Mean Reversion in Macroeconomic Models

John B. Donaldson
Columbia University

Rajnish Mehra
Arizona State University
Luxembourg School of Finance, NBER and NCAER

Abstract

We explore the property of mean reversion in stock prices and returns within a class of dynamic stochastic general equilibrium macroeconomic models. Our objective is to understand the macroeconomic structures responsible for mean reversion and to gain insight regarding the observed difficulty in detecting mean reversion in actual return and price data.

We thank Sergio Villar for his superb programming assistance and for his insights on Propositions 4.2 and 4.4
1. Introduction

It is well known that tests for mean reversion in financial asset return series at frequencies of three to five years yield extremely weak results because of the lack of sufficient non-overlapping historical data. There nevertheless remains substantial interest in the phenomenon of mean reversion, (and its implications for return predictability) because of its implications for long term investor portfolio composition, as well as its significance for the efficient markets hypothesis.

In this note we revisit the notion of mean reversion in a very simple macroeconomic context. Recognizing that traded equity securities represent ownership of a substantial portion of the capital stock of the United States, we seek to explore the mean reverting property of equity returns within the context of standard DSGE macroeconomic models parameterized to replicate the patterns in macroeconomic time series found in United States macroeconomic data.

Within this general context our study has two interlocking objectives:

1. To clarify the three principal characterizations of mean reversion in use in the finance literature and study the relationships among them. We also propose an alternative notion of “mean reversion,” one that provides an intuitive measure of one series being “more strongly mean-reverting” than another.

2. To understand the underlying macroeconomic mechanisms that promote or discourage mean reversion in an economy’s capital stock, equity price, equity rate of return and equity premium series. In particular, this objective calls us to understand the relative influence of various model features (and the actual phenomena they represent) in promoting or confounding mean reversion in the
aforementioned financial quantities. We consider, for example, the influence of persistence in total factor productivity shocks.\footnote{In doing so we extend the pioneering work of Basu and Vinod (1994) to richer macroeconomic settings.}


One challenge facing this literature is the absence of adequate data for discriminating statistical tests. The modeling environment we consider sidesteps the paucity of actual return data by allowing us to generate arbitrarily long return series, and thus undertake our statistical analysis with a high degree of precision.\footnote{An abundance of model-generated data allows us to discriminate between the weak evidence of mean reversion due to insufficient data or due to fundamental underlying macroeconomic mechanisms.}

As we will see, for standard DSGE models, mean reversion in equity returns as an equilibrium phenomenon is the exception rather than the rule: mean aversion is the more generic property (we precisely define these terms in the next section).\footnote{DSGE abbreviates “dynamic stochastic general equilibrium.}  \footnote{This assertion holds for all characterizations of mean reversion found in the literature – see Section 2.} Accordingly, if the DSGE model class does represent the structure of the real economy in some fundamental way, this result makes less puzzling the difficulty in recovering strong proof of mean reversion in the data: we have no strong theoretical reason to suggest its existence, at least as customarily defined.

An outline of the paper is as follows. In section 2 we detail the simple overarching model framework of our study, identify the three definitions of mean
reversion found in the literature, and partially characterize their interrelationships. In Section 3, these definitions are then applied to the analysis of the most basic dynamic macroeconomic model. An alternative measure of mean reversion to those already present in the literature is presented. In Section 4 we add additional features to the model and study the resulting implications for the strength of mean reversion/aversion in model-generated equity return and equity premium data. Section 5 concludes.

Our analysis is both analytical and, where useful, computational as based on wide ranging numerical simulations.

2. Modeling Perspective and the Introductory Paradigm

Our initial focus will be macroeconomic models for which the fundamental underlying structure is based on the one good stochastic growth model with “planning” representation:

$$\max E \left( \sum_{t=0}^{\infty} \beta^t u \left( \tilde{c}_t, 1 - \tilde{n}_t \right) \right)$$

s.t. 
$$c_t + i_t \leq y_t = f \left( k_t, n_t \right) \tilde{\lambda}_t$$

$$k_{t+1} = \left( 1 - \Omega \right) k_t + i_t$$

$$k_0 \equiv 1, \text{ given.}$$

$$\tilde{\lambda}_{t+1} \sim G \left( \tilde{\lambda}_{t+1}; \lambda_t \right).$$

Adopting the customary notation, $$u \left( c_t, 1 - n_t \right)$$ represents the representative agent’s period utility function defined over his period $$t$$ consumption $$c_t$$ and leisure, $$(1 - n_t)$$, where $$n_t$$ is labor supplied. $$f \left( k_t, n_t \right) \lambda_t$$ denotes the representative firm’s CRS production function of capital stock $$k_t$$ and labor supplied with $$\left\{ \tilde{\lambda}_t \right\}$$ the stochastic total factor productivity shock. The probability distribution
function for \( \{\lambda_{t+1}\} \) conditional on \( \lambda_t \) is denoted \( G(\lambda_{t+1}; \lambda_t) \) and is assumed known to the representative agent.\(^5\) Lastly, \( \beta \) denotes the representative agent’s subjective time discount factor and \( \Omega \) is the period depreciation rate.

As the previous notation suggests, the state variables for this economy are \( k_t \) and \( \lambda_t \). Under standard assumptions, problem (1) has a solution; that is:

(i) continuous, time-invariant consumption \( c_t = c(k_t, \lambda_t) \), investment \( i_t = i(k_t, \lambda_t) \) and labor service \( n_t = n(k_t, \lambda_t) \) functions exist that solve problem (1), and

(ii) a unique invariant probability measure on the state variable pair \( (k_t, \lambda_t) \) exists to which the joint stochastic process on \( (k_t, \lambda_t) \) converges weakly and which describes its long run behavior. With these attributes we say that the joint process on \( (k_t, \lambda_t) \) is stationary.\(^6\) \(^7\) As a result the stochastic processes

\(^5\) When \( \{\lambda_t\} \) is not i.i.d., we typically model it to be of the form \( \lambda_t = e^{x_t} \) where \( x_t \) is an A.R.-1 process.

\(^6\) Our notion of stationarity for a discrete time Markov process \( \{x_t\} \) is as follows; Let \( s, t \) be arbitrary time indices and \( X \) the state space with \( \hat{x} \in X \), and \( B \subseteq X, B \) a subset. Define \( P(s, \hat{x}, t, B) = \text{Prob}(x_t = B; x_s = \hat{x}) \). Then for any integer \( u \), if \( P(s + u, \hat{x}, t + u, B) = \text{Prob}(x_{t+u} = \hat{x}; x_{t+u} \in B) = P(s, \hat{x}, t, B) \), the Markov process is said to be stationary. The same Markov process possesses an invariant distribution \( \hat{G}(\cdot) \) on \( X \) if and only if for any \( B \subseteq X \),

\[
\text{Prob}(x_{t+u} \in B) = \int_{x \in X} P(x_{t+1} \in B; x_t = x) \hat{G}(dx).
\]

All the stochastic processes analyzed in this article are Markov, stationary and possess unique invariant distributions defined on compact sets.

\(^7\) The details behind these assertions can be found in the literature. Part (i) is entirely standard. As for part (ii), the stochastic kernel, the expression \( P(s, \hat{x}, t, B) \) in footnote 3 can be shown to be increasing, order reversing and to satisfy the “Feller Property.” By Theorem 3.2 in Kamihigashi and Stachurski (2014) a unique stationary probability distribution exists with the indicated properties.
governing investment, \( i(k_t, \lambda_t) \), consumption, \( c(k_t, \lambda_t) \), labor service, \( n(k_t, \lambda_t) \) and output, \( y_t = y(k_t, n(k_t, \lambda_t)) \lambda_t \) are also stationary. The same investment and consumption functions arising as the solution to (1) coincide with the aggregate investment and consumption functions arising from an analogous decentralized market economy in recursive competitive equilibrium, a fact well known to the literature; see, e.g., Prescott and Mehra (1980), Brock (1982), or Danthine and Donaldson (2015).

These decentralization schemes for (1) may be generalized to accommodate an implied financial market where risk free debt and equity are competitively traded.\(^8\) Under this expanded interpretation the period \( t \) dividend satisfies

\[
d_t = f(k_t, n_t)\lambda_t - w_t n_t - i_t \quad (4i)
\]

while the ex dividend aggregate equity price, \( p_t^e \), is identified with the capital stock:

\[
p_t^e = k_{t+1}. \quad (4ii)
\]

In (4i) \( w_t \) denotes the competitive wage rate. In equilibrium it satisfies

\[
w_t = f_2(k_t, n_t)\lambda_t.
\]

Accordingly,

\[
1 + r_{t+1}^e = \frac{p_{t+1}^e + d_{t+1}}{p_t^e} \quad (5)
\]

\[
= k_{t+2} + f(k_{t+1}, n_{t+1})\lambda_{t+1} - n_{t+1} f_2(k_{t+1}, n_{t+1})\lambda_{t+1} - i_{t+1}
\]

\[
= k_{t+1}
\]

---

\(^8\) Under one such interpretation the financial market can be regarded as “complete.”

