Matching a Ramsey Model to a Descriptive Statistical Model*

Thomas J. Sargent
New York University

Ziyue Yang
Australian National University

January 1, 2026

Abstract

This paper uses a two-stage specification search to study growth miracles in China, Singapore, and South Korea. The first stage estimates a scalar additive-functional model that decomposes per capita GDP into deterministic trend, martingale, and stationary components. Bayesian posterior estimates indicate highly persistent transient growth gaps and substantial volatility. The second stage uses the method of simulated moments to calibrate a stochastic Ramsey growth model that replicates key moments of the estimated statistical model. Specification searches over the long-run growth parameter ν reveal that fixing $\nu = 0.02$ to match US growth systematically understates high-growth outcomes for China, while estimating ν separately for each economy improves the fit of both statistical and structural models. Our analysis illustrates how priors and likelihoods jointly determine posterior distributions and inferences about parameters.

1 Introduction

This paper presents statistical representations of three "growth miracles" that are outcomes of "specification searches" in the spirit of Leamer (1978). A statistical model is a joint probability distribution for a vector of random variables of interest that is indexed by a vector of parameters. A manifold of statistical models is swept out as we vary the parameter vector within a set of possible values. In the Bayesian tradition, we will endow ourselves with a subjective prior probability distribution over the parameters, then view the observables as draws from the marginal joint probability distribution of the variables of interest. We can make inferences about unknown parameters by computing the probability distribution of the parameters conditional on the observed variables of interest, i.e., by computing a "posterior probability distribution".

The random variables that interest us in this paper are generated by a univariate stochastic process of per capita GDP growth rates. We have organized our specification search to reflect a distinction

*We thank Lawrence Kotlikoff for helpful comments on an earlier version.

between "description" and "explanation" cast in terms of whether a statistical model's parameters are "incompletely" or "fully" interpreted in terms of objects intelligible to an economic theorist. We thus follow Koopmans (1947) in distinguishing between, on the one hand, purely descriptive "Kepler stage" statistical models whose parameters are just useful data-compression devices that are not interpretable in terms of objects appearing in economic theories and, on the other hand, "structural" Newton-stage models whose parameters pin down the preferences, technologies, and information flows underlying an economic model. Section 2 describes the per capita GDP data that we aspire to describe in Section 3 and to explain in Section 4. Section 3 describes our descriptive "Kepler stage" model, an additive functional model of stochastic growth, while Section 4 describes our "Newton stage" structural model, a Brock and Mirman (1972) stochastic optimal growth model.[1]

Because statistically evaporating transient dynamics in per capital GDP are an important part of the "growth miracles" that we study, we use statistical models that don't satisfy workhorse assumptions of stationarity and ergodicity that underlie important properties of frequentist and Bayesian inferences about parameters. Consequently, we build non-stationarity into the class of statistical models within which we confine our specification searches and inferences about parameters. In doing this, we take advantage of the insights presented by Kohn and Ansley (1985, 1986), Hansen and Sargent (2013, ch. 8, app. A), and others.[2]

Fears of misspecification pervade an econometrician's specification searches and a decision maker's use of a statistical model. Section 5 briefly describes how some econometricians have expressed their concerns about misspecifications and organized responses to them. Section 6 offers concluding remarks.

2 Growth Facts

We study GDP growth in three East Asian economies from the 1960s to the 2020s: China, Singapore, and South Korea. For each economy i, let $\mathrm{GDPpc}_{i,\tau}$ denote real GDP per capita in year τ. We use real GDP per capita (2015 US dollars) from the World Bank.[3] Define the log level

$$\ell_{i,\tau} = \log\left(\mathrm{GDPpc}_{i,\tau}\right).$$

Figure 1 plots five-year moving averages of annual growth rates $\ell_{i,\tau} - \ell_{i,\tau-1}$ for the three East Asian economies, with a horizontal line showing the long-run US growth rate calculated as the average growth rate over 1961–2023.

[1] Kydland and Prescott (1982) calibrated a version of this model and called it a "real business cycle model".

[2] See `https://python.quantecon.org/ar1_bayes.html` for an elementary computational presentation.

[3] Specifically, we use the `NY.GDP.PCAP.KD` series from the World Bank's World Development Indicators database (`https://data.worldbank.org/`).

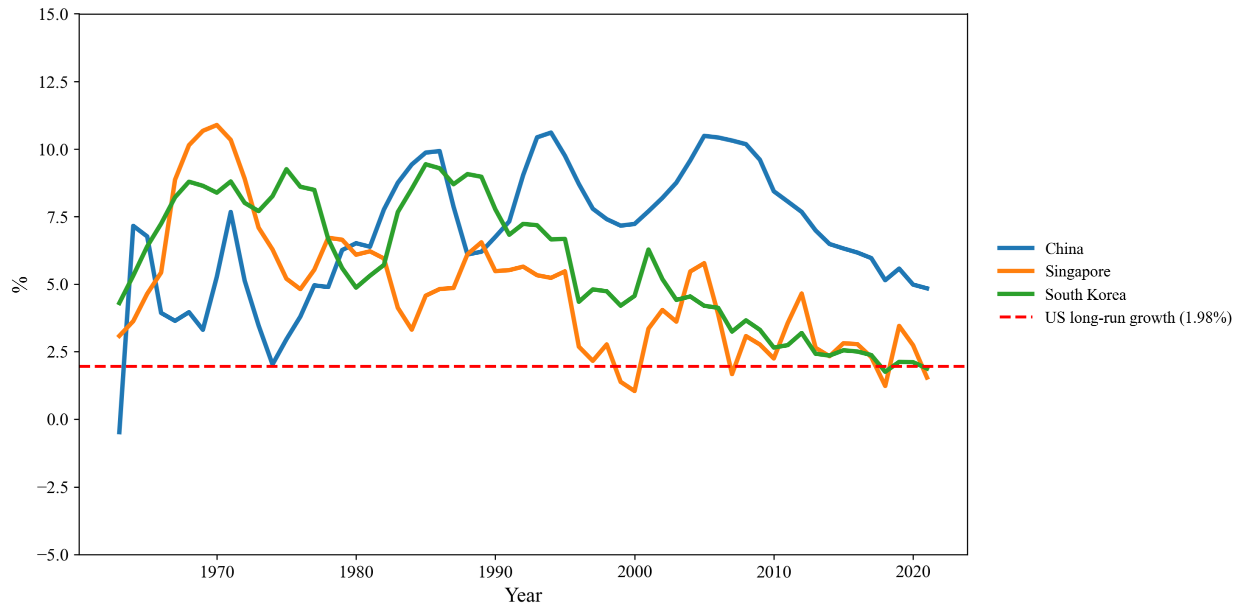

Figure 1: Per-capita GDP growth (5-year moving averages)

The East Asian economies experienced sustained high growth rates over multi-decade periods, far exceeding the long-run US growth rate but eventually converging towards it. In the following section, we develop a statistical model that captures these growth patterns.

3 A Scalar Additive-Functional Growth Model

For a chosen economy and a starting year τ^* (the "miracle start"), we define a normalized log level series. For notational convenience, we suppress the economy subscript i in what follows, with the understanding that all parameters and variables are economy-specific:

$$y_t := \ell_{\tau^*+t} - \ell_{\tau^*}, \qquad t = 0, 1, \ldots, T,$$

where T is the sample length, so $y_0 = 0$ and y_t is the cumulative log growth since the start of the episode. The one-period growth rate between t and $t+1$ is

$$\Delta y_{t+1} := y_{t+1} - y_t, \qquad t = 0, 1, \ldots, T-1.$$

The model links these growth rates to states and shocks via the equations

$$x_{t+1} = \varphi x_t + B z_{t+1}, \tag{1}$$

$$\Delta y_{t+1} = \nu + x_t + F z_{t+1}, \tag{2}$$

where x_t is a latent "growth gap" that captures how far growth is above or below the stationary long-run trend, φ is the persistence parameter with $0 < \varphi < 1$, and ν measures the long-run growth rate for the economy. $B > 0$ and $F > 0$ are volatility parameters governing state and growth innovations, respectively, and $z_{t+1} \sim \mathcal{N}(0, 1)$ are i.i.d. standard normal shocks. Conditional on the realized initial state x_0, the path $(x_t, \Delta y_{t+1})$ is determined by (φ, B, F, ν) and the innovations (z_{t+1}).

This model is deliberately modest: it has one state x_t and one shock z_{t+1}. Yet it is rich enough to model the stylized growth patterns observed in Figure 1. We discuss the estimation of this model using Bayesian methods below. In the baseline specification (Section 3.2), we fix $\nu = 0.02$ to match typical estimates of long-run per-capita GDP growth in the United States, and later we consider an alternative specification in which ν is estimated jointly with (φ, B, F, x_0) in Section 3.5.

3.1 Likelihood of Descriptive Model

The model's likelihood is derived from the conditional distribution of growth, Δy_{t+1}, which depends on a latent state, x_t, and parameters $\psi = (\varphi, B, F, \nu, x_0)$. From (2), this distribution is:

$$\Delta y_{t+1} \mid x_t, \psi \sim \mathcal{N}(\nu + x_t, F^2). \tag{3}$$

While the state evolves according to the structural shocks in (1), for estimation it is more convenient to use an alternative representation. We can express the state updates in terms of the observed innovations in growth. First, define the one-step-ahead forecast error as

$$\varepsilon_{t+1} \coloneqq \frac{\Delta y_{t+1} - (\nu + x_t)}{F}.$$

Then, the state recursion becomes

$$x_{t+1} = \varphi x_t + B\varepsilon_{t+1}.$$

This representation generates the same likelihood function as the structural model defined by (1) and (2). Its advantage is for MCMC estimation: given the parameters ψ, the path of latent states (x_t) is fully determined by the data (Δy_{t+1}) without requiring auxiliary sampling of the structural shocks. This avoids the need to sample the states as separate parameters.

