and the theoretically predicted point is only 4.4% of the distance between the SCE and the theoretical point. Fig. 5 shows that most runs end in escapes along the direction $(1,-5)'$, which is the theoretically predicted one for this approach.

TABLE 1. Distance between the average of simulated escape points and the theoretically predicted escape points, expressed as a percentage of the distance between the SCE and the theoretically predicted point.

ϵ	Dynamic model			Static model	
	d1, %	d2, %	d3, %	d1, %	d3, %
$2 \cdot 10^{-5}$	75.45	86.72	99.99		
$3 \cdot 10^{-5}$	81.56	73.65	72.25	24.82	244.93
$5 \cdot 10^{-5}$	88.50	70.69	40.95	29.37	229.23
$1 \cdot 10^{-4}$	93.04	57.73	20.39	40.91	189.36
$2 \cdot 10^{-4}$	92.78	45.68	22.75	54.73	141.64
$4 \cdot 10^{-4}$	95.83	35.26	7.61	67.41	97.85
$1 \cdot 10^{-3}$	99.65	30.98	4.35	81.02	50.82
$1 \cdot 10^{-2}$	99.39	127.86	11.87	93.78	6.77

To support further our claim that $\epsilon = 0.001$ is not low enough to guarantee sufficient averaging, we have performed simulation runs for smaller values of ϵ . Consider Fig. 9, which plots a histogram of escape directions from the cylinder, for a 1000 simulations with $\epsilon = 2 \cdot 10^{-5}$. Comparison with Fig. 5 shows that one indeed observes an accumulation of escape directions towards the theoretically predicted direction of -6.5, but there is still a long way to go: A full 14% of escapes occur in the bin centered on -5 with the width of 0.25, and only 66% of escape directions' tangents are in bins with centers below -5. Given that theoretically one expects a mode at -6.5, we calculated the share of escapes in the bin centered at -6 and below it. Only 9.6% of escapes fall in this category, which is still much better than less than 1% observed for $\epsilon = 1 \cdot 10^{-3}$.

In contrast, the performance of the second way of selecting ∂D does not improve as ϵ decreases.¹⁹ At $\epsilon = 1 \cdot 10^{-3}$, only 41% of escapes are at -5, and 26% escape at the “correct” -4.5 or above (recall that the second way predicts a escape direction of -4).

¹⁹The reason for the failure of the second way for smaller ϵ is clear: for large ϵ , the majority of escapes are concentrated along \tilde{v}_1 , and we used a point on \tilde{v}_1 as a starting point in deriving the trajectory which, after being projected into the $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ plane, became the separating curve. As a result, we are relatively confident that this curve is a good description of the true 27-dimensional separating surface for majority of escaping trajectories. As ϵ decreases, more and more escapes start to take place away from \tilde{v}_1 (and thus, away from the initial point used to derive the separating curve), and the curve stops working as a result. We could use a point on the vector $T_1 \bar{G}^{1/2} \xi$ derived in Appendix B as the initial point for deriving the separating curve. We believe that this trajectory, projected into the $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ plane, would have worked well for very small ϵ , when almost all escapes do occur in the direction prescribed by \bar{G} .

However, the percentage of “correct” escapes does not increase as ϵ falls but fluctuates between 25% and 28% for $\epsilon = 2 \cdot 10^{-4}$, $4 \cdot 10^{-4}$, and $2 \cdot 10^{-5}$. A large number of escapes occur below -5 and below -6 (59% and 9.7% respectively at $\epsilon = 2 \cdot 10^{-5}$). In other words, the distribution of the directions of escape from the region bounded by the separating curve resembles the distribution of escapes from the cylinder, even though theoretically we expected them to diverge.

4.3 Analytical Results vs. Simulations: Escape Time

To see how the proposed approach works in terms of predicting the mean escape time, we exploited the relation (17). It implies that a plot of $\ln E_x(\tau^\epsilon)$ vs. $1/\epsilon$ is a straight line with the slope equal to the rate of convergence \bar{T} . Note that τ^ϵ , first escape time, is given in continuous time units of the approximating diffusion (16), and is approximately equal to ϵ times the expected number of iterations of the discrete time process (10) needed to observe the first escape. Fig. 8 shows that the straight line is not observed: The slope decreases as one moves towards higher $1/\epsilon$ (lower ϵ), and seems to converge asymptotically to the solid line with the slope \bar{T} only for the lowest considered values of ϵ .²⁰

Table 2 shows the theoretically predicted values of the slope \bar{T} for the first and second way of selecting ∂D as well as empirically observed slopes (calculated at the slope of the line connecting the last two observations in Fig. 8 and its analog for the separating curve). For the first (second) way, the empirical slope almost converges to (undershoots) the theoretically predicted value for $\epsilon = 2 \cdot 10^{-5}$. As explained in footnote 25, our method is likely to produce the values of \bar{T} which are *higher* than the true ones, which explains observed undershooting.

²⁰Though the limiting characteristics of mean escape time predicted by the large deviation theory do not hold true for economically plausible region of gain ϵ , the empirical distribution of escape times follows the theoretically predicted exponential distribution even for “large” ϵ : Fig. 4 shows that the logarithm of the empirical cumulative distribution function, $\ln[\Pr(\tau^\epsilon \geq T)]$, is approximately linear in T , as expected for the exponential distribution.

TABLE 2. A comparison of the theoretically derived value of the action functional and empirically observed slope of $\ln E_x(\tau^\epsilon)$ vs. $1/\epsilon$ line.

ϵ		Dynamic model		Static model
		Way 1	Way 2	Way 1
Simulations	$2 \cdot 10^{-5}$	$2.21 \cdot 10^{-5}$	$2.14 \cdot 10^{-5}$	$1.96 \cdot 10^{-4}$
	$3 \cdot 10^{-5}$	$3.67 \cdot 10^{-5}$	$3.19 \cdot 10^{-5}$	
	$5 \cdot 10^{-5}$	$6.60 \cdot 10^{-5}$	$4.92 \cdot 10^{-5}$	
	$1 \cdot 10^{-4}$	$1.31 \cdot 10^{-4}$	$8.46 \cdot 10^{-5}$	
	$2 \cdot 10^{-4}$	$2.49 \cdot 10^{-4}$	$1.77 \cdot 10^{-4}$	
	$4 \cdot 10^{-4}$	$4.90 \cdot 10^{-4}$	$3.79 \cdot 10^{-4}$	
	$1 \cdot 10^{-3}$	$8.14 \cdot 10^{-4}$	$6.76 \cdot 10^{-4}$	
Theory		$2.01 \cdot 10^{-5}$	$3.01 \cdot 10^{-5}$	$3.20 \cdot 10^{-4}$

The theoretical formula for mean escape time in the third way of characterising escape dynamics in the model can be derived using the formula for mean exit time of one-dimensional Brownian motion, see Karatzas and Shreve (1991, Eq. 5.62, p. 345). We derive the mean escape time for the projection of the process $d\varphi_t = \sqrt{\epsilon}\Sigma^{1/2}(\bar{\theta}) dW_t$ onto the most probable direction of escape \tilde{v}_1 , the “largest” eigenvector of Σ . The formula for the mean escape time, derived in Appendix C, is given as $E\tau^\epsilon = \frac{rad^2}{\epsilon\lambda}$, where λ is the largest eigenvalue of Σ , and rad is the distance between the SCE and the point where the “largest” eigenvector of Σ crosses the cylinder described above. In Table 3, we compare this formula’s predictions with averages from the simulations. The formula performs very well, especially for $\epsilon \in [0.001; 0.01]$, the range used previously by CWS and others. Indeed, when we plot average simulated escape time vs. $1/\epsilon^2$ in Fig. 10,²¹ we see a straight line with the slope approximately equal to $\frac{rad^2}{\lambda}$ for a large range of ϵ . The formula starts to lose precision once one moves to lower ϵ , *i.e.* into the region where the averaging is better as the system spends more periods in the neighborhood of SCE, and so the large deviations theory becomes more applicable to characterizing mean escape times.