\(^9\) In a related study, Lansing (2015) refers to this dividend expression as the “macroeconomic dividend.”
\[
\frac{k_{t+2} + k_{t+1} f_1(k_{t+1}, n_{t+1}) \lambda_{t+1} - i_{t+1}}{k_{t+1}} \quad \text{(by CRS)}
\]

\[
= f_1(k_{t+1}, n_{t+1}) \lambda_{t+1} + (1 - \Omega)
\]

where \( r^c_{t+1} = f_1(k_{t+1}, n_{t+1}) \lambda_{t+1} - \Omega \) denotes the net return on unlevered equity from the “end of period t” to the “end of period t+1.” \(^{10}\)

The period price, \( p^f_t \), of a risk-free bond paying one unit of consumption in period t+1 irrespective of the realized state is

\[
p^f_t = p^f(k_t, \lambda_t) = \beta \int \frac{u_1(c_{t+1}^{-1})c_{t+1}^{1-n}}{u_1(c_t^{-1})^{1-n_t}} \ dG(\lambda_{t+1}; \lambda_t)
\]

with the risk-free rate \( r^f_t = r^f(k_{t-1}, \lambda_{t-1}) \) satisfying \( 1 + r^f_t = 1 / p^f_t \). Accordingly, the equity premium is defined by \( r^p_t = r^e_t - r^f_t \).

As continuous bounded functions of the economy’s state variables, \( p^f_t \), \( p^f(k_t, \lambda_t) \), \( r^c(k_t, \lambda_t) \) and \( r^f(k_t, \lambda_t) \) are also stationary stochastic processes. In addition, the capital-output ratio, \( \left\{ \frac{k_t}{y_t} \right\} \), the growth rate of output, \( \left\{ g^p_t \right\} = \left\{ \frac{y_t}{y_{t-1}} \right\} \), and the share of income to capital, \( \left\{ \frac{r^c_t k_t}{y_t} \right\} \), represent the ratios of strictly positive stationary stochastic processes and thus are stationary as well. These latter quantities will be relevant for our discussion to follow.

2.1 Mean Reversion

Our objective is to determine if the equilibrium rate of return series for model economies of type (1) are mean reverting. There is, unfortunately, no

---

\(^{10}\) Identification (6) does not necessarily hold in more elaborate models; see Section 4.
unique meaning attached to the expression “mean reversion.” Historically, mean reversion has been identified with the concept of stationarity, but this is surely inadequate since any i.i.d. process is stationary but not mean-reverting in any discriminating sense of that word. Within the finance literature there appear to be three distinct candidate time series characteristics identified with “mean reversion.” They are as follows, expressed in terms of an arbitrary stationary stochastic process \( \{ \tilde{x}_t \} \).

A stationary stochastic process \( \{ \tilde{x}_t \} \) is said to be mean reverting if and only if:

I. \( \text{cov}(\tilde{x}_t, \tilde{x}_{t-1}) < 0 \)

II. \( \frac{\text{var}(\tilde{x}_t + \tilde{x}_{t+1} + \ldots + \tilde{x}_{t+j})}{j+1} < \text{var} \tilde{x}_t \), for any \( j \geq 1 \).

Property I is used by Guvenen (2009) and Lansing (2014). Property II was first proposed in Summers (1986) and used in, e.g., Poterba and Summers (1988) and Mukherji (2011) for their discussions of mean reversion in stock price and rate of return series.\(^{11}\)

III. for any time integers \( 0 \leq r < s < t < u \),

\[
\text{cov}(x_s - x_r, x_u - x_t) < 0. \tag{8}
\]

Definition III, to our knowledge first proposed in Exley et al. (2009), is a comment about sequential changes in the values of the stochastic process \( \{ \tilde{x}_t \} \) rather than a statement about the statistical properties of the values themselves. Interpreting \( \{ \tilde{x}_t \} = \{ \tilde{k}_t \} = \{ \tilde{p}_t \} \), it states that increases in the price of capital over a particular interval of time will generally be followed by reductions in the price.

\(^{11}\) It stands in specific contrast to the analogous property of a random walk where

\[
\text{var}(\tilde{x}_t + \tilde{x}_{t-1} + \ldots + \tilde{x}_{t+j-1}) = j \text{var}(\tilde{x}_t).
\]
in future time intervals. As such, it represents a fundamentally different sense of mean reversion than properties I and II. In each of the definitions if the identifying inequality is reversed, the series is said to be mean averting.

The relationship between Definitions I and II is partially captured in Proposition 2.1.

**Proposition 2.1** Let \( \{\tilde{x}_t\} \) be a stationary stochastic process with an ergodic probability distribution. With respect to that distribution, statistical properties I and II detailed above are related according to

a) \( \text{II} \Rightarrow \text{I} \)

b) If \( \left| \text{cov}(\tilde{x}_t, \tilde{x}_{t+1}) \right| > \sum_{s=2}^{\infty} \text{cov}(\tilde{x}_t, \tilde{x}_{t+s}) \) for all \( j \) \( (9) \) then \( \text{I} \Rightarrow \text{II} \).

Proof: See the Appendix.

Property III can be guaranteed if certain sufficient conditions are satisfied. This is the subject of the following Lemma and Proposition.

**Lemma 2.1:** Consider arbitrary time indices \( 0 < r < s < t < u \). A stochastic process \( \{\tilde{x}_t\} \) is mean reverting by Property III if and only if

\[
\text{var}(\tilde{x}_u - \tilde{x}_r) + \text{var}(\tilde{x}_s - \tilde{x}_r) < \text{var}(\tilde{x}_u - \tilde{x}_r) + \text{var}(\tilde{x}_s - \tilde{x}_r). \quad (10)
\]

Proof: See Exley et al. (2004).

Let us next make the identification

\[
v(h - k) \equiv \text{def} \ \text{var}(\tilde{x}_h - \tilde{x}_k) \text{ for any time integers } h > 0, k > 0. \]

This allows a simple presentation of the following proposition:

**Proposition 2.2:** If \( v(\ ) \) is concave then \( \{\tilde{x}_t\} \) is mean reverting.

Proof: See Exley et al. (2004).
As regards the empirical literature on mean reversion, characterization III is rarely employed. Accordingly, we first focus principally on characterizations I – II; in particular, are they satisfied in equilibrium macroeconomic models?

3. Variations on a Classical Example

Here we restrict problem (1) by requiring that \( u(c_t) = \ell n(c_t) \) while

\[
y_t = f(k_t, n_t) \tilde{\lambda}_t = \tilde{k}_t^\alpha \tilde{\lambda}_t \quad \text{where labor} \ n_t \ \text{is fixed at} \ n_t = 1. \quad \text{It is again widely known that the optimal policy functions assume the form}
\]

\[
c_t = c(k_t, \lambda_t) = (1 - \alpha \beta) y_t \quad \text{and} \quad (11)
\]

\[
i_t = k_{t+1} = \alpha \beta y_t = \alpha \beta k_t^\alpha \tilde{\lambda}_t, \quad \text{while} \quad (12)
\]

\[
d_t = \alpha k_t^\alpha \tilde{\lambda}_t - \alpha \beta k_t^\alpha \tilde{\lambda}_t = \alpha \beta k_t^\alpha \tilde{\lambda}_t. \quad (13)
\]

By recursive substitution

\[
k_t = \left( (\alpha \beta)^{1 + \alpha + \alpha^2 + \ldots + \alpha^{t-1}} k_0 \prod_{s=0}^{t-1} \lambda_s^{\alpha^{t-s}} \right). \quad (14)
\]

With \( \{\tilde{\lambda}_t\} \) again an i.i.d. process, Danthine and Donaldson (1981) and Hopenhayn and Prescott (1992) have shown that the derived stochastic process on capital stock is stationary and that there exists a corresponding ergodic probability distribution which captures its long run behavior.

Let us first explore this model as regards characterization I.

**Proposition 3.1:** For Model (1), specialized to imply (11) and (12), the equity price and dividend series are mean-averting by characterization I.\(^{13}\)

**Proof:** The price relationship follows from a double application of Jensen’s inequality:

\[^{12}\] This identification of the dividend assumes that investment comes out of capital’s share.
\[^{13}\] This result was first presented in Basu and Vinod (1994) and Basu and Samanta (2001) in a slightly less general setting. We extend their explorations with a different goal in mind.
$$\text{cov}(\tilde{p}_{t-1}, \tilde{p}_t) = \text{cov}(\tilde{k}_t, \tilde{k}_{t+1}) = \text{cov}(\tilde{k}_t, \alpha \beta \tilde{k}_t \tilde{\lambda}_t)$$

$$= \alpha \beta E(\tilde{\lambda}_t) \left\{ E(\tilde{k}_t^{1+\alpha}) - E(\tilde{k}_t) E(\tilde{k}_t^\alpha) \right\}$$

$$> \alpha \beta E(\tilde{\lambda}_t) \left\{ E(\tilde{k}_t^{1+\alpha}) - E(\tilde{k}_t) \left( E(\tilde{k}_t) \right)^\alpha \right\}$$

$$= \alpha \beta E(\tilde{\lambda}_t) \left\{ E(\tilde{k}_t^{1+\alpha}) - \left( E(\tilde{k}_t) \right)^{1+\alpha} \right\} > 0 .$$

By a similar derivation, $\text{cov}(d_t, d_{t+1}) > 0 . \quad \blacksquare$

Equity prices are thus unambiguously mean averting under characterization I. This result confirms that the notions of mean reversion (under I) and stationarity (the existence of a long run ergodic probability distribution on capital stock – the equity price – to which the economy converges) are fundamentally different, and that the distinction arises in the simplest equilibrium macroeconomic models. By Theorem 2.1, equity prices are similarly mean averting under Property II. Furthermore, these observations are generic in the sense expressed by the following result:

**Proposition 3.2:** Consider any equilibrium model of the form (1) for which the equilibrium investment function $g(k, \lambda)$ is continuous and increasing in the capital stock. Suppose also that the price of equity and the level of the capital stock coincide (no costs of adjustment). Then the price of equity will be mean averting.