3.2 Priors

In the baseline specification, we place independent priors on the parameters $\tilde{\psi} = (\varphi, B, F, x_0)$, with $\nu = 0.02$ fixed:

$$\begin{aligned} \varphi &\sim \mathrm{Beta}(\alpha_\varphi, \beta_\varphi), & (\alpha_\varphi, \beta_\varphi) &= (1, 1), \\ B &\sim \mathrm{HalfNormal}(\sigma_B), & \sigma_B &= 1, \\ F &\sim \mathrm{HalfNormal}(\sigma_F), & \sigma_F &= 1, \\ x_0 &\sim \mathcal{N}(\mu_{x_0}, \sigma_{x_0}^2), & (\mu_{x_0}, \sigma_{x_0}) &= (0, 0.5). \end{aligned} \tag{4}$$

The prior for φ is a flat prior on $(0, 1)$, thereby restricting attention to stationary growth gaps consistent with our observations in Figure 1. The priors for B and F are weakly informative half-normal priors, and the prior for x_0 centers the initial growth deviation at zero with a large standard deviation, which allows the data to inform the estimation. One alternative specification of x_0 fixes it such that $\nu + x_0$ matches the first observed growth rate in the sample; we discuss this alternative in Appendix A. Another alternative specification treats ν as an unknown parameter to be estimated using MCMC instead of being fixed at 0.02. We explore this specification in Section 3.5.

Another modeling choice concerns the starting year τ^* for each economy. The choice of τ^* determines the initial growth gap x_0, which in turn influences the estimation of (φ, B, F). For China, we choose $\tau^* = 1984$, since it marks the formal shift of reform and opening-up policies to urban and enterprise sectors documented in the "Decision on Reform of the Economic Structure" issued by the Third Plenum of the 12th Central Committee (Central Committee of the Communist Party of China, 1984; Xinhua News Agency, 1984). This year saw the expansion of special economic zones and the 1980s peak of real GDP growth. For Singapore, we choose $\tau^* = 1966$, the year after its independence from Malaysia. This year marked the start of Singapore as a sovereign nation, during which it implemented strategic economic policies that laid the foundation for its rapid growth and development. For South Korea, we choose $\tau^* = 1963$, the first year of Park Chung-hee's presidency, which initiated a series of ambitious five-year economic development plans that transformed South Korea's economy. This year also marked the beginning of a sustained period of rapid industrialization and export-led growth.

3.3 Posterior sampling

Posterior inference is carried out using the No-U-Turn Sampler (NUTS) (Hoffman et al., 2014), an adaptive variant of Hamiltonian Monte Carlo implemented in the `NumPyro` library (Phan et al., 2019).

For each of the three economies (China, Singapore, and South Korea), the scalar model (1)–(2) is estimated using the procedure described in Algorithm 1, holding $\nu = 0.02$ fixed. Posterior summaries

Algorithm 1: Run configuration for Bayesian estimation

Require:

Data: $(y_t)_{t=0}^T$ with $y_0 = 0$

Priors: (φ, B, F, x_0) as in (4), with $\nu = 0.02$ fixed

Sampler settings: number of chains = 4, warm-up iterations = 5000, post-warm-up samples per chain = 5000

1: For each economy and specification, form growth rates $\Delta y_{t+1} = y_{t+1} - y_t$.

2: Run NUTS with the specified priors and sampler settings.

3: Collect posterior draws and compute summaries for (φ, B, F, x_0).

(medians and 95% HDIs) for (φ, B, F, x_0) are reported in Table 1 and in the prior–posterior plots in Figure 2.

In Table 1, we see that the posterior medians of φ are all close to 1, indicating that growth gaps are highly persistent across the three economies. The posterior medians of F lie between roughly 0.025 (China) and 0.039 (Singapore), which is substantial relative to the long-run growth rate of 0.02. The reported values for x_0 range between about 6% and 9%, capturing the initial growth deviations above the long-run mean observed in Figure 1. Relative to the diffuse priors in Figure 2, the posterior distributions are much more concentrated for (φ, B, F, x_0), signaling that the data are quite informative about the persistence, volatility of growth gaps, and initial growth rate.

Table 1: Bayesian posterior medians and 95% credible intervals for (φ, B, F, x_0) with ν fixed at 0.02 and x_0 estimated.

Economy	Start year	φ	B	F	x_0
China	1984	0.929	0.0176	0.0246	0.086
		(0.810, 1.000)	(0.0070, 0.0300)	(0.0190, 0.0310)	(0.037, 0.137)
Singapore	1966	0.928	0.0025	0.0391	0.093
		(0.813, 0.989)	(0.0000, 0.0100)	(0.0320, 0.0470)	(0.044, 0.144)
South Korea	1963	0.970	0.0053	0.0324	0.061
		(0.912, 1.000)	(0.0000, 0.0130)	(0.0270, 0.0390)	(0.014, 0.101)

Figure 3 shows that the posterior-median predictive distribution with ν fixed at 0.02 tracks the broad movements in growth for each economy well. For each economy, we fix (φ, B, F, x_0) at their posterior medians, simulate many artificial (y_t) paths from the estimated scalar model, convert them to growth rates, and at each horizon take quantiles across simulations. The dark and light shaded regions therefore plot the 50% and 90% predictive bands for annual growth.

In all three episodes, the realized growth path typically stays within the 50% posterior predictive band and almost always within the wider 90% band, while gradually converging toward the long-run US growth rate. For all three economies, the posterior predictive bands capture both the initial high-growth phase and the subsequent slowdown.

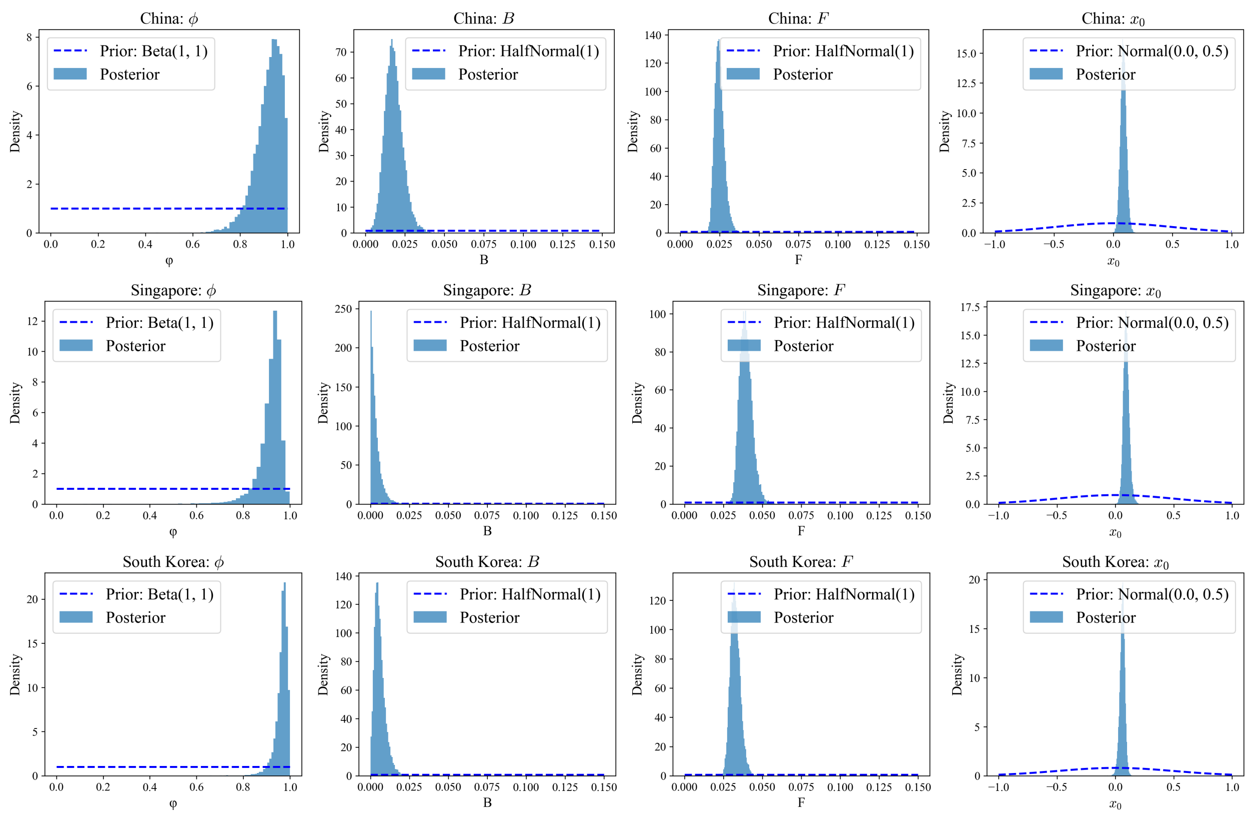

Figure 2: Prior and posterior distributions of (φ, B, F, x_0) for China (top), Singapore (middle), and South Korea (bottom).

