Department of Policy and Planning Sciences

Discussion Paper Series

No.1354

On the Computation of Detection Error Probabilities under Normality Assumptions

by

Masakatsu OKUBO

May 2018

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

On the Computation of Detection Error Probabilities under Normality Assumptions *

Masakatsu Okubo †

University of Tsukuba

Abstract

This note describes a simple method for computing the exact value of detection error probabilities under log-consumption models with i.i.d. Gaussian errors. The method is applicable to a class of models widely used in the literature, including the random walk, trend-stationary, long-run risk, and idiosyncratic risk models. The advantage of the method is more evident in applications where the overall detection error probability, defined as the average of two kinds of detection error probabilities, is computed many times.

Keywords: Asset Pricing; Detection error probability; Model misspecification; Multiplier preferences

JEL Classification: D81; E21; G12

---

*First draft, December 2014; Revised, January 2018.

†Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8573, Japan; Tel.: +81-29-853-5369; Fax: +81-29-853-5070; E-mail: okubo@sk.tsukuba.ac.jp

1 Introduction

Hansen and Sargent (2008a) and Barillas et al. (2009) used detection error probabilities to demonstrate that a moderate amount of concern about model misspecification under multiplier preferences can substitute for an implausibly high level of risk aversion. The computation of these detection error probabilities is under the assumption that the log consumption streams an agent faces in an endowment economy follow a random walk or trend-stationary process with i.i.d. Gaussian errors. The computational procedure relies entirely on simulation. In this note, we show that it is possible to compute the detection error probabilities using the cumulative distribution function under a class of models widely used in the literature, including the random walk, trend-stationary, long-run risk, and idiosyncratic risk models.

Under the random walk and trend-stationary models, Djeutem (2014) was the first to show that detection error probabilities can be calculated in a closed form. However, this note extends these results and makes the following unique contributions. First, it demonstrates that there are closed-form solutions for detection error probabilities if the value function is linear in i.i.d. Gaussian shocks, which also holds for a particular class of long-run and idiosyncratic risk models.[1] Thus, it provides a generalization of the formula in two directions.[2] Second, it presents a method for calculating standard errors for the overall detection error probability using the delta method.

The advantages of our result described here are twofold. The first is that it more quickly and easily provides the exact value of the detection error probabilities and enables us to test for their statistical significance unlike the existing simulation-based method. The second is that it enables us to reveal analytically their properties and therefore facilitates our interpretation. Our method, being based on a closed-form solution, is also useful if the overall detection error probability must be computed many times, which holds for the asset-pricing applications in Hansen and Sargent (2008a) and Barillas et al. (2009).

[1] The intuition for this is given in footnote 3 using a simple static setting.

[2] Our proof differs from that of Djeutem (2014) in several respects and includes a correction of his proof.

This note is organized as follows. Section 2 briefly reviews the framework and computation procedure proposed by Hansen and Sargent (2008a) and Barillas et al. (2009). Section 3 explains our approach and shows how it is applicable to their asset-pricing applications. Section 4 discusses the extensions and some limitations of our formulas. All proofs are in the separate appendix.

2 The Framework and Computation Procedure

Hansen and Sargent (2008a) and Barillas et al. (2009) used the finding that risk-sensitive preferences and multiplier preferences are observationally equivalent to reinterpret the quantitative finding of Tallarini (2000) concerning the risk aversion parameter. The risk-sensitive preferences are a special case of the recursive preferences suggested by Epstein and Zin (1989) and Weil (1990), in which the intertemporal elasticity of substitution is fixed at unity:

$$U_t = c_t - \beta\theta \ln\left(E_t\left[\exp\left(-\frac{U_{t+1}}{\theta}\right)\right]\right), \tag{1}$$

where c_t is log consumption and $\beta \in (0,1)$ is a discount factor. The parameter θ represents a measure of risk aversion

$$\theta = -\frac{1}{(1-\beta)(1-\gamma)}, \tag{2}$$

where γ is a coefficient of relative risk aversion (RRA).

From the viewpoint of multiplier preferences, this parameter θ can be interpreted as the degree of an agent's concern about model misspecification. The detection error probabilities are used to quantify the degree to which the agent fears model misspecification. To illustrate the calibration method, let model A be an approximating model (a reference model), and let model B be a worst-case model associated with θ^{-1} (an alternative model in proximity to model A). Let p_A denote the probability that a likelihood-ratio test selects model B when model A generates the data. Define p_B similarly as the probability that selects model A when model B generates the data. Finally, define the overall detection error probability $p(\theta^{-1})$ by $p(\theta^{-1}) \equiv \frac{1}{2}(p_A + p_B)$.

In Hansen and Sargent (2008a) and Barillas et al. (2009), model A is assumed to be the following random walk and trend-stationary models

$$c_t = \mu + c_{t-1} + \sigma_\epsilon \epsilon_t, \tag{3}$$

$$c_t = \zeta + \mu t + z_t,\ z_t = \rho z_{t-1} + \sigma_\epsilon \epsilon_t,\ |\rho| < 1, \tag{4}$$

where $\epsilon_t \sim$ i.i.d.$N(0,1)$. The corresponding worst-case model (model B) is then given by

$$c_t = \mu + \sigma_\epsilon w_{RW} + c_{t-1} + \sigma_\epsilon \epsilon_t,\ w_{RW} \equiv -\sigma_\epsilon/\theta(1-\beta), \tag{5}$$

$$c_t = \mu_1 + \mu_2 t + \sigma_\epsilon w_{TS} + \rho c_{t-1} + \sigma_\epsilon \epsilon_t,\ w_{TS} \equiv -\sigma_\epsilon/\theta(1-\rho\beta), \tag{6}$$

where $\mu_1 \equiv \zeta(1-\rho) + \rho\mu$ and $\mu_2 \equiv (1-\rho)\mu$. The procedure for calibrating the detection error probabilities developed by Hansen and Sargent (2008a) and Barillas et al. (2009) (henceforth, the BHS procedure) proceeds as follows.

1. Set the values of θ^{-1}, β, ζ, μ, ρ, and σ_ϵ. Simulate a path of length T for c_t using model A. Calculate the log-likelihood ratio, $\ln(L_A/L_B)$, to perform a test for distinguishing model A from model B. The test selects model A if $\ln(L_A/L_B) > 0$ and model B if $\ln(L_A/L_B) < 0$. Perform this test many times by simulating a large number of paths under model A, and count the fraction of $\ln(L_A/L_B) < 0$

$$p_A \equiv \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right) < 0\right) \approx \frac{\#\ln(L_A/L_B) < 0}{\#\text{simulations}}. \tag{7}$$

2. Simulate a large number of paths of length T for c_t using model B. Perform the log-likelihood ratio test, and count the fraction of $\ln(L_A/L_B) > 0$

$$p_B \equiv \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right) > 0\right) \approx \frac{\#\ln(L_A/L_B) > 0}{\#\text{simulations}}. \tag{8}$$

3. Calculate the overall detection error probability $p(\theta^{-1})$.

4. Repeat steps 1–3 for different values of θ^{-1} to obtain a graph of the overall detection error probability versus θ^{-1} (i.e., a detection error probability function).

The number of simulations for each computation of p_A and p_B is 100,000 or 500,000 in the BHS procedure (see Barillas et al. (2009, p. 2405) and Hansen and Sargent (2008a, p. 320)), so that the total number of simulations required is 200,000 or 1,000,000 to obtain *one* value of the overall detection error probability $p(\theta^{-1})$.

3 Simplification of the Procedure

Let $\Phi(\cdot)$ be the standard normal cumulative distribution function. The following proposition states that we can compute $p(\theta^{-1})$ without relying on simulation under the random walk and trend-stationary models with i.i.d. Gaussian errors. To our knowledge, Djeutem (2014) has already noted this claim, but in a different context and form.

Proposition 1.

