Robustifying Learnability

Robert Tetlow*
Peter von zur Muehlen

*Division of Research and Statistics, Federal Reserve Board, Washington, D.C., 20051 USA

E-mail: rtetlow@frb.gov http://www.roberttetlow.com

Disclaimer!

- The views in this presentation are those of the authors only and are not necessarily shared by the members of the Board of Governors or the staff.

Motivation

- Model uncertainty is widely accepted.
- Agents may not to have RE but are rational and can learn.
- But: No assurance they use LS correctly or even use LS at all.

Objectives

- **Introduce uncertainty in the models agents use.**
- **Devise procedures for maximizing prospect that an economy converges in expectations in the presence of model uncertainty…that is, robustify learnability**

Methodology

- We take the *learnability* literature…
- …and marry it to the *robust control* literature
- …to come up with tools to choose policy rules that render a model's *actual law of motion* locally robust to misspecification

Contribution of this paper

- Policy makers can act to minimize consequences of these errors.
- Not knowing where mistakes arise, and to protect against worst cases, the authority uses methods of *structured* robust control.

Performance metrics

- *Not* concerned with loss function minimization
- *Mostly* concerned with ensuring convergence of learning by agents to REE: E-stability.

A few of the pertinent references

Learning & determinacy literature*:*

- Bullard and Mitra (2001) JME, (2003).
- Evans and Honkapohja (EH) *Learning and Expectations in Macroeconomics* (2001).

Control literature:

- Zames (1966), Zhou-Doyle-Glover (1996).
- Onatski and Stock (2002), Zhou and Doyle, *Essentials of Robust Control* (1998), Tetlow and vzM (2001).

Literature related to this paper*:*

- *Evans & McGough (2004),EH&Marimon (MD,2001)*

Determinacy and Learning

- Plausible policy rules can be unstable under learning. (Bullard & Mitra, JME 2002 and Evans&Honkapohja 2003).
- Determinacy and learnability are not the same.

A General Linear Framework

$$y_t = \alpha + M E_t^* y_{t+1} + N y_{t-1} + P v_t \quad with \quad v_t = \rho v_{t-1} + \varepsilon_t$$

$$\begin{pmatrix} E_t^* y_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} \alpha M^{-1} \\ 0 \end{pmatrix} + \begin{pmatrix} M^{-1} & -NM^{-1} \\ I & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} - \begin{pmatrix} M^{-1} \\ 0 \end{pmatrix} P v_t + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \varepsilon_t$$

$$E_t^* Y_{t+1} = A + B Y_t + C v_t + D \varepsilon_t$$

- Unique REE if the Blanchard & Kahn conditions hold

Perceived Law of Motion

- Perceived law of motion (PLM): the model agents estimate using LS

$$Y_t = a_t + b_t Y_{t-1} + c_t v_t.$$

- Actual law of motion (ALM):
- Agents substitute expectations formed with PLM into the linear model.

Mapping from PLM to ALM

- **Resulting mapping**

$$(a,b,c) = T(a,b,c)$$

- **is unique if there is a fixed point for which b has all roots inside the unit circle.**

E-stability

- If eigenvalues of Jacobian of ODE

$$\frac{d}{d\tau}(a,b,c)=T(a,b,c)-(a,b,c)$$

- have real parts <1. See Evans and Honkapohja (2001).

Model Perturbation

- Begin with a PLM (omit intercept and let $X_t = [Y_t, v_t]$):

$$X_t = \begin{pmatrix} b & c \\ 0 & \rho \end{pmatrix} X_{t-1} + \varepsilon_t$$

$$\equiv \Pi \cdot X_{t-1} + \varepsilon_t$$

- Now perturb this model:

$$X_t = [\Pi + \Delta_W] \cdot X_{t-1} + \varepsilon_t$$

Perturbation Operator

- where, with W_1 and W_2 as scaling matrices

$$\Delta_W = W_1 \Delta W_2$$

Δ is a diagonal matrix with norm:

$$\|\Delta\|_\infty = \sqrt{\sup_\omega \max_{eig}[\Delta'(e^{-i\omega})\Delta(e^{i\omega})]} < r < \infty$$