**Proof:** Direct application of the FKG inequality (Fortunin et al. (1971)). See Appendix 2. $\blacksquare$

Many of the macroeconomic models to be considered in this paper satisfy the conditions of the above proposition. In light of Proposition 3.2, if theory has much to say about economic reality, it is not particularly surprising that empirical studies have found, at best, weak evidence of mean reversion in equity prices (see the excellent discussion and literature review in Spierdijk and Bikker (2012)).
We next make statements regarding mean reversion in the equity return series for this model.

**Proposition 3.3**: For Model (1), specialized by (11) and (12)

\[
\text{corr}(\tilde{r}_t, \tilde{r}_{t+1}) \leq 0; \text{ i.e., equity returns are mean reverting by }
\]

Property I. The same is true of the equity premium.

**Proof**:

\[
\text{cov} \left( \tilde{r}_t, \tilde{r}_{t+1} \right) = \text{cov} \left( \alpha \hat{k}_{t-1} \tilde{\lambda}_t, \alpha \left[ \alpha \beta \hat{k}_{t} \tilde{\lambda}_t \right]^{a-1} \tilde{\lambda}_{t+1} \right)
\]

\[
= \alpha^2 (\alpha \beta)^{a-1} E \left( \tilde{\lambda}_{t+1} \right) \left( E \left( \tilde{g}_{t}^{a-1} \tilde{\lambda}^a_t \right) - E \left( \tilde{g}_{t}^{a-1} \tilde{\lambda}^a_t \right) \right)
\]

\[
= \alpha^2 (\alpha \beta)^{a-1} E \left( \tilde{\lambda}_{t+1} \right) \left( E \left( \tilde{g}_{t}^{a-1} \right) E \left( \tilde{\lambda}^a_t \right) - E \left( \tilde{g}_{t}^{a-1} \right) E \left( \tilde{\lambda}^a_t \right) \right).
\]

(15)

We wish first to explore the following constituents of expression (15):

\[
E \left( \tilde{g}_{t}^{a-1} \right) \text{ vs. } E \left( \tilde{g}_{t}^{a-1} \right) E \left( \tilde{g}_{t}^{a-1} \right) \text{ vs. } E \left( \tilde{g}_{t}^{a-1} \right) E \left( \tilde{g}_{t}^{a-1} \right).
\]

These expressions are of the general form

\[
E \left( \tilde{g}_{t}^{\gamma_0+\gamma_1} \right) \text{ and } E \left( \tilde{g}_{t}^{\gamma_0} \right) E \left( \tilde{g}_{t}^{\gamma_1} \right)
\]

where \( \gamma_0 < 0, \gamma_1 < 0 \). Define \( x = \tilde{g}_{t}^{\gamma_0} \), and \( g(x) = x^{(\gamma_1/\gamma_0)} \).

Since \( (\gamma_1 / \gamma_0) > 0 \), \( g(x) \) is an increasing function of \( x \), and \( g(x) = \tilde{g}_{t}^{\gamma_1} \).

Thus, \( E \left( \tilde{g}_{t}^{\gamma_0+\gamma_1} \right) = E \left( \tilde{x} g(\tilde{x}) \right) > E \left( \tilde{x} \right) E \left( g(\tilde{x}) \right) = E \left( \tilde{g}_{t}^{\gamma_0} \right) E \left( \tilde{g}_{t}^{\gamma_1} \right) \) by the FKG inequality. Accordingly,

\[
E \left( \tilde{g}_{t}^{a-1} \right) \geq E \left( \tilde{g}_{t}^{a-1} \right) E \left( \tilde{g}_{t}^{a-1} \right).
\]

We may thus conclude that expression (14)

\[
< \alpha^2 (\alpha \beta)^{a-1} \left( E \left( \tilde{\lambda}_{t+1} \right) \left( E \left( \tilde{\lambda}^a_t \right) - E \left( \tilde{\lambda}^a_t \right) \right) \right)
\]

\[
\leq \alpha^2 (\alpha \beta)^{a-1} \left( E \left( \tilde{\lambda}_{t+1} \right) \left( E \left( \tilde{\lambda}^a_t \right) - E \left( \tilde{\lambda}^a_t \right) \right) \right)
\]

12
by Jensen’s inequality since $g(\lambda) = \lambda^\alpha$, $0 < \alpha < 1$ is a concave function of $\lambda$.

$$\leq \alpha^2 (\alpha \beta)^{\alpha-1} E(\tilde{\lambda}_{t+1}) \left\{ \left( E\left( \tilde{\lambda}_t^\alpha \right) \right)^\alpha - E\left( \tilde{\lambda}_t \right) \left( E\left( \tilde{\lambda}_t \right) \right)^{\alpha-1} \right\},$$

again by Jensen’s inequality, since $g(\lambda) = \lambda^{\alpha-1}$ is a convex function of $\lambda$.

$$= \alpha^2 (\alpha \beta)^{\alpha-1} E(\tilde{\lambda}_{t+1}) \left\{ \left( E\left( \tilde{\lambda}_t^\alpha \right) \right)^\alpha - E\left( \tilde{\lambda}_t \right) \left( E\left( \tilde{\lambda}_t \right) \right)^{\alpha-1} \right\} = 0.14$$

Clearly it is the concavity of the production function $(\alpha - 1) < 0$ that gives mean reversion (defined by I) in equity returns, a fact first observed in Basu and Vinod (1994).

Propositions 3.1 – 3.3 give cause for reflection. If the notion of mean reversion is intended to capture the property that above average values of a stochastic process must regularly be followed by below average values, then all the series considered thus far, $\{p_t^c\}$, $\{d_t\}$, $\{r_t^c\}$, and $\{r_t^p\}$ qualify: each follows a stationary stochastic process that converges weakly and each possesses a unique, irreducible ergodic set. Accordingly, the commonplace definitions of mean reversion in the literature (properties I – II) seem to discriminate artificially: both $\{r_t^c\}$ and $\{p_t^c\} = \{k_t\}$ revert to their respective means, yet behave inconsistently as regards definitions I and II.

3.1. An Alternative Measurement

These observations suggest that “mean reversion,” at least as it is often characterized, is not an especially informative equilibrium concept. For every model considered in this paper, the equilibrium series $\{p_t^c\}$, $\{r_t^c\}$, $\{r_t^f\}$ and $\{r_t^p\}$ are all ergodic irreducible stochastic processes. Accordingly, if any of these series

14 We thank Awi Federgruen for bringing the FKG inequality to our attention.
assumes values above its mean value, this experience must inevitably be reversed in time. Yet, for the simple models considered in this paper, none of our three mean reversion characterizations is uniformly satisfied.

Thus we are in a sense back to square one: what is this property we call mean reversion? Informally we think of a mean reverting series as one whose value “crosses its mean ‘fairly frequently’.” As we have shown, the requirement that \( \text{corr}(x_t, x_{t+1}) < 0 \) is, however, too restrictive to be exclusively identified with this concept of mean reversion. A less restrictive notion is needed. Implicit in this comment is the desire also to have a simple measure by which one series can be said to be more highly mean reverting than another.

Let us recall to this discussion an old notion of a stochastic process’s average time to crossing (ACT): the average number of periods for which a stochastic process uniformly exceeds or uniformly falls short of its mean value. Under this concept an economic time series is mean reverting if and only if its ACT is finite. Within the family of models under consideration in this paper, it is also natural to say that a stochastic series \( \{\hat{x}_t\} \) is more highly mean reverting than a stochastic series \( \{\hat{y}_t\} \) if and only if the ACTs \( \text{ACT}_{\{\hat{x}_t\}} < \text{ACT}_{\{\hat{y}_t\}} \). Below we list the ACTs for the model considered in the earlier sections of this paper.

### Table 3.1

**Model 1: Mean Crossing Times**<sup>(i)</sup>

\[ \alpha = .36, \beta = .96 \text{ (annual)}, \sigma_t^2 = .00712 \]

\[
\begin{array}{cccccc}
\{p_t\} & \{d_t\} & \{r_t^e\} & \{r_t^f\} & \{r_t^p\} \\
\rho = 0 & 2.47 & 2.47 & 1.64 & 2.61 & 1.74 \\
\end{array}
\]

<sup>(i)</sup> ACTs for this panel are measured in years
The outcomes presented in Table 3.1 are both specific to this model and
generic to more general versions. The fact that the MCT for the dividend and the
price of equity is the same follows from the fact that the former is proportional to
the latter. The fact of the relatively shorter MCTs for \( \{ r^e_i \} \) and \( \{ r^p_i \} \) follows from
their comparatively high volatility vis-à-vis \( \{ k_i \} \); this property is widely observed
in more general versions. By the same token (and with similar generality) the risk
free rate series is relatively more slow moving with a resulting higher MCT.

Model (1) falls short, however, of a full-fledged business cycle model on
many dimensions. In particular, none of the aggregate series is sufficiently
persistent vis-à-vis the data. In the next section we remedy this particular
shortcoming and explore the consequences.