(i) For the random walk drift model, the detection error probabilities p_A and p_B are given by

$$p_A = \Phi\left(-\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\beta)}\right) \quad \text{and} \quad p_B = 1 - \Phi\left(\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\beta)}\right). \tag{9}$$

(ii) For the trend-stationary model, they are

$$p_A = \Phi\left(-\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right) \quad \text{and} \quad p_B = 1 - \Phi\left(\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right). \tag{10}$$

The overall detection error probability $p(\theta^{-1})$ is equal to p_A.

A proof for this proposition is in Appendix A. In the proof, the key is that if the value function U_t is *linear* in random shocks ϵ_t, then a likelihood ratio $g(\epsilon_{t+1}) \equiv \hat{\pi}(\epsilon_{t+1})/\pi(\epsilon_{t+1})$ can be expressed as the exponential of a linear function of ϵ_{t+1}. Here, $\pi(\epsilon_{t+1})$ is a conditional density of a sequence of random shocks $\{\epsilon_{t+1}\}$, and $\hat{\pi}(\epsilon_{t+1})$ is some other density in proximity to $\pi(\epsilon_{t+1})$ (i.e., a distorted density). By this result, the log-likelihood ratio $\ln(L_A/L_B)$ takes the familiar form under the AR(1) structure. Using this and the normality assumption of the shocks ϵ_t, it is shown that the detection error probability p_A in the BHS procedure represents the cumulative distribution function of a standard normal random variable (constructed from the i.i.d. Gaussian shocks ϵ_t).[3] Given this result, the representation for p_B follows from the symmetry of the standard normal distribution.

When $\theta^{-1} = 0$ (i.e., model A and model B are identical), it is easy to confirm from the formulas that $p(\theta^{-1}) = 0.5$ because of $\Phi(0) = 0.5$. Also, our formulas establish that the overall detection error probability is a decreasing function of θ^{-1}, other things being equal. These are consistent with both the claim and simulation-based finding in Hansen and Sargent (2008a) and Barillas et al. (2009). In addition, our formulas reveal that the overall detection error probability is a decreasing function of two variables. One is the sample size $\sqrt{T}$. This means that the agent can distinguish between the approximating model and the worst-case model more easily given more data (i.e., a longer history of the economy). The other is the volatility parameter σ_ϵ of the consumption processes. A higher volatility also makes it easier for the agent to distinguish between the two models, so that the model detection errors become lower.

To illustrate the use of our result in the asset-pricing application, we apply estimates of the random walk and trend-stationary models and the values of β and γ given in Barillas et al. (2009). They estimated μ, σ_ϵ, ρ, and ζ using maximum likelihood (ML) methods and US quarterly consumption data from 1948:2 to 2006:4 ($T = 235$). Panels A and B of Table 1 summarize the ML estimates and parameter values.[4] Their calibration results indicate that overall detection error probabilities between 0.01 and 0.05 succeed in achieving the Hansen–Jagannathan bounds. However, Barillas et al. (2009) do not reveal their exact value.

Panel C of Table 1 presents the calculation results of the overall detection error probability

[3] The intuition of the proof is the following. To see the idea clearly, consider a simplified static structure. Note that the likelihood ratio $g(\epsilon)$ takes the form, $g(\epsilon) \equiv \hat{\pi}(\epsilon)/\pi(\epsilon) = \exp(-U/\theta)/E[\exp(-U/\theta)]$. Then the detection error probability p_A is $p_A = \text{Prob}(\text{select model B}|\text{model A generated the data}) = \text{Prob}(\ln g^*(\epsilon) < 0|\pi(\epsilon))$, where $g^*(\epsilon) \equiv 1/g(\epsilon)$. (This inversion is merely for maintaining consistency with L_A/L_B and is not essential.) If the value function U is linear in ϵ, say, $U = a_0 + a_1\epsilon$, then $p_A = \text{Prob}(\epsilon < -(\theta/a_1)\ln(E[\exp(-(a_1/\theta)\epsilon)])|\pi(\epsilon)) = \text{Prob}(\epsilon < -a_1/2\theta|\pi(\epsilon))$, so that the distribution function $\Phi(\cdot)$ can be used because of $\epsilon \sim N(0,1)$. Note that while this static-case derivation conveys our idea, our proof is needed in the dynamic setting that we treated.

[4] While Hansen and Sargent (2008a, p. 321) reported that the value of γ that achieves the Hansen–Jagannathan bounds is around 250 for the trend-stationary model, Barillas et al. (2009, p. 2406) pointed out that it is only about 75, despite both using US data from almost the same period. This large difference in γ between the two studies arises only for the case of the trend-stationary model. According to a preliminary investigation based on our formulas, it seems difficult to corroborate the claim of Barillas et al. (2009) for $\gamma = 75$. Hence, we adopt $\gamma = 250$ for the trend-stationary model.

based on our formulas (9) and (10), where we employ the MATLAB function `normcdf` as the cumulative distribution function $\Phi(\cdot)$. The first step of our procedure is to find the value of θ^{-1} corresponding to the value of γ that attains the Hansen–Jagannathan bounds. This is $\theta^{-1} = (1-0.995)(50-1) = 0.245$ for $\gamma = 50$ in the random walk case and $\theta^{-1} = (1-0.995)(250-1) = 1.245$ for $\gamma = 250$ in the trend-stationary case. The second step is to determine the overall detection error probability for this γ by substituting the value of θ^{-1} obtained and the ML estimates into our formulas. It is $p(\theta^{-1}) = 0.0302$ for the random walk model and $p(\theta^{-1}) = 0.0277$ for the trend-stationary model.[5] Of these, at least the former is significantly different from zero (see Appendix D for the calculation of the standard errors).

4 Extensions

This section discusses what types of consumption processes have a closed-form solution for the detection error probabilities. We focus here on two models. One is a simple version of the long-run risk model of Bansal and Yaron (2004), which has been studied in Hansen (2007) and Hansen and Sargent (2008b, 2010). The other is the model in which log individual consumption has both aggregate and idiosyncratic risk components, which has been considered in De Santis (2007) and Ellison and Sargent (2015).

4.1 Long-Run Risk

A simple version of the long-run risk models used in Hansen (2007, Example 2) and Hansen and Sargent (2008b, 2010) is given by

$$\begin{aligned} c_{t+1} - c_t &= \mu + z_t + \sigma_\epsilon \epsilon_{t+1}, \\ z_{t+1} &= \rho z_t + \sigma_z \epsilon_{t+1}, \quad \epsilon_{t+1} \sim \text{i.i.d.} N(0,1). \end{aligned} \tag{11}$$

As described in Appendix B, this long-run risk model can be expressed as an ARIMA(1,1,1) process. The following proposition therefore states that we can derive a closed-form representation of the detection error probabilities under the ARIMA(1,1,1) model for log consumption.

[5] To confirm the validity of our results, we plotted the overall detection error probabilities against various values of θ^{-1} for the random walk and trend-stationary models using formulas (9) and (10) (see Appendix E). This figure is consistent with Figure 2(a) in Barillas et al. (2009).

Proposition 2.

For the long-run risk model, the detection error probabilities p_A and p_B are given by

$$p_A = \Phi\left(-\frac{\sqrt{T-1}}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{1}{\theta}\left\{\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right\}\right) \tag{12}$$

and

$$p_B = 1 - \Phi\left(\frac{\sqrt{T-1}}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{1}{\theta}\left\{\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right\}\right). \tag{13}$$

The overall detection error probability $p(\theta^{-1})$ is equal to p_A.

A proof for this proposition is in Appendix B. Note that the simple long-run risk model (11) is a special case of the multivariate state model of Hansen et al. (2008) where differences in log consumption are a linear function of a state vector $\mathbf{x}$ that follows a first-order vector autoregression: $c_{t+1} - c_t = \mu_c + U_c\mathbf{x}_t + \lambda_0\mathbf{w}_{t+1},\ \mathbf{x}_{t+1} = G\mathbf{x}_t + H\mathbf{w}_{t+1},\ \mathbf{w}_{t+1} \sim \text{i.i.d.} N(\mathbf{0}, I)$ (here, the notation follows theirs). Our proof remains valid for a multivariate state case if all elements of the row vector U_c are one and the matrix G is the diagonal one with the same element, say, ρ (note that the model has the same structure as that of (11) in this case).