²¹Observe that τ^ϵ , first escape time, is given in continuous time units of the approximating diffusion (16), and is approximately equal to ϵ times the expected number of iterations of the discrete time process (10) needed to observe the first escape, see the beginning of this subsection for an explanation. Therefore, we divide the “continuous-time mean escape time” $E\tau^\epsilon$ by ϵ to get the “mean number of periods before escape”.

TABLE 3. A comparison of the theoretically derived values of expected escape time and empirically observed average escape times

ϵ	Dynamic model		Static model	
	Simulations	Theory, $\frac{rad^2}{\epsilon^2\lambda}$	Simulations	Theory, $\frac{rad^2}{\epsilon^2\lambda}$
$2 \cdot 10^{-5}$	$1.10 \cdot 10^5$	$1.26 \cdot 10^5$		
$3 \cdot 10^{-5}$	$5.10 \cdot 10^4$	$5.61 \cdot 10^4$	$4.38 \cdot 10^7$	$6.97 \cdot 10^5$
$5 \cdot 10^{-5}$	$1.88 \cdot 10^4$	$2.02 \cdot 10^4$	$1.98 \cdot 10^6$	$2.51 \cdot 10^5$
$2 \cdot 10^{-4}$	$1.26 \cdot 10^3$	$1.26 \cdot 10^3$	$1.56 \cdot 10^5$	$6.27 \cdot 10^4$
$4 \cdot 10^{-4}$	336.96	315.65	4928.00	3919.27
$1 \cdot 10^{-3}$	64.59	50.50	733.57	627.08
$2 \cdot 10^{-3}$	21.49	12.63	189.98	156.77
$3 \cdot 10^{-3}$	12.50	5.61	87.00	69.68
$4 \cdot 10^{-3}$	8.77	3.16	52.08	39.19
$5 \cdot 10^{-3}$	6.79	2.02	34.39	25.08
$6 \cdot 10^{-3}$	5.99	1.40	24.76	17.42
$7 \cdot 10^{-3}$	4.98	1.03	19.14	12.80
$8 \cdot 10^{-3}$	4.49	0.79	15.02	9.80
$9 \cdot 10^{-3}$	4.12	0.62	13.32	7.74
$1 \cdot 10^{-2}$	3.70	0.51	11.16	6.27

Simulation results for very low ϵ tell a consistent story: one can use large deviations theory estimates for escape time only for gain values $\epsilon \lesssim 2 \cdot 10^{-5}$; even $\epsilon = 2 \cdot 10^{-5}$ is still not low enough to observe the limiting behavior predicted by the theory. Note that this is true for both continuous-time approach developed here and discrete-time approach used by CWS.²² As for the continuous-time approach, one can avoid the non-applicability of the large deviation theory for economically interesting values of ϵ by disregarding the mean dynamics altogether and following the third way of deriving the escape dynamics, which works well for such ϵ .

For the lowest ϵ considered here, the large deviations theory just begins to work. Is it possible to claim that empirically observed disinflations are, indeed, escapes from the SCE generated by the model? At $\epsilon = 2 \cdot 10^{-5}$, the average number of simulation periods needed to observe the first escape is about $1 \cdot 10^5$ (for cylinder) and $1.5 \cdot 10^5$ (for the separating curve). Add to these numbers the time needed to travel to the induction hypothesis plane $\tilde{\gamma}_1 = 0$ (of order 10^4) and recall that the time period in the Phelps model could not be much lower than a quarter. In this economy, one would wait, on average, twenty thousand years or longer for the low inflation episode caused

²²Even if one believes that our calculations overestimate the rate of convergence because of the arguments presented in footnote 25 and the true \bar{T} is closer to the numbers given in CWS, the empirically observed slope at $\epsilon = 0.001$ is $40 \div 80$ times larger than predicted. This result means that both discrete- and continuous-time approaches fail at least for $\epsilon \sim 0.001$.

by adaptive learning with a constant gain equal to $2 \cdot 10^{-5}$. It is immediately obvious that the region of ϵ values, for which large deviations theory estimates of mean escape time start to be applicable in the dynamic model of CWS, is far removed from those ϵ values which lead to simulated escapes at empirically interesting times, such as 65 periods at $\epsilon = 1 \cdot 10^{-3}$.

5 Discussion

5.1 Better Averaging for Larger ϵ

As we have shown in the dynamic model of CWS at empirically relevant values of the constant gain parameter, the mean escape time can be easily characterized using a simple formula for the expected exit time of one-dimensional Brownian motion, while the large deviation theory predictions of mean escape time do not hold. This is due to three facts. A large value of u leads to very large λ_1/λ_2 , the ratio of the two largest eigenvalues of the inverse second moments matrix at SCE. This fact makes the SRA dynamics almost one-dimensional. Second, both static and dynamic models of CWS have a very specific phase portrait: despite potentially global stability of the SCE, the region where the mean dynamics points back to it is exceedingly small. Third, the dominant eigenvalue λ_1 is huge, which means that any stochastic deviation in the “only” direction is strongly magnified. A combination of these three features of the SRA in the dynamic CWS model makes “escape” easy and puts very tight requirements on the values of ϵ used in the constant gain learning algorithm. In particular, this means that values of ϵ commonly used in the literature are not small enough to guarantee enough time for averaging, and therefore, methods of characterizing the mean escape time, based on the theory of large deviations which relies on the mean (averaged) dynamics, are not expected to work well in this particular setting. This argument is applicable to both the continuous-time approach and to the discrete-time approach used in CWS and elsewhere. Nevertheless, a version of the continuous-time approach which disregards the mean dynamics, and in fact does not rely on the large deviations theory to characterize the mean escape time, provides a very good fit to the simulations.

Lack of time for averaging near the SCE leads to the relative failure of our first and second ways in deriving the properties of escapes. This failure could be tracked down to the limited applicability of (11) and of approximating diffusion (16). With better averaging, both large deviations theory characteristics of mean escape time and

continuous-time approximation (16) improve. We have three suggestions on how to achieve better averaging in this model for the empirically interesting region of ϵ .

First, the matrix \overline{R}^{-1} should be more balanced: A lower value of λ_1/λ_2 effectively increases the dimensionality of the problem and expands the volume of the state space available to the system, thus increasing the expected escape time and producing better averaging. In the current model, a more balanced second moments matrix means lower value of u . Second, having a stronger drift towards SCE under mean dynamics (this probably implies a larger region of immediate attraction to the SCE) will help to achieve better averaging around SCE and thus ensure that (16) approximates (6) reasonably well. We conjecture that the smallness of the region where the mean dynamics points towards SCE is due to the fact that at the SCE, the learning is not well specified: some of the regressors are perfectly collinear in both the static and dynamic models of CWS. In other words, forcing the agents to use better specified learning might help. Third, reducing the magnitude of λ_1 while keeping fixed the SCE’s region of attraction increases the time spent by the system in this region and provides better averaging, at a potential cost of a higher expected escape time; as a result, the region of the constant gain parameter ϵ where large deviations theory’s asymptotic predictions start to be valid is more likely not to include empirically interesting magnitudes of ϵ .