4. Adding Persistence in the Productivity Disturbance

Consistent with the macroeconomics DSGE literature in this section we
slightly specialize the production technology to be of the form \( y_t = \left( \hat{k}_t \right)^{\alpha} e^{\hat{\lambda}} \) and
compare two cases: \( \hat{\lambda}_t \sim N(0, \sigma^2_{\hat{\lambda}}) \), i.i.d. and \( \hat{\lambda}_{t+1} = \rho \hat{\lambda}_t + \hat{\varepsilon}_t \) where \( \hat{\varepsilon}_t \) is i.i.d.,
\( \hat{\varepsilon}_t \sim N(0, \sigma^2_{\varepsilon}) \). In either case the decision rules take the same form as (11) – (12).

We rely on numerical simulations of (11), (12) to give some measure of the
relevant magnitudes. The results are presented in Table 4.1 below.
Table 4.1
Model 2: Autocorrelations, $p^e_t, r^e_t$.

\[ \ddot{y}_t = k_t^e e^{\delta_t} \]
\[ \ddot{\lambda}_{t+1} = \rho \lambda_t + \ddot{\xi}_t, \ddot{\xi}_t \sim N(0, \sigma^2_{\ddot{\xi}}), \]

Panel A
First order autocorrelations: various $\rho$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\rho = 0$</th>
<th>$\rho = .2$</th>
<th>$\rho = .4$</th>
<th>$\rho = .6$</th>
<th>$\rho = .8$</th>
<th>$\rho = .95$</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr($p^e_t, p^e_{t+1}$)</td>
<td>.36</td>
<td>.52</td>
<td>.66</td>
<td>.79</td>
<td>.90</td>
<td>.97</td>
</tr>
<tr>
<td>corr($r^e_t, r^e_{t+1}$)</td>
<td>-.31</td>
<td>-.18</td>
<td>-.04</td>
<td>.09</td>
<td>.23</td>
<td>.33</td>
</tr>
<tr>
<td>corr($r^p_t, r^p_{t+1}$)</td>
<td>-.20</td>
<td>-.06</td>
<td>.08</td>
<td>.20</td>
<td>.31</td>
<td>.36</td>
</tr>
</tbody>
</table>

Panel B
Autocorrelations at various lags, $\rho = .95$

<table>
<thead>
<tr>
<th>$j$</th>
<th>$j = 1$</th>
<th>$j = 2$</th>
<th>$j = 3$</th>
<th>$j = 5$</th>
<th>$j = 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr($p^e_t, p^e_{t+j}$)</td>
<td>.97</td>
<td>.93</td>
<td>.88</td>
<td>.79</td>
<td>.71</td>
</tr>
<tr>
<td>corr($r^e_t, r^e_{t+j}$)</td>
<td>.33</td>
<td>.08</td>
<td>.00</td>
<td>-.04</td>
<td>-.04</td>
</tr>
<tr>
<td>corr($r^p_t, r^p_{t+j}$)</td>
<td>.36</td>
<td>.11</td>
<td>.04</td>
<td>-.00</td>
<td>-.00</td>
</tr>
</tbody>
</table>

(i) For these simulations, $\alpha = .36, \beta = .99$, and $\sigma^2_{\ddot{\xi}} = .00712$ typical choices in the macro literature. In more complex DSGE models, $\sigma^2_{\ddot{\xi}}$ is chosen to allow the SD of detrended output in the model to match its counterpart in the data.

(ii) Statistics based on time series of length 100,000.

While the results of Table 4.1–Panel A certainly respect the implications of Propositions 3.1 and 3.3 for the $\rho = 0$ case, the conclusion is not robust: adding sufficient persistence to the random productivity disturbance unexpectedly causes the return series to become mean averting. We formalize this observation in the following proposition:
Proposition 4.1: Consider the model defined by (11), (12) with shock process and production technology specialized to $y_t = k_t^\alpha e^{\hat{\lambda}_t}$ where $\hat{\lambda}_{t+1} = \rho \lambda_t + \xi_{t+1}$, $\{\xi_{t+1}\}$ i.i.d. $N(0, \sigma^2)$. Then a sufficient condition for $\{\tilde{r}_t^c\}$ to be mean averting by Definition I is that $\alpha + \rho > 1$.

Proof:

\[
\text{cov} \left( \tilde{r}_t^c, \tilde{r}_{t+1}^c \right) = \text{cov} \left( \alpha \tilde{k}_t^{\alpha-1} e^{\hat{\lambda}_t}, \alpha \left[ \alpha \beta \tilde{k}_t^{\alpha} e^{\hat{\lambda}_t} \right]^{\alpha-1} e^{\rho \hat{\lambda}_t + \xi_{t+1}} \right)
\]

\[
= \alpha^2 (\alpha \beta)^{\alpha-1} \text{cov} \left( \tilde{k}_t^{\alpha-1} e^{\hat{\lambda}_t}, \tilde{k}_t^{\alpha-2} e^{\hat{\lambda}_t (\alpha+\rho) + \xi_{t+1}} \right)
\]

\[
= \alpha^2 (\alpha \beta)^{\alpha-1} \left\{ E \left( \tilde{k}_t^{\alpha-1} e^{\hat{\lambda}_t (\alpha+\rho) + \xi_{t+1}} \right) \right\}
\]

\[
= \alpha^2 (\alpha \beta)^{\alpha-1} \left\{ E \left( \tilde{k}_t^{\alpha-1} \right) E \left( e^{\hat{\lambda}_t (\alpha+\rho) + \xi_{t+1}} \right) \right\}
\]

\[
> \alpha^2 (\alpha \beta)^{\alpha-1} E \left( \tilde{k}_t^{\alpha-1} \right) \left\{ E \left( e^{\hat{\lambda}_t (\alpha+\rho) + \xi_{t+1}} \right) \right\}
\]

since Proposition 3.3 \( E \left( \tilde{k}_t^{\alpha-1} \right) > E \left( \tilde{k}_t^{\alpha-2} \right) E \left( \tilde{k}_t^{\alpha-1} \right) \)

\[
= \alpha^2 (\alpha \beta)^{\alpha-1} E \left( \tilde{k}_t^{\alpha-1} \right) E \left( \tilde{k}_t^{\alpha-2} \right) \left\{ E \left( e^{\hat{\lambda}_t (\alpha+\rho) + \xi_{t+1}} \right) \right\}
\]

\[
= \alpha^2 (\alpha \beta)^{\alpha-1} E \left( \tilde{k}_t^{\alpha-1} \right) E \left( \tilde{k}_t^{\alpha-2} \right) \left\{ E \left( e^{\hat{\lambda}_t (\alpha+\rho) + \xi_{t+1}} \right) \right\}.
\]

With the first three terms all positive, the sign of $\left( r_t^c, r_{t+1}^c \right)$ is determined by the right most term.
Since \( \varepsilon_t \sim N(0, \sigma^2) \) for all \( t \), \( \tilde{\lambda}_t \) is normal with some mean and variance \( \hat{\mu}_t \) and \( \hat{\sigma}^2_t \).

Accordingly, the right-most term can be expressed as

\[
\left\{ e^{(\alpha+\rho)\hat{\mu}_t + \frac{1}{2}(\alpha+\rho)^2\hat{\sigma}^2_t} - e^{\frac{1}{2}\sigma^2_t}\left( e^{(\alpha+\rho)\hat{\mu}_t + \frac{1}{2}(\alpha+\rho)^2\sigma^2_t} \right) \right\} \\
= e^{(\alpha+\rho)\hat{\mu}_t} \left\{ e^{\frac{1}{2}(\alpha+\rho)^2\sigma^2_t} - e^{\frac{1}{2}\sigma^2_t}\left( e^{\frac{1}{2}(\alpha+\rho)^2\sigma^2_t} \right) \right\} \\
= e^{(\alpha+\rho)\hat{\mu}_t} \left\{ e^{\frac{1}{2}(\alpha+\rho)^2\sigma^2_t} - e^{\frac{1}{2}(\alpha+\rho)^2}\left( e^{\frac{1}{2}(\alpha+\rho)^2\sigma^2_t} \right) \right\} > 0.
\]

since \(-2(\alpha + \rho) + 2 < 0.\)

Proposition 4.1 introduces persistence in the productivity disturbances into Model 1 in a simple way typical of the DSGE literature. The conclusion is that if these productivity disturbances are sufficiently persistent, returns will be mean averting. As a consequence, if a model of this sort even is to come close to matching the observed persistence in output, equity returns will surely be mean averting at least by Definition I. (As per Proposition 2.1, returns will be similarly mean averting by Property II for sufficiently high persistence.) Since Proposition 4.1 is not an equivalence, however, mean aversion is not necessarily observed for any \( \rho \) such that \( \alpha + \rho \geq 1 \). If \( \alpha = .36 \), for example, the passage from mean reverting to mean averting equity returns roughly occurs at \( \rho \approx .48 \), yielding \( \text{cov}(\tilde{r}_t^c, \tilde{r}_{t+1}^c) \approx .0051 \).