- We want *r* as large as possible

Augmented Feedback Loop

- We can write the model as

$$\begin{pmatrix} X_{t+1} \\ p_t \end{pmatrix} = \begin{pmatrix} \Pi & W_1 \\ W_2 & 0 \end{pmatrix} \begin{pmatrix} X_t \\ h_t \end{pmatrix}$$

$$h_t = \Delta p_t$$

- *Transfer function from h to X and p:*

$$\begin{pmatrix} X_t \\ p_t \end{pmatrix} \equiv \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} h_t$$

$$G_1 = (IL^{-1} - \Pi)^{-1} W_1; G_2 = W_2 (IL^{-1} - \Pi)^{-1} W_1$$

Small Gain Theorem

$$\|\Delta\|_{\infty} < 1/r \;\; iff \;\; \|G_2(s)\|_{\infty} < r$$

- i.e. stabilize G_2 and you stabilize the entire system
- The object is to find the largest singular value of
- such that $I - G_2\Delta$ is not invertible

Structured Singular Value

$$\mu(\phi,\omega) = \min\left\{\begin{matrix}\overline{\sigma}[\Delta(e^{i\omega})] : \Delta \in D_r, \\ \det[I - G_2(\omega)\Delta(e^{i\omega})] = 0\end{matrix}\right\}^{-1}$$

$$\overline{\mu}(\phi^*) = \inf_{\phi} \sup_{\omega \in [0,2\pi]} \mu(\phi,\omega)$$