4.2 Idiosyncratic Risk

Following Ellison and Sargent (2015), we assume the following value function recursion:

$$U_t = c_t^i - \beta\theta \ln\left(E_t\left[\exp\left(-\frac{U_{t+1}}{\theta}\right)\right]\right). \tag{14}$$

Here, log individual consumption c_t^i has aggregate and idiosyncratic risk components that follow random walk processes:

$$\begin{aligned} c_t^i &= c_t + \delta_t^i, \\ \Delta c_t &= \sqrt{\epsilon} w_{1t}, \\ \Delta \delta_t^i &= \sqrt{\epsilon} w_{2t}, \end{aligned} \tag{15}$$

where

$$\begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix} \sim N\left(\begin{bmatrix} g - \tau_1^2/2 \\ -\tau_2^2/2 \end{bmatrix}, \begin{bmatrix} \tau_1^2 & 0 \\ 0 & \tau_2^2 \end{bmatrix}\right).$$

A noteworthy point for our purpose is that this specification can be rewritten as a random walk model for log individual consumption: $c_{t+1}^i = c_t^i + \sqrt{\epsilon}(w_{1t+1} + w_{2t+1})$. Assuming that the aggregate and idiosyncratic shocks w_{1t} and w_{2t} are i.i.d. as in the previous cases, we can show the following.

Proposition 3.

For the model with idiosyncratic risk, the detection error probabilities are given by

$$p_A = \Phi\left(-\frac{\sqrt{T}}{2}\frac{\sqrt{\epsilon(\tau_1^2+\tau_2^2)}}{\theta(1-\beta)}\right) \quad \textit{and} \quad p_B = 1-\Phi\left(\frac{\sqrt{T}}{2}\frac{\sqrt{\epsilon(\tau_1^2+\tau_2^2)}}{\theta(1-\beta)}\right). \tag{16}$$

The overall detection error probability $p(\theta^{-1})$ is equal to p_A.

A proof for this proposition is in Appendix C. This proposition can be regarded as a generalization of Proposition 1 because $\sqrt{\epsilon(\tau_1^2+\tau_2^2)}$ corresponds to the square root of the variance of the error term in the random walk model for log individual consumption.

Acknowledgments

This note was originally part of a paper presented at the 2014 Fall Meeting of the Japanese Economic Association held on October 11–12, 2014. The author is grateful to Masao Ogaki and anonymous referees for useful comments and constructive suggestions. This work was supported by a grant-in-aid from the Japan Society for the Promotion of Science (JSPS KAKENHI Grant No. JP15K03391).

References

Bansal, Ravi, and Amir Yaron (2004), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance* 59, 1481–1509.

Barillas, Francisco, Lars Peter Hansen, and Thomas J. Sargent (2009), "Doubts or Variability?," *Journal of Economic Theory* 144, 2388–2418.

De Santis, Massimiliano (2007), "Individual Consumption Risk and the Welfare Cost of Business Cycles," *American Economic Review* 97, 1488–1506.

Djeutem, Edouard (2014), "Model Uncertainty and the Forward Premium Puzzle," *Journal of International Money and Finance* 46, 16–40.

Ellison, Martin, and Thomas J. Sargent (2015), "Welfare Cost of Business Cycles with Idiosyncratic Consumption Risk and a Preference for Robustness," *American Economic Journal: Macroeconomics* 7(2), 40–57.

Epstein, Larry G., and Stanley E. Zin (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57, 937–969.

Hansen, Lars Peter (2007), "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," *American Economic Review* 97, 1–30.

Hansen, Lars Peter, and Thomas J. Sargent (2008a), *Robustness*, Princeton, NJ: Princeton University Press.

Hansen, Lars Peter, and Thomas J. Sargent (2008b), "Fragile Beliefs and the Price of Model Uncertainty," Unpublished manuscript, University of Chicago and New York University.

Hansen, Lars Peter, and Thomas J. Sargent (2010), "Fragile Beliefs and the Price of Uncertainty," *Quantitative Economics* 1, 129–162.

Hansen, Lars Peter, John C. Heaton, and Nan Li (2008), "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy* 116, 260–302.

Tallarini, Thomas D., Jr. (2000), "Risk-Sensitive Real Business Cycles," *Journal of Monetary Economics* 45, 507–532.

Weil, Philippe (1990), "Nonexpected Utility in Macroeconomics," *Quarterly Journal of Economics* 105, 29–42.

Table 1

A Computational Example of Detection Error Probabilities Based on an Alternative Method

	Model	
	Random walk	Trend stationary
A. Barillas et al. (2009) ML estimates		
μ	0.00495	0.00418
	(0.0003)	(0.0003)
σ_ϵ	0.0050	0.0050
	(0.0002)	(0.0002)
ρ	—	0.980
	—	(0.010)
B. Barillas et al. (2009) setting of parameters		
T	235	235
β	0.995	0.995
γ	50	250
C. Detection error probability		
θ^{-1}	0.2450	1.2450
$p(\theta^{-1})$	0.0302	0.0277
	(0.0051)	(0.0490)

Note: Standard errors in parentheses. The estimates and standard errors for μ, σ_ϵ, and ρ in Panel A are from Table 2 in Barillas et al. (2009). The values of T, β, and γ in Panel B are reported in Barillas et al. (2009). See footnote 4 for the choice of γ. The value of θ^{-1} in Panel C is calculated using $\theta^{-1} = (1 - \beta)(\gamma - 1)$ derived from equation (2). The overall detection error probability $p(\theta^{-1})$ is calculated using formulas (9) and (10).

Appendix for "On the Computation of Detection Error Probabilities under Normality Assumptions" *

Masakatsu Okubo †

University of Tsukuba

*First draft, December 2014; Revised, September 2017.

†Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8573, Japan; Tel.: +81-29-853-5369; Fax: +81-29-853-5070; E-mail: okubo@sk.tsukuba.ac.jp

A. Proof of Proposition 1

The result (i) for the random walk model is a special case of (ii) for the trend-stationary model. Therefore, we provide only the proof of (ii) below.[1] Consider equation (4) reproduced as

$$c_t = \zeta + \mu t + z_t,\ z_t = \rho z_{t-1} + \sigma_\epsilon \epsilon_t,\ \epsilon_t \sim \text{i.i.d.}N(0,1). \tag{A1}$$

As z_t has an AR(1) structure, the (average) log-likelihood function for a sample of $t = 1, 2, \ldots, T$ takes the form

$$\ln L = \frac{1}{T}\ln f(c_1) + \frac{1}{T}\sum_{t=2}^{T}\ln f(c_t|c_{t-1}). \tag{A2}$$

The density $f(c_1)$ under model A is obtained by writing (A1) at $t = 1$ as $c_1 = \zeta + \mu + z_1$ and $z_1 = \rho z_0 + \sigma_\epsilon \epsilon_1$. Assuming the initial condition $z_0 = 0$, it follows that $z_1 = \sigma_\epsilon \epsilon_1$, so that $c_1 = \zeta + \mu + \sigma_\epsilon \epsilon_1$. Therefore, the logarithm of the density $f(c_1)$ is

$$\ln f(c_1) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln \sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2}(c_1 - \zeta - \mu)^2. \tag{A3}$$

To obtain the conditional density $f(c_t|c_{t-1})$ under model A, rewrite (A1) by substituting out z_t as

$$c_t = \mu_1 + \mu_2 t + \rho c_{t-1} + \sigma_\epsilon \epsilon_t, \tag{A4}$$

where $\mu_1 \equiv \zeta(1-\rho) + \rho\mu$ and $\mu_2 \equiv (1-\rho)\mu$. Because of $\sigma_\epsilon \epsilon_t \sim \text{i.i.d.}N(0, \sigma_\epsilon^2)$, the logarithm of the conditional density $f(c_t|c_{t-1})$ is given by

$$\ln f(c_t|c_{t-1}) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln \sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2}(c_t - \mu_1 - \mu_2 t - \rho c_{t-1})^2. \tag{A5}$$

Substituting (A3) and (A5) into (A2), the log-likelihood function under model A is

$$\ln L_A = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln \sigma_\epsilon^2 - \frac{1}{T}\frac{1}{2\sigma_\epsilon^2}(c_1 - \zeta - \mu)^2 - \frac{1}{T}\sum_{t=2}^{T}\frac{1}{2\sigma_\epsilon^2}(c_t - \mu_1 - \mu_2 t - \rho c_{t-1})^2. \tag{A6}$$

[1] When $\rho = 1$, it is possible to simplify the proof by beginning with the first-differenced form of the model: $\Delta c_{t+1} = \mu + \sigma_\epsilon \epsilon_{t+1}$. However, when $\rho < 1$, this approach is not valid. The proof based on the first-difference form is described in Appendix C for a more general case that includes the random walk model (3) as a special case.