None of these results is surprising in the least: the process on the disturbance component \( \{\lambda_t\} \) is itself highly mean averting as the following proposition makes clear. Cogley and Nason (1995) emphasize the close relationship of the productivity process to the derived properties of DSGE models’ state variables.
Proposition 4.2: Consider a stochastic process of the form
\[ \tilde{x}_t = \rho \tilde{x}_{t-1} + \tilde{\varepsilon}_t, \]
where \( \{\tilde{\varepsilon}_t\} \) is i.i.d. with variance \( \sigma^2 \). Define a new
stochastic process by
\[ \tilde{\lambda}_t = e^{\tilde{x}_t}. \]
Then,
\[ \text{cov}(\tilde{\lambda}_t, \tilde{\lambda}_{t+1}) = e^{\\left[\frac{\sigma^2(1+\rho^2+1)}{2(1-\rho^2)}\right]} \left( e^{\left[\frac{-\rho^2}{1-\rho^2}\right]} - 1 \right) \]
and, if \(-1 < \rho < 1\),
\[ \text{cov}(\tilde{\lambda}_t, \tilde{\lambda}_{t+1}) \begin{cases} > 0 & \text{if } \rho > 0 \\ < 0 & \text{if } \rho < 0 \end{cases}. \]

Proof: See the Appendix.

By Proposition 4.2, the shock process to our production technology is mean
reverting by Property I only if \( \rho < 0 \), an attribution antithetical to its empirical
counterpart, the Solow residual, and this feature drags mean aversion into equity
returns as well. In this light, it would be useful to understand what additional
model features allow high persistence in aggregate series (as the data reveals) to
be compatible with mean reversion in equity returns and the equity premium. For
example, Guvenen (2009) reports \( \text{corr}(r^e_t, r^e_{t+j}) = -0.03, -0.03, -0.02, -0.02, -0.02 \) for
lags, respectively, of \( j = 1, 2, 3, 5 \) and 7 years.

It remains to see if equity prices or returns are mean reverting by
Definition III. Following the protocol of the prior sections we first explore the
case where \( \{\tilde{\lambda}_t\} \) are i.i.d. Proposition 2.2 informs us that the \( \text{var}(k_t) \) and the
\( \text{var}(r^e_t) \) will be the critical factors in the analysis. By (14), the expression for the
variance of the equity price becomes:
\[
\text{var } \tilde{p}_t^* = \text{var } \tilde{k}_t = \left[ (\alpha \beta)^{1 + \alpha + \alpha^2 + \ldots + \alpha^{t-1}} \right]^2 \left[ k_0^{\alpha} \right]^2 \text{var } \left( \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha(s-1)} \right)
\]

where \( \text{var } \left( \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha(s-1)} \right) = \text{var } \left( \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha} \right) \) by the i.i.d. assumption on the \( \{ \tilde{\lambda} \} \). By the same analysis,

\[
\text{var } r_t^* = \text{var } \alpha k_0^{\alpha-1} \lambda_t , \text{ where }
\]

\[
r_t^* = \alpha \left( \frac{\alpha \beta}{1 - \alpha} \right)^{1-\alpha} k_0^{\alpha} \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha} \tilde{\lambda} - 1
\]

\[
= \alpha \tilde{\lambda} \left( \frac{\alpha \beta}{1 - \alpha} \right)^{-1} k_0^{\alpha-1} \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha(s-1)} - 1.
\]

Accordingly,

\[
\text{var } \tilde{r}_t^* = \alpha^2 E \tilde{\lambda}_t^2 \left[ \left( \frac{\alpha \beta}{1 - \alpha} \right)^{1-\alpha} k_0^{\alpha} \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha(s-1)} \right]^2 \left[ \prod_{s=0}^{t-1} E \left( \tilde{\lambda}^{\alpha(s-1)} \right) \right]^2
\]

\[- \alpha^2 \left( E \tilde{\lambda}_t \right)^2 \left[ \left( \frac{\alpha \beta}{1 - \alpha} \right)^{1-\alpha} k_0^{\alpha} \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha(s-1)} \right]^2 \left[ \prod_{s=0}^{t-1} E \left( \tilde{\lambda}^{\alpha(s-1)} \right) \right]^2 \]

\[
= \alpha^2 \left[ \left( \frac{\alpha \beta}{1 - \alpha} \right)^{-2} k_0^{2\alpha} \prod_{s=0}^{t-1} \tilde{\lambda}^{\alpha(s-1)} \right] \left[ \prod_{s=0}^{t-1} E \left( \tilde{\lambda}^{\alpha(s-1)} \right) \right]^2
\]

\[- \prod_{s=0}^{t-1} E \left( \tilde{\lambda}^{\alpha(s-1)} \right)^2 \right],
\]

\[(17)\]

We also continue to specialize the productivity to be of the business-cycle-literature-inspired form \( \{ e^\lambda \} \), where \( \tilde{\lambda}_{t+1} = \rho \lambda_t + \tilde{\varepsilon}_{t+1}, \{ \tilde{\varepsilon}_t \} \) i.i.d. \( N(0, \sigma^2_\varepsilon) \). For Property III, we have a combination of analytical and numerical results.

**Proposition 4.3**: Consider Model 1 specialized to (11) and (12), with shock process as per above but with \( \rho = 0 \). Then,

\[\text{var } (\tilde{x} \tilde{y}) = (E \tilde{x}^2) E \tilde{y}^2 - (E \tilde{x})^2 (E \tilde{y})^2\]

---

\[\text{Here we use the following property of independent random variables:}\]

\[\text{var } (\tilde{x} \tilde{y}) = (E \tilde{x}^2) E \tilde{y}^2 - (E \tilde{x})^2 (E \tilde{y})^2\]
(i) The capital price series and dividend series are mean averting by Property III;
(ii) The return on equity and equity premium series are mean reverting by Property III.
Proof: See the Appendix.
These results are entirely consistent with those obtained for our earlier analysis using Properties I (and, by implication, Property II).
A check of the proof of Proposition 4.3 reveals that concavity in production ($\alpha < 1$) is, once again, the overriding guarantor of mean reversion, though in a somewhat indirect way arising as it does not through the productivity disturbance but through the repeated influence of the capital share and discount factor parameters on capital accumulation.
As shock persistence increases ($\rho > 0$), however, our earlier results are confounded: Table 4.2 summarizes the results of extensive numerical simulations that compute $\text{corr}\left(\ddot{x}_s - \ddot{x}_r, \ddot{x}_u - \ddot{x}_r\right)$ for a wide class of \{r, s, t, u\} where $r < s < t < u$, and $\{\ddot{x}_i\}$ chosen from $\{r^s, \{r^p\}, \{p^s\}, \{d\}$.
### Table 4.2

**Correlations: Various Series**

Simulation Results for $s = r+i$, $t = s+j$, $u = t+k$

$i, j, k \in \{1, 2, 3\}$

$\alpha = .36, \beta = .96, \sigma^2 = .00712, \rho = .95$

<table>
<thead>
<tr>
<th>Series Correlation</th>
<th>Range of Values across all $i, j, k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) $\text{corr}(p_{s}^c - p_{r}^c, p_{u}^c - p_{t}^c) &lt; 0$</td>
<td>$(-.08, +.07)$</td>
</tr>
<tr>
<td>(ii) $\text{corr}(r_{s}^c - r_{r}^c, r_{u}^c - r_{t}^c) &lt; 0$</td>
<td>$(-.18, .00)$</td>
</tr>
<tr>
<td>(iii) $\text{corr}(r_{s}^p - r_{r}^p, r_{u}^p - r_{t}^p) &lt; 0$</td>
<td>$(-.15, .00)$</td>
</tr>
</tbody>
</table>

The switch (from the conclusions of Proposition 3.5) occurs at around $\rho = .6$ for $\alpha = .36, \beta = .96$, and $\sigma^2 = .00712$, our customary business cycle parameters. While these results are consistent with those concerning $\{r_{t}^c\}$ and $\{r_{t}^p\}$ for Properties I and II, they are distinctly different for the $\{p_{t}^c\}$ and $\{\hat{d}_{t}\}$ series, a fact that accounts for our earlier comment that Property III represents a fundamentally different measurement from either Property I or II.

What is the effect of shock persistence on MCT? For the model of Table 4.2, we observe Table 4.3, which is simply an expanded version of Table 3.1.
Table 4.3

Model 1: Average Crossing Time for Various Parameter Values \(^{(i)}\)

\[ \alpha = .36, \beta = .96 \text{ (annual), } \sigma_t^2 = .00712 \]

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( { p'_t } )</th>
<th>( { d_t } )</th>
<th>( { r'_t } )</th>
<th>( { r_t } )</th>
<th>( { r^p_t } )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.47</td>
<td>2.47</td>
<td>1.64</td>
<td>2.61</td>
<td>1.74</td>
</tr>
<tr>
<td>.95</td>
<td>14.10</td>
<td>14.10</td>
<td>2.54</td>
<td>2.60</td>
<td>2.60</td>
</tr>
</tbody>
</table>

\(^{(i)}\) ACTs for this panel are measured in years

Not surprisingly, the MCTs for all the reported series increase with greater shock persistence, with the greatest increase among the slow moving series \( \{ p'_t \} \) and \( \{ d_t \} \). Indeed (and obviously) these series are “crossing their mean values” on a regular basis despite their all being mean averting by Property I. This observation suggests to us that for economic models of the sort considered here, MCT is the “better” measurement in the sense of giving a more intuitive picture of what is going on.

Yet, there must be some indirect connection of Property I to the MCT concept. One possibility is suggested by:

**Proposition 4.4:**

Consider \( \{ \hat{x}_t \} \), \( \{ \hat{y}_t \} \) as stochastic processes related to the model specialized at the start of Section 4, but corresponding to distinct parameter choices (e.g., different values of \( \rho \)). Then

\[ \text{corr}(x_t, x_{t+1}) > \text{corr}(y_t, y_{t+1}) \]

if and only if \( MCT_{\{x_t\}} > MCT_{\{y_t\}} \).

**Proof:** To be supplied.

We close Section 4 with a short summary of what we have learned and it is this: First, persistence in the productivity disturbance generically overturns the
results where there is independence. Consistent with all three of our definitions, equity returns appear necessarily to be mean reverting only in the presence of low persistence productivity disturbances. While it is already well known that mean reversion in returns is compatible with mean aversion in prices (Spierdijk and Bikker (2012)), it is surprising to find this compatibility in such an economically simple context. Proposition 4.1 further suggests that this particular phenomenon is well likely to be pervasive across many DSGE formulations, implying that the search for mean reversion in equity returns and the premium is unlikely to be fruitful – if the present family of models has anything to say about actual economies. To say it differently, we find it unsurprising that evidence for mean revisersion in equity returns and the premium is weak (see Table 4.4).

**Table 4.4**

**First Order Autocorrelations: Annual Returns on the S&P_{500}, Various Historical Periods**

<table>
<thead>
<tr>
<th>Historical Period</th>
<th>corr($r_t^{S_{500}}, r_{t+1}^{S_{500}}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900 – 2014</td>
<td>-.011</td>
</tr>
<tr>
<td>1900 – 2012</td>
<td>-.011</td>
</tr>
<tr>
<td>1926 – 1996</td>
<td>.289</td>
</tr>
<tr>
<td>1952 – 2006</td>
<td>.276</td>
</tr>
<tr>
<td>1952 – 2014</td>
<td>-.088</td>
</tr>
</tbody>
</table>

Second, as clearly demonstrated by our example, weak Markovian stationarity and mean reversion in economic time series as commonly characterized (Properties I and II) are not equivalent concepts; in particular, the former does not imply the latter, a fact that is demonstrated here in the simplest possible macro model (in the case, namely, of the capital stock price).
5. Conclusion

We close by revisiting our objectives going forward. The results presented in Table 4.1 clearly demonstrate an absence of mean reversion – as it is conventionally defined (Property 1) for high persistence productivity disturbances.\textsuperscript{16} It is also the case, however, that the elaborate constructs of Guvenen (2009) and Bansal and Yaron (2004) do yield mild mean reversion by Property I (near zero but negative autocorrelations in $\{\tilde{r}_t^c\}$ and $\{\tilde{r}_t^p\}$) even when persistence in the productivity shock is high. What accounts for the difference? It is to this topic that we turn in a companion paper. In particular, we explore a number of additional model features and assess their cumulative contributions in generating mean reversion in equity returns.\textsuperscript{17}

\textsuperscript{16} From a business cycle perspective, a reasonable matching of the stylized facts demands high persistence in the productivity disturbance (see Cogley and Nason (1995)). It is in this sense that the high persistence cases are the only ones of empirical relevance.

\textsuperscript{17} All the while we must restrict our explorations to models which simultaneously are able to replicate the most basic stylized business cycle facts.
References


Appendix 1

A. Proof of Proposition 2.1

a. Property II $\Rightarrow$ Property I

By Property II,
\[ \text{var}(\hat{x}_t + \hat{x}_{t+1}) = \text{var}(\hat{x}_t) + \text{var}(\hat{x}_{t+1}) + 2\text{cov}(\hat{x}_t, \hat{x}_{t+1}) < 2\text{var}(\hat{x}_t) \]
\[ = \text{var}(\hat{x}_{t+1}) + 2\text{cov}(\hat{x}_t, \hat{x}_{t+1}) < \text{var}(\hat{x}_t). \]

Since $\text{var}(\hat{x}_{t+1}) = \text{var}(\hat{x}_t)$,
\[ \text{cov}(\hat{x}_t, \hat{x}_{t+1}) < 0, \]
and Property I holds.

b. Property I together with the covariance condition (9) implies Property II. The proof is by induction.

Let $j = 1$, $\text{var}(\hat{x}_t + \hat{x}_{t+1}) = \text{var}(\hat{x}_t) + \text{var}(\hat{x}_{t+1}) + 2\text{cov}(\hat{x}_t, \hat{x}_{t+1}) < 2\text{var}(\hat{x}_t)$ by Property I and the fact that $\text{var}(\hat{x}_t) = \text{var}(\hat{x}_{t+1})$

Let $j = 2$, $\text{var}(\hat{x}_t + \hat{x}_{t+1} + \hat{x}_{t+2})$
\[ = \sum_{j=0}^{2} \text{var}(\hat{x}_{t+j}) + 2\text{cov}(\hat{x}_t, \hat{x}_{t+1}) + 2\text{cov}(\hat{x}_t, \hat{x}_{t+2}) + 2\text{cov}(\hat{x}_{t+1}, \hat{x}_{t+2}) \]
by Property I
\[ 2\text{cov}(\hat{x}_t, \hat{x}_{t+1}) < 0 \] and by condition (9),
\[ 2\text{cov}(\hat{x}_t, \hat{x}_{t+1}) + 2\text{cov}(\hat{x}_t, \hat{x}_{t+2}) < 0 \]
by Property I, $2\text{cov}(\hat{x}_{t+1}, \hat{x}_{t+2}) < 0$. Therefore,
\[
\text{var}(\tilde{x}_t + \tilde{x}_{t+1} + \tilde{x}_{t+2}) < \sum_{j=0}^{2} \text{var}(\tilde{x}_{t+j}) = 3 \text{var}(\tilde{x}_t).
\]

Suppose, by Property I and condition (9), that
\[
\text{var}(\tilde{x}_t + \tilde{x}_{t+1} + ... + \tilde{x}_{t+j-1}) < j \text{var}(\tilde{x}_t).
\]

To show \(\text{var}(\tilde{x}_t + \tilde{x}_{t+1} + ... + \tilde{x}_{t+j}) < (j+1)\text{var}(\tilde{x}_t)\).

\[
\begin{align*}
\text{var}(\tilde{x}_t + \tilde{x}_{t+1} + \tilde{x}_{t+2} + ... + \tilde{x}_{t+j}) &= \text{var}(\tilde{x}_t + \tilde{x}_{t+1} + ... + \tilde{x}_{t+j-1}) + \text{var}(\tilde{x}_{t+j}) \\
&\quad + 2\sum_{s=0}^{j-2} \text{cov}(\tilde{x}_{t+s}, \tilde{x}_{t+j}) + 2\text{cov}(\tilde{x}_{t+j-1}, \tilde{x}_{t+j}) \\
&< (j)\text{var}(\tilde{x}_t) + \text{var}(\tilde{x}_{t+j}) \\
&\quad + 2\sum_{s=0}^{j-2} \text{cov}(\tilde{x}_{t+s}, \tilde{x}_{t+j}) + 2\text{cov}(\tilde{x}_{t+j-1}, \tilde{x}_{t+j}) \\
&< (j+1)\text{var}(\tilde{x}_t)
\end{align*}
\]

(by induction)

by Property I, \(\text{cov}(\tilde{x}_{t+j-1}, \tilde{x}_{t+j}) < 0\).

Thus by condition (9),
\[
+ 2\sum_{s=0}^{j-2} \text{cov}(\tilde{x}_{t+s}, \tilde{x}_{t+j}) + 2\text{cov}(\tilde{x}_{t+j-1}, \tilde{x}_{t+j}) < 0.
\]

Therefore, since \(\text{var}(\tilde{x}_{t+j}) = \text{var}(\tilde{x}_t)\)

\[
\begin{align*}
\text{var}(\tilde{x}_t + \tilde{x}_{t+1} + ... + \tilde{x}_{t+j}) &< (j+1)\text{var}(\tilde{x}_t).
\end{align*}
\]

Basically, for I \(\Rightarrow\) II, the persistence of the series must rapidly decline.

B. Proof of Proposition 3.2

\[
\begin{align*}
\text{cov}(\tilde{k}_t, \tilde{k}_{t+1}) &= \text{cov}(\tilde{k}_t, g(\tilde{k}_t, \tilde{x}_t))
\end{align*}
\]
By FKG’s inequality, for every \( \lambda_i \) the conditional covariance \( \text{cov}(k_i, k_{i+1}; \lambda_i) \) is positive. Thus \( \text{cov}(\tilde{k}_i, \tilde{k}_{i+1}) \geq 0 \).

C. Proof of Proposition 4.2

Knowing that for \( y \sim N(0, V) \), \( E[\exp(y)] = \frac{y}{2} \) and for the AR(1) process,

\[
\text{Var}(x_t) = \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right);
\]

\[
\text{Cov}(\lambda_i, \lambda_{i+1}) = \text{Cov}(\exp(x_i), \exp(\rho x_i + \varepsilon_{i+1}))
\]

\[
= E[\exp(x_i) \exp(\rho x_i + \varepsilon_{i+1})] - E[\exp(x_i)]E[\exp(\rho x_i + \varepsilon_{i+1})]
\]

\[
= E[\exp((\rho + 1)x_i + \varepsilon_{i+1})] - E[\exp(x_i)]E[\exp(\rho x_i)]E[\exp(\varepsilon_{i+1})]
\]

\[
= E[\exp(\varepsilon_{i+1})][E[\exp((\rho + 1)x_i)] - E[\exp(x_i)]E[\exp(\rho x_i)]
\]

\[
= \exp \left( \frac{\sigma^2}{2} \right) \exp \left( \frac{1}{2} (\rho + 1)^2 \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) - \exp \left( \frac{1}{2} \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) \exp \left( \frac{1}{2} \rho^2 \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right)
\]

\[
= \exp \left( \frac{\sigma^2}{2} \right) \exp \left( \frac{1}{2} (\rho^2 + 2\rho + 1) \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) - \exp \left( \frac{1}{2} (\rho^2 + 1) \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right)
\]

\[
= \exp \left( \frac{\sigma^2}{2} \right) \exp \left( \frac{1}{2} (\rho^2 + 1) \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) \left[ \exp \left( \rho \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) - 1 \right]
\]

\[
= \exp \left( \frac{\sigma^2}{2} + \frac{1}{2} (\rho^2 + 1) \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) \left[ \exp \left( \rho \sigma^2 \left( \frac{1 - \rho^2}{1 - \rho^2} \right) \right) - 1 \right]
\]
\[
\text{Let } \{\lambda_t\} \text{ be i.i.d. with } E\lambda_t = \mu_\lambda \text{ and } \sigma^2_\lambda = \sigma^2_\lambda. \text{ We will inductively describe}
\]

\[
\text{var}\left(\prod_{s=0}^{t} \lambda_s\right) \text{ using the following relationship for independent random variables } \tilde{x}, \tilde{y}:
\]

\[
\text{var}(\tilde{x} \cdot \tilde{y}) = (\mu_x)^2 \sigma^2_y + (\mu_y)^2 \sigma^2_x + \sigma^2_x \sigma^2_y
\]

\[
\text{var}(\tilde{\lambda}_0 \cdot \tilde{\lambda}_t) = (\mu_{\lambda_0})^2 \sigma^2_{\lambda_t} + (\mu_{\lambda_0})^2 \sigma^2_{\lambda_0} + \sigma^2_{\lambda_0} \sigma^2_{\lambda_t}
\]

\[
= 2\mu_0^2 \sigma^2 + (\sigma_\lambda)^4
\]

\[
\text{var}(\tilde{\lambda}_0 \cdot \tilde{\lambda}_1 \cdot \tilde{\lambda}_2) = \text{var}\left((\tilde{\lambda}_0 \cdot \tilde{\lambda}_1) \cdot \tilde{\lambda}_2\right)
\]

\[
= (\mu_{\lambda_0} \cdot \mu_{\lambda_1})^2 \sigma^2_{\lambda_2} + (\mu_{\lambda_1})^2 \text{[var}\left(\tilde{\lambda}_0, \tilde{\lambda}_1\right)\text{]} + \text{[var}\left(\tilde{\lambda}_0, \tilde{\lambda}_1\right)\text{][var}\lambda_2]\]

\[
= (\mu_{\lambda_1})^4 \sigma^2_{\lambda_2} + (\mu_{\lambda_1})^2 + \sigma^2_{\lambda_2} \text{var}\left(\tilde{\lambda}_0 \tilde{\lambda}_1\right)
\]

\[
= (\mu_{\lambda_1})^4 \sigma^2_{\lambda_2} + (\mu_{\lambda_1})^2 + \sigma^2_{\lambda_2} \text{var}\left(\tilde{\lambda}_0 \tilde{\lambda}_1\right)
\]

\[
\text{Thus } \text{Cov}(\lambda_{t+1}) = \exp\left(\frac{\sigma^2}{2} \left(1 + (\rho^2 + 1) \left(\frac{1 - \rho^{2t}}{1 - \rho^2}\right)\right)\right) \exp\left(\rho \sigma^2 \left(\frac{1 - \rho^{2t}}{1 - \rho^2}\right)\right) - 1.
\]

Now, clearly, the first element is positive since it is an exponent. Further, since

\[
\sigma^2 \left(\frac{1 - \rho^{2t}}{1 - \rho^2}\right) > 0,
\]

we know that the second element is a strictly increasing function of \(\rho\), reaching a value of zero at \(\rho = 0\). Therefore, for \(\rho < 0\), the expression is negative, while for \(\rho > 0\) the expression is positive.
\[ \text{var}(\tilde{\lambda}_0, \tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3) = \text{var}((\tilde{\lambda}_0, \tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3)) \]
\[ = \left[ \mu_{\lambda} \mu_{\lambda} \mu_{\lambda} \right]^2 \sigma_{\lambda}^2 + \left( \mu_{\lambda} \right)^2 \cdot \text{var}(\tilde{\lambda}_0, \tilde{\lambda}_1, \tilde{\lambda}_2) \]
\[ = \left[ \left( \mu_{\lambda} \right)^2 \right]^2 \sigma_{\lambda}^2 + \text{var}(\tilde{\lambda}_0, \tilde{\lambda}_1, \tilde{\lambda}_2) \left[ \left( \mu_{\lambda} \right)^2 + \sigma_{\lambda}^2 \right] \]

This is sufficient to obtain the pattern:
\[ x_i = \text{var} \left( \prod_{s=0}^{t} \lambda_s \right) = \left( \mu_{\lambda} \right)^{2t} \sigma_{\lambda}^2 + \left[ \mu_{\lambda}^2 + \sigma_{\lambda}^2 \right] x_{t-1}. \]

2. Proof of Lemma 2.2

a. Suppose that \( (1 - \theta)^2 \left( 1 + \sigma_{\lambda}^2 \right) > 1 \).

\[ \text{var} k_i = \left( \left( 1 - \theta \right) A \right)^{2t} k_0^2 x_{t-1} \text{ where } x_i \text{ follows} \]
\[ x_i = \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right] x_{t-1} \]
\[ = \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right] \left\{ \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right] x_{t-2} \right\} \]
\[ = \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right] \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right]^2 x_{t-2} \]
\[ = \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right] \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right]^2 \left\{ \sigma_{\lambda}^2 + \left[ 1 + \sigma_{\lambda}^2 \right] x_{t-3} \right\} \]
\[ x_i = \sigma_{\lambda}^2 \sum_{s=0}^{t-1} \left[ 1 + \sigma_{\lambda}^2 \right]^s + \left[ 1 + \sigma_{\lambda}^2 \right] \sigma_{\lambda}^2 \text{ etc.} \]

Thus, \( x_i = \sigma_{\lambda}^2 \sum_{s=0}^{t-1} \left[ 1 + \sigma_{\lambda}^2 \right]^s \) and
\[ \text{var} k_i = \sigma_{\lambda}^2 k_0^2 \left( \left( 1 - \theta \right)^2 A^2 \right)^{t} \left\{ \sum_{s=0}^{t-1} \left[ 1 + \sigma_{\lambda}^2 \right]^s \right\}. \]

Suppose \( (1 - \theta)^2 A^2 \left( 1 + \sigma_{\lambda}^2 \right) > 1. \)

Then \( \text{var} k_i = \sigma_{\lambda}^2 k_0^2 \left( (1 - \theta)^2 A^2 \right)^{t} \left\{ \sum_{s=0}^{t-1} \left[ 1 + \sigma_{\lambda}^2 \right]^s \right\} \)
\[ \sigma^2 k_0^2 \left[ \left( 1 - \theta \right)^2 A^2 \right]^{t} \left[ 1 + \sigma^2 \right]^{t-1} \]

\[ = \sigma^2 k_0^2 \left[ \left( 1 - \theta \right)^2 A^2 \left( 1 + \sigma^2 \right) \right]^{t} \left[ \frac{1}{1 + \sigma^2} \right] \]

If \( \left( 1 - \theta \right)^2 A^2 \left( 1 + \sigma^2 \right) > 1 \), then \( \text{var} k \rightarrow \infty \) as \( t \rightarrow \infty \).

b. Suppose that \( \left( 1 - \theta \right)^2 A^2 \left( 1 + \sigma^2 \right) < 1 \).

As previously, \( \text{var} k = \sigma^2 k_0^2 \left[ \left( 1 - \theta \right)^2 A^2 \right]^{t} \left\{ \sum_{s=0}^{t-1} \left[ 1 + \sigma^2 \right]^{t} \right\} \)

\[ < \sigma^2 k_0^2 \left[ \left( 1 - \theta \right)^2 A^2 \right]^{t} \left( 1 + \sigma^2 \right)^t (t) \]

\[ = \sigma^2 k_0^2 \left[ \left( 1 - \theta \right)^2 A^2 \left( 1 + \sigma^2 \right) \right]^{t} (t) \]

\( \text{var} k \rightarrow 0 \) as \( t \rightarrow \infty \) provided.

\[ \sum_{s=0}^{t} \left[ \left( 1 + \sigma^2 \right)^2 A^2 \left( 1 + \sigma^2 \right) \right]^{s} (t) < \infty \]

Since all terms in the series are positive, we may employ the ratio test:

Let \( \text{RATIO}_t = \left[ (1 - \theta)^2 A^2 \left( 1 + \sigma^2 \right) \right]^{t+1} \left( t + 1 \right) \)

\[ \left[ (1 - \theta)^2 A^2 \left( 1 + \sigma^2 \right) \right]^{t} \left( t \right) \]

\[ = \left[ (1 - \theta)^2 A^2 \left( 1 + \sigma^2 \right) \right] \left( \frac{t+1}{t} \right) \]

\( \lim_{t \rightarrow \infty} \text{RATIO}_t = \left[ (1 - \theta)^2 A^2 \left( 1 + \sigma^2 \right) \right] < 1 \), and the sum above is finite.

c. Suppose that \( \left( 1 - \theta \right)^2 A^2 \left( 1 + \sigma^2 \right) = 1 \).
\[\text{var } k_i = \sigma^2 \lambda_0^2 \left[ (1 - \theta)^2 A^2 \right] \left\{ \sum_{s=0}^{t-1} \left[ 1 + \sigma^2 \right]^s \right\} \]

\[= \sigma^2 \lambda_0^2 \sum_{s=0}^{t-1} \left\{ \begin{array}{l}
\frac{(1 + \sigma^2)^s}{(1 + \sigma^2)^{t-s}} \left[ (1 - \theta)^2 A^2 \right]^s
\end{array} \right\} \]

\[= \sigma^2 \lambda_0^2 \left\{ \begin{array}{l}
\frac{1}{(1 + \sigma^2)^{t-s}}
\end{array} \right\} \]

\[= \sigma^2 \lambda_0^2 \left\{ \begin{array}{l}
1 + \sum_{s=1}^{t-1} \frac{1}{(1 + \sigma^2)^s}
\end{array} \right\} \]

\[= \sigma^2 \lambda_0^2 \left\{ \begin{array}{l}
1 + \frac{1}{(1 + \sigma^2)^{t-1}}
\end{array} \right\} \]

Thus as \( t \to \infty \)

\[\text{var } k_i \to \sigma^2 \lambda_0^2 \left( 1 - \frac{1}{\sigma^2} \right), \text{ where the convergence is monotone increasing.} \]

Arguments for the other series are similar.

3. Proof of Proposition 4.3.

We apply Proposition 2.2. Accordingly, our interest is in the behavior of \( \text{var}(\hat{k}_i - \hat{k}_s), t - s > 0 \), as a function of \( (t - s) \). It is sufficient to study the concavity/convexity of \( (\hat{k}_i - \hat{k}_s) \) as a function of \( t \). Certainly,
\[
\text{var}(\tilde{k}_i - \tilde{k}_s) = \text{var}(\tilde{k}_i) + \text{var}(\tilde{k}_s) - 2\text{cov}(\tilde{k}_i, \tilde{k}_s)
\]
\[
= \sigma^2 \lambda_0^i \left[ (1 - \theta)^2 A^2 \right] \sum_{j=0}^{t-1} \left( 1 + \sigma^2 \right)^j 
\]
\[
+ \sigma^2 \lambda_0^i \left[ (1 - \theta)^2 A^2 \right] \sum_{j=0}^{s-1} \left( 1 + \sigma^2 \right)^j 
\]
\[
- 2k_0^2 \text{cov} \left[ \left( (1 - \theta)^t A \right) \Pi_{j=0}^{t-1} \lambda_j, \left( (1 - \theta)^s A \right) \sum_{j=0}^{s-1} \lambda_j \right] 
\]
\[
\equiv \text{identification } f(t) + g(s) - h(t,s), \text{ and}
\]
\[
\frac{\partial^2 (\tilde{k}_i - \tilde{k}_s)}{\partial t^2} = f_{11}(t) - h_{11}(t,s). \text{ Below we explore the signs of the}
\]
latter two terms. First consider \( f_{11}(t) \).

\[
f(t) = \sigma^2 \lambda_0^i \left[ (1 - \theta)^2 A^2 \right] \sum_{j=0}^{t-1} \left( 1 + \sigma^2 \right)^j
\]
\[
= \sigma^2 \lambda_0^i \sum_{s=0}^{t-1} \left[ (1 - \theta)^2 A^2 \right] \left( 1 + \sigma^2 \right)^s \left( 1 + \sigma^2 \right)^{t-s}
\]
\[
= \sigma^2 \lambda_0^i \sum_{s=0}^{t-1} \frac{1}{\left( 1 + \sigma^2 \right)^{t-s}}
\]
\[
= \sigma^2 \lambda_0^i \left[ (1 - \theta)^2 A^2 \left( 1 + \sigma^2 \right) \right] \sum_{s=0}^{t} \frac{1}{\left( 1 + \sigma^2 \right)^s}
\]
\[
= \sigma^2 \lambda_0^i \left[ (1 - \theta)^2 A^2 \left( 1 + \sigma^2 \right) \right] \sum_{s=1}^{t} \frac{1}{\left( 1 + \sigma^2 \right)^s}
\]
\[
= \frac{1}{\sigma^2 \lambda} \left[ 1 - \frac{1}{\left( 1 + \sigma^2 \right)^t} \right]
\]
\[
(1 - \theta)^2 \lambda^2 (1 + \sigma^2) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2)} \frac{1}{1 - e^{-\lambda \eta(1 + \sigma^2)}} ,
\]
where we omit the \( k_0^2 \) term as it is not relevant to our calculation.

Thus,

\[
f_i(t) = \left( \ln \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2)} \left( 1 - e^{-\lambda \eta(1 + \sigma^2)} \right)
\]

\[
+ \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2)} \left( \ln(1 + \sigma^2) \right) e^{-\lambda \eta(1 + \sigma^2)}
\]

\[
= \left( \ln \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2) + \ln(1 + \sigma^2)}
\]

\[
- \left( \ln \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2)}
\]

\[
+ \left( \ln(1 + \sigma^2) \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2)}
\]

\[
f_{i1}(t) = \left( \ln \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) \right)^2 \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2) + \ln(1 + \sigma^2)}
\]

\[
- \left( \ln \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) \right) \left( \ln \left( (1 - \theta)^2 A^2 \right) \right) \left( (1 - \theta)^2 A^2 \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2) + \ln(1 + \sigma^2)}
\]

\[
+ \left( \ln(1 + \sigma^2) \right) \left( \ln \left( (1 - \theta)^2 A^2 \right) \right) \left( (1 - \theta)^2 A^2 \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2) + \ln(1 + \sigma^2)}
\]

if \( \left( (1 - \theta)^2 A^2 (1 + \sigma^2) \right) = 1 \)

\[
f_{i1}(t) = \left( \ln(1 + \sigma^2) \right) \left( \ln \left( (1 - \theta)^2 A^2 \right) \right) \left( (1 - \theta)^2 A^2 \right) \epsilon^{\lambda \eta(1 - \theta)^2 A^2 (1 + \sigma^2) + \ln(1 + \sigma^2)} < 0
\]

if \( \left( (1 - \theta)^2 A^2 \right) = 1, f_{i1}(t) > 0 \).
Now consider $h_{11}(t, s)$.

$$h(t, s) = -2 \left\{ E \left[ \left( (1 - \theta) A \right)^{t-1} k_0 \prod_{j=0}^{t-1} \tilde{\lambda}_j \right] \left( (1 - \theta) A \right)^s \prod_{j=0}^{s-1} \tilde{\lambda}_j \right\]$$

$$- E \left[ \left( (1 - \theta) A \right)^{t-1} k_0 \prod_{j=0}^{t-1} \tilde{\lambda}_j \right] E \left[ \left( (1 - \theta) A \right)^s \prod_{j=0}^{s-1} \tilde{\lambda}_j \right] \right\}$$

$$= -2 \left( 1 - \theta A \right)^{t+s} k_0^2 \left\{ E \left[ \prod_{j=0}^{t-1} \tilde{\lambda}_j \right] \prod_{j=0}^{s-1} \tilde{\lambda}_j \right\]$$

$$-E \left[ \prod_{j=0}^{t-1} \tilde{\lambda}_j \right] E \left[ \prod_{j=0}^{s-1} \tilde{\lambda}_j \right] \right\}$$

$$= -2 \left[ (1 - \theta) A \right]^{t+s} k_0^2 \left\{ E \left( \tilde{\lambda}_2 \right)^s E \left( \tilde{\lambda} \right)^{t-s} - \left( E \left( \tilde{\lambda}_2 \right) \right)^s E \left( \tilde{\lambda} \right)^{t-s} \right\}$$

$$= -2 \left[ (1 - \theta) A \right]^{t+s} k_0^2 \left( 1 - \theta A \right)^{t+s} (E \tilde{\lambda})^{t-s}$$

The sign of $h_{11}(t, s)$ is the same as the sign of $\hat{h}(t, s)$ where

$$\hat{h}(t, s) = -\left[ (1 - \theta) A E \tilde{\lambda} \right]^t$$

$$-\hat{h}(t, s) = -e^{(t-t_0)(1-\theta)A E \tilde{\lambda}}$$

$$-h_{11}(t, s) = -\ell n \left[ (1 - \theta) A E \tilde{\lambda} \right] e^{(t-t_0)(1-\theta)A E \tilde{\lambda}}$$

$$-h_{11}(t, s) = -\left( \ell n \left[ (1 - \theta) A E \tilde{\lambda} \right] \right)^2 e^{(t-t_0)(1-\theta)A E \tilde{\lambda}} < 0 \text{ in all cases.}$$
Thus,

\[
\left[(1 - \theta)^2 A^2 (1 + \sigma_w^2)\right] = 1,
\]

\[
f_{\lambda_1} (t) - h_{\lambda_1} (t, s) < 0 \quad \text{and} \quad \{\check{k}_t\} \text{ is mean reverting.}
\]

If \[
\left[(1 - \theta)^2 A^2\right] = 1, \text{ then } h_{\lambda_1} (t) = 0 \text{ but } f_{\lambda_1} (t) > 0, \text{ and } \{\check{k}_t\}\{\check{p}_t\} \text{ is mean averting.}
\]

The other series are even more tiresome.