5.2 A Static Model vs. A Dynamic Model

In order to test some of these conjectures, we repeated our analysis for the static model of CWS. In contrast to the dynamic model analyzed previously, the government’s beliefs do not take into account lagged inflation and unemployment rates, and the vector X_{n-1} contains only the constant.²³ We expect the large deviations theory to become applicable for larger values of ϵ than with the dynamic model because the phase portraits of the mean dynamics in the dynamic (in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ coordinates) and static models are very similar, but the largest “static” eigenvalue of \overline{R}^{-1} equals 26.09 and is much smaller than 3084 for the dynamic model. On the other hand, the ratio of the first two eigenvalues, λ_1/λ_2 , equals 7561 compared to 106. In other words, the

²³The government’s problem (3) can be solved explicitly for the policy function $h(\gamma)$. A preliminary analysis of the static CWS model was performed in Evans and Honkapohja (2001, Section 14.4). In particular, the matrix V derived on p. 358 is nothing else but the Gramian \overline{G} of the 2-dimensional problem which discards 3 elements of the covariance matrix. As stated by Evans and Honkapohja (2001), one could do so because the matrix $D_{\theta p}(\bar{\theta})$ is block-lower-triangular and the matrix $\Sigma(\bar{\theta})$ is block-diagonal. In order to preserve continuity with our analysis of the dynamic model, we analyze the full 5-dimensional static problem, with two elements of the vector γ and 3 elements of the variance-covariance matrix R .

static model is even more one-dimensional than the dynamic one, but the system is expected to spend more time in the neighborhood of the SCE before the first escape and thus achieve better averaging than in the dynamic model.²⁴

Based on the discussion in the previous Section, we do not consider the second way of selecting ∂D and use only the first and the third one. The cylinder is chosen so that it intersects with the “largest” eigenvector \tilde{v}_1 at $\gamma_1 = -0.975$, using the same reasoning as in the dynamic model case. Theoretically predicted directions of escape out of the cylinder are $(1, -7.5)'$ and $(1, -5)'$ for the first and third way, respectively.

On several dimensions, the first way produces more favorable results when applied to the static model. First, at $\epsilon = 1 \cdot 10^{-3}$, only 22% of escapes happen in the -5 bin, compared to 56% for the dynamic model. At $\epsilon = 2 \cdot 10^{-4}$, this number drops to 6.7% for the static model, but still stands at 49% in the dynamic one. At $\epsilon = 3 \cdot 10^{-5}$ in the static model a full 34% of the escapes happen at the direction -7 or below and 100% of them are below -5. As is clear from Table 1, the distance to the escape point theoretically predicted by the first way declines very fast as ϵ decreases in the static model, but it decreases only slightly in the dynamic one. Finally, already at $\epsilon = 1 \cdot 10^{-4}$, the empirically observed slope of the $\ln E_x(\tau^\epsilon)$ vs. $1/\epsilon$ line is lower than the theoretically predicted one, see Table 2.²⁵

As for the third way, we see that it works well in predicting the point of escape (see Table 1) and in predicting the mean escape time (see Table 3). However, the simplified continuous-time approximation (without drift term) and the simple mean escape time formula based on it are valid for the region of ϵ shifted to a larger ϵ compared to the one for the dynamic model, see Fig. 10. This is explained by better averaging for larger ϵ in the static version of the model.

²⁴This prediction is borne out by the simulations: at $\epsilon = 1 \cdot 10^{-2}$, $1 \cdot 10^{-3}$, and $2 \cdot 10^{-4}$, the average escape times out of the cylinder with intersects \tilde{v}_1 at -0.985 are 3.7, 64.6, and 1255 periods for the dynamic model, while for the static one the corresponding numbers are 5.6, 265.6, and 6903 periods, respectively.

²⁵Recall that \bar{I} , which determines this slope, is called quasipotential. It measures the energy needed to get the system out of region D around the point O . The reason one needs to spend energy at all is the drift, or non-stochastic component of the diffusion, which points back towards O . We have selected our region D in such a way that the drift pointing inwards becomes very weak near the boundary ∂D . However, as the approximation to (6) is made at point O , it overestimates the strength of the drift which has to be overcome in order to cross ∂D , and so overestimates \bar{I} . The argument from Dembo and Zeitouni (1998), cited in Section 2.1, persuades us that this upward bias is not likely to be very large.

5.3 Comparisons with CWS

Comparing the predictions of our most successful (third) way of deriving the most probable escape point and the mean escape time *for an economically plausible region of gain values* in both dynamic and static versions of the CWS model with the CWS predictions, we see that our estimates of *magnitudes* of mean escape times fit simulations very well whereas the mean escape times' *limiting behavior* predicted by the large deviations theory used by CWS is not confirmed by simulations. The reason for this failure (and for the failure of our first and second ways of describing the mean escape time for economically interesting gain values) is bad averaging as discussed above.

In terms of the escape point, all three ways used in this paper have the same escape point as understood by CWS:²⁶ escaping trajectories hit the “induction hypothesis” plane very close to the point prescribed by the largest eigenvalue’s eigenvector of \bar{R}^{-1} , the inverse second moments matrix of beliefs evaluated at the SCE. We explain this by the fact that this “escape” is for the most part a *deterministic* movement along the mean dynamics trajectory, and all such trajectories are very close to each other, see Fig. 2.

We have to address three technical considerations that could potentially influence our comparison. The first is “Kushner critique”: can one guarantee that the escape dynamics generated by a continuous–time approximating process is a valid approximation of the escape dynamics of the original discrete–time learning process? Kushner (1984) proposed that this question could be answered by checking whether the action functional of a discrete–time process converges to the action functional of its continuous–time approximation. Checking the convergence is very hard to do taking into account that the action functional for a discrete–time process in the CWS model depends on a numerically derived function. Therefore, the only way of judging the performance of the large deviations theory approach based on a continuous–time approximation of the discrete–time SRA is to compare the predictions of this approach to the predictions of the discrete–time approach and to simulation results.

Second, the CWS approach ignores the cross effects of the second moments matrix R for the dynamic version of the model due to numerical complications, while we fully take them into account. We estimate that the influence of this assumption is not large as the matrix Σ is close to being block diagonal, see footnote 14 for a discussion. We believe that if one were to implement the CWS approach without ignoring the cross

²⁶Escape out of a cylinder with radius 66 times as large as in our paper.

effects of R , the results obtained would be more valid than those obtained here. This path, however, might be blocked by computational complications.

Third, the discrete-time version of large deviations theory does not contain theoretical results for the most probable point of escape and mean escape time in the case of unbounded (for example, Gaussian) shocks, see CWS, Theorem 5.3. The result which is available for unbounded shocks is that “the probability of observing an escape episode is exponentially decreasing in the gain with the rate given by the minimized value of the cost function \bar{S} ”, see CWS, p. 13. The extent to which this lack of theoretical results influences numerical predictions is not generally known, but has been shown to be small in some situations: CWS use both normal (unbounded) and binomial (bounded) shocks in the static model and obtain the values of \bar{S} (\bar{I} in our notation) which are numerically very close to each other; predicted most probable escape paths also are very similar.

6 Conclusion

We extended a continuous-time approach for the analysis of escape dynamics in economic models with adaptive constant gain learning. Foundations of this approach were laid down by Kasa (2004), Evans and Hohkapohja (2001, Ch. 14), and Williams (2001). This approach is based on applying the results of FW’s continuous-time version of large deviations theory to the diffusion approximation of the original discrete-time learning process.

When applied to the Phelps problem of government controlling inflation using an approximate Phillips curve, deriving escape dynamics characteristics for the “original” diffusion approximation with a drift term did not generate results compatible with the results of simulations in terms of the mean escape time. This is due to the limited validity of such an approximation for economically plausible values of the constant gain parameter ϵ . Limited validity of the approximation, in turn, is caused by bad averaging in the CWS model. To account for high “noise-to-signal” ratio near the SCE, we used a “modified” diffusion approximation without the drift term and the formula for the mean exit time of one-dimensional Brownian motion, rather than limiting characteristics provided by the theory of large deviations. We managed to predict the values of mean escape times with high precision for the “modified” approximation.

All our ways of deriving escape dynamics characteristics work well in predicting

the final point of escape: Escape occurs in the small neighborhood of the SCE, then mean dynamics moves the system along the largest eigenvector of the inverse second moments matrix evaluated at the SCE towards the “induction hypothesis” plane. As for predicting the most probable point of “initial” escape out of a small neighborhood of the SCE, the ways based on an “original” continuous–time approximation did not work well for an economically plausible region of ϵ because the approximation has limited validity at such ϵ . Considering an escape out of a separating surface on which the mean dynamics changes its direction, thus taking into account some information about the behavior of the mean dynamics away from the SCE, provides better results for the point of “initial” escape.

As another result of this paper, we express reservations regarding the applicability of escape dynamics theory for the characterization of mean escape time for economically plausible values of gain in both versions of the CWS model. We show that for the region of gain values used in the economics literature, simple considerations and formulae work much better than large deviations theory’s results. This, again, is explained by bad averaging for a relatively large ϵ in this model.

We suggest two changes to help the approaches based on large deviation theory work to better in terms of characterizing the mean escape time for the model and gain values considered: to set a lower mean unemployment rate, in order to construct a more balanced second moments matrix, and to use better specified learning of agents. The same changes will help our “original” continuous–time approximation to become valid for larger gain values. In general, one has to look for economically sensible models with better averaging for economically plausible gain values in order to apply large deviations theory characteristics of the mean escape time.

Finally, we believe that utilizing a continuous–time approximation can be used to analyze escape dynamics in more complicated models, where it is not possible to derive analytical characteristics of escape dynamics in discrete time. For example, the model with a dynamic Phillips curve can be studied as a possible extension of the approach proposed here. The question, however, remains whether large deviations theory predictions of mean escape time would be valid in this model for economically plausible ϵ , or whether one would have to employ something resembling our reliance on the “modified” approximation and the mean exit time result for one–dimensional Brownian motion. This is the focus of our current research.

References

- Binmore, K. and Samuelson, L.: 1997, Muddling through: Noisy equilibrium selection, *Journal of Economic Theory* **74**, 235–265.
- Bray, M. M.: 1982, Learning, estimation, and the stability of rational expectations equilibria, *Journal of Economic Theory* **26**, 318–339.
- Bray, M. M. and Savin, N. E.: 1986, Rational expectations equilibria, learning, and model specification, *Econometrica* **54**(5), 1129–60.
- Bullard, J. B. and Cho, I.-K.: 2005, Escapist policy rules, *Journal of Economic Dynamics and Control* **29**(11), 1841–1865.
- Bullard, J. and Mitra, K.: 2002, Learning about monetary policy rules, *Journal of Monetary Economics* **49**(6), 1105–29.
- Cho, I.-K. and Kasa, K.: 2003, Learning dynamics and endogenous currency crises. Manuscript, Simon Fraser University.
- Cho, I.-K., Williams, N. and Sargent, T. J.: 2002, Escaping nash inflation, *Review of Economic Studies* **69**(1), 1–40.
- Dahleh, M., Dahleh, M. A. and Verghese, G.: 2004, Lectures on dynamic systems and control. MIT lecture notes.
- Dembo, A. and Zeitouni, O.: 1998, *Large Deviations Techniques and Applications*, Springer, New York Berlin Heidelberg.
- Dupuis, P. and Kushner, H. J.: 1989, Stochastic approximation and large deviations: Upper bounds and w.p.1 convergence, *SIAM Journal on Control and Optimization* **27**, 1108–1135.
- Evans, G. W. and Honkapohja, S.: 2001, *Learning and Expectations in Macroeconomics*, Princeton University Press.
- Evans, G. W. and Honkapohja, S.: 1994a, Learning, convergence, and stability with multiple rational expectations equilibria, *European Economic Review* **38**, 1071–1098.

- Evans, G. W. and Honkapohja, S.: 1994b, On the local stability of sunspot equilibria under adaptive learning rules, *Journal of Economic Theory* **64**, 142–161.
- Evans, G. W. and Honkapohja, S.: 1995, Local convergence of recursive learning to steady states and cycles in stochastic nonlinear models, *Econometrica* **63**, 195–206.
- Evans, G. W. and Honkapohja, S.: 2001, *Learning and Expectations in Macroeconomics*, Princeton University Press, Princeton, NJ.
- Evans, G. W. and Honkapohja, S.: 2003, Adaptive learning and monetary policy design, *Journal of Money, Credit, and Banking* **35**(6), 1045–72.
- Fourgeaud, C., Gouieroux, C. and Pradel, J.: 1986, Learning procedures and convergence to rationality, *Econometrica* **54**(4), 845–68.
- Freidlin, M. I. and Wentzell, A. D.: 1998, *Random Perturbations of Dynamical Systems*, second edn, Springer-Verlag, New York.
- Kandori, M., Mailath, G. J. and Rob, R.: 1993, Learning, mutation, and long run equilibria in games, *Econometrica* **61**, 29–56.
- Karatzas, I. and Shreve, S. E.: 1991, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, Berlin and Heidelberg.
- Kasa, K.: 2004, Learning, large deviations, and recurrent currency crises, *International Economic Review* **45**, 141–173.
- Kushner, H. J.: 1984, Robustness and approximation of escape times and large deviations estimates for systems with small noise effects, *SIAM Journal of Applied Mathematics* **44**, 160–182.
- Marcet, A. and Sargent, T. J.: 1989, Convergence of least squares learning mechanisms in self-referential linear stochastic models, *Journal of Economic Theory* **48**(2), 337–68.
- Marimon, R.: 1997, Learning from learning in economics, in D. M. Kreps and K. F. Wallis (eds), *Advances in Economics and Econometrics: Theory and Applications.*, Vol. 1, Cambridge University Press, Cambridge, chapter 9, pp. 278–315.
- Sargent, T. J.: 1999, *The Conquest of American Inflation*, Princeton University Press.

Williams, N.: 2001, *Escape Dynamics in Learning Models*, PhD thesis, University of Chicago.

Williams, N.: 2002, Stability and long run equilibrium in stochastic fictitious play. Manuscript, Princeton University.

Williams, N.: 2003, Adaptive learning and business cycles. Manuscript, Princeton University.

Williams, N.: 2004, Small noise asymptotics for a stochastic growth model, *Journal of Economic Theory* **119**(2), 271–298.

A Large deviations theory

Action functional for a diffusion $d\varphi_t = A\varphi_t dt + \sqrt{\epsilon} B dW_t$ is defined as

$$I_{0T}(\varphi) = \inf_{\{\dot{\varphi}_t = A\varphi_t + Bg_t\}} \frac{1}{2} \int_0^T |g_t|^2 dt \quad (\text{see Dembo-Zeitouni, p. 214}).$$

The results on the mean exit time and dominant escape path are given in Dembo and Zeitouni (1998, Theorem 5.7.11).

Consider the system $dx_t^\epsilon = b(x_t^\epsilon) dt + \sqrt{\epsilon} \sigma(x_t^\epsilon) dW_t$, $x_t^\epsilon \in \mathbb{R}^d$, $x_0^\epsilon = x$.

Assumption A1 The unique stable equilibrium point in D (open, bounded domain) of the d -dimensional ODE $\dot{x}_t = b(x_t)$ is at $O \in D$ and $x_0 \in D \implies \forall t > 0$, $x_t \in D$ and $\lim_{t \rightarrow \infty} x_t = O$.

Assumption A2 All the trajectories of the deterministic ODE $\dot{x}_t = b(x_t)$ starting at $x_0 \in \partial D$ converge to O as $t \rightarrow \infty$.

Assumption A3 $\bar{I} \stackrel{\text{def}}{=} \inf_{y \in \partial D} I(O, y) < \infty$.

Assumption A4 An $M < \infty$ exists such that, for all $\rho > 0$ small enough and all x, y with $|x - z| + |y - z| \leq \rho$ for some $z \in \partial D \cup \{O\}$, there is a function u satisfying $\|u\| < M$ and $\varphi_{T(\rho)} = y$, where $\varphi_t = x + \int_0^t b(\varphi_s) ds + \int_0^t \sigma(\varphi_s) u_s ds$ and $T(\rho) \rightarrow 0$ as $\rho \rightarrow 0$.

Definition A1 $\tau^\epsilon \stackrel{\text{def}}{=} \inf \{t > 0 : x_t^\epsilon \in \partial D\}$.