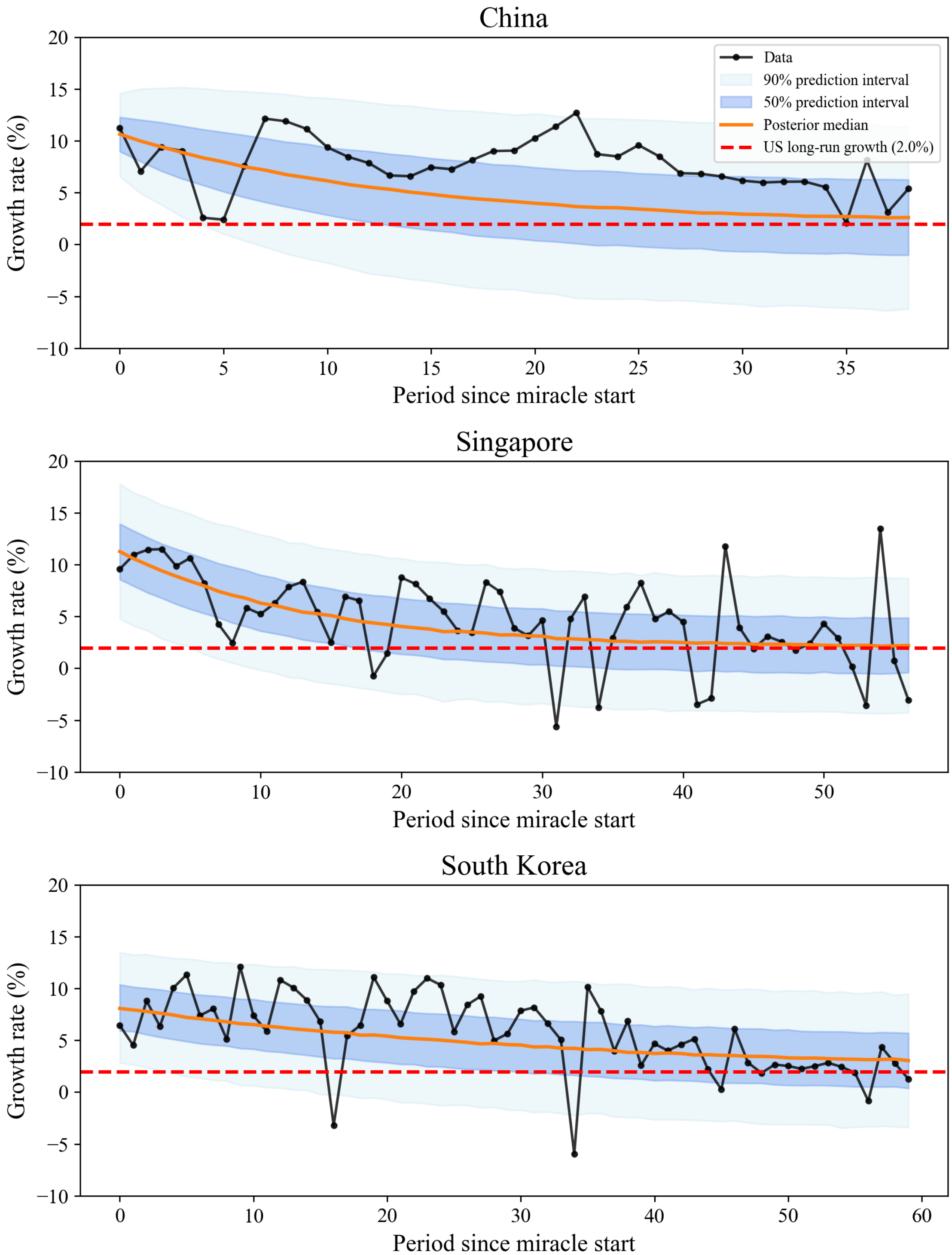

Figure 3: Posterior-median predictive bands for China (top), Singapore (middle), and South Korea (bottom). The solid black line shows realized growth; shaded areas show 50% and 90% prediction intervals; the dashed horizontal line marks the long-run US growth rate.

3.4 Decomposition of growth paths

In this section, we follow the discrete-time variant of Hansen (2012) to decompose the observed path of (y_t) specified in the additive-functional model in (1)–(2) into trend, martingale, stationary, and constant components. We use the posterior-median estimates from the fixed-ν scalar gap specification in the previous section, summarized in Table 1. This specification imposes a common steady-state growth rate $\nu = 0.02$ across economies and therefore yields a clean separation of 2% deterministic trend growth from the stochastic components of growth that vary across economies and contribute to their growth miracles.

The following proposition summarizes the decomposition:

Proposition 3.1 (Additive-Functional Decomposition)**.** *Suppose* (x_t, y_t) *satisfy the scalar additive-functional model* (1)–(2) *with* $|\varphi| < 1$*, and let* $\{z_{t+1}\}$ *be the i.i.d. shocks in* (1)*. Define*

$$g \coloneqq \frac{1}{1-\varphi}, \qquad H \coloneqq F + \frac{B}{1-\varphi}. \tag{5}$$

Then y_t *admits the decomposition*

$$y_t = \underbrace{t\nu}_{\text{Trend }(T_t)} + \underbrace{\sum_{j=1}^{t} H z_j}_{\text{Martingale }(M_t)} \underbrace{-g x_t}_{\text{Stationary }(S_t)} + \underbrace{(g x_0 + y_0)}_{\text{Constant}}. \tag{6}$$

Proof. See Appendix B. □

We plot the decomposition of y_t for each economy in Figure 4, using the posterior median estimates of (φ, B, F, x_0) from the fixed-ν specification (Table 1). For each episode we use the normalized log level $y_t = \ell_{\tau^*+t} - \ell_{\tau^*}$, so that $y_0 = 0$ and all paths are measured relative to the log level at the miracle start year. The black line in the figure shows the data y_t. The martingale component M_t (magenta line) accumulates the shocks and captures the permanent, stochastic part of growth. The stationary component S_t (green line) is a function of the latent gap x_t and converges to zero over time, representing transitory deviations from trend. The trend component T_t (red dashed line) grows deterministically at the common long-run growth rate $\nu = 0.02$ and traces the contribution of the constant-growth term. Given the normalization $y_0 = 0$, the constant term in (6) is absorbed into the initial level of S_t and is therefore not plotted separately.

Across all three economies, the decomposition displays a common structure once the normalization is taken into account. With $y_0 = 0$, (6) implies $y_t = T_t + M_t + S_t - S_0$, so the rise of y_t reflects deterministic trend growth together with (i) the gradual run-down of an initial gap, captured by $S_t - S_0$, and (ii) the accumulation of shocks with permanent effects, captured by the martingale

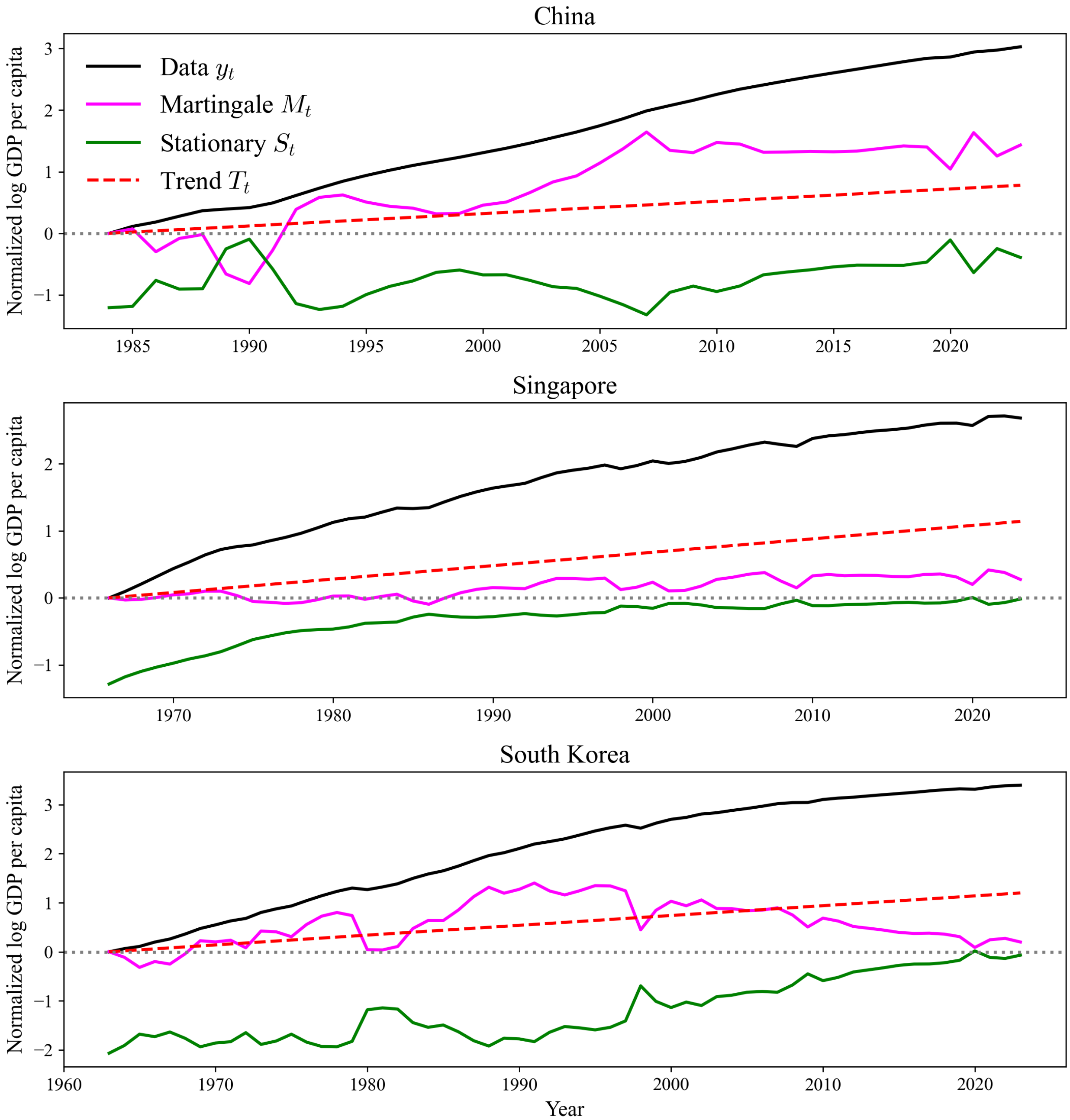

Figure 4: Decomposition of normalized log GDP per capita y_t into martingale (M_t, magenta), stationary (S_t, green), and deterministic trend (T_t, red dashed) components for China (top), Singapore (middle), and South Korea (bottom), using posterior median estimates of (φ, B, F, x_0) from the fixed-ν scalar additive-functional model in Section 3.3 (Table 1). The black line shows the observed normalized series y_t.

M_t. In each episode $S_0 = -gx_0$ is negative when $x_0 > 0$, and S_t drifts back toward (and fluctuates around) zero as the gap closes. The martingale component typically rises through the high-growth phase and then flattens when growth slows, accounting for the persistent gap between the realized path and what would be implied by deterministic trend growth alone.

3.5 Long-run growth rate as unknown parameter

The fixed-ν specification imposes a common long-run growth rate across all economies. In this section, we "describe as completely as possible the mapping from priors into posteriors" (Leamer, 1978, pp. 15–16).

To allow for heterogeneous trends, we also estimate the scalar gap model, treating ν as an unknown parameter with a diffuse prior $\nu \sim \mathcal{N}(0.02, 0.5^2)$, and the same priors for (φ, B, F, x_0) as in (4). The likelihood and state recursions are unchanged; only the parameter vector is enlarged to $(\varphi, B, F, x_0, \nu)$.

Posterior medians and 95% credible intervals for $(\varphi, B, F, x_0, \nu)$ under this specification are reported in Table 2. Relative to the fixed-ν case, the posterior medians of ν are roughly 0.3% for South Korea, 2.9% for Singapore, and 7.5% for China, with wide credible intervals for all three economies. The change in the median estimate of ν also alters the persistence of x_t. For China, φ is lowered from 0.929 in the fixed-ν specification to 0.521 in the estimated-ν specification. These results illustrate how the mapping from prior to posterior is shaped by beliefs about the long-run growth pattern.

The wide credible intervals for ν in Table 2 illustrate Leamer's point that the mapping from prior to posterior is a deliberate modeling choice. For South Korea, the 95% interval $(-0.433, 0.183)$ includes implausible negative values—a consequence of pairing a diffuse prior with data that provide little information to distinguish trend from persistence. One could impose a more informative prior on ν to rule out negative values, or fix ν at a benchmark like 2%. Both approaches encode beliefs about long-run growth; what matters is transparency about how those beliefs shape posteriors.

Table 2: Posterior medians and 95% credible intervals for $(\varphi, B, F, x_0, \nu)$ with ν estimated.

Economy	Start year	φ	B	F	x_0	ν
China	1984	0.521	0.0138	0.0229	0.040	0.075
		(0.105, 1.000)	(0.0060, 0.0240)	(0.0180, 0.0290)	(−0.022, 0.105)	(0.023, 0.108)
Singapore	1966	0.909	0.0033	0.0398	0.092	0.029
		(0.423, 1.000)	(0.0000, 0.0140)	(0.0320, 0.0490)	(−0.016, 0.271)	(−0.168, 0.077)
South Korea	1963	0.986	0.0053	0.0323	0.082	0.003
		(0.793, 1.000)	(0.0000, 0.0130)	(0.0270, 0.0390)	(−0.132, 0.499)	(−0.433, 0.183)

Figure 5 displays posterior-median predictive bands for the estimated-ν specification. Allowing ν to move reduces uncertainty around the median, which is also consistent with lower B and F estimates in Table 2.

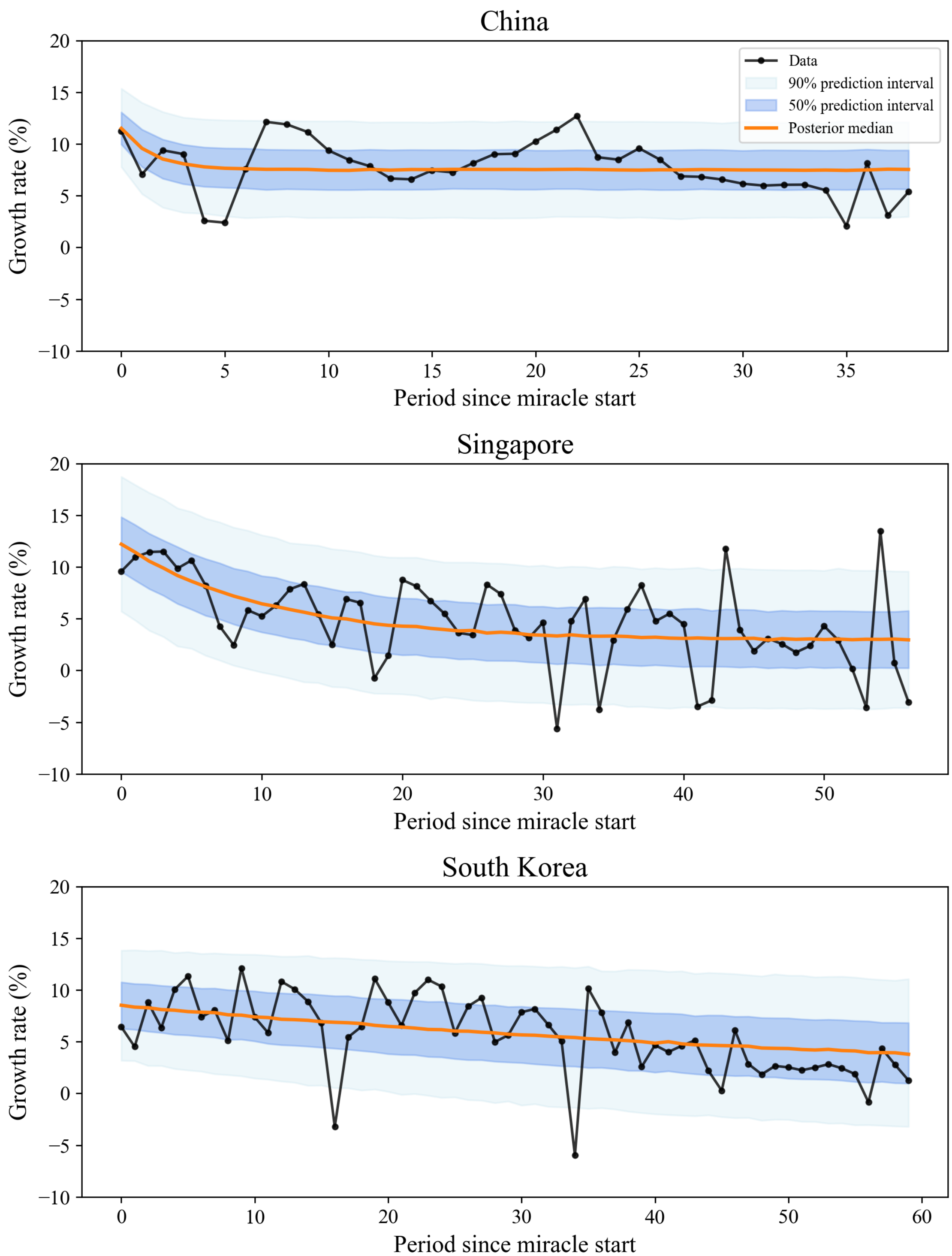

Figure 5: Posterior-median predictive bands for the scalar gap model with ν estimated. The solid black line shows realized growth; shaded areas show 50% and 90% prediction intervals.

4 Inferring Parameters of a Ramsey Model

In this section, we connect a version of the stochastic Ramsey–Cass–Koopmans model formulated by Brock and Mirman (1972) to our section 3 statistical model. Section 4.1 introduces the technology, preferences, shocks. and planner's problem for the Ramsey model, together with a scaling for growth that delivers a stationary optimal growth formulation susceptible to dynamic programming. Section 4.2 then shows how to use the Method of Simulated Moments to choose structural parameters so that the Ramsey model's simulated growth paths match key moments implied by the scalar additive–functional model estimated via Algorithm 1. This provides a mapping from the parameters of the scalar additive–functional model to the structural primitives in the Ramsey model.

4.1 Environment

Time is discrete with $t = 0, 1, \ldots, T$. Let Y_t, C_t, and K_t denote per-capita output, consumption, and capital.

Output is produced according to a Cobb–Douglas technology

$$Y_t = A_t \xi_t K_t^{\alpha}, \qquad \alpha \in (0, 1), \tag{7}$$

where A_t is an exogenous level of total factor productivity (TFP), ξ_t is a stationary multiplicative shock, and K_t is the capital stock. For later use, let $f(k) := k^{\alpha}$ so that $Y_t = A_t \xi_t f(K_t)$.

The deterministic TFP component follows

$$A_{t+1} = \mu A_t, \qquad \mu > 0, \tag{8}$$

so that $\log A_t$ grows at rate $\log \mu$.