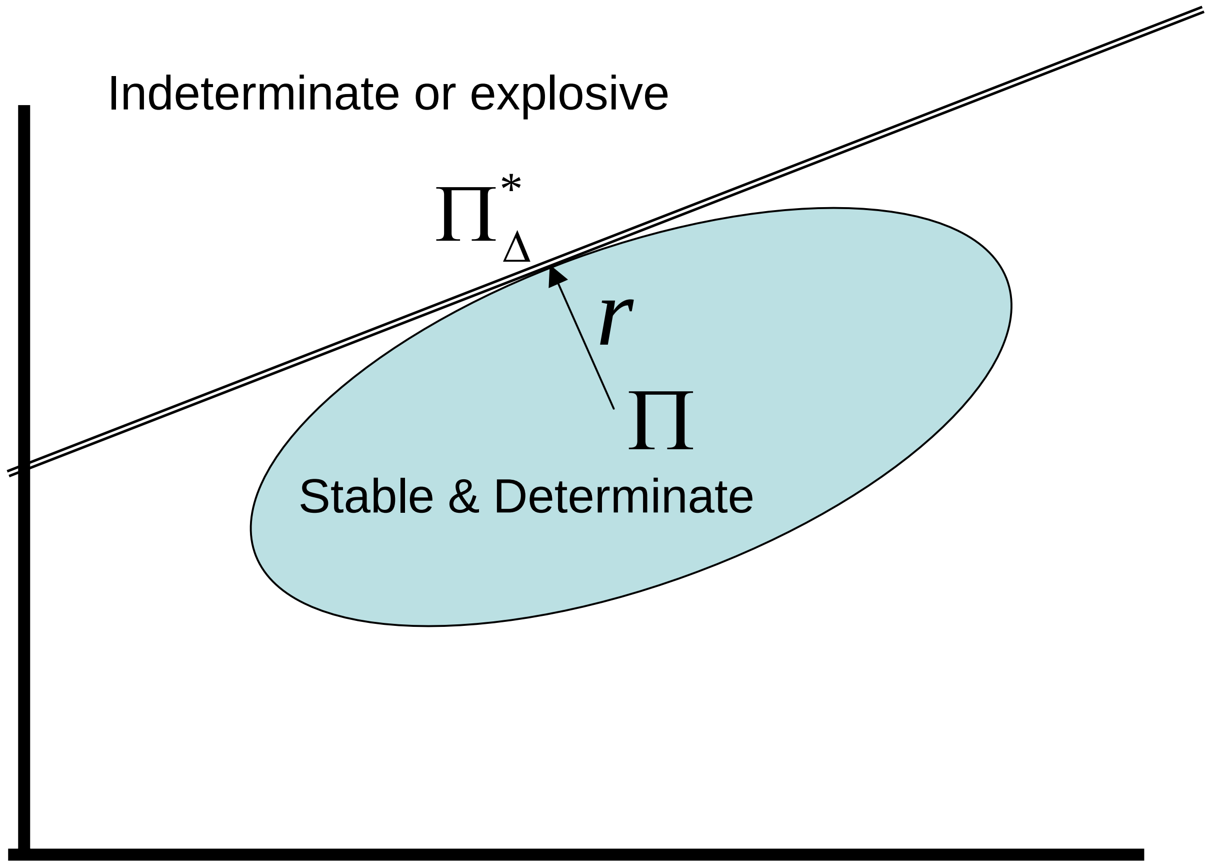
$$r \approx 1/\overline{\mu}$$

r = radius of allowable perturbations

$\overline{\sigma}$ = maximum singular value

Schematic representation of robust learnability

maximum perturbation space around the reference PLM that keeps the model determinate and stable



The NKB model

$$x_t = E_t^* x_{t+1} - 1/\sigma (r_t - r_t^n - E_t^* \pi_{t+1})$$

$$\pi_t = \beta E_t^* \pi_{t+1} + \kappa x_t$$

$$r_t^n = \rho r_{t-1}^n + \varepsilon_t$$

…plus one of three policy rules

1. Lagged-data rule

$$r_t = \phi_x x_{t-1} + \phi_\pi \pi_{t-1} + \phi_r r_{t-1}$$

2. Contemporaneous-data rule

$$r_t = \phi_x x_t + \phi_\pi \pi_t + \phi_r r_{t-1}$$

3. Forecast-based rule

$$r_t = \phi_x E_t x_{t+1} + \phi_\pi E_t \pi_{t+1} + \phi_r r_{t-1}$$

Information Protocol

The central bank knows

$$\sigma, \kappa, \beta, \rho$$

Agents observe

$$x_{t-1}, \pi_{t-1}, r_{t-1}, r^n_{t-1}$$

The PLM

- We assume that agents estimate a VAR in x_t, π_t, r_t, r_t^n
- In most cases our PLMs are over-parameterized.
- But these PLMs converge to the MSV solution in learning.

Scaling the perturbations

- We allow all 16 coefficients in the VAR to be perturbed.
- Each perturbation is scaled by its standard deviation in the VAR estimated prior to the experiments.

No Sunspots!

- Learnability is made robust conditional on establishment of a unique saddle-point equilibrium.

Bullard and Mitra (2003)