Theorem A1 (Dembo and Zeitouni (1998, Theorem 5.7.11))

Assume A1-A4

(a) For all $x \in D$ and all $\delta > 0$,

$$\lim_{\epsilon \rightarrow 0} P_x \left(e^{(\bar{I} + \delta)/\epsilon} > \tau^\epsilon > e^{(\bar{I} - \delta)/\epsilon} \right) = 1.$$

Moreover, for all x ,

$$\lim_{\epsilon \rightarrow 0} \epsilon \ln E_x(\tau^\epsilon) = \bar{I}.$$

(b) If $N \subset \partial D$ is a closed set and $\inf_{z \in N} I(O, z) > \bar{I}$, then for any $x \in D$,

$$\lim_{\epsilon \rightarrow 0} P_x(x_{\tau^\epsilon}^\epsilon \in N) = 0.$$

In particular, if exists $z^* \in \partial D$ such that $I(O, z^*) < I(O, z)$ for all $z \neq z^*$, $z \in \partial D$, then

$$\forall \delta > 0, \forall x \in D, \lim_{\epsilon \rightarrow 0} P_x(|x_{\tau^\epsilon}^\epsilon - z^*|) = 1.$$

Part a) of the theorem characterizes the escape probability and the mean escape time, and part b) gives the dominant escape path.

B Minimizing the action functional

We need to solve the problem

$$\begin{aligned} & \min \frac{1}{2} \int_0^{\bar{T}} \|u_t\|^2 dt, \\ & \text{subject to} \\ & \dot{\varphi}_t = A\varphi_t + Bu_t, \\ & \varphi(0) = 0, \varphi(\bar{T}) \in \partial D \end{aligned}$$

We know that matrix B is singular, and therefore, the system (A, B) is not reachable. Such systems are usually converted into the *standard form* as follows:²⁷ change the coordinates so that $\varphi = Tz$:

$$z = T^{-1}\varphi = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$

where dimension of z_1 is r , dimension of the reachable subspace. In the new coordinates, we get

$$\begin{aligned} T\dot{z} &= ATz + Bu, \text{ or} \\ \dot{z} &= \underbrace{T^{-1}AT}_{\bar{A}}z + \underbrace{T^{-1}B}_{\bar{B}}u. \end{aligned}$$

The matrix T is constructed as $[T_1 | T_2]$, where T_1 consists of columns that form the basis of the reachable subspace (it is convenient to select columns of T to be the [orthonormal] basis of Gramian G in initial problem. T_1 are the columns corresponding to the nonzero eigenvalues of G , therefore they constitute the basis of reachable subspace). Columns of T_2 form (orthonormal) basis of the complement to the reachable subspace. By construction, matrix T is invertible.

Let us find the structure of \bar{A} and \bar{B} . Look at $AT = A[T_1 | T_2] = [T_1 | T_2]\bar{A}$. Reachable space is invariant for all controls including $u = 0$, therefore range of $[T_1 | T_2]\bar{A}$ should not include vectors from T_2 . This could be achieved if

$$\bar{A} = \begin{bmatrix} \bar{A}_1 & \bar{A}_{12} \\ 0 & \bar{A}_2 \end{bmatrix}.$$

Similarly, no control should push the system out of reachable subspace, this means that

$$B = [T_1 | T_2]\bar{B} = [T_1 | T_2] \begin{bmatrix} \bar{B}_1 \\ 0 \end{bmatrix}.$$

²⁷See the details in Dahleh, Dahleh, and Verghese (2003) and Boyd (2004).

With this, we can now write our system as

$$\dot{z} = T^{-1}ATz + T^{-1}Bu = \begin{bmatrix} \bar{A}_1 & \bar{A}_{12} \\ 0 & \bar{A}_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \bar{B}_1 \\ 0 \end{bmatrix} u.$$

As we are interested in the movement from initial point $z(0) = 0$, unreachable dimensions z_2 cannot influence dynamics of z_1 :

$$\dot{z}_1 = \bar{A}_1 z_1 + \bar{B}_1 u.$$

This system is called *reachable subsystem* of the original one.

Let us find \bar{A}_1 and \bar{B}_1 . $T^{-1}AT = \bar{A}$, or $\begin{bmatrix} T_1^T \\ T_2^T \end{bmatrix} [A] \begin{bmatrix} T_1 & T_2 \end{bmatrix} = \begin{bmatrix} \bar{A}_1 & \bar{A}_{12} \\ 0 & \bar{A}_2 \end{bmatrix}$. Therefore, $T_1^T AT_1 = \bar{A}_1$. For \bar{B}_1 , $T^{-1}B = \bar{B} \Rightarrow \begin{bmatrix} T_1^T \\ T_2^T \end{bmatrix} B = \begin{bmatrix} \bar{B}_1 \\ 0 \end{bmatrix} \Rightarrow T_1^T B = \bar{B}_1$.

In the new variables our problems transforms into

$$\begin{aligned} \min I_{0T} &= \frac{1}{2} \int_0^{\bar{T}} \|u_t\|^2 dt, \\ &\text{subject to} \\ \dot{z}_1 &= \bar{A}_1 z_1 + \bar{B}_1 u, \\ z_1(0) &= 0, \quad T_1 z_1(\bar{T}) \in \partial D. \end{aligned}$$

(Note that z_2 stays zero under our dynamics, therefore $Tz = \begin{bmatrix} T_1 & T_2 \end{bmatrix} \begin{bmatrix} z_1 \\ 0 \end{bmatrix} = T_1 z_1$).

This problem is easily solved, as the system (\bar{A}_1, \bar{B}_1) is reachable by construction. Standard result is that $I = \frac{1}{2} z_{1,des}^T \cdot \bar{G}^{-1} \cdot z_{1,des}$, where \bar{G} is Gramian in the reduced problem, given as a solution of the matrix Lyapunov equation $\bar{A}_1 \bar{G} + \bar{G} \bar{A}_1' + \bar{B}_1 \bar{B}_1' = 0$, and $T_1 z_{1,des} \in \partial D$.

In case when the set ∂D is surface of the cylinder — a sphere of radius R in γ space, $\|\gamma\| = R$, and no restrictions in space of components of R , — the problem of minimizing action functional becomes

$$\begin{aligned} \min & \frac{1}{2} z_1^T \cdot \bar{G}^{-1} \cdot z_1, \\ \text{s. t.} & (I_{27}^6 T_1 z_1)^T \cdot I_{27}^6 T_1 z_1 = R^2, \end{aligned}$$

where I_{27}^6 is 27×27 zero matrix with 6×6 identity matrix in the upper left corner. After defining $v = \bar{G}^{-1/2} \cdot z_1$, it is straightforward to get solution $\phi_{des} = \pm \frac{R}{\lambda_1} T_1 \bar{G}^{1/2} \xi$, where ξ is the unit eigenvector of matrix $\Gamma = \bar{G}^{1/2} T_1^T I_{27}^6 T_1 \bar{G}^{1/2}$ corresponding to the largest eigenvalue, $(\lambda_1)^2$. Note that if matrix Σ is block-diagonal, as in the static model, the eigenvector ξ coincides with the “largest” eigenvector of \bar{G} , and eigenvalue λ_1 coincides with the largest eigenvalue of \bar{G} .

For the problem when ∂D is given numerically, one has to find $\min_{z_1 \in \partial D} \frac{1}{2} z_1^T \cdot \overline{G}^{-1} \cdot z_1$. All the points on the boundary are given parametrically by 2-dimensional function $\varepsilon(t)$, where t is the index number of the point in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. Write $Fz_1 = \varepsilon(t)$, where F transforms 13 dimensions of z_1 into 2-dimensional space $(\tilde{\gamma}_1, \tilde{\gamma}_2)$. Define $v = \overline{G}^{-\frac{1}{2}} z_1$. Then the problem becomes

$$\begin{aligned} & \min \frac{1}{2} \|v\|^2 \\ & \text{s. t.} \\ & F\overline{G}^{\frac{1}{2}} v = \varepsilon(t) \end{aligned}$$

and the solution of this problem is $v_t = \text{pinv} \left(F\overline{G}^{\frac{1}{2}} \right) \cdot \varepsilon(t)$, and $\bar{t} = \arg \min \left\| \text{pinv} \left(F\overline{G}^{\frac{1}{2}} \right) \cdot \varepsilon(t) \right\|$. $\varepsilon(\bar{t})$ is the predicted point of escape in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. To transform this point into original 27-dimensional space of beliefs, use the following transformation:

$$z_{1\bar{t}} = \overline{G}^{\frac{1}{2}} \cdot \text{pinv} \left(F\overline{G}^{\frac{1}{2}} \right) \cdot \varepsilon(\bar{t}) \implies \varphi = T_1^T \overline{G}^{\frac{1}{2}} \cdot \text{pinv} \left(F\overline{G}^{\frac{1}{2}} \right) \cdot \varepsilon(\bar{t}).$$

For the third way problem (disregarding the mean dynamics), we set $A = 0$ and, using the general result above, get $I = \frac{1}{2} z_{1,des}^T \cdot \overline{G}^{-1} \cdot z_{1,des}$. In this case \overline{G} is defined for arbitrary time \overline{T} as $\overline{G}^{-1} = \left(\overline{B}_1 \overline{B}_1' \right)^{-1} \frac{1}{\overline{T}} = \left(T_1^T B B^T T_1 \right)^{-1} \frac{1}{\overline{T}} = T_1 \left(B B^T \right)^{-1} T_1^T \frac{1}{\overline{T}} = T_1 \left(\Sigma(\bar{\theta}) \right)^{-1} T_1^T \frac{1}{\overline{T}}$, where T_1 is the basis of spectral decomposition of $\Sigma(\bar{\theta})$, and at the same time the orthonormal basis of the reachable subspace. For any \overline{T} the solution of the problem on the cylinder is expressed by the formula for escape out of the cylinder written above, where instead of \overline{G} one uses $T_1 \Sigma(\bar{\theta}) T_1^T$. In the model of CWS the resulting direction almost coincides with the eigenvector corresponding to the largest eigenvalue of $\Sigma(\bar{\theta})$.

C Formula for the Third Way of Deriving Mean Escape Time

The third way of deriving escape dynamics characteristics is based on the “modified” continuous-time approximation without drift term, $d\varphi_t = \sqrt{\varepsilon} \Sigma^{1/2}(\bar{\theta}) dW_t$. To find the projection of the process on the most probable direction of escape \tilde{v}_1 , the “largest” eigenvector of Σ , we multiply the above expression by this eigenvector from the left. The resulting diffusion is $d\varphi_t^{\text{projection}} = \sqrt{\varepsilon \lambda} dW_t$, where λ is the largest eigenvalue of Σ . Then we use formula for the mean exit time for one-dimensional Brownian motion in Karatzas and Shreve (1991, Eq. 5.62, p. 345). For a process $Y_t = x +$

$\int_0^t \sigma(Y_s) dW_s$, the mean exit time $T_{a,b}(x) = \inf \{t \geq 0; Y_t \notin (a, b)\}$ is expressed as

$$ET_{a,b}(x) = \int_a^b \frac{(\min(x,y)-a)(b-\max(x,y))}{b-a} \cdot \frac{2dy}{\sigma^2(y)}.$$
 In our case x , the starting point of the projection of the process of deviations from the SCE, is zero, $\sigma(y)$ is replaced by $\sqrt{\epsilon\lambda}$, and the interval (a, b) is given by $(-rad, rad)$, where rad is the distance between the SCE and the point where the “largest” eigenvector of Σ crosses the cylinder used in the first and third way of deriving the escape dynamics. After plugging these values into the expression for the mean exit time and evaluating the integral, we get the formula for the mean escape time: $E\tau^\epsilon = \frac{rad^2}{\epsilon\lambda}$.

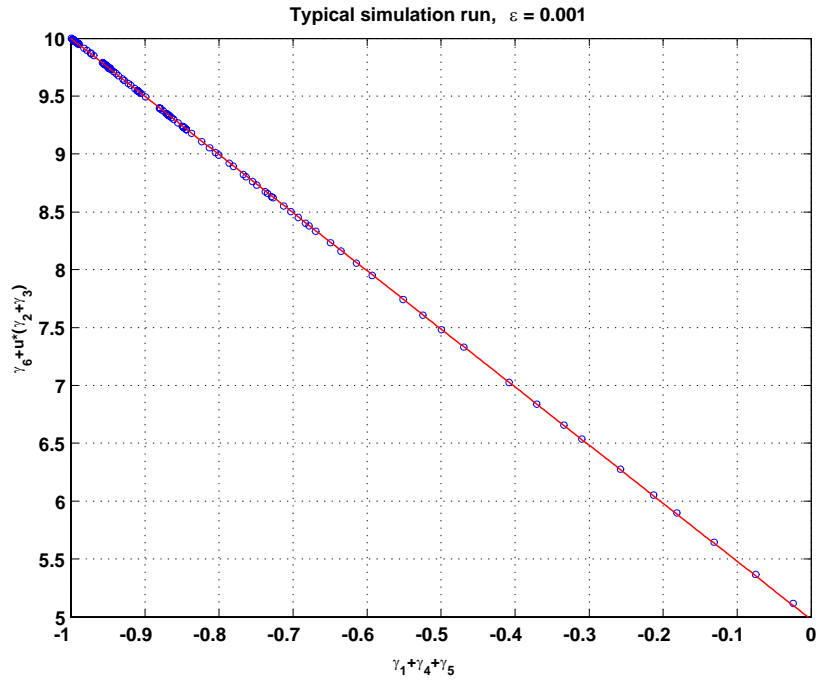


Figure 1:

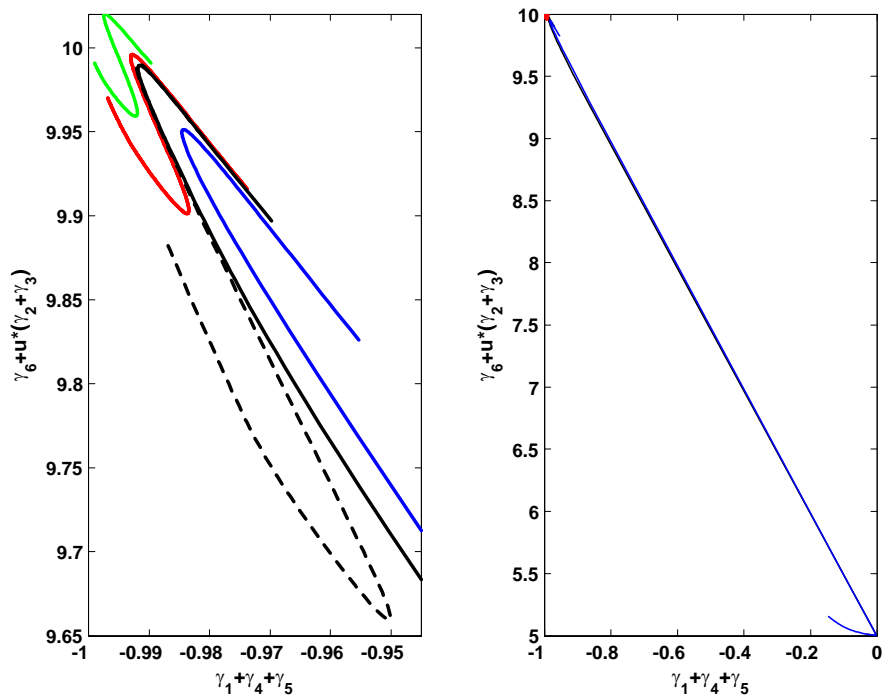


Figure 2: The mean dynamics trajectories

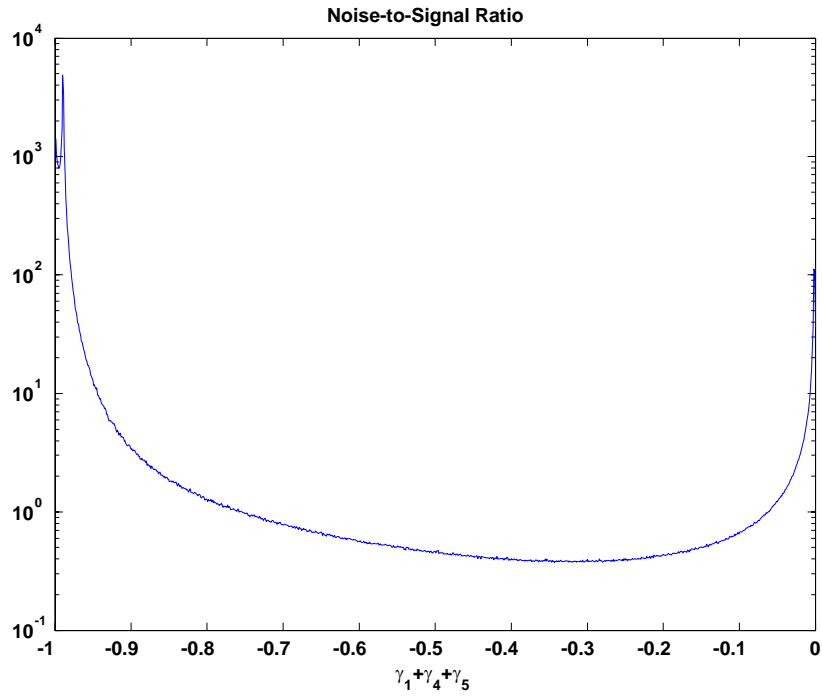


Figure 3:

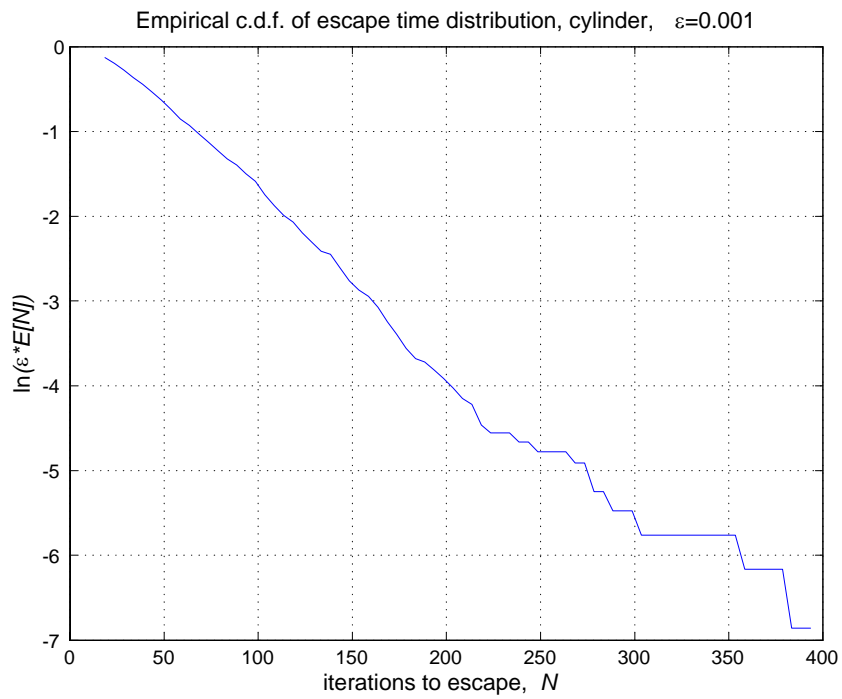


Figure 4:

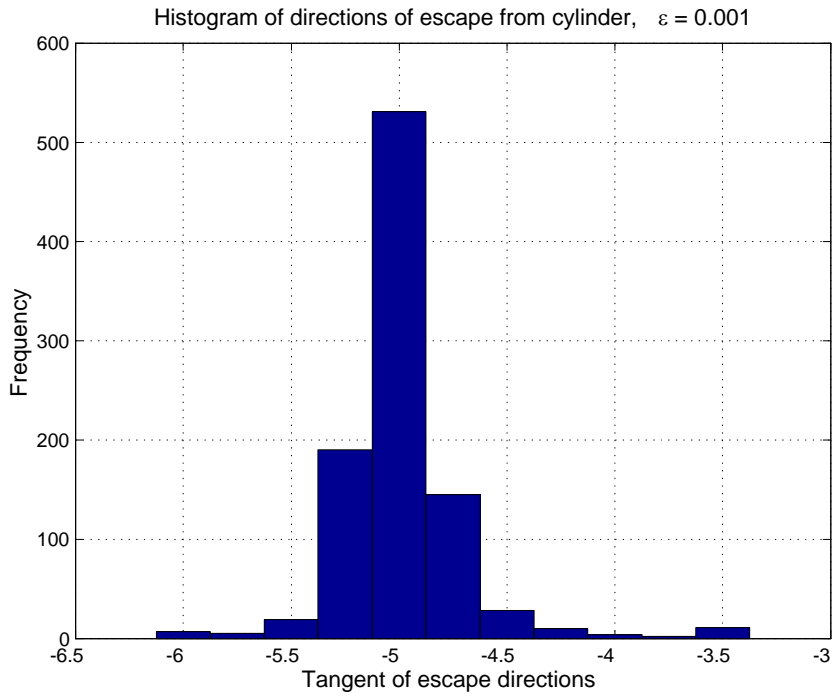


Figure 5:

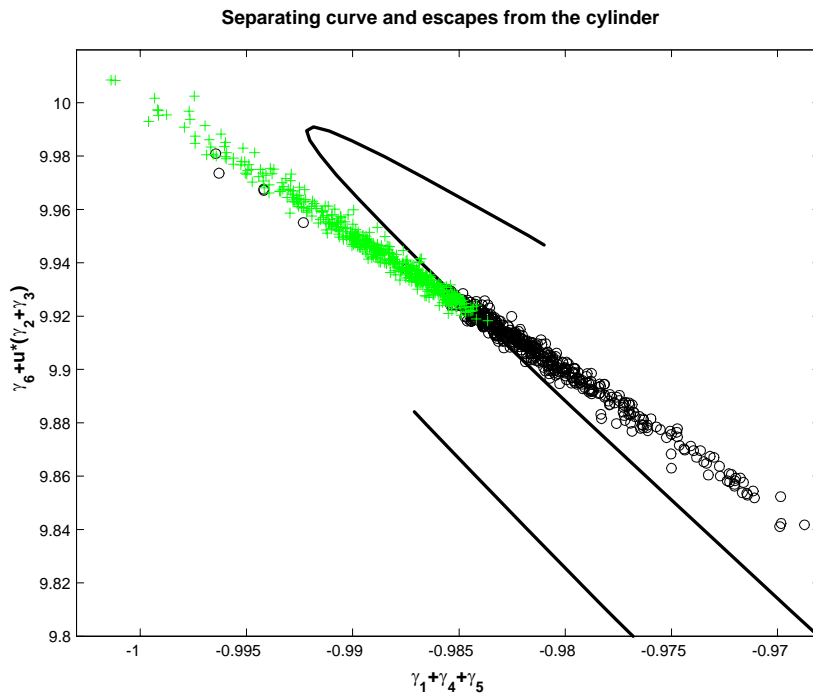


Figure 6:

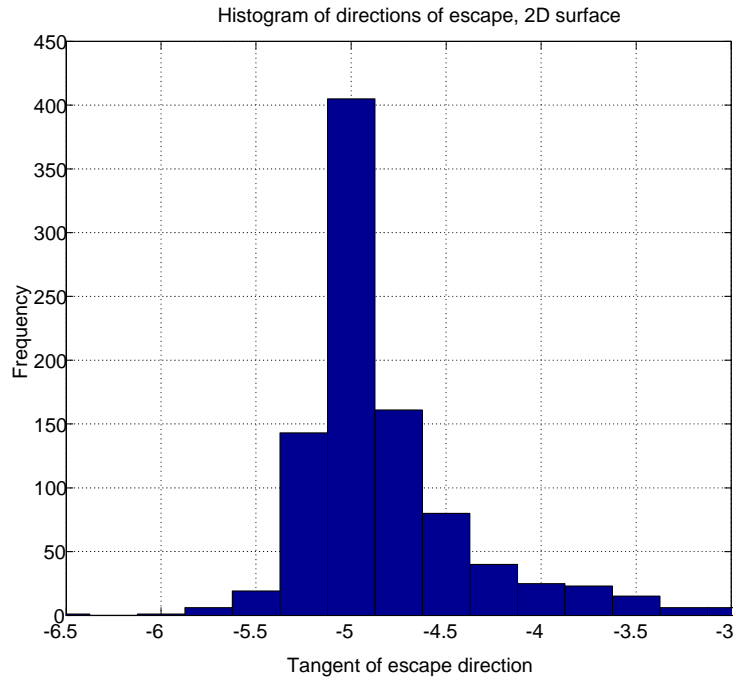


Figure 7:

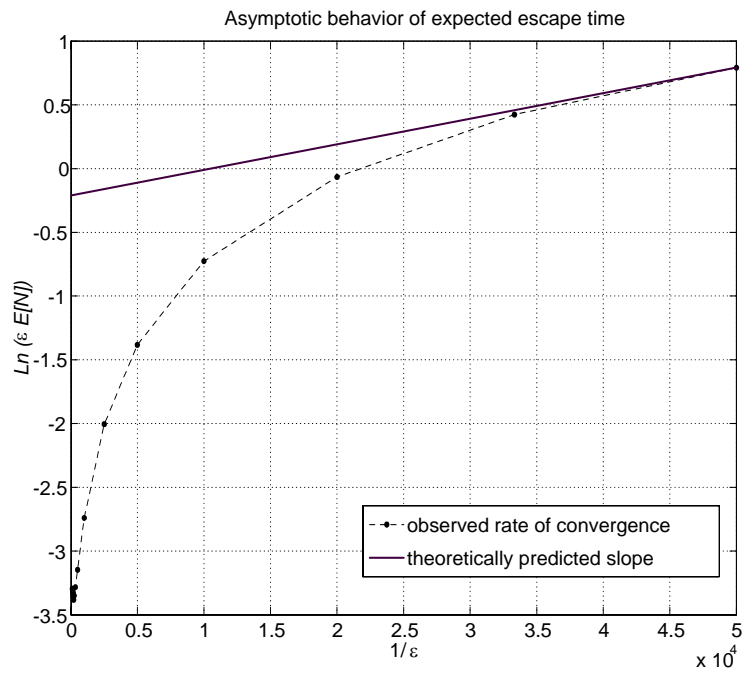


Figure 8:

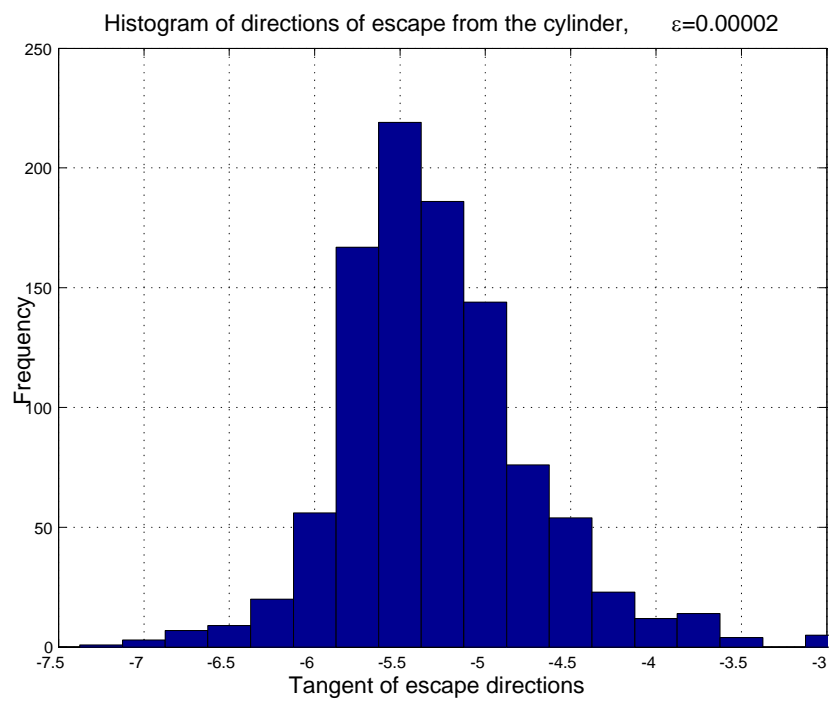


Figure 9:

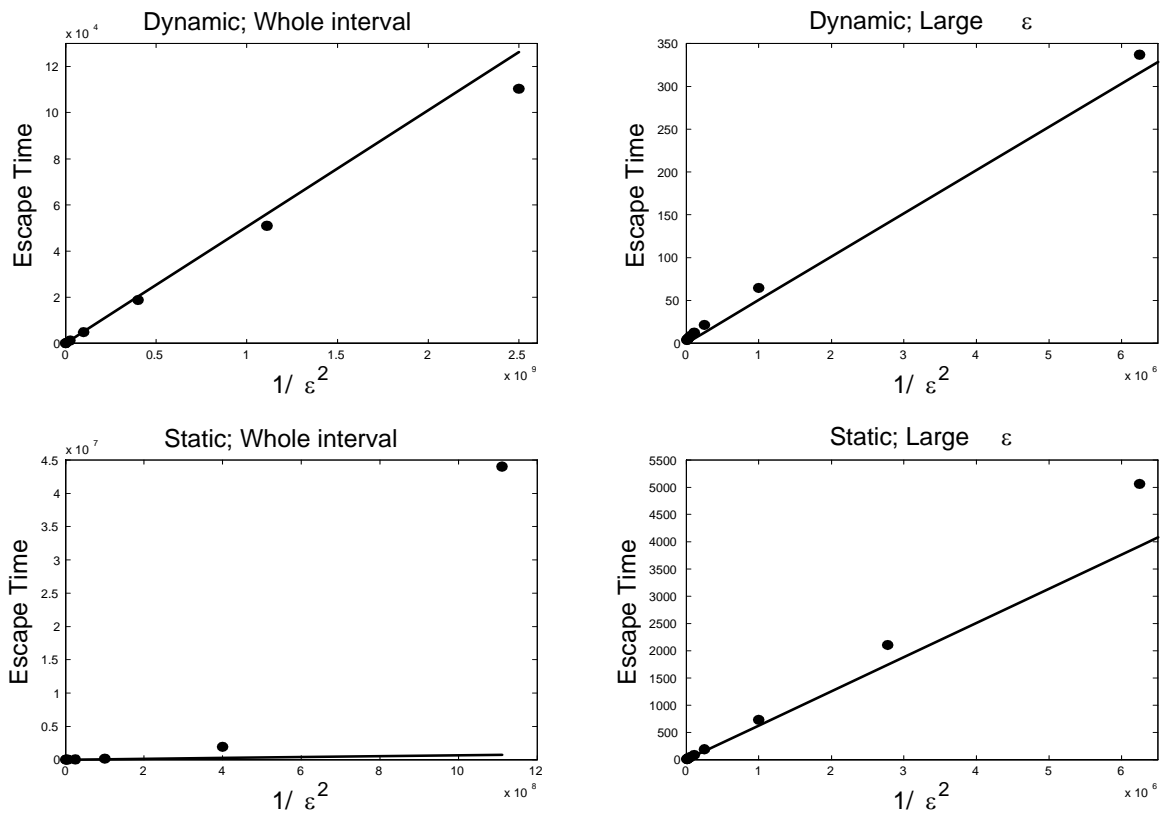


Figure 10:

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the following institutional grants:

- Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2009),
- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

(c) Dmitri Kolyuzhnov, Anna Bogomolova, Sergey Slobodyan, 2006

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by

Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and

Economics Institute (EI), Academy of Sciences of the Czech Republic

CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.

Printed by CERGE-EI, Prague

Subscription: CERGE-EI homepage: <http://www.cerge-ei.cz>

Editors: Directors of CERGE and EI

Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298

ISBN 80-7343-080-0 (Univerzita Karlova v Praze, CERGE)

ISBN 80-7344-069-5 (Národohospodářský ústav AV ČR, Praha)



CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
<http://www.cerge-ei.cz>