The multiplicative productivity shock is i.i.d. log-normal with unit mean:

$$\log \xi_t \sim \mathcal{N}\left(-\tfrac{1}{2}\sigma_{\xi}^2,\ \sigma_{\xi}^2\right), \tag{9}$$

so that $\mathbb{E}[\xi_t] = 1$ for all t.

Capital accumulates according to

$$K_{t+1} = (1 - \delta) K_t + I_t, \quad \delta \in (0, 1], \tag{10}$$

where I_t denotes investment. The resource constraint is given by $C_t + I_t \leqslant Y_t$.

A representative household chooses consumption to maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(C_t), \quad \beta \in (0,1), \tag{11}$$

where u takes the form of CRRA utility:

$$u(x) = \frac{x^{1-\gamma} - 1}{1-\gamma}, \quad \gamma > 0,\ \gamma \neq 1, \tag{12}$$

and $u(x) = \log x$ in the limiting case $\gamma = 1$.

It is convenient to work with stationary variables by dividing out the deterministic TFP trend. We define detrended variables

$$k_t := \frac{K_t}{A_t^{1/(1-\alpha)}}, \quad c_t := \frac{C_t}{A_t^{1/(1-\alpha)}}, \quad \tilde{y}_t := \frac{Y_t}{A_t^{1/(1-\alpha)}}. \tag{13}$$

Using $Y_t = A_t \xi_t f(K_t)$ and the normalization above, we obtain

$$\tilde{y}_t = \frac{Y_t}{A_t^{1/(1-\alpha)}} = \xi_t f(k_t).$$

Using $A_{t+1} = \mu A_t$ and the capital-accumulation equation (10),

$$k_{t+1} = \frac{K_{t+1}}{A_{t+1}^{1/(1-\alpha)}} = \frac{(1-\delta)K_t + I_t}{\mu^{1/(1-\alpha)} A_t^{1/(1-\alpha)}} = \frac{(1-\delta)k_t + \xi_t f(k_t) - c_t}{\mu^{1/(1-\alpha)}}. \tag{14}$$

Equivalently, the detrended resource constraint is

$$c_t + \mu^{1/(1-\alpha)} k_{t+1} = \xi_t f(k_t) + (1-\delta)k_t. \tag{15}$$

With the normalization $C_t = c_t A_t^{1/(1-\alpha)}$ and $A_t = \mu^t A_0$, the planner's objective can be written, up to a constant factor, as

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \tilde{\beta}^t u(c_t), \quad \text{s.t.} \quad c_t + \mu^{1/(1-\alpha)} k_{t+1} = \xi_t f(k_t) + (1-\delta)k_t, \quad k_{t+1} \geqslant 0,$$

where $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ is applied to per-efficiency-unit consumption and

$$\tilde{\beta} := \beta \mu^{(1-\gamma)/(1-\alpha)} \tag{16}$$

is an effective discount factor that absorbs TFP growth. This is the same stochastic optimal growth framework studied in standard optimal growth treatments (see e.g., Brock and Mirman (1972)), with state variable k_t, control c_t, and i.i.d. productivity shocks ξ_t.

4.2 Moments we Match

We now show how to connect that structural model to the scalar additive-functional representation estimated in Section 3. We adopt a method of simulated moments (MSM) (McFadden, 1989; Pakes and Pollard, 1989): choose θ so that key statistics of the Ramsey model's simulated growth rates match the corresponding moments implied by our Bayesian scalar additive-functional model.

Using the estimated-ν (Table 2) estimates of the scalar additive-functional model described in Section 3, we obtain posterior medians for each economy:

$$\hat{\psi} = (\hat{\varphi}, \hat{B}, \hat{F}, \hat{\nu}, \hat{x}_0). \tag{17}$$

These parameters have the following interpretations: $\hat{\nu}$ is the long-run growth rate, $\hat{\varphi}$ measures the persistence of deviations from trend growth, $\hat{B}$ governs the contribution of transient state x_t shocks to growth volatility, $\hat{F}$ is the volatility of the innovation in growth conditional on the state, and $\hat{x}_0$ is the initial deviation of growth from its long-run mean. In the "fixed-ν" specification, we replace $\hat{\nu}$ by the common benchmark value $\nu = 0.02$ and use the corresponding posterior medians from Table 1.

For each country, we compress $(\hat{\varphi}, \hat{B}, \hat{F}, \hat{\nu}, \hat{x}_0)$ into three scalar target moments. Under the scalar transient gap model, the posterior-median path of expected growth is

$$\Delta y_{t+1}^{\text{med}} = \hat{\nu} + \hat{\varphi}^t \hat{x}_0, \qquad t = 0, \ldots, T-1,$$

so the sample average of this path,

$$\hat{m}_1 := \frac{1}{T} \sum_{t=0}^{T-1} \left(\hat{\nu} + \hat{\varphi}^t \hat{x}_0\right),$$

serves as the target for the mean growth rate. The scalar gap model also implies an approximately stationary unconditional standard deviation of one-period growth. For values of φ sufficiently below unity, the innovation variance converges to

$$\hat{m}_2 := \sigma_{\Delta y} = \sqrt{\frac{\hat{B}^2}{1 - \hat{\varphi}^2} + \hat{F}^2}, \tag{18}$$

which we use as the volatility target. Finally, the initial expected growth rate

$$\hat{m}_3 := \hat{\nu} + \hat{x}_0$$

is the target for the Ramsey model's first-period growth rate.

For any candidate structural parameter vector $\theta = (\mu, \sigma_\xi, K_0)$, where μ is the TFP growth factor, σ_ξ is the productivity shock volatility, and K_0 is the initial capital, we solve the nonlinear Ramsey model using Optimistic Policy Iteration (OPI) (Sargent and Stachurski, 2025). This yields optimal policy functions for consumption and capital accumulation. We fix the time-invariant structural parameters at the following values: capital share $\alpha = 0.33$, discount factor $\beta = 0.96$, risk aversion $\gamma = 2.0$, and depreciation rate $\delta = 0.03$. We also set $A_0 = 1$.

We then simulate $R = 1000$ independent paths of length T, each starting from initial capital K_0 and drawing i.i.d. log-normal productivity shocks $\xi_t \sim \text{LogNormal}(-\sigma_\xi^2/2, \sigma_\xi^2)$. For the r-th simulated path, we compute normalized log output

$$y_t^{\text{sim},(r)}(\theta) = \log Y_t^{(r)}(\theta) - \log Y_0^{(r)}(\theta), \qquad t = 0, 1, \ldots, T,$$

and the corresponding growth rates $\Delta y_{t+1}^{\text{sim},(r)}(\theta) = y_{t+1}^{\text{sim},(r)}(\theta) - y_t^{\text{sim},(r)}(\theta)$.

From the simulated growth rates, we compute three key moments averaged across the R replications:

$$\bar{m}_1(\theta) = \frac{1}{R}\sum_{r=1}^{R}\frac{1}{T}\sum_{t=0}^{T-1}\Delta y_{t+1}^{\text{sim},(r)}(\theta) \qquad \text{(mean growth)}, \tag{19}$$

$$\bar{m}_2(\theta) = \frac{1}{R}\sum_{r=1}^{R}\text{std}\big(\Delta y_{t+1}^{\text{sim},(r)}(\theta)\big) \qquad \text{(growth volatility)}, \tag{20}$$

$$\bar{m}_3(\theta) = \frac{1}{R}\sum_{r=1}^{R}\Delta y_1^{\text{sim},(r)}(\theta) \qquad \text{(initial growth)}. \tag{21}$$

For each country, we collect the simulated and target moments into vectors $\bar{m}(\theta) = (\bar{m}_1(\theta), \ldots, \bar{m}_3(\theta))'$ and $\hat{m} = (\hat{m}_1, \ldots, \hat{m}_3)'$. The MSM criterion is

$$Q(\theta) = \big(\bar{m}(\theta) - \hat{m}\big)' W \big(\bar{m}(\theta) - \hat{m}\big), \tag{22}$$

where W is a diagonal positive-definite weighting matrix. In the first stage, we set W equal to the identity (equal weights on the three moments). In the second stage we estimate the variance of each moment at the first-stage estimate and set W to a diagonal matrix with entries $1/\widehat{\text{Var}}(\bar{m}_i)$, following standard two-step MSM practice.

The moment matching estimator is

$$\hat{\theta} \in \arg\min_{\theta \in \Theta} Q(\theta), \tag{23}$$

where Θ is a compact parameter space for the structural parameters. In the implementation we use a bound-constrained quasi-Newton method (L-BFGS-B) to minimize $Q(\theta)$ for each country.

We report MSM estimates under two different calibrations of the scalar additive-functional model in Table 3. One ("fixed-ν") uses targets based on the posterior medians $(\hat{\varphi}, \hat{B}, \hat{F}, \hat{x}_0)$ with $\hat{\nu} = 0.02$, and another ("estimated-ν") uses targets based on the posterior medians $(\hat{\varphi}, \hat{B}, \hat{F}, \hat{\nu}, \hat{x}_0)$ from the estimated-ν scalar gap specification. This allows us to compare how much of the observed growth miracles can be rationalized by a Ramsey model disciplined by a US-style 2% trend versus one that inherits heterogeneous long-run growth rates from the Bayesian scalar gap model.

The quality of the fit is illustrated in Figures 6 and 7, which compare realized growth rates to the distribution implied by the Ramsey model. In Figure 6, the China fit is struggling to reconcile the low average growth rate implied by the fixed $\nu = 0.02$ with the high initial growth rates and volatility observed in the data. Under the fixed-ν calibration, the model matches the volatility of growth but systematically understates the frequency of high-growth outcomes for China. Allowing for country-specific ν shifts the Ramsey model's growth distributions to the right, bringing the model-implied means and upper tails much closer to the empirical histograms for these economies, while leaving the fit for Singapore and South Korea largely unchanged. This serves as an indication either that China has not yet converged to a steady state with US-like growth, or that China has a higher long-run growth rate compared to the 2% benchmark.

Leamer (1978) called a process that lets evidence guide a specification search "Sherlock Holmes inference". He warned that such post data model construction can promote overconfidence if the data used to generate the model is also used to help evaluate its performance. In this spirit, we report both fixed-ν and estimated-ν calibrations as a transparent sensitivity check: the fixed-ν case imposes an *ex ante* trend-growth discipline, while the systematic China misfit motivates a targeted relaxation to country-specific ν. Presenting both calibrations makes explicit how assumptions about long-run growth map into the Ramsey MSM parameter estimates. In the spirit of (Leamer, 1978, p. 15), this mapping from assumptions to parameter estimation complements our Section 3.5 mapping from priors to posteriors.

5 Approximations and Misspecifications

Scientists who participate in specification searches of course know that all models are imperfect approximations to a hidden "truth" that governs their observations.[4] It is natural to think about

[4] See Weinberg (2015, ch. 12), White (1982, 1994) and White and Hong (1999).

Table 3: MSM estimates of Ramsey model parameters under the fixed-ν and estimated-ν calibrations. Entries report TFP growth $g_\mu := \mu - 1$ (percent per year), productivity-shock volatility σ_ξ, initial capital K_0, and the minimized MSM objective value for each economy. The depreciation rate is fixed at $\delta = 0.03$ in all cases.

	g_μ (%)		σ_ξ		K_0	
Economy	fixed-ν	estimated-ν	fixed-ν	estimated-ν	fixed-ν	estimated-ν
China	2.39	4.62	0.036	0.018	0.848	0.823
Singapore	2.08	2.42	0.025	0.026	0.731	0.695
South Korea	2.90	3.62	0.027	0.032	1.350	1.301

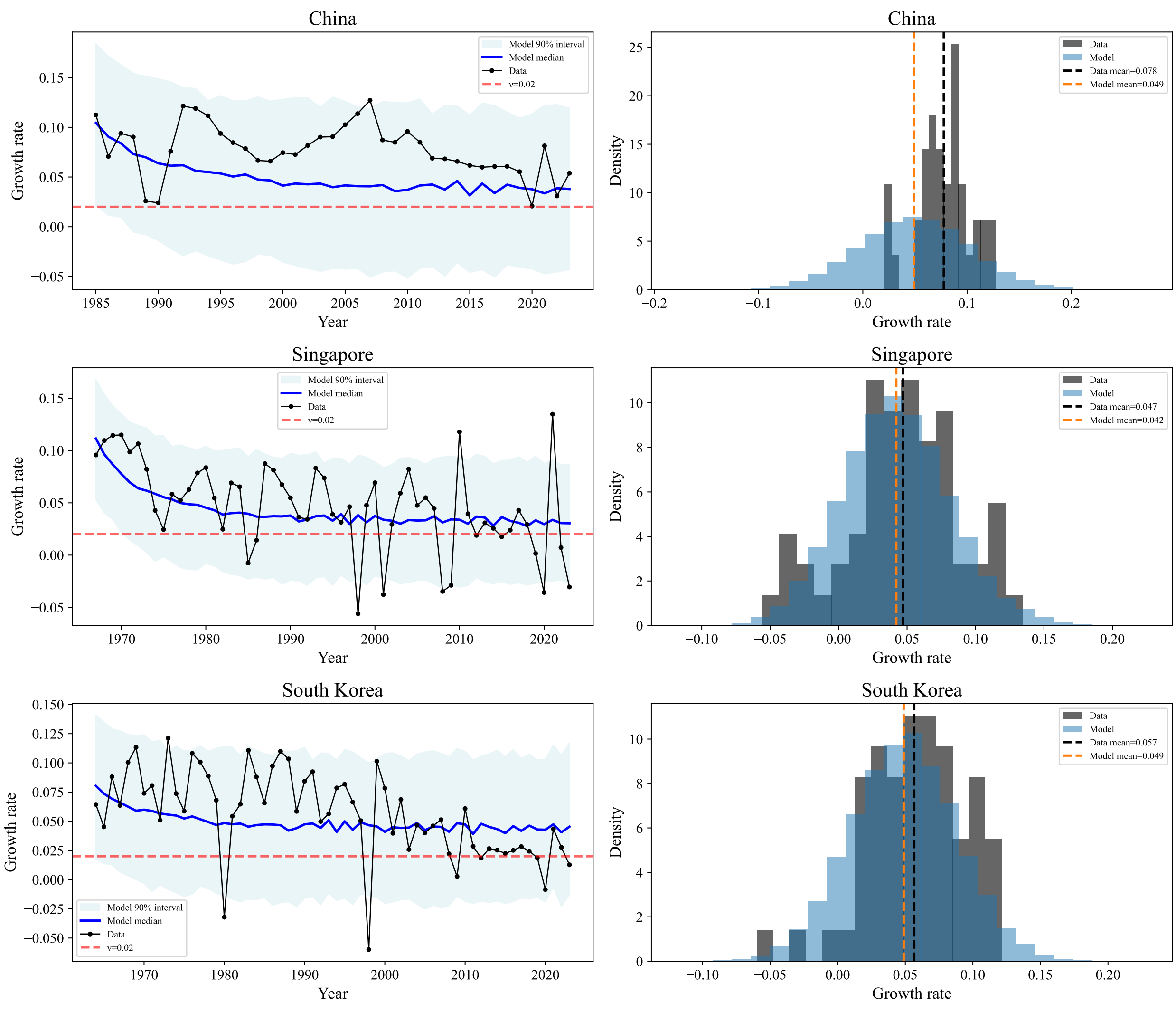

Figure 6: Distribution of annual growth rates implied by the Ramsey model under the fixed-ν MSM calibration (histograms) compared with empirical growth rate histograms for each economy.

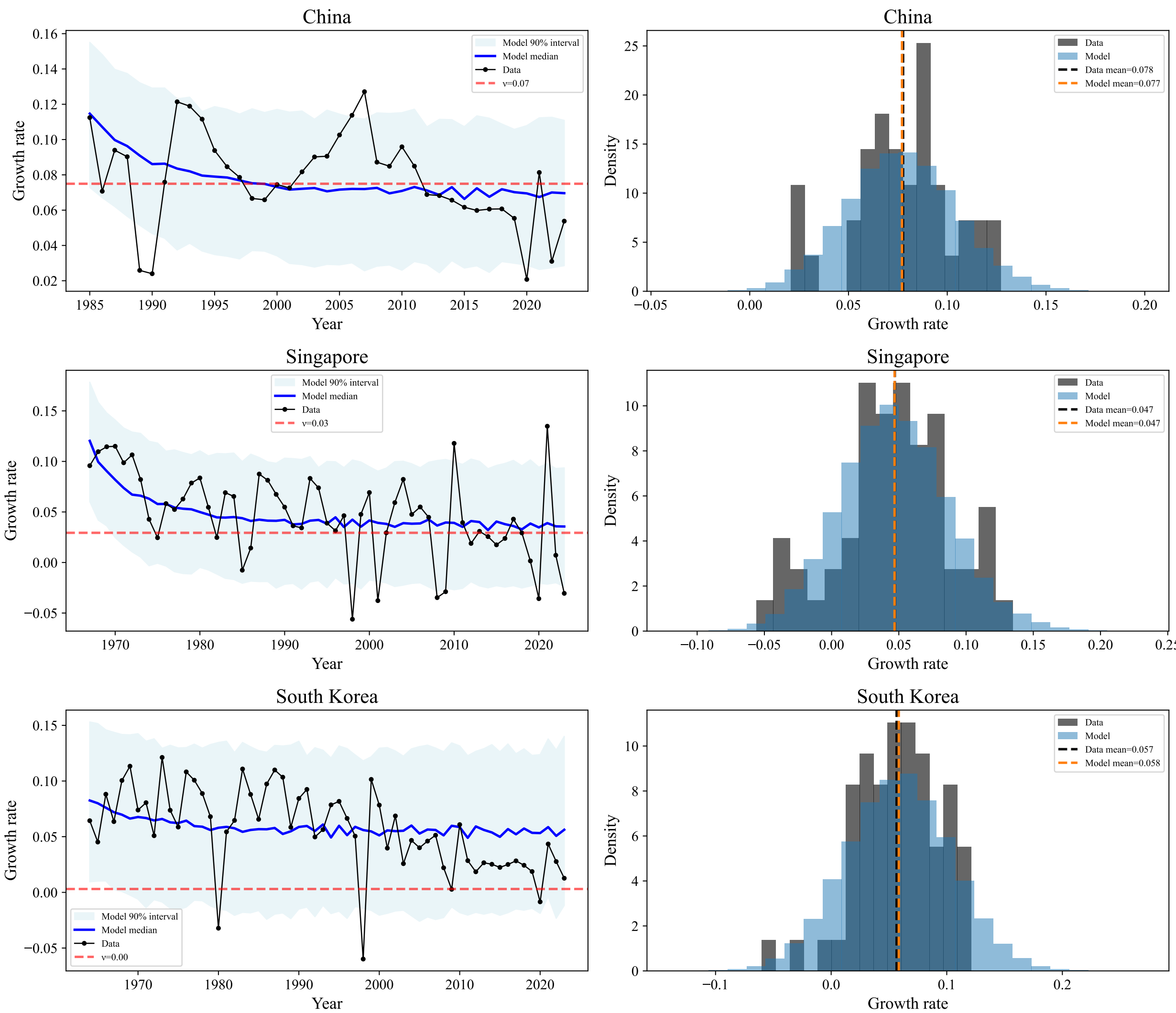

Figure 7: Distribution of annual growth rates implied by the Ramsey model under the estimated-ν MSM calibration compared with empirical growth rate histograms. Relative to Figure 6, the model places more mass at higher growth rates for China and South Korea, reflecting the higher Bayesian posterior medians for ν.

some of the approaches to specification searches analyzed by Leamer (1978) in the light of Christopher Sims's work on finite-dimensional-parameter approximations to infinite-dimensional statistical models.[5] Sims challenged a Bayesian's sharp distinction between a likelihood function and a prior distribution.[6] Thus, Sims challenged Hansen and Sargent's work that imposed the strongly overridentifying cross-equation restrictions brought by the rational expectations hypothesis Hansen and Sargent (1980, 1981). Sims noted that in specifying their likelihood function, Hansen and Sargent had, from Sims's broader perspective, in effect put overly informative and implausible prior distributions on parameters of a much higher-dimensional parameter models that Sims favored.[7] Partly in response to Sims's criticisms, Hansen and Sargent (2022, 2024) proposed a structure that perturbs the two basic building blocks of a Bayesian approach, the likelihood function and the prior. Hansen and Sargent imagine a setting in which a decision maker distrusts both of those components. The decision maker expresses that distrust by considering a set of likelihood functions and a set of priors. The decision maker uses a set of priors to describe what Hansen and Sargent call "ambiguity". Hansen and Sargent's decision makers use the set likelihood functions to describe what Hansen and Sargent call "misspecification". Hansen and Sargent proceed to study axiomatic foundations for preferences that distinguish a decision maker's aversion to ambiguity from its aversion to model misspecification.

These issues vitally affect a rational expectations theorist.[8] The rational expectations assumption imputes a common joint distribution to all agents inside a model, thereby excluding disagreements about statistical models and parameter values as well as any concerns about model misspecifications. The artificial agents who live inside a rational expectations model don't do specification searches. But econometricians like Lars Peter Hansen who have constructed and estimated rational expectations models have expressed ample doubts about both likelihood functions and priors. Hansen (2014) describes research that makes the agents inside his models more like himself and econometricians like White (1994) who have concerns about misspecifications. Hansen tells how doing that has helped him understand some behavior of asset pricing that rational expectations models had struggled to explain.

In the context of this paper, the choice between fixed-ν and estimated-ν specifications illustrates

[5] See Sims (1971, 1972, 1974).

[6] Leamer makes a closely related point:

> . . . if the prior could be uniquely determined, there would be a unique interpretation of the data, but ambiguity in the choice of prior implies ambiguity in the posterior distribution. In the case of data-selection searches, if the data distribution could be taken as given, the data would imply a unique likelihood function. But just as it is impossible unambiguously to select a prior, so too is it impossible unambiguously to select a data distribution. Not only must the interpretation of the data evidence thus remain elusive, but also the data evidence itself must be defined imprecisely. (Leamer, 1978, p. 260)

[7] See Figure 1 on page 128 of Del Negro et al. (2006) for a graphical illustration of what Sims seems to have had in mind.

[8] They also affect a game theorist committed to the "Harsanyi doctrine" that imposes a "common priors" assumption.

Leamer's mapping from priors to posteriors. Fixing $\nu = 0.02$ encodes a prior belief about long-run convergence to US growth rates; estimating ν with a diffuse prior lets the data speak more freely, though at the cost of wider posterior uncertainty. Both specifications give plausible statistical representations of the data, and each leads to a different mapping from statistical to structural parameters in the Ramsey MSM step. The framework of Hansen and Sargent suggests that a cautious analyst might want to evaluate decisions under a range of probability specifications.

6 Concluding Remarks

This paper implements a two-stage specification search to represent growth miracles in China, Singapore, and South Korea. The first stage estimates a scalar additive-functional model via Bayesian methods. The second stage uses the method of simulated moments to select parameters of a stochastic Ramsey growth model that replicate key features of the estimated statistical model.

The scalar additive-functional model decomposes each growth path into deterministic trend, martingale, and stationary components. For all three countries, the martingale component accounts for sustained deviations from deterministic trend while the stationary component describes a transient growth gap that dissipates over time. In the fixed-ν baseline, posterior estimates of the dynamics of the transient component indicate high persistence (φ near unity) and substantial volatility of this transient component.

We explored two specifications of the long-run growth parameter ν. Fixing $\nu = 0.02$ imposes a common steady-state growth rate over three countries, while estimating ν allows heterogeneous trends. For China, the estimated-ν specification yields a posterior median of 7.5% annual growth, substantially above the US benchmark. This difference matters for the MSM calibration: the fixed-ν Ramsey model systematically understates high-growth outcomes for China, while the estimated-ν calibration aligns model and data distributions more closely.

The analysis illustrates Leamer's point that specification searches involve dual sources of uncertainty. The choice of prior distributions generates posterior uncertainty about (φ, B, F, x_0), while the choice between fixed and estimated ν alters both the statistical model and the implied structural parameters. Presenting both calibrations makes this mapping transparent.

Several extensions merit consideration. The scalar state restriction limits the model's capacity to represent richer dynamics. The MSM step matches only three moments per economy; alternative moment sets or full-information methods could tighten the connection between our descriptive and structural statistical models. Similar methods could be applied to other growth episodes and to declining economies. More generally, it would be interesting to try to use additive functional statistical models like those used in section 3 to approximate transition paths generated by deterministic models like Kotlikoff et al. (2024), Benzell et al. (2024).

References

Seth G. Benzell, Laurence J. Kotlikoff, Maria Kazakova, Guillermo Lagarda, Kristina Nesterova, Victor Yifan Ye, and Andrey Zubarev. The future of global economic power. *Journal of Policy Modeling*, 46(2):395–412, 2024. doi: 10.1016/j.jpolmod.2024.03.004. Working paper versions also available from NBER.

William A Brock and Leonard J Mirman. Optimal economic growth and uncertainty: the discounted case. *Journal of Economic Theory*, 4(3):479–513, 1972.

Central Committee of the Communist Party of China. Decision on reform of the economic structure, October 1984. URL https://w.wiki/GdJW. Chinese-language text transcription; 2025-12-14.

Marco Del Negro, Frank Schorfheide, Frank Smets, and Rafael Wouters. On the fit of new keynesian models. *Journal of Business & Economic Statistics*, 25(2):123–143, 2006.

Lars Peter Hansen. Dynamic valuation decomposition within stochastic economies. *Econometrica*, 80(3):911–967, 2012.

Lars Peter Hansen. Nobel lecture: Uncertainty outside and inside economic models. *Journal of Political Economy*, 122(5):945–987, 2014.

Lars Peter Hansen and Thomas J. Sargent. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control*, 2:7–46, 1980.

Lars Peter Hansen and Thomas J. Sargent. Linear rational expectations models for dynamically interrelated variables. *Rational Expectations and Econometric Practice*, pages 127–156, 1981.

Lars Peter Hansen and Thomas J. Sargent. *Recursive Models of Dynamic Linear Economies*. Princeton University Press, Princeton, 2013. ISBN 9780691042770.

Lars Peter Hansen and Thomas J. Sargent. Structured ambiguity and model misspecification. *Journal of Economic Theory*, 199:105275, 2022. ISSN 0022-0531.

Lars Peter Hansen and Thomas J. Sargent. Robust filtering with a robust kalman filter. *Journal of Applied Econometrics*, 39(3):501–517, 2024.

Matthew D Hoffman, Andrew Gelman, et al. The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623, 2014.

Robert Kohn and C. F. Ansley. Efficient estimation and prediction in time series regression models. *Biometrika*, 72(3):694–697, 1985.

Robert Kohn and C. F. Ansley. Estimation, prediction, and interpolation for ARIMA models with missing data. *Journal of the American Statistical Association*, 81(395):751–761, 1986.

Tjalling C Koopmans. Measurement without theory. *The Review of Economics and Statistics*, 29 (3):161–172, 1947.

Laurence Kotlikoff, Felix Kubler, Andrey Polbin, and Simon Scheidegger. Can today's and tomorrow's world uniformly gain from carbon taxation? *Journal of Environmental Economics and Management*, 123:102926, 2024. ISSN 0095-0696. doi: https://doi.org/10.1016/j.jeem.2024.102926. URL https://www.sciencedirect.com/science/article/pii/S0095069624000130.

Finn E. Kydland and Edward C. Prescott. Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–1370, 1982.

Edward E. Leamer. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, 1978.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026, 1989.

Ariel Pakes and David Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, pages 1027–1057, 1989.

Du Phan, Neeraj Pradhan, and Martin Jankowiak. Numpyro: A Pyro-based probabilistic programming library on JAX. https://github.com/pyro-ppl/numpyro, 2019.

Thomas J. Sargent and John Stachurski. *Dynamic Programming: Finite States*. Cambridge University Press, 2025.

Christopher A. Sims. Comments on: "a decision-theoretic approach to the problem of testing a null hypothesis". *The Annals of Mathematical Statistics*, 42(5):1740–1742, 1971.

Christopher A. Sims. The role of approximate prior restrictions in distributed lag estimation. *The Annals of Mathematical Statistics*, 43(1):326–341, 1972.

Christopher A. Sims. Distributed lags. *The Annals of Mathematical Statistics*, 2(1):1–22, 1974.