Noting that the difference between model A and model B is that the mean of ϵ_t shifts from 0 to w_{TS}, the log-likelihood function under model B is

$$\ln L_B = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma_\epsilon^2 - \frac{1}{T}\frac{1}{2\sigma_\epsilon^2}(c_1-\zeta-\mu-\sigma_\epsilon w_{TS})^2 - \frac{1}{T}\sum_{t=2}^{T}\frac{1}{2\sigma_\epsilon^2}(c_t-\mu_1-\mu_2 t-\rho c_{t-1}-\sigma_\epsilon w_{TS})^2. \tag{A7}$$

Thus, the log-likelihood ratio for the trend-stationary model is

$$\begin{aligned}\ln\left(\frac{L_A}{L_B}\right) &= -\frac{1}{T}\left[\frac{1}{2\sigma_\epsilon^2}(c_1-\zeta-\mu)^2 + \sum_{t=2}^{T}\frac{1}{2\sigma_\epsilon^2}(c_t-\mu_1-\mu_2 t-\rho c_{t-1})^2\right] \\ &\quad + \frac{1}{T}\left[\frac{1}{2\sigma_\epsilon^2}(c_1-\zeta-\mu-\sigma_\epsilon w_{TS})^2 + \sum_{t=2}^{T}\frac{1}{2\sigma_\epsilon^2}(c_t-\mu_1-\mu_2 t-\rho c_{t-1}-\sigma_\epsilon w_{TS})^2\right].\end{aligned} \tag{A8}$$

The detection error probability when model A generates log consumption c_t is obtained by substituting $c_1-\zeta-\mu=\sigma_\epsilon\epsilon_1$ for $t=1$ and $c_t-\mu_1-\mu_2 t-\rho c_{t-1}=\sigma_\epsilon\epsilon_t$ for $t=2,\ldots,T$:

$$\begin{aligned}p_A &= \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right)<0\right), \\ &= \text{Prob}\left(-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{2\sigma_\epsilon^2}(\sigma_\epsilon\epsilon_t)^2 + \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2\sigma_\epsilon^2}[\sigma_\epsilon(\epsilon_t-w_{TS})]^2<0\right), \\ &= \text{Prob}\left(\frac{1}{T}\sum_{t=1}^{T}(-w_{TS}\epsilon_t)+\frac{1}{2}w_{TS}^2<0\right), \\ &= \text{Prob}\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t<-\frac{1}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right), \\ &= \text{Prob}\left(Z<-\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right),\end{aligned} \tag{A9}$$

where $Z\equiv(1/\sqrt{T})\sum_{t=1}^{T}\epsilon_t$. On the other hand, the detection error probability when model B generates log consumption c_t is obtained by substituting $c_1-\zeta-\mu=\sigma_\epsilon w_{TS}+\sigma_\epsilon\epsilon_1$ for $t=1$ and $c_t-\mu_1-\mu_2 t-\rho c_{t-1}=\sigma_\epsilon w_{TS}+\sigma_\epsilon\epsilon_t$ for $t=2,\ldots,T$:

$$\begin{aligned}p_B &= \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right)>0\right), \\ &= \text{Prob}\left(-\frac{1}{T}\sum_{t=1}^{T}\frac{1}{2\sigma_\epsilon^2}[\sigma_\epsilon(\epsilon_t+w_{TS})]^2 + \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2\sigma_\epsilon^2}(\sigma_\epsilon\epsilon_t)^2>0\right), \\ &= \text{Prob}\left(\frac{1}{T}\sum_{t=1}^{T}(-w_{TS}\epsilon_t)-\frac{1}{2}w_{TS}^2>0\right), \\ &= \text{Prob}\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t>\frac{1}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right), \\ &= \text{Prob}\left(Z>\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right).\end{aligned} \tag{A10}$$

Because $\epsilon_t \sim$ i.i.d.$N(0,1)$, $(1/T)\sum_{t=1}^{T}\epsilon_t \sim N(0,1/T)$, so that $Z \sim N(0,1)$. Thus, using the standard normal cumulative distribution function $\Phi(\cdot)$, (A9) and (A10) can be written as

$$p_A = \Phi\left(-\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right) \quad \text{and} \quad p_B = 1-\Phi\left(\frac{\sqrt{T}}{2}\frac{\sigma_\epsilon}{\theta(1-\rho\beta)}\right). \tag{A11}$$

From the symmetry of the standard normal distribution, it follows that $p_A = p_B$, so that $p(\theta^{-1}) \equiv \frac{1}{2}(p_A + p_B) = p_A$.

B. Proof of Proposition 2

To prove this proposition, we must specify the worst-case model (model B) for the long-run risk model (model A). This requires two steps: first, the derivation of the value function, and second, the derivation of the distorted density. Before proceeding, we need to introduce some pieces of notation, following Hansen and Sargent (2008) and Barillas et al. (2009). Let $\pi(\epsilon_t)$ be conditional densities of a sequence of random shocks $\{\epsilon_t\}$, and let $\hat{\pi}(\epsilon_t)$ be some other density in proximity to $\pi(\epsilon_t)$, which we call the distorted density. Consider the following value function recursion:

$$U_t = c_t - \beta\theta \ln\left(E_t\left[\exp\left(\frac{-U_{t+1}}{\theta}\right)\right]\right). \tag{B1}$$

The first step is to solve for U_t under the long-run risk model. Guess the value function to be $U_t = k_0 + k_1 c_t + k_2 z_t$. Using equation (11), the value function at $t+1$ is

$$U_{t+1} = k_0 + k_1(\mu + c_t) + (k_1 + k_2\rho)z_t + (k_1\sigma_\epsilon + k_2\sigma_z)\epsilon_{t+1}. \tag{B2}$$

Substitute (B2) into $E_t[\exp(-U_{t+1}/\theta)]$ to obtain

$$\begin{aligned} E_t\left[\exp\left(\frac{-U_{t+1}}{\theta}\right)\right] &= \exp\left(\frac{-\{k_0 + k_1(\mu + c_t) + (k_1 + k_2\rho)z_t\}}{\theta}\right) \\ &\quad \times E_t\left[\exp\left(\frac{-(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}\epsilon_{t+1}\right)\right]. \end{aligned} \tag{B3}$$

Then take logs of both sides of (B3):

$$\begin{aligned} \ln\left(E_t\left[\exp\left(\frac{-U_{t+1}}{\theta}\right)\right]\right) &= \frac{-\{k_0 + k_1(\mu + c_t) + (k_1 + k_2\rho)z_t\}}{\theta} \\ &\quad + \ln\left(E_t\left[\exp\left(\frac{-(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}\epsilon_{t+1}\right)\right]\right). \end{aligned} \tag{B4}$$

Thus, recursion (B1) is

$$U_t = c_t + \beta\{k_0 + k_1(\mu + c_t) + (k_1 + k_2\rho)z_t\} - \beta\theta \ln\left(E_t\left[\exp\left(\frac{-(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}\epsilon_{t+1}\right)\right]\right). \quad \text{(B5)}$$

Using the properties of the lognormal distribution, (B5) can be further rewritten as

$$\begin{aligned} U_t &= c_t + \beta\{k_0 + k_1(\mu + c_t) + (k_1 + k_2\rho)z_t\} - \beta\theta\frac{(k_1\sigma_\epsilon + k_2\sigma_z)^2}{2\theta^2}, \\ &= \beta(k_0 + k_1\mu) - \beta\frac{(k_1\sigma_\epsilon + k_2\sigma_z)^2}{2\theta} + (1 + \beta k_1)c_t + \beta(k_1 + k_2\rho)z_t. \end{aligned} \quad \text{(B6)}$$

Matching the coefficients in $U_t = k_0 + k_1 c_t + k_2 z_t$, we obtain

$$k_0 = \frac{1}{1-\beta}\left[\frac{\beta\mu}{1-\beta} - \frac{\beta}{2\theta}\left(\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right)^2\right], k_1 = \frac{1}{1-\beta}, k_2 = \frac{\beta}{(1-\beta\rho)(1-\beta)}. \quad \text{(B7)}$$

The second step is to derive the distorted density $\hat{\pi}(\epsilon_t)$. To do this, we use the following result, as shown by Hansen and Sargent (2008) and Barillas et al. (2009).

$$g(\epsilon_{t+1}) \equiv \frac{\exp\left(\frac{-U_{t+1}}{\theta}\right)}{E_t\left[\exp\left(\frac{-U_{t+1}}{\theta}\right)\right]} = \frac{\hat{\pi}(\epsilon_{t+1})}{\pi(\epsilon_{t+1})}. \quad \text{(B8)}$$

Using (B2) and (B3), $g(\epsilon_{t+1})$ can be written as

$$g(\epsilon_{t+1}) = \frac{\exp\left(\frac{-(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}\epsilon_{t+1}\right)}{E_t\left[\exp\left(\frac{-(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}\epsilon_{t+1}\right)\right]}. \quad \text{(B9)}$$

Noting that the denominator of (B9) is equal to $\exp((k_1\sigma_\epsilon + k_2\sigma_z)^2/2\theta^2)$ by the properties of the lognormal distribution, (B9) can be rewritten as

$$\begin{aligned} g(\epsilon_{t+1}) &= \exp\left(\frac{-(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}\epsilon_{t+1} - \frac{1}{2}\frac{(k_1\sigma_\epsilon + k_2\sigma_z)^2}{\theta^2}\right), \\ &= \exp\left(w_{LR}\epsilon_{t+1} - \frac{1}{2}w_{LR}^2\right), \end{aligned} \quad \text{(B10)}$$

where $w_{LR} \equiv -(k_1\sigma_\epsilon + k_2\sigma_z)/\theta$. Because $g(\epsilon_{t+1}) = \hat{\pi}(\epsilon_{t+1})/\pi(\epsilon_{t+1})$ and $\pi(\epsilon_{t+1})$ denotes the density of the standard normal random variable ϵ_{t+1}, the distorted density $\hat{\pi}(\epsilon_{t+1})$ is

$$\begin{aligned} \hat{\pi}(\epsilon_{t+1}) &= \pi(\epsilon_{t+1})\exp\left(w_{LR}\epsilon_{t+1} - \frac{1}{2}w_{LR}^2\right), \\ &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\epsilon_{t+1}^2\right)\exp\left(w_{LR}\epsilon_{t+1} - \frac{1}{2}w_{LR}^2\right), \\ &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(\epsilon_{t+1} - w_{LR})^2\right), \end{aligned} \quad \text{(B11)}$$

which implies that $\epsilon_{t+1} \sim$ i.i.d.$N(w_{LR}, 1)$. Using (B7), the mean w_{LR} is

$$\begin{aligned} w_{LR} &\equiv -\frac{(k_1\sigma_\epsilon + k_2\sigma_z)}{\theta}, \\ &= -\frac{1}{\theta}\left(\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right). \end{aligned} \tag{B12}$$

Now we turn to the proof of formulas (12) and (13) for the detection error probabilities. Rewrite the long-run risk model (model A) in the following form by substituting the first equation into the second equation in (11) and shifting time by one period:

$$\Delta c_{t+1} = \mu^* + \rho\Delta c_t + \epsilon^*_{t+1} + \psi^*\epsilon^*_t, \tag{B13}$$

where $\mu^* \equiv (1-\rho)\mu$, $\psi^* \equiv \sigma_z/\sigma_\epsilon - \rho$, and $\epsilon^*_t \equiv \sigma_\epsilon\epsilon_t \sim$ i.i.d.$N(0, \sigma^2_\epsilon)$. Then, the worst-case model (model B) is given by

$$\Delta c_{t+1} = \mu^* + \alpha^* w_{LR} + \rho\Delta c_t + \epsilon^*_{t+1} + \psi^*\epsilon^*_t, \tag{B14}$$

where $\alpha^* \equiv \sigma_\epsilon(1+\psi^*)$. This is because the constant term of model A shifts by $\alpha^* w_{LR}$ with the change of the mean of ϵ_t from 0 to w_{LR}.[2] Given that model A and model B take the form of an ARIMA(1,1,1) process, we can form the (conditional) likelihood function provided that Δc_1 and $\epsilon^*_1 = 0$ are taken as given (see Hamilton (1994, Ch. 5)).

Taking Δc_1 and $\epsilon^*_1 = 0$ as given, (B13) at $t = 1$ is $\Delta c_2 = \mu^* + \rho\Delta c_1 + \epsilon^*_2$, so that $\Delta c_2 | (\Delta c_1, \epsilon^*_1 = 0) \sim N(\mu^* + \rho\Delta c_1, \sigma^2_\epsilon)$. For $t = 2, \ldots, T-1$, it follows that $\Delta c_{t+1} | (\Delta c_t, \ldots, \Delta c_1, \epsilon^*_1 = 0) \sim N(\mu^* + \rho\Delta c_t + \psi^*\epsilon^*_t, \sigma^2_\epsilon)$. Thus, the (conditional) likelihood function is given by

$$\begin{aligned} & f(\Delta c_T, \ldots, \Delta c_2 | \Delta c_1, \epsilon^*_1 = 0) \\ &= f(\Delta c_2 | \Delta c_1, \epsilon^*_1 = 0) \times \prod_{t=2}^{T-1} f(\Delta c_{t+1} | \Delta c_t, \ldots, \Delta c_1, \epsilon^*_1 = 0) \\ &= \frac{1}{\sqrt{2\pi\sigma^2_\epsilon}} \exp\left[-\frac{1}{2\sigma^2_\epsilon}(\Delta c_2 - \mu^* - \rho\Delta c_1)^2\right] \\ &\qquad \times \prod_{t=2}^{T-1} \frac{1}{\sqrt{2\pi\sigma^2_\epsilon}} \exp\left[-\frac{1}{2\sigma^2_\epsilon}(\Delta c_{t+1} - \mu^* - \rho\Delta c_t - \psi^*\epsilon^*_t)^2\right]. \end{aligned} \tag{B15}$$

[2] Consider $\epsilon_t + w_{LR}$, where $\epsilon_t \sim$ i.i.d.$N(0, 1)$. Replacing the error term ϵ_t of model A by this, we have

$$\begin{aligned} \Delta c_{t+1} &= \mu^* + \rho\Delta c_t + \sigma_\epsilon(\epsilon_{t+1} + w_{LR}) + \psi^*\sigma_\epsilon(\epsilon_t + w_{LR}), \\ &= \mu^* + \sigma_\epsilon(1+\psi^*)w_{LR} + \rho\Delta c_t + \sigma_\epsilon\epsilon_{t+1} + \psi^*\sigma_\epsilon\epsilon_t. \end{aligned}$$

Define the (average) log-likelihood function as

$$
\begin{aligned}
\ln L &\equiv \frac{1}{T-1}\ln f(\Delta c_T,\ldots,\Delta c_2|\Delta c_1,\epsilon_1^*=0),\\
&= \frac{1}{T-1}\ln f(\Delta c_2|\Delta c_1,\epsilon_1^*=0)+\frac{1}{T-1}\sum_{t=2}^{T-1}\ln f(\Delta c_{t+1}|\Delta c_t,\ldots,\Delta c_1,\epsilon_1^*=0).
\end{aligned}
\tag{B16}
$$

Then the log-likelihood function under model A is

$$
\begin{aligned}
\ln L_A &= -\frac{1}{2}\ln 2\pi-\frac{1}{2}\ln\sigma_\epsilon^2-\frac{1}{T-1}\frac{1}{2\sigma_\epsilon^2}(\Delta c_2-\mu^*-\rho\Delta c_1)^2\\
&\quad -\frac{1}{T-1}\sum_{t=2}^{T-1}\frac{1}{2\sigma_\epsilon^2}(\Delta c_{t+1}-\mu^*-\rho\Delta c_t-\psi^*\epsilon_t^*)^2.
\end{aligned}
\tag{B17}
$$

Noting that the constant term shifts by $\alpha^* w_{LR}$ as in equation (B14), the log-likelihood function under model B is

$$
\begin{aligned}
\ln L_B &= -\frac{1}{2}\ln 2\pi-\frac{1}{2}\ln\sigma_\epsilon^2-\frac{1}{T-1}\frac{1}{2\sigma_\epsilon^2}(\Delta c_2-\mu^*-\alpha^* w_{LR}-\rho\Delta c_1)^2\\
&\quad -\frac{1}{T-1}\sum_{t=2}^{T-1}\frac{1}{2\sigma_\epsilon^2}(\Delta c_{t+1}-\mu^*-\alpha^* w_{LR}-\rho\Delta c_t-\psi^*\epsilon_t^*)^2.
\end{aligned}
\tag{B18}
$$

Thus, the log-likelihood ratio is given by

$$
\begin{aligned}
\ln\left(\frac{L_A}{L_B}\right) &= -\frac{1}{T-1}\left[\frac{1}{2\sigma_\epsilon^2}(\Delta c_2-\mu^*-\rho\Delta c_1)^2+\sum_{t=2}^{T-1}\frac{1}{2\sigma_\epsilon^2}(\Delta c_{t+1}-\mu^*-\rho\Delta c_t-\psi^*\epsilon_t^*)^2\right]\\
&+\frac{1}{T-1}\left[\frac{1}{2\sigma_\epsilon^2}(\Delta c_2-\mu^*-\alpha^* w_{LR}-\rho\Delta c_1)^2+\sum_{t=2}^{T-1}\frac{1}{2\sigma_\epsilon^2}(\Delta c_{t+1}-\mu^*-\alpha^* w_{LR}-\rho\Delta c_t-\psi^*\epsilon_t^*)^2\right].
\end{aligned}
\tag{B19}
$$

Note that $\epsilon_{t+1}^*=\sigma_\epsilon\epsilon_{t+1}$ and $\alpha^*=\sigma_\epsilon\{\sigma_z/\sigma_\epsilon+(1-\rho)\}$. Substituting $\Delta c_2-\mu^*-\rho\Delta c_1=\epsilon_2^*$ and $\Delta c_{t+1}-\mu^*-\rho\Delta c_t-\psi^*\epsilon_t^*=\epsilon_{t+1}^*$ for $t=2,\ldots,T-1$ into (B19), the log-likelihood ratio under model A is

$$
\begin{aligned}
\ln\left(\frac{L_A}{L_B}\right) &= -\frac{1}{T-1}\sum_{t=1}^{T-1}\frac{1}{2\sigma_\epsilon^2}\epsilon_{t+1}^{*2}+\frac{1}{T-1}\sum_{t=1}^{T-1}\frac{1}{2\sigma_\epsilon^2}(\epsilon_{t+1}^*-\alpha^* w_{LR})^2,\\
&= \frac{1}{T-1}\sum_{t=1}^{T-1}\left[-\frac{1}{\sigma_\epsilon^2}\alpha^* w_{LR}\epsilon_{t+1}^*+\frac{1}{2\sigma_\epsilon^2}\alpha^{*2}w_{LR}^2\right],\\
&= \frac{1}{T-1}\sum_{t=1}^{T-1}-\left\{\frac{\sigma_z}{\sigma_\epsilon}+(1-\rho)\right\}w_{LR}\epsilon_{t+1}+\frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon}+(1-\rho)\right\}^2 w_{LR}^2.
\end{aligned}
\tag{B20}
$$

Alternatively, substituting $\Delta c_2-\mu^*-\rho\Delta c_1=\epsilon_2^*+\alpha^* w_{LR}$ and $\Delta c_{t+1}-\mu^*-\rho\Delta c_t-\psi^*\epsilon_t^*=$

$\epsilon^*_{t+1} + \alpha^* w_{LR}$ for $t = 2, \ldots, T-1$ into (B19), the log-likelihood ratio under model B is

$$\begin{aligned}
\ln\left(\frac{L_A}{L_B}\right) &= -\frac{1}{T-1}\sum_{t=1}^{T-1}\frac{1}{2\sigma_\epsilon^2}(\epsilon^*_{t+1} + \alpha^* w_{LR})^2 + \frac{1}{T-1}\sum_{t=1}^{T-1}\frac{1}{2\sigma_\epsilon^2}\epsilon^{*2}_{t+1}, \\
&= \frac{1}{T-1}\sum_{t=1}^{T-1}\left[-\frac{1}{\sigma_\epsilon^2}\alpha^* w_{LR}\epsilon^*_{t+1} - \frac{1}{2\sigma_\epsilon^2}\alpha^{*2}w_{LR}^2\right], \qquad \text{(B21)} \\
&= \frac{1}{T-1}\sum_{t=1}^{T-1} - \left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\} w_{LR}\epsilon_{t+1} - \frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}^2 w_{LR}^2.
\end{aligned}$$

Using (B20) and (B21), the detection error probabilities under model A and model B are, respectively,

$$\begin{aligned}
p_A &= \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right) < 0\right), \\
&= \text{Prob}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}(-w_{LR}\epsilon_{t+1}) + \frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\} w_{LR}^2 < 0\right), \\
&= \text{Prob}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\epsilon_{t+1} < -\frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{k_1\sigma_\epsilon + k_2\sigma_z}{\theta}\right), \\
&= \text{Prob}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\epsilon_{t+1} < -\frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{1}{\theta}\left\{\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right\}\right), \\
&= \text{Prob}\left(Z < -\frac{\sqrt{T-1}}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{1}{\theta}\left\{\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right\}\right),
\end{aligned} \qquad \text{(B22)}$$

and

$$\begin{aligned}
p_B &= \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right) > 0\right), \\
&= \text{Prob}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}(-w_{LR}\epsilon_{t+1}) - \frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\} w_{LR}^2 > 0\right), \\
&= \text{Prob}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\epsilon_{t+1} > \frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{k_1\sigma_\epsilon + k_2\sigma_z}{\theta}\right), \\
&= \text{Prob}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\epsilon_{t+1} > \frac{1}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{1}{\theta}\left\{\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right\}\right), \\
&= \text{Prob}\left(Z > \frac{\sqrt{T-1}}{2}\left\{\frac{\sigma_z}{\sigma_\epsilon} + (1-\rho)\right\}\frac{1}{\theta}\left\{\frac{1}{1-\beta}\sigma_\epsilon + \frac{\beta}{(1-\beta\rho)(1-\beta)}\sigma_z\right\}\right),
\end{aligned} \qquad \text{(B23)}$$

where $Z \equiv (1/\sqrt{T-1})\sum_{t=1}^{T-1}\epsilon_{t+1} \sim N(0,1)$. These give equations (12) and (13).

C. Proof of Proposition 3

To prove this proposition, we need to specify the worst-case model (model B). This requires two steps, as in the proof of Proposition 2. Guess the value function to be $U_t = k_0 + k_1 c_t^i$. Note that equation (15) can be written as

$$\begin{aligned} c_{t+1}^i &= c_t^i + \sqrt{\epsilon}(w_{1t+1} + w_{2t+1}), \\ &= c_t^i + \epsilon_{t+1}, \end{aligned} \tag{C1}$$

where $\epsilon_{t+1} \equiv \sqrt{\epsilon}(w_{1t+1} + w_{2t+1})$. Then $U_{t+1} = k_0 + k_1 c_t^i + k_1 \epsilon_{t+1}$. Substitute this into equation (14) to obtain

$$U_t = \beta k_0 + (1 + \beta k_1) c_t^i - \beta\theta \ln\left(E_t\left[\exp\left(-\frac{1}{\theta}k_1\epsilon_{t+1}\right)\right]\right). \tag{C2}$$

Using the property of the lognormal distribution, it can be verified that

$$\ln\left(E_t\left[\exp\left(-\frac{1}{\theta}k_1\epsilon_{t+1}\right)\right]\right) = -\frac{k_1}{\theta}\sqrt{\epsilon}\left(g - \frac{\tau_1^2 + \tau_2^2}{2}\right) + \frac{k_1^2}{\theta^2}\epsilon\frac{\tau_1^2 + \tau_2^2}{2}. \tag{C3}$$

Matching the coefficients in $U_t = k_0 + k_1 c_t^i$ after substituting (C3) into (C2), we obtain

$$k_0 = \frac{\beta}{(1-\beta)^2}\left[\sqrt{\epsilon}\left(g - \frac{\tau_1^2 + \tau_2^2}{2}\right) - \frac{1}{\theta(1-\beta)}\frac{\epsilon(\tau_1^2 + \tau_2^2)}{2}\right],\ k_1 = \frac{1}{1-\beta}. \tag{C4}$$

To derive the distorted density $\hat{\pi}(\epsilon_{t+1})$, we use (B8) again. Note that

$$\exp\left(-\frac{U_{t+1}}{\theta}\right) = \exp\left(-\frac{1}{\theta}(k_0 + k_1 c_t^i)\right)\exp\left(-\frac{k_1}{\theta}\epsilon_{t+1}\right). \tag{C5}$$

Using (C5) and the property of the lognormal distribution, we have

$$g(\epsilon_{t+1}) = \frac{\exp\left(-\frac{k_1}{\theta}\epsilon_{t+1}\right)}{E_t\left[\exp\left(-\frac{k_1}{\theta}\epsilon_{t+1}\right)\right]} = \exp\left(w_{IR}\epsilon_{t+1} - w_{IR}B - w_{IR}^2\frac{C}{2}\right), \tag{C6}$$

where

$$w_{IR} \equiv -\frac{k_1}{\theta},\ B \equiv \sqrt{\epsilon}\left(g - \frac{\tau_1^2 + \tau_2^2}{2}\right),\ C \equiv \epsilon(\tau_1^2 + \tau_2^2). \tag{C7}$$

Therefore, it follows from (B8) that the ratio of densities is

$$\frac{\hat{\pi}(\epsilon_{t+1})}{\pi(\epsilon_{t+1})} = \exp\left(w_{IR}\epsilon_{t+1} - w_{IR}B - w_{IR}^2\frac{C}{2}\right). \tag{C8}$$

Here $\epsilon_{t+1} \sim$ i.i.d.$N(B, C)$, so that the density $\pi(\epsilon_{t+1})$ is

$$\pi(\epsilon_{t+1}) = \frac{1}{\sqrt{2\pi C}} \exp\left[-\frac{1}{2C}\left(\epsilon_{t+1} - B\right)^2\right]. \tag{C9}$$

Substituting (C9) into (C8) and rearranging terms, we obtain

$$\begin{aligned} \hat{\pi}(\epsilon_{t+1}) &= \pi(\epsilon_{t+1}) \exp\left(w_{IR}\epsilon_{t+1} - w_{IR}B - w_{IR}^2 \frac{C}{2}\right), \\ &= \frac{1}{\sqrt{2\pi C}} \exp\left[-\frac{1}{2C}\left(\epsilon_{t+1} - (B + w_{IR}C)\right)^2\right]. \end{aligned} \tag{C10}$$

Thus, the approximating model (model A) is

$$c_{t+1}^i = c_t^i + \epsilon_{t+1},\ \epsilon_{t+1} \sim \text{i.i.d.} N(B, C). \tag{C11}$$

The worst-case model (model B) is given by

$$c_{t+1}^i = c_t^i + w_{IR}C + \epsilon_{t+1},\ \epsilon_{t+1} \sim \text{i.i.d.} N(B, C). \tag{C12}$$

Now we turn to the derivation of the formulas for the detection error probabilities. From $E(\Delta c_{t+1}^i) = B$ and $\text{Var}(\Delta c_{t+1}^i) = C$ under model A, it follows that the log-likelihood function under model A is

$$\ln L_A = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln C - \frac{1}{T}\sum_{t=0}^{T-1} \frac{1}{2C}(\Delta c_{t+1}^i - B)^2. \tag{C13}$$

From $E(\Delta c_{t+1}^i) = w_{IR}C + B$ and $\text{Var}(\Delta c_{t+1}^i) = C$ under model B, the log-likelihood function under model B is

$$\ln L_B = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln C - \frac{1}{T}\sum_{t=0}^{T-1} \frac{1}{2C}(\Delta c_{t+1}^i - (B + w_{IR}C))^2. \tag{C14}$$

Thus, the log-likelihood ratio is

$$\ln\left(\frac{L_A}{L_B}\right) = \frac{1}{T}\sum_{t=0}^{T-1}\left[-\frac{1}{2C}(\Delta c_{t+1}^i - B)^2 + \frac{1}{2C}(\Delta c_{t+1}^i - (B + w_{IR}C))^2\right]. \tag{C15}$$

Substituting $\Delta c_{t+1}^i = \epsilon_{t+1}$ into (C15), we have

$$\begin{aligned} \ln\left(\frac{L_A}{L_B}\right) &= \frac{1}{T}\sum_{t=0}^{T-1}\left[-\frac{1}{2C}(\epsilon_{t+1} - B)^2 + \frac{1}{2C}(\epsilon_{t+1} - (B + w_{IR}C))^2\right], \\ &= \frac{1}{T}\sum_{t=0}^{T-1}(-w_{IR}\epsilon_{t+1}) + w_{IR}B + \frac{1}{2}w_{IR}^2 C. \end{aligned} \tag{C16}$$

Therefore, the detection error probability under model A is

$$\begin{aligned} p_A &= \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right) < 0\right), \\ &= \text{Prob}\left(\frac{1}{T}\sum_{t=0}^{T-1}\epsilon_{t+1} < B - \frac{k_1}{2\theta}C\right), \\ &= \text{Prob}\left(Z < -\frac{\sqrt{T}k_1}{2\theta}\sqrt{C}\right), \\ &= \text{Prob}\left(Z < -\frac{\sqrt{T}}{2}\frac{\sqrt{\epsilon(\tau_1^2+\tau_2^2)}}{\theta(1-\beta)}\right), \end{aligned} \tag{C17}$$

where $Z \equiv (\frac{1}{T}\sum_{t=0}^{T-1}\epsilon_{t+1} - B)/\sqrt{C/T} \sim N(0,1)$.

Alternatively, substituting $\Delta c_{t+1}^i = w_{IR}C + \epsilon_{t+1}$ into (C15), we have

$$\begin{aligned} \ln\left(\frac{L_A}{L_B}\right) &= \frac{1}{T}\sum_{t=0}^{T-1}\left[-\frac{1}{2C}(\epsilon_{t+1} - (B - w_{IR}C))^2 + \frac{1}{2C}(\epsilon_{t+1} - B)^2\right], \\ &= \frac{1}{T}\sum_{t=0}^{T-1}(-w_{IR}\epsilon_{t+1}) + w_{IR}B - \frac{1}{2}w_{IR}^2C. \end{aligned} \tag{C18}$$

Therefore, the detection error probability under model B is

$$\begin{aligned} p_B &= \text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right) > 0\right), \\ &= \text{Prob}\left(\frac{1}{T}\sum_{t=0}^{T-1}\epsilon_{t+1} > B + \frac{k_1}{2\theta}C\right), \\ &= \text{Prob}\left(Z > \frac{\sqrt{T}k_1}{2\theta}\sqrt{C}\right), \\ &= \text{Prob}\left(Z > \frac{\sqrt{T}}{2}\frac{\sqrt{\epsilon(\tau_1^2+\tau_2^2)}}{\theta(1-\beta)}\right). \end{aligned} \tag{C19}$$

These give equation (16).

D. Calculation of Standard Errors

D.1 Random-Walk Case

The overall detection error probability $p(\theta^{-1})$ for the random-walk case can be written as a function of σ_ϵ:

$$g(\sigma_\epsilon) = \Phi\left(-\frac{\sqrt{T}}{2}(\gamma-1)\sigma_\epsilon\right). \tag{D1}$$

Let $\hat{\sigma}_\epsilon$ be the maximum likelihood (ML) estimator of σ_ϵ and let $\sigma_{\epsilon 0}$ be its true value. Applying the univariate delta method, we obtain

$$\sqrt{T}\left(g(\hat{\sigma}_\epsilon) - g(\sigma_{\epsilon 0})\right) \to_d N\left(0, \{g'(\sigma_{\epsilon 0})\}^2\text{Var}(\sigma_{\epsilon 0})\right). \tag{D2}$$

The standard error for $p(\widehat{\theta^{-1}})$ is therefore given by

$$
\begin{aligned}
se\left(p(\widehat{\theta^{-1}})\right) &= se\left(g(\hat{\sigma}_\epsilon)\right), \\
&= \sqrt{\frac{1}{T}\{g'\left(\hat{\sigma}_\epsilon\right)\}^2 \mathrm{Var}(\hat{\sigma}_\epsilon)},
\end{aligned} \tag{D3}
$$

where

$$
g'(\hat{\sigma}_\epsilon) = -\frac{\sqrt{T}}{2}(\gamma-1)\cdot\Phi'\left(-\frac{\sqrt{T}}{2}(\gamma-1)\hat{\sigma}_\epsilon\right). \tag{D4}
$$

D.2 Trend-Stationary Case

The overall detection error probability $p(\theta^{-1})$ for the trend-stationary case can be regarded as a function of $\boldsymbol{\theta} \equiv (\rho, \sigma_\epsilon)'$:

$$
g(\boldsymbol{\theta}) = \Phi\left(-\frac{\sqrt{T}}{2}(1-\beta)(\gamma-1)\frac{\sigma_\epsilon}{1-\beta\rho}\right) \tag{D5}
$$

Let $\hat{\boldsymbol{\theta}}$ be the ML estimator of $\boldsymbol{\theta}$ and let $\boldsymbol{\theta}_0$ be its true value. Let $G(\boldsymbol{\theta}) \equiv \partial g(\boldsymbol{\theta})/\partial\boldsymbol{\theta}'$. Applying the multivariate delta method, we obtain

$$
\sqrt{T}\left(g(\hat{\boldsymbol{\theta}}) - g(\boldsymbol{\theta}_0)\right) \to_d N\left(0, G(\boldsymbol{\theta}_0)\Omega_0 G(\boldsymbol{\theta}_0)'\right), \tag{D6}
$$

where

$$
\Omega_0 \equiv \begin{bmatrix} \mathrm{Var}(\rho_0) & 0 \\ 0 & \mathrm{Var}(\sigma_{\epsilon 0}) \end{bmatrix}. \tag{D7}
$$

The standard error for $p(\theta^{-1})$ is therefore given by

$$
\begin{aligned}
se\left(p(\widehat{\theta^{-1}})\right) &= se\left(g(\hat{\boldsymbol{\theta}})\right), \\
&= \sqrt{\frac{1}{T}\left(\hat{g}_1^2 \mathrm{Var}(\hat{\rho}) + \hat{g}_2^2 \mathrm{Var}(\hat{\sigma}_\epsilon)\right)}, \\
&= \sqrt{\hat{g}_1^2\left(se(\hat{\rho})\right)^2 + \hat{g}_2^2\left(se(\hat{\sigma}_\epsilon)\right)^2},
\end{aligned} \tag{D8}
$$

where

$$
\hat{g}_1 \equiv \frac{\partial g(\hat{\boldsymbol{\theta}})}{\partial\rho} = -\frac{\sqrt{T}}{2}(1-\beta)(\gamma-1)\frac{\beta}{(1-\beta\hat{\rho})^2}\hat{\sigma}_\epsilon\cdot\Phi'\left(-\frac{\sqrt{T}}{2}(1-\beta)(\gamma-1)\frac{\hat{\sigma}_\epsilon}{1-\beta\hat{\rho}}\right), \tag{D9}
$$

$$
\hat{g}_2 \equiv \frac{\partial g(\hat{\boldsymbol{\theta}})}{\partial\sigma_\epsilon} = -\frac{\sqrt{T}}{2}(1-\beta)(\gamma-1)\frac{1}{1-\beta\hat{\rho}}\cdot\Phi'\left(-\frac{\sqrt{T}}{2}(1-\beta)(\gamma-1)\frac{\hat{\sigma}_\epsilon}{1-\beta\hat{\rho}}\right). \tag{D10}
$$

E. Graphs of the Overall Detection Error Probability

Figure 1 plots the overall detection error probabilities against various values of θ^{-1} for the random walk (solid line) and trend-stationary (dashed line) models using formulas (9) and (10), in order to confirm the validity of our results based on the cumulative distribution function. This figure is consistent with Figure 2(a) in Barillas et al. (2009, p. 2406).

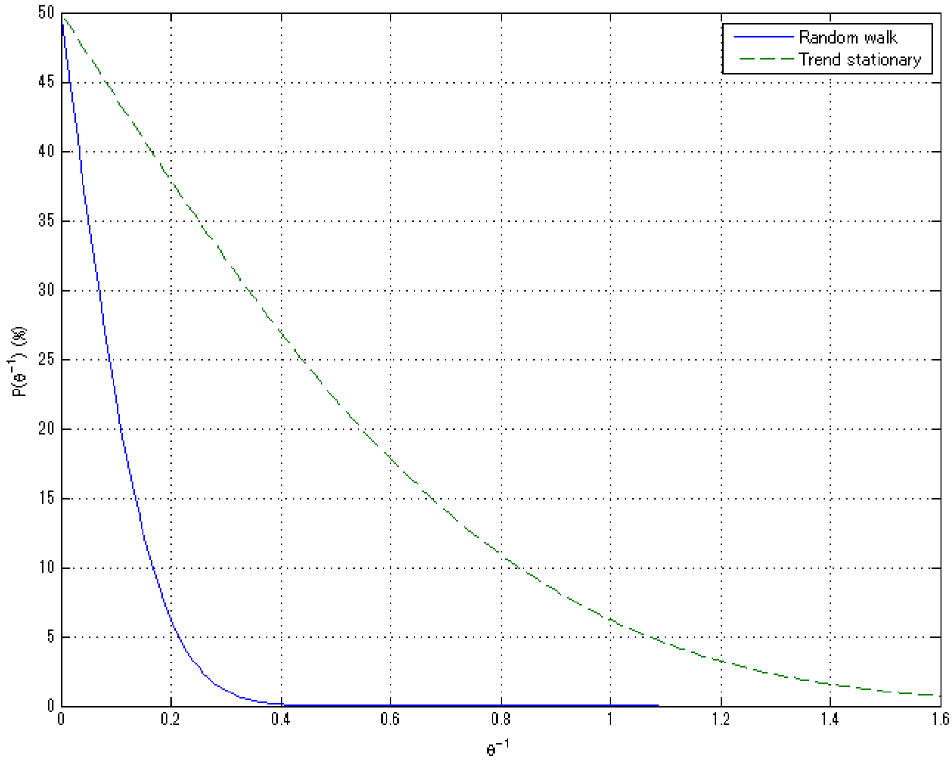

Figure 1: Detection Error Probability versus the Inverse of the Penalty Parameter

References

Barillas, Francisco, Lars Peter Hansen, and Thomas J. Sargent (2009), "Doubts or Variability?," *Journal of Economic Theory* 144, 2388–2418.

Hamilton, James D. (1994) *Time Series Analysis*, Princeton, NJ: Princeton University Press.

Hansen, Lars Peter, and Thomas J. Sargent (2008), *Robustness*, Princeton, NJ: Princeton University Press.