Asymmetric Cycles

by

Boyan Jovanovic

April 12, 2005

ABSTRACT

I estimate a model in which new technology entails random adjustment needs. Rapid adjustments may cause measured productivity to decline. The slowdowns persist because adjustment is costly and, hence, protracted. The model explains both the "steepness" and the "deepness" asymmetry of cycles. Adjustment costs amount to about 14 percent of output, and technological inefficiency to about 28 percent. Firms abandon technologies long before they are perfected – current-practice TFP is 20 percent below its maximal level.

KEYWORDS: Business cycles, technology.

JEL CLASSIFICATION NUMBER: E3

AUTHOR INFORMATION:

Boyan Jovanovic
Department of Economics
New York University
269 Mercer Street
New York, N.Y. 10003
USA
bj2@nyu.edu

Acknowledgement: I would like to thank G. Barlevy, S. Braguinsky, X. Chen, J. Greenwood, J. Imbs, A. Jao, P. Krusell, L. Veldkamp, and G. Violante for comments, and the NSF (Grant No. 0241350) for support. Thanks to J. Glachant for comments and to A. Gavazza for doing the estimation and correcting many errors. A special thank you to P. Fève for his helpful discussion at the 2003 Toulouse Macroeconomics Conference.

1 Introduction

This paper presents a model that generates growth and cycles. The growth and the cycles are the result of the adoption of technologies of uncertain quality. A technology's "quality" refers to how closely the technology's needs match the economy's input endowments. Since technology needs are unpredictable and since adoption is irreversible, adoption of new technologies can lead output to decline.

Figure 1 documents negative skewness in GDP and industrial production using residuals from a linear trend. The GDP series is from NIPA tables and industrial production from OECD's *Main Economic Indicators*. The two spikes to the right in the GDP plot are due to the war period, and if these observations are dropped, the p-value drops to .000016. In HP-filtered data, the evidence for negative skewness of industrial production becomes stronger, and for GDP it is highly significant (with a p-value of 0.0165) in the post-war period, but not significant for the period as a whole. Sichel (1993), McQueen and Thorley (1993) and Kontolemis (1997) provide more evidence on skewness in these two series.

The model generates cycles because firms adopt technologies the exact character of which they do not know. A firm must use any technology that it adopts for at least one period. How well that technology fits the firm's asset endowments is revealed only after the fact. The mismatch is distributed symmetrically but the *cost* of that mismatch is quadratic so that maximal losses from technology adoption exceed the maximal gains. The firm's growth rate is therefore negatively skewed, and the firm's output spends more periods above trend than below it. For example, suppose a firm computerizes its administrative operations. This raises output if the workers can easily use the new programs. If they cannot, output falls until the workers can be trained, or until the firm can find suitable replacements. The outcome depends on the match between technology and skills. The key assumption is that, while the match value is distributed symmetrically, a good match raises output by less than a bad match reduces it, so that the distribution of the firm's output itself is negatively skewed. So is aggregate output if the technology is adopted simultaneously by many firms.

The model has an analytic solution. The parameter estimates imply that as a fraction of actual output, adjustment costs are about 14 percent, and technological inefficiency is about 28 percent. These numbers fall in booms and rise in recessions. By "technological inefficiency," I mean the fraction of a given technology's learning curve that remains unexploited at the moment at which the technology is abandoned in favor of a better one. Equilibrium is Pareto optimal.

In related work, Ramey and Ramey (1991) also assume technological commitment so that output may decline when new technology is adopted. Chalkley and Lee (1998) and Veldkamp (2002) explain asymmetry by assuming that firms can more quickly detect negative shocks than positive ones. In their models investment responds asymmetrically to a symmetric exogenous TFP process the changes in which