Steven Weinberg. *To Explain the World: The Discovery of Modern Science*. Harper, New York, 2015.

Halbert White. Maximum likelihood estimation of misspecified models. *International Economic Review*, 23(1):1–25, February 1982.

Halbert White. *Estimation, Inference and Specification Analysis*. Number 22 in Econometric Society Monographs. Cambridge University Press, New York, 1994. ISBN 0-521-25280-6.

Halbert White and Yongmiao Hong. M testing using finite and infinite dimensional parameter estimators. *Journal of Econometrics*, 91(2):195–228, 1999.

Xinhua News Agency. Decision on reform of the economic structure. China Reform and Opening-up Database (ReformData), October 1984. URL https://www.reformdata.org/1984/1020/21123.shtml. Chinese-language text; dated 1984-10-20; accessed 2025-12-14.

A Fixing initial x_0

One alternative to the prior for x_0 is to hold it fixed rather than estimate it. We do so by choosing x_0 to match the initial observed growth rate under the assumption of zero initial innovation, which is achieved by

$$x_0 = \Delta y_1 - \nu,$$

using equation (2) with $z_1 = 0$, so that only (φ, B, F) are treated as parameters to be estimated. Re-estimating the model under this restriction yields posterior summaries reported in Figure 8. For China and Singapore, the fixed value of x_0 lies close to the posterior median under the estimated-x_0 specification.

For South Korea, by contrast, the posterior links x_0 much more tightly to (φ, F). The early part of the growth path can be rationalized either by a relatively large and persistent initial gap (large x_0, φ close to one, moderate F) or by a smaller initial gap combined with more transitory shocks (smaller x_0, lower φ, larger F). Imposing $x_0 = \Delta y_1 - \nu$ therefore amounts to conditioning on a particular value of a variable that is strongly correlated with (φ, F) and excludes part of the "large and persistent gap" region of the joint posterior. The conditional posterior under fixed x_0 correspondingly shifts probability mass toward the alternative "more transitory shocks" story, which appears in Figure 8 as the additional bump in the posterior for φ and the secondary mode in the posterior for F for South Korea.

B Proof of Additive Functional Decomposition

Proof of Proposition 3.1. Recall the model

$$x_{t+1} = \varphi x_t + B z_{t+1}, \tag{24}$$

$$\Delta y_{t+1} := y_{t+1} - y_t = \nu + x_t + F z_{t+1}, \tag{25}$$

with $|\varphi| < 1$, $B > 0$, $F > 0$, and (z_t) i.i.d. standard normal. Define

$$g := \frac{1}{1-\varphi}, \qquad H := F + \frac{B}{1-\varphi}.$$

We first show that, with these choices of g and H, each increment of y_t can be written as

$$\Delta y_{t+1} = \nu + H z_{t+1} - g x_{t+1} + g x_t, \tag{26}$$

for all $t \geqslant 0$. Starting from the right-hand side of (26) and substituting the state recursion (24) for

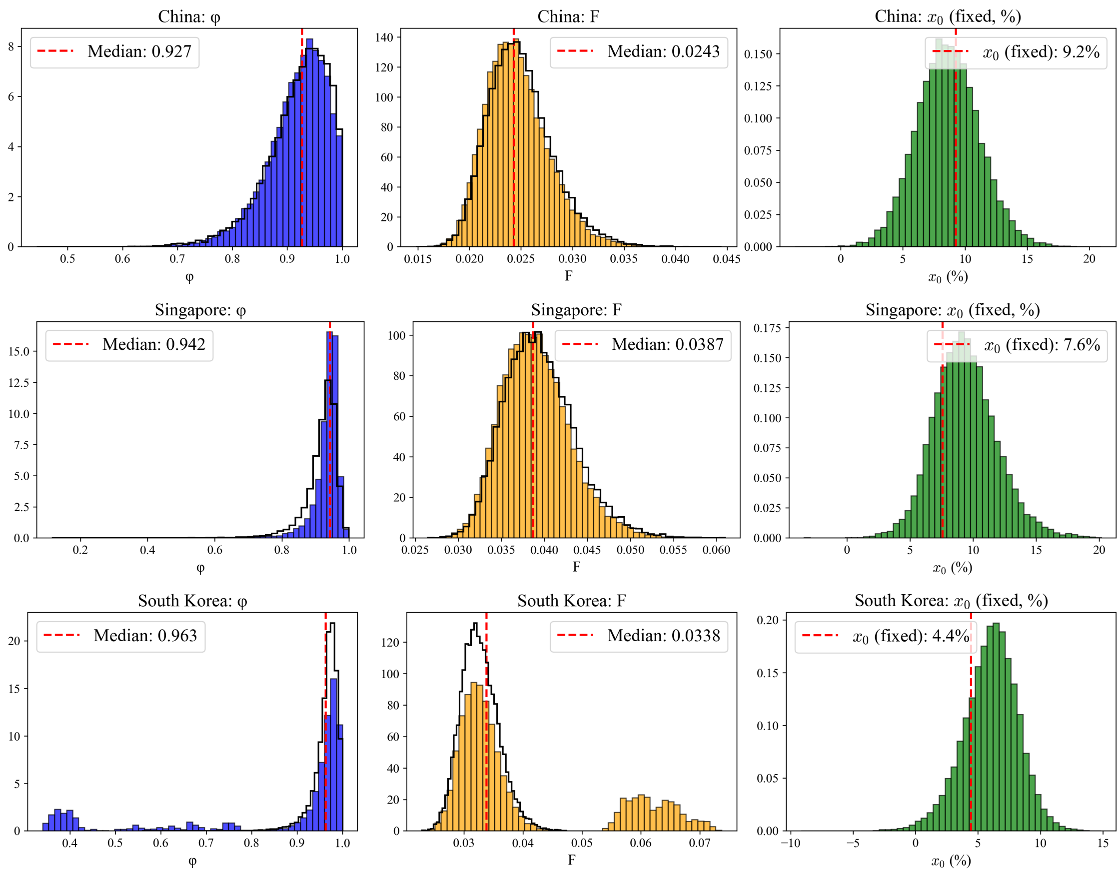

Figure 8: Posterior distributions of (φ, B, F, x_0) under the estimated- and fixed-x_0 specifications for China (top), Singapore (middle), and South Korea (bottom). For each economy, colored histograms show the posteriors under the fixed-x_0 specification; black step histograms show the corresponding posteriors when x_0 is estimated. In the right panels, the histogram is the posterior for x_0 under the estimated-x_0 specification and the dashed line marks the fixed value.

x_{t+1}, we have

$$\begin{aligned} \nu + Hz_{t+1} - gx_{t+1} + gx_t &= \nu + Hz_{t+1} - g(\varphi x_t + Bz_{t+1}) + gx_t \\ &= \nu + \big(H - gB\big)z_{t+1} + g(1-\varphi)x_t. \end{aligned}$$

By the definitions of g and H,

$$g(1-\varphi) = \frac{1}{1-\varphi}(1-\varphi) = 1, \qquad H - gB = F + \frac{B}{1-\varphi} - \frac{B}{1-\varphi} = F.$$

Substituting these identities into the last expression makes the right-hand side equal to $\nu + x_t + Fz_{t+1}$, which is exactly Δy_{t+1} by (25). Hence (26) holds.

Now sum (26) over $t = 0, \ldots, T-1$:

$$\begin{aligned} y_T - y_0 &= \sum_{t=0}^{T-1} \Delta y_{t+1} \\ &= \sum_{t=0}^{T-1} \big[\nu + Hz_{t+1} - gx_{t+1} + gx_t\big] = \nu T + H \sum_{t=0}^{T-1} z_{t+1} - g \sum_{t=0}^{T-1} x_{t+1} + g \sum_{t=0}^{T-1} x_t. \end{aligned}$$

The last two sums telescope:

$$-g \sum_{t=0}^{T-1} x_{t+1} + g \sum_{t=0}^{T-1} x_t = -gx_T + gx_0.$$

Thus

$$y_T - y_0 = \nu T + H \sum_{t=0}^{T-1} z_{t+1} - gx_T + gx_0,$$

or, after reindexing the shock sum,

$$y_T = T\nu + \sum_{j=1}^{T} Hz_j - gx_T + (gx_0 + y_0),$$

which is the decomposition in (6) with t in place of T.

To interpret the components, let

$$T_t := t\nu, \qquad M_t := \sum_{j=1}^{t} Hz_j, \qquad S_t := -gx_t, \qquad t \geqslant 0.$$

Work with the filtration $\mathcal{F}_t := \sigma\{x_0, z_1, \ldots, z_t\}$. Since $\{z_{t+1}\}$ are i.i.d. with $\mathbb{E}[z_{t+1}|\mathcal{F}_t] = 0$, we have

$$\mathbb{E}[M_{t+1} \mid \mathcal{F}_t] = \mathbb{E}\Big[\sum_{j=1}^{t} Hz_j + Hz_{t+1} \;\Big|\; \mathcal{F}_t\Big] = M_t + H\,\mathbb{E}[z_{t+1} \mid \mathcal{F}_t] = M_t.$$

Hence (M_t) is an $(\mathcal{F}_t)$-martingale. Under the assumptions, the process (x_t) is a stable AR(1) driven by i.i.d. shocks; under $|\varphi| < 1$ there exists a unique stationary distribution. The remaining term $gx_0 + y_0$ is constant over time.

Collecting terms, we obtain the stated decomposition of y_t into a deterministic trend T_t, a martingale M_t, a stationary component S_t, and a constant term. This proves Proposition 3.1. □