FIGURE 2. Forward Expectations

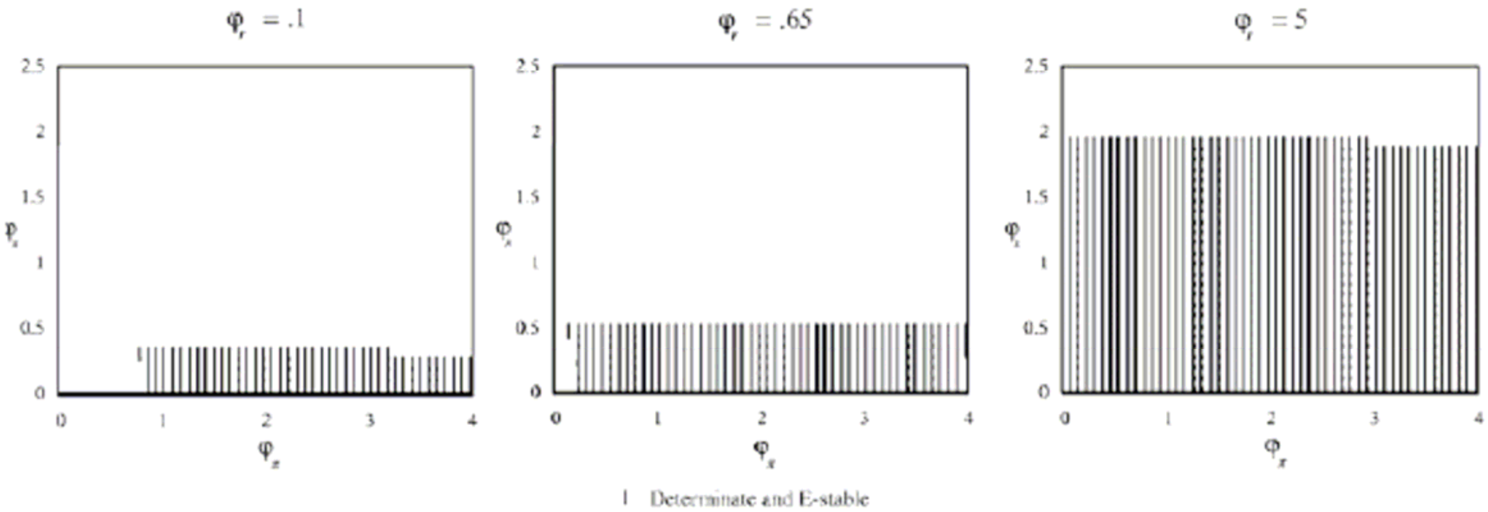


Table 1 : **Contemporaneous data rules**

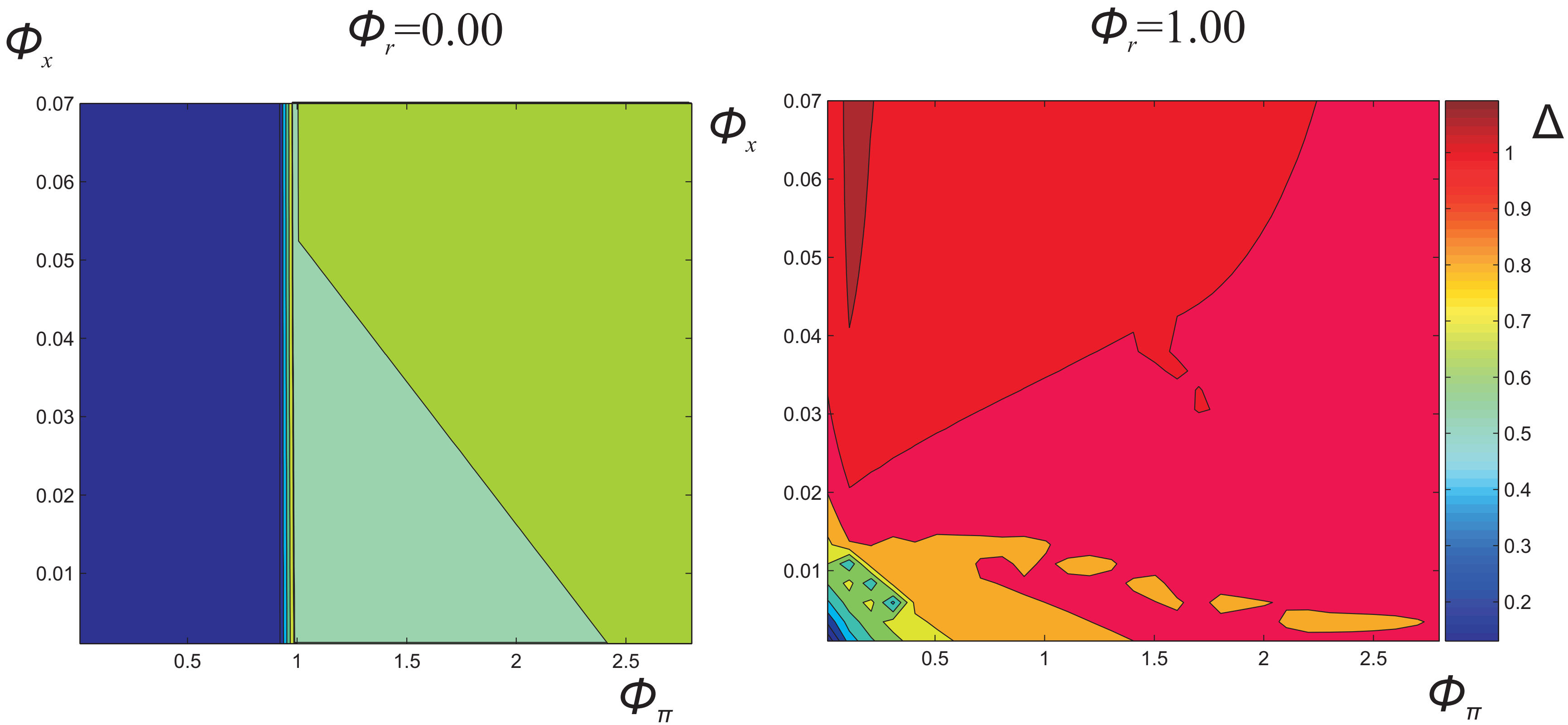
	Rule coefficients				
rule	x_t	π_t	r_{t-1}	**radius**	**loss**
optimized	**0.053**	**0.995**	**1.12**	**1.06**	**3.63**
robust	**0.052**	**1.21**	**1.41**	**1.13**	**3.70**

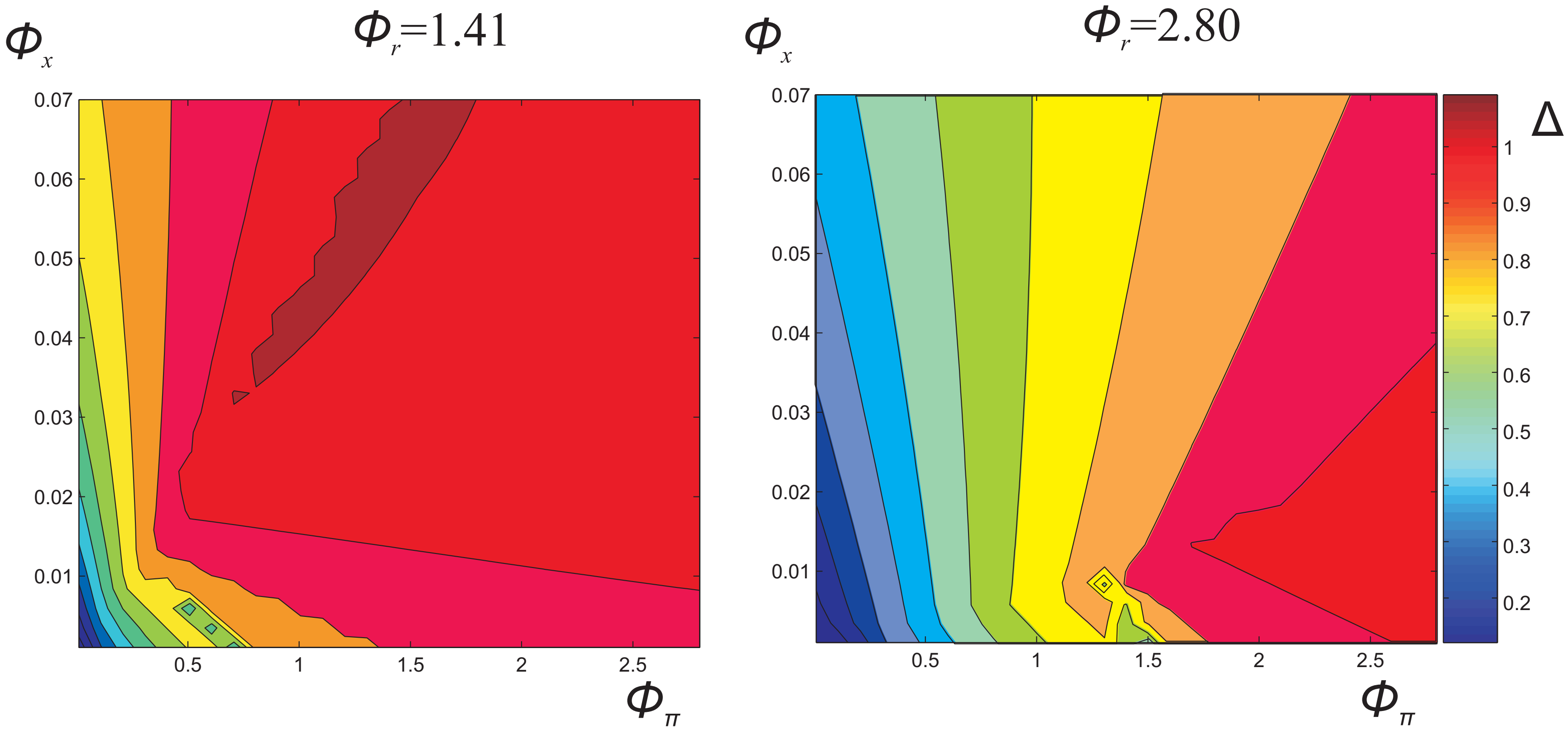
- Does high policy inertia necessarily imply learnability under misspecification?
- The contour maps in the next slides suggest: not exactly.
- Difference between

  --- robust learnability

  --- size of learnable space.

Robustness contours
Contemporaneous data rules



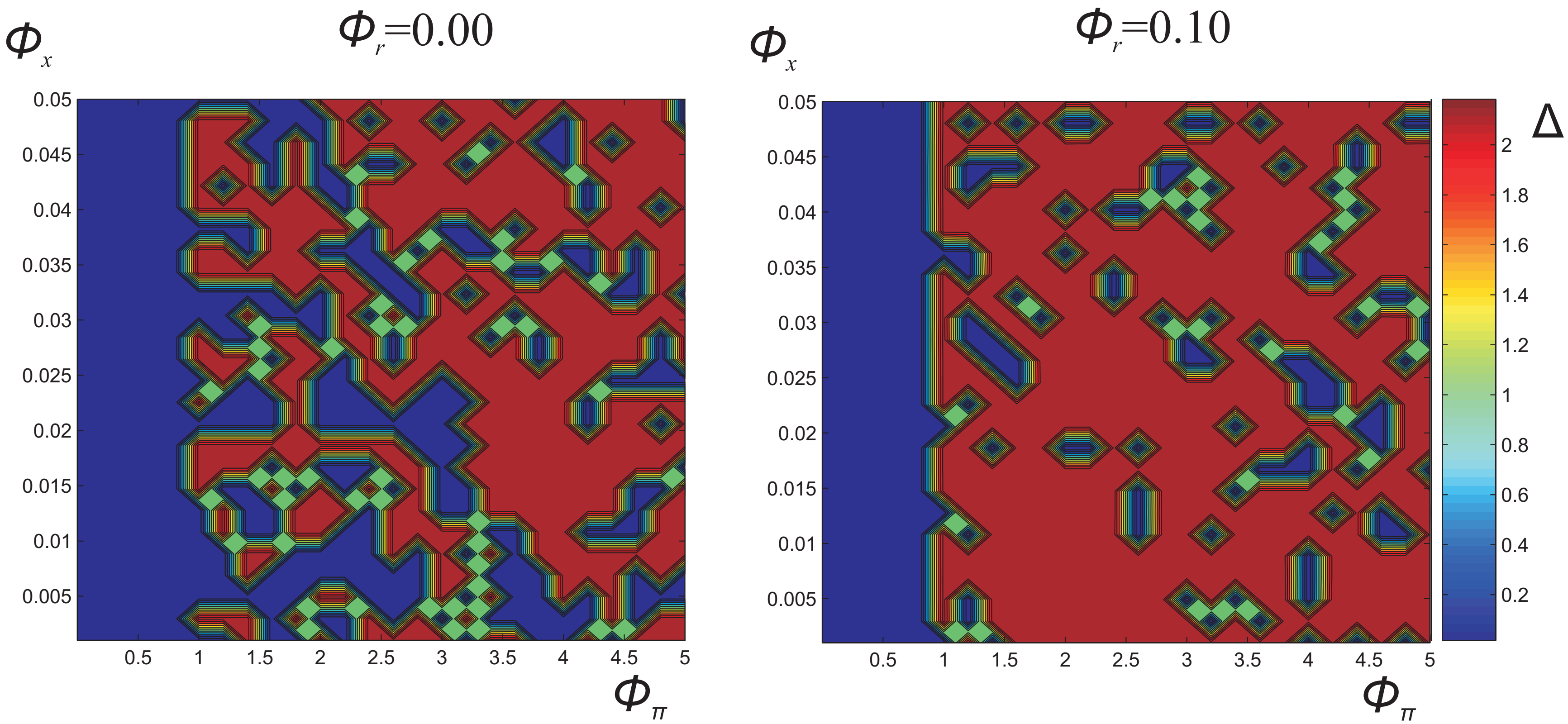


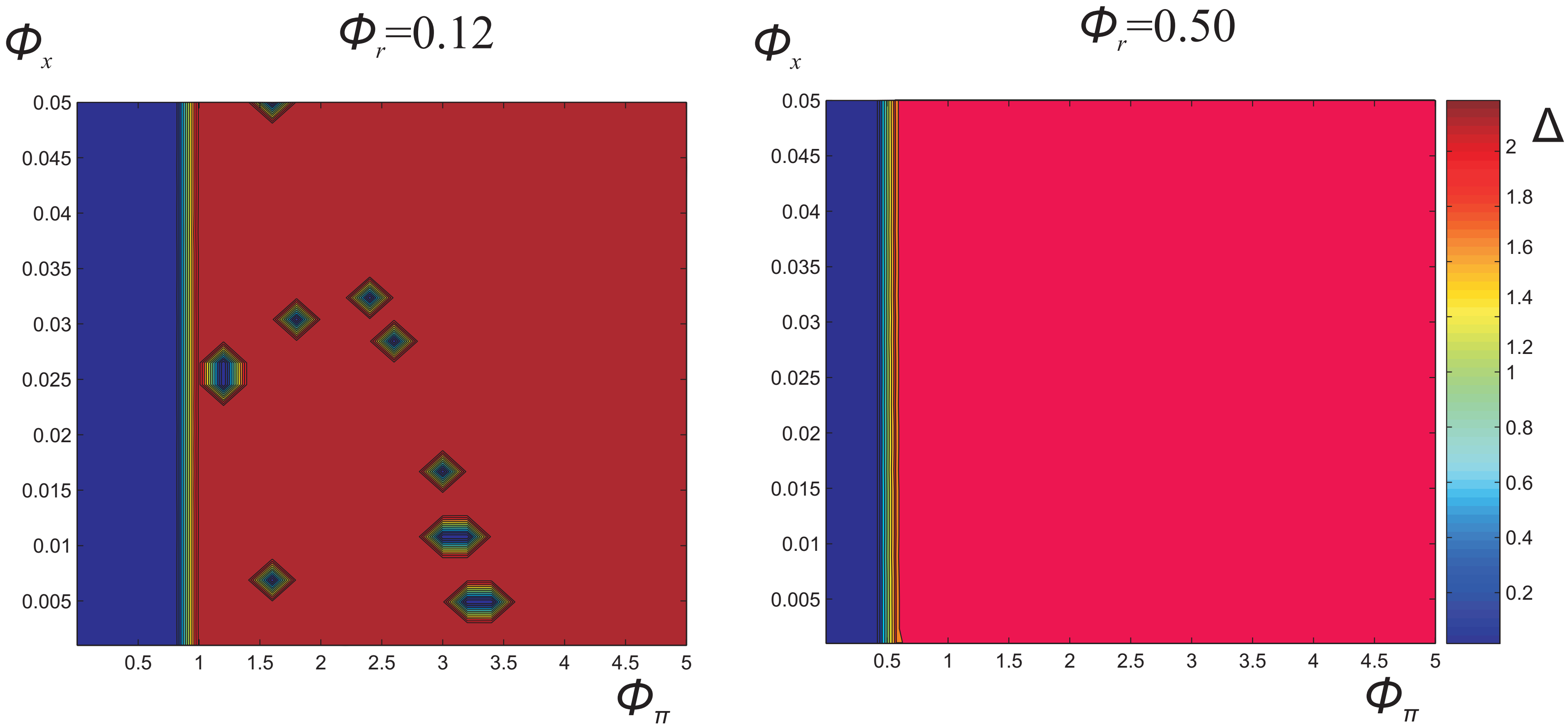
Φ_r=1.41
Φ_x
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.5
1
1.5
2
2.5
Φ_π
Φ_r=2.80
Φ_x
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.5
1
1.5
2
2.5
Φ_π
Δ
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2

Table 2 : **Forecast-based rules**

	Rule coefficients				
rule	x_t	π_t	r_{t-1}	**radius**	**loss**
optimized	**0.29**	**0.99**	**1.32**	**0.88**	**3.63**
robust	**0.04**	**2.80**	**0.10**	**2.32**	**4.43**

Robustness contours forecast-based rule




Φ_x
Φ_r=0.12
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5
Φ_π
Φ_x
Φ_r=0.50
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0.5 1 1.5 2 2.5 3 3.5 4 4.5 5
Φ_π
Δ
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2

Conclusions

- **Policy is about more than minimizing a loss function.**
- **If agents form expectations by recursive learning in mis-specified models, policy can facilitate learning to achieve a REE.**
- **We have identified and described tools to do this using robust control theory.**
- **A robustly learnable rule is not the same as rule that has a wide learnable space in a given model.**