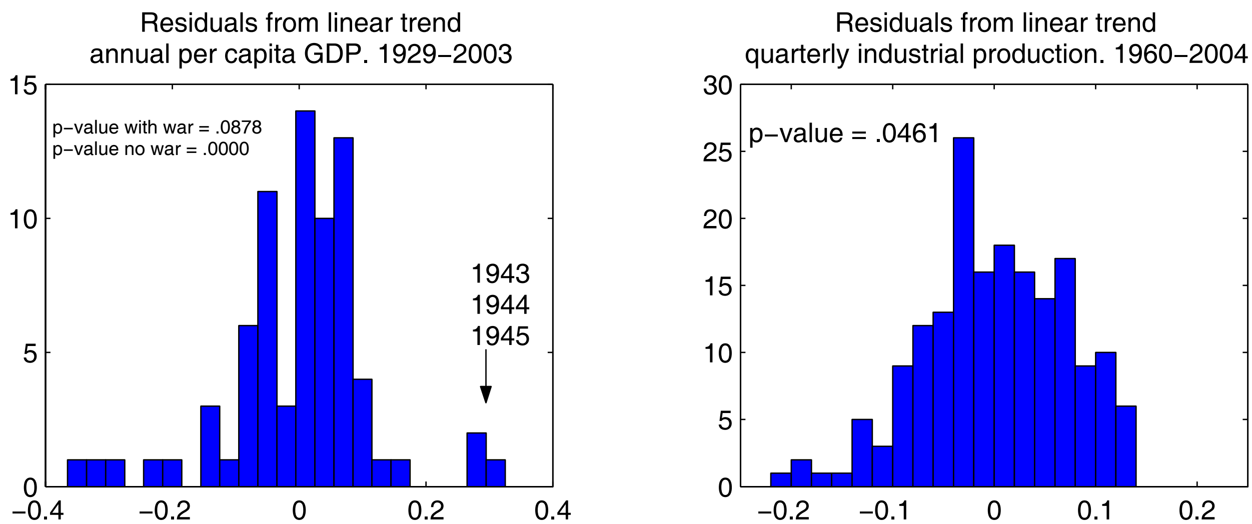


Figure 1: Asymmetry in U.S. GDP and industrial production

firms can infer and track more easily when their investment is high. Thus a negative TFP shock has a quicker impact than a positive shock would in a slump. Klenow (1998) explains asymmetry by assuming that each technology improves with use. Upon an update, productivity drops sharply and then recovers gradually. In Acemoglu and Scott (1997) a firm may invest in a project that reduces its productivity today because investment raises the future productivity of investment, and this may lead to sharper downturns than upturns.

Plan of paper.—Section 2 presents the optimal growth model and solves for the optimal policy. Section 3 analyzes the model's empirical implications. Section 4 presents the decentralized equilibrium and shows that it is optimal. Section 5 discusses the assumption of technological commitment that drives the results. Section 6 concludes the paper.

2 Model

Let us start with the single agent, "Crusoe," optimal growth version of the model. His preferences over consumption sequences (c_t) are

$$E_0 \left\{ \sum_0^\infty \beta^t \ln c_t \right\}.$$

Potential output, y^p, is

$$y^p = \exp \left\{ A - \frac{\lambda}{2} (s_A - h')^2 \right\}.$$

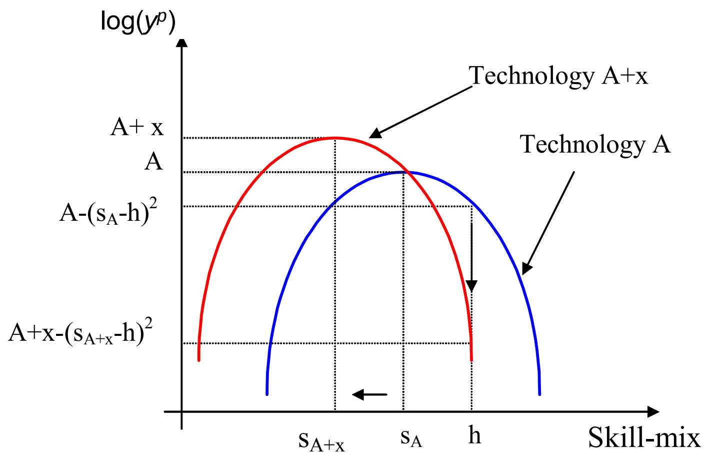


Figure 2: Consequences of technological upgrading

Here A is the technology, h' is the skill mix, and s_A is the skill-mix ideal for technology A.

Adoption of technology.—Adopting a new technology is free. Crusoe can raise his technology by any amount, $x \geq 0$, so that starting today at A, tomorrow's technology is

$$A' = A + x.$$

Crusoe must use technology A' for at least one period. But A' makes unpredictable demands on the skill mix. Assume that

$$s_{A'} = s_A + x\varepsilon,$$

where $\varepsilon \sim F(\varepsilon)$ is *time* specific and i.i.d., having mean zero and variance σ^2. Once ε is drawn, $s_{A'}$ becomes an invariant skill requirement for technology A'. Crusoe chooses A' before seeing ε, and he can*not* return to technologies that he used in the past. I shall discuss this no-recall assumption in detail in Section 5.

Adjustment of h.—Crusoe starts the period with h. Before producing, he can change it to $h' \equiv h + \Delta$ at a cost of

$$C(y^p, \Delta) \equiv \left[1 - \exp\left\{-\frac{\theta}{2}\Delta^2\right\}\right] y^p.$$

Crusoe's net output therefore is

$$\begin{aligned} y^p - C(y^p, \Delta) &= \exp\left\{A - \frac{\lambda}{2}(u-\Delta)^2 - \frac{\theta}{2}\Delta^2\right\} \\ &\equiv y(u, \Delta, A), \end{aligned}$$

where

$$u = s_A - h$$

is the gap between ideal skill, s_A, and actual skill, h.

Output changes.—Figure 2 shows how output may fall as a result of technological change. Suppose $\lambda = 2$, and suppose that Crusoe starts with (A, s_A, h). Then $\ln y^p = A - (s_A - h)^2$. Upon adopting technology $A + x$, $\ln y^p$ drops to $A + x - (s_{A+x} - h)^2$ because s_A has moved in the unfavorable direction, i.e., away from h. Crusoe can soften this drop by adjusting h to the left, but if θ is high, a drop will still take place.

The figure also shows why the distribution of output growth is negatively skewed. Suppose that instead of moving in the unfavorable direction, i.e., to the left, s_{A+x} had in fact moved towards h by the same amount. The way things are drawn, this would have resulted in the happy outcome of $s_{A+x} = h$, and log TFP would have risen to $A + x$. This rise is smaller than the drop that we discussed in the previous paragraph. And since a movement of s_A to the right is just as likely as a movement to the left, $\ln y^p$ must have a longer tail on the left than on the right.

2.1 The optimal growth problem

We shall deny Crusoe the opportunity to diversify his technological portfolio. While avoiding this issue here, we shall face it squarely in Section 3 where we shall give firms the option of choosing different technologies; they will end up rejecting that option for reasons that I shall explain at the end of Section 3.

Crusoe's state is the pair (u, A), and his decisions are (x, Δ). He has no assets other than h and A, and he simply consumes his output. His Bellman equation is

$$V(u, A) = \max_{x,\Delta}\left\{\ln y(u, \Delta, A) + \beta \int V(u + x\varepsilon - \Delta, A + x)\, dF(\varepsilon)\right\}. \qquad (1)$$

The solution to (1), derived in Appendix 1, can be summarized as follows:

Proposition 1 *The policy functions are*

$$x = \frac{1}{\theta\sigma^2(1-\beta)(1-\alpha)}, \qquad (2)$$

and

$$\Delta = (1-\alpha)\, u, \qquad (3)$$

where

$$\alpha = \frac{1}{2\beta}\left\{1+\beta+\frac{\lambda}{\theta}-\sqrt{\left(1+\beta+\frac{\lambda}{\theta}\right)^2-4\beta}\right\} \tag{4}$$

is the fraction of the gap that Crusoe leaves open. The solution for V is

$$V(u,A) = \frac{A}{1-\beta}-\frac{1}{2}\theta(1-\alpha)u^2+C, \tag{5}$$

where

$$C = \frac{\beta}{1-\beta}\left(\frac{x}{1-\beta}-\left[\frac{\psi}{1-\beta\alpha^2}\right]x^2\sigma^2\right), \quad \textit{and} \quad \psi = \frac{1}{2}\left(\lambda\alpha^2+\theta(1-\alpha)^2\right). \tag{6}$$

The first component of the optimal policy, x, is the average growth of the frontier, whereas Δ is the adjustment of skills. Let us comment on each effect.

Average growth.—As stated in (2) the growth of A, i.e., x, is a constant, independent of the technological imbalance u. The reason is that the expected marginal cost of raising x does not depend on u: Since ε has a zero mean, a rise in x is as likely to reduce u as raise it, (which is why in (22) of the proof in the Appendix the term $\int \varepsilon\Delta dF$ vanishes). We would get x to depend on u if the cost were $\frac{\theta}{2}\Delta^\rho$ with $\rho \neq 2$. If $\rho > 2$, x would fall with u, and if $\rho < 2$, it would rise with u. Thus x goes to infinity if $\theta \to 0$, or $\lambda \to 0$, or $\sigma^2 \to 0$, or $\beta \to 1$. Conversely, x goes to zero if σ^2 or θ become large.[1] These results are intuitive except, perhaps for the fact that (2) and (4) imply that

$$\lim_{\beta\to 0} x = \left(\frac{1}{\lambda}+\frac{1}{\theta}\right)\frac{1}{\sigma^2} > 0.$$

That is, even with infinite discounting, investment in x remains positive; this is because its gains and losses both occur in the future, only *after* adoption.

Adjustment to technology shocks.—Note that α, which is defined in (4) and which enters the other four equations, is homogeneous of degree zero in λ and θ. Since $1-\alpha$ is the fraction of the gap, u, that is closed, we should think of α as the persistence of the cycle in that it is the fraction of the skill gap that remains open. It is the persistence with which output tends to stay below its frontier. By L'Hôpital's rule,

$$\lim_{\beta\to 0}\alpha = \frac{\theta}{\lambda+\theta},$$

[1] Comin (2000) has argued that the productivity slowdown of the 70s and 80s was caused by a rise in technological uncertainty in the 1970's which raised the demand for less productive but more flexible capital. The present model get a similar effect from a rise in σ^2 that reduces x and, hence, TFP. The recent TFP-growth revival is, in terms of the model, consistent with the evidence that the variance of aggregate shocks has declined.

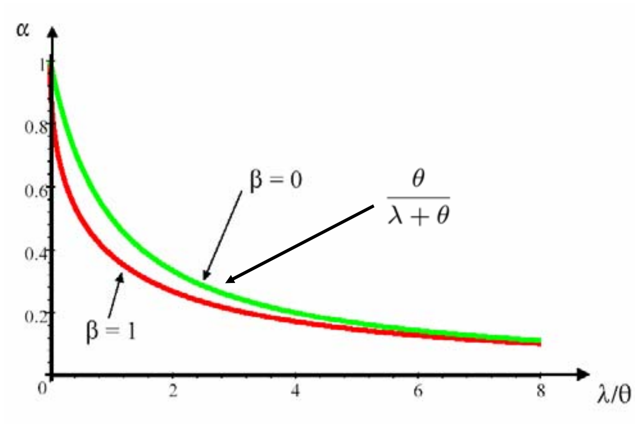


Figure 3: The nature of α

which says that persistence depends on how high the cost of adjustment, θ, is relative to the cost, λ, of technological imbalance. This value is bounded away from unity for the same reason, mentioned above, that x remains positive as $\beta \to 0$. An increase in patience lowers this persistence:

$$\lim_{\beta \to 1} \alpha = \frac{1}{2}\left(2 + \frac{\lambda}{\theta} - \sqrt{\left(2 + \frac{\lambda}{\theta}\right)^2 - 4}\right).$$

The quantitative effect is small, however, as Figure 3 shows. The Figure displays a plot of these two limits as a function of λ/θ. The vertical axis measures the persistence of the cycle. As $\lambda \to 0$, the cost of a technological imbalance goes to zero, and it is not worth incurring any adjustment cost to redress it, and so $\alpha \to 1$. Conversely, as $\theta \to 0$, adjustment becomes free so that $\alpha \to 0$ and the gap is closed completely. Figure 6 in the Appendix shows α in three dimensions.

The value function.—Let's focus on how V depends on the technological imbalance u. The coefficient of u^2 is $-\frac{1}{2}\theta(1-\alpha)$. The costlier h is to adjust (i.e., the higher is θ), the more u^2 reduces Crusoe's lifetime utility. Holding θ fixed, on the other hand, the higher are the costs of technological imbalance (i.e., the higher is λ), the more, again, does u^2 reduces Crusoe's utility.

3 The (y_t) process

Sichel (1993) distinguishes "Steep" asymmetry from "Deep" asymmetry, both of which are present in the (y_t).process. Figure 4 portrays these two concepts in a stylized way.

Steep asymmetry.—The three panels on the left of Figure 4 depict a steepness asymmetry for the detrended variable y_t, portrayed in the top panel. In the *General Theory*, Keynes had argued that business cycles were of this type. Steepness implies no asymmetry of the frequency distribution of y_t as illustrated in panel 3 of Figure 4. Instead, it implies negative skewness of the distribution of growth rates (bottom panel) which will motivate Propositions 2 and 3. A steep-asymmetric series is time irreversible.

Deep asymmetry.—The three panels on the right of Figure 4 depict a deepness asymmetry for y_t. Deepness requires, roughly, that booms last longer than recessions. It implies a negative skewness in the frequency distribution of y_t which will motivate Proposition 4, but no skewness in the distribution of growth rates. As drawn, the deep-asymmetric series is time reversible.

Both asymmetries seem to be in the data. Steep asymmetry is found in GDP and industrial production in the U.S., and among the G7 countries it is apparent in Canada, Germany and the U.K. (Kontolemis 1997). Deep asymmetry is found in U.S. industrial production (Sichel 1993 and McQueen and Thorley 1993) and in that of most G7 economies (Kontolemis 1997).

3.1 The (y_t) process

Suppose that $C(y^p, \Delta)$ consists entirely of foregone output. The log of measured output then is

$$\begin{aligned} \ln y_t &= A - \frac{\lambda}{2}\alpha^2 u_t^2 - \frac{\theta}{2}(1-\alpha)^2 u_t^2 \\ &= A_0 + xt - \psi u_t^2, \end{aligned} \tag{7}$$

where, since $u' = u - \Delta + x\varepsilon$,

$$u_{t+1} = \alpha u_t + x\varepsilon_{t+1}. \tag{8}$$

Since α is between zero and one, u_t is stationary. Therefore $\ln y$ is trend-stationary, the trend and the long-run rate of output growth being x.

Let us now describe the two types of asymmetry – deep and steep – that the model exhibits.

3.2 Steepness asymmetry

Although $C(y^p, \Delta)$ is symmetric in Δ, when θ is large adjustment costs are effectively highly asymmetric in the following sense: Adjusting h towards its technologically ideal value is costly and will therefore take time, but output will fall sharply whenever an unlucky draw of s_A occurs. So we should expect steepness asymmetry to arise when θ is large.

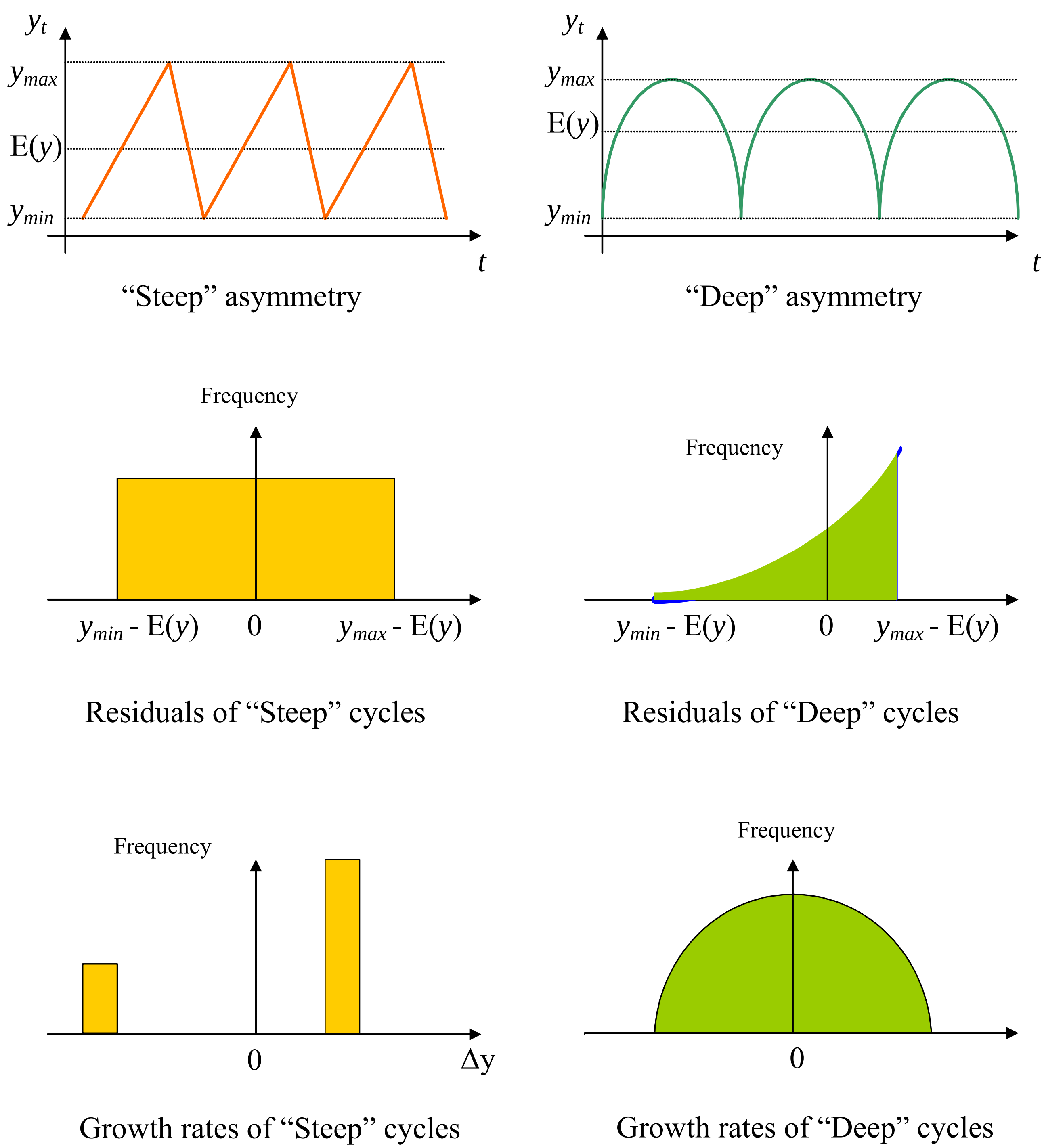


Figure 4: Two kinds of asymmetry

In Figure 3, a large θ is seen to imply a large value of α. Therefore we should expect steepness asymmetry for large α. And this is indeed what happens.

Proposition 2 *Condition on y_0. If ε_t is symmetric and uni-modal, as $\alpha \to 1$, the distribution of $\Delta \ln y_t$ is negatively skewed for each $t > 0$.*

Proof. When $\alpha \to 1$,

$$\Delta \ln y_t \to x - \psi \left(u_{t+1}^2 - u_t^2\right) = x - \psi x^2 \left(\varepsilon_{t+1}^2 + \frac{2}{x}\varepsilon_{t+1} u_t\right)$$

Now because ε is symmetric around zero, The distribution of $\varepsilon_{t+1} u_t$ conditional on u_0 is the same as its distribution conditional on $-u_0$. So we may just adopt the convention that $u_0 = +\sqrt{u^2} = +\sqrt{\frac{1}{\psi}\left(\ln y_0 - A_0\right)^2}$. Therefore as $\alpha \to 1$,

$$\Delta \ln y_t \to x - \psi x^2 \left(\varepsilon_{t+1}^2 + \frac{2}{x}\varepsilon_{t+1}\left[\sqrt{\frac{1}{\psi}\left(\ln y_0 - A_0\right)^2} + \sum_{j=1}^{t} \varepsilon_j\right]\right).$$

Therefore $\Delta \ln y_t$ becomes the sum of two random variables; one symmetric and the other, $\psi x^2 \varepsilon_{t+1}^2$, with a longer left tail. ■

Given the model's estimates (presented in section 3.4) the relevant value of α is indeed close to unity. On the other hand, (y_t) loses its steepness asymmetry as α becomes small:

Proposition 3 *Regardless of the distribution of ε, as $\alpha \to 0$, $\Delta \ln y_t$ becomes symmetric. In particular,*

$$\Delta \ln y_t \to x - \psi x^2 \left(\varepsilon_{t+1}^2 - \varepsilon_t^2\right),$$

a symmetric random variable.

Proof. Taking the limit in (8) and substituting into (7) the claim follows. ■

Since α is estimated to be .95, this result seems to be empirically uninteresting. A small α arises in the model when λ is large relative to θ.

This shows that we can have steepness asymmetry, i.e., a slow uptake to the boom even with a single representative firm. In contrast, Lippi and Reichlin (1994) argue that diffusion lags help explain the sluggish impulse responses to technology shocks. It's just the modeling choice of whether we put adjustment costs on the intensive or the extensive margin

3.3 Deepness asymmetry

If ε is symmetric and uni-modal, the model generates deepness asymmetry for all $\alpha < 1$, and this stems directly from the quadratic loss function of technological imbalance which precludes miracles and allows only disasters:

Proposition 4 *If ε_t is symmetric and uni-modal, $\Delta \ln y_t - A_0 - xt$ is negatively skewed for all $\alpha \in [0, 1)$.*

Proof. For $\alpha < 1$, the stationary distribution of u_t exists and is uni-modal and symmetric, and the claim follows. ■

Conditioning on y_0 does not now suffice for proving the result for $\alpha = 1$ because the conditional distribution of u_t is then symmetric only when $\ln y_0 = A_0$.

3.3.1 The case where ε is normal

If ε_t is normally distributed, the stationary distribution of u_t is also normal with mean zero and variance

$$\sigma_u^2 = \frac{x^2 \sigma^2}{1 - \alpha^2}.$$

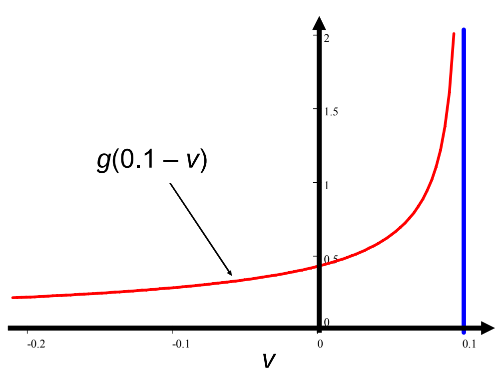
Now, the stationary distribution of the square of a standard normal variate, is $\chi^2_{(1)}$. Denote by v the square of such a variable, i.e.,

$$v = \left(\frac{\sqrt{1 - \alpha^2}}{x\sigma} u \right)^2 .$$

Then v has a Chi-squared distribution with 1 degree of freedom:

$$z^{-\frac{1}{2}} \frac{1}{2\pi} \exp\left(-\frac{1}{2} z^{\frac{1}{2}} \right) \equiv g(z),$$

for $z \geq 0$.



THE DENSITY OF A $\chi^2_{(1)}-$ DISTRIBUTED RESIDUAL

Even when ε is not normally distributed, the distribution of u is likely to be approximately normal because u is a linear combination of independent variates. If ε has a mode at zero, then the stationary distribution of u inherits this property and exaggerates it. The higher is the mode of u, the more skewed will be the distribution of u^2.

3.4 The probability of "technological regress"

Output and TFP are the same thing in the model. The probability of technological regress is $1-\gamma$, where $\gamma \equiv \Pr(\Delta \ln y_t > 0)$ is given by

$$\begin{aligned}
\gamma &= \Pr\left(u_{t+1}^2 < \frac{x}{\psi} + u_t^2\right) \\
&= \Pr\left(\sqrt{\frac{x}{\psi} + u_t^2} > u_{t+1} > -\sqrt{\frac{x}{\psi} + u_t^2}\right) \\
&= \Pr\left(\frac{1}{x}\sqrt{\frac{x}{\psi} + u_t^2} - \frac{\alpha}{x}u_t > \varepsilon_{t+1} > -\frac{1}{x}\sqrt{\frac{x}{\psi} + u_t^2} - \frac{\alpha}{x}u_t\right) \\
&= \int \left[F\left(\sqrt{\frac{1}{\psi x} + \frac{u^2}{x^2}} - \frac{\alpha}{x}u\right) - F\left(-\sqrt{\frac{1}{\psi x} + \frac{u^2}{x^2}} - \frac{\alpha}{x}u\right)\right] dP(u).
\end{aligned}$$

3.4.1 The case where ε is normal

When $\varepsilon \sim N(0,1)$, we can solve for γ analytically because we then have $u \sim N\left(0, \frac{1}{1-\alpha^2}\right)$, so that

$$\gamma = \int_{-\infty}^{\infty} \left[\Phi\left(\sqrt{\frac{1}{\psi x} + \frac{u^2}{x^2}} - \frac{\alpha}{x}u\right) - \Phi\left(-\sqrt{\frac{1}{\psi x} + \frac{u^2}{x^2}} - \frac{\alpha}{x}u\right)\right] \frac{e^{-\frac{u^2\left(1-\alpha^2\right)}{2}}}{\sqrt{2\pi\left(1-\alpha^2\right)^{-1}}} du$$

where Φ is the standard normal cumulative integral. Using the parameter estimates given in section 3.4, $\psi = 59.3$. Substituting also the estimated values for x and α, we find that $\gamma = 0.97$. Therefore at the estimated parameter values, the model predicts that TFP will fall about 3% of the time.

Figure 5 plots the probability of an output drop conditional on u, $\Pr(\Delta y < 0 \mid u)$, as well as the density of u multiplied by 2 so as to make the two curves similar in scale. The figure shows that when when technological gap is close to zero, output is quite likely to fall. But as the gap opens up, this probability falls dramatically. This is because the variance of the innovation, ε, is scaled by x which is small and, secondarily, because α is close to unity.

Now, annual TFP growth in the U.S. has been negative in 11 of the 53 years spanning the period 1948-2001.[2] Compared to that number, 3% is an underestimate.

[2] BLS series on Multifactor Productivity, Private Non-Farm Business.

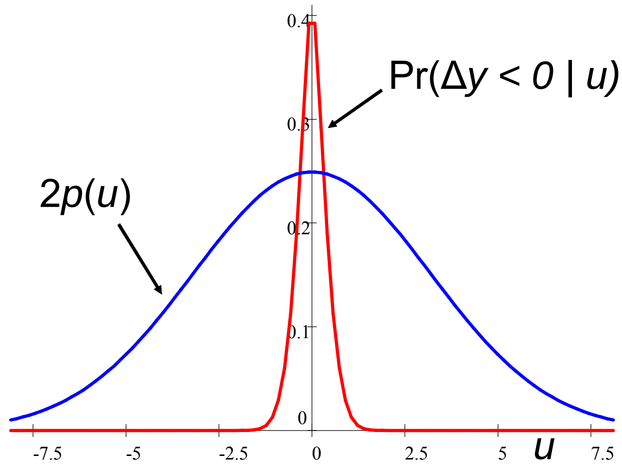


Figure 5: THE CONDITIONAL PROBABILITY OF OUTPUT DROPS AND THE DENSITY OF u IN THE NORMAL CASE

Figure 5 sheds some light on why the model under-predicts the number of falls. The point is that having y_t fall often is not the same as having a fat left tail , which from the previous Figure we know a normal ε does generate; y_t could be way below trend and still be rising. Indeed, since y_t is stationary, that is precisely when output will be rising. It appears, then, than a normal ε has too much mass in its tails relative to its variance. In any case, it is not clear what the right percentage of drops for TFP is – King and Rebelo (1999) claim that when one allows for variable capacity utilization, the percentage of dates at which TFP declines can be pushed close to zero (seemingly to under 1% according to the simulations they report in Figure 14). On those grounds 3% would be an overestimate.

3.5 Estimates

The parameters are β, σ, λ, and θ. We shall work with annual data and set $\beta = 0.95$, and estimate σ, λ, and θ. A glance at the model shows that λ and θ can be identified only in units of σ^2. Therefore we shall set $\sigma^2 = 1$. The parameter estimates are based on (7), i.e., on

$$\ln y_t = A_0 + xt - \psi u_t^2,$$

and, in view of (8), on

$$u_{t+1}^2 = \alpha^2 u_t^2 + x^2 \varepsilon_{t+1}^2 + 2\alpha x \varepsilon_{t+1} u_t. \tag{9}$$

The second Appendix describes the estimation procedure in detail. Table 1 reports the estimates:

Parameter	x	α	λ	θ	$\rho = \frac{\lambda}{\theta}$	$\psi\sigma_u^2$
Estimate	.0241	.9504	84.7	16733	.0051	.36
(S.E.)	(.0006)	(.027)	(24.5)	(8978)	(.0042)	(.12)

Table 1: *Parameter Estimates*

The estimate of $\psi E\left(u^2\right) = \psi\sigma_u^2$ implies that output is $e^{-0.36} = 0.70$ of its maximal level, i.e., it is 30 percent below what it would be if $h = s_A$ so that efficiency was maximal and no adjustment was needed. This number is precisely estimated but its breakdown into the two categories – efficiency and adjustment cost – is not. The point estimates imply that one third of the 30 percent comes from the adjustment cost, and two thirds from technological inefficiency.[3] As a fraction of actual output, adjustment costs are about 14 percent and technological inefficiency about 28 percent. To put it differently, firms abandon technologies long before they are perfected – efficiency is on average 20 percent below its best-possible level.

The adjustment-cost estimate of 14 percent of output is slightly less than the typical estimate of the cost of adjusting the physical capital stock. In a Hayashi type of constant-returns-to-scale setup with adjustment costs quadratic, when Tobin's Q is at its historical average of 1.5, adjustment costs roughly equal the investment rate itself. If, as is usually assumed, costs depend on gross investment, they are a few percentage points higher than our estimate.

We may think of h as organization capital that the firm owns and that is costly to adjust, as Prescott and Visscher (1980) argue. The empirical suggestion, then, is that sluggish productivity arises because firms find it costly to adjust their internal practices in response to technological needs. In the Prescott-Visscher model as in the present one, organization capital was a non-hierarchical variable. In the spirit of their model we can think of s_A as the correct factor proportions between their two production tasks. They had an unchanging s_A, however, in that these desired factor proportions were fixed over time. Here, the proportions adjust in response to technological needs, and they adjust most in recessions. In support of such a view, Sepulveda (2002) finds that on-the-job training is countercyclical.

4 Decentralizing the optimum

The decentralized model has two markets: A market for output, and a market for firm's shares which also are the only assets available to households. No markets exist for either A or h. Since A was free for Crusoe, it is natural that markets for it should not exist in the decentralized setting. On the other hand, h was for Crusoe costly to adjust, and we shall assume that h is owned by firms; one may think of h as organization capital in the sense of Prescott and Visscher (1980).

[3] Of course, adjustment of h is essential for growth. If $x > 0$ and if we shut down the adjustment of h by setting $\alpha = 1$, σ_u^2 would be infinite and efficiency, $\exp\left(-\frac{1}{2}\lambda\sigma_u^2\right)$ would be zero.

Assume a continuum of firms of measure one and an equal number of shareholders. Equilibrium then requires that each shareholder hold exactly one share. Shareholders have no other income.

A firm pays a dividend of $y(u, \Delta(u), A) \equiv \delta(u, A)$, where $\Delta(u)$ is defined in (3) as the solution to Crusoe's problem; we shall need to verify that it also solves the firm's problem. Let $p(u, A)$ be the price of owning a representative firm in state (u, A). This is also the price of a firm' shares.

The savings decision.—If it owns n shares, a shareholder's wealth is $[\delta(u, A) + p(u, A)] n$. With n' denoting the number of shares it will carry into the next period, the shareholder's budget constraint therefore is

$$p(u, A) n' + c = [\delta(u, A) + p(u, A)] n.$$

The consumer's state is (n, u, A). Since $u' = \alpha u + x\varepsilon$, the transition function for u is

$$\Phi(u', u) = F\left(\frac{u' - \alpha u}{x}\right),$$

where F is the C.D.F. of ε. Writing $U(c)$ instead of the less intuitive $\ln c$ for his utility function, a shareholder's Bellman equation is

$$w(n, u, A) = \max_{n'} \left\{ U(\delta(u, A) n + p(u, A)[n - n']) + \beta \int w(n', u', A + x) \, d\Phi(u', u) \right\}.$$

The first-order condition is

$$U'(c) p(u, A) = \beta \int w_1 d\Phi. \tag{10}$$

The envelope theorem gives

$$w_1 = U'(\delta(u, A)) [\delta(u, A) + p(u, A)]$$

which, upon an update and a substitution into (10) gives

$$U'(\delta(u, A)) p(u, A) = \beta \int U'(\delta(u', A + x)) [\delta(u', A + x) + p(u', A + x)] \, d\Phi,$$

which is the analog of equation (6) of Lucas (1978). Evidently, then, $p(u, A)$ satisfies the familiar equation

$$p(u, A) = \beta \int \frac{U'(\delta(u', A + x))}{U'(\delta(u, A))} [\delta(u', A + x) + p(u', A + x)] \, d\Phi \tag{11}$$

This is an implicit function for the *ex*-dividend price of a representative firm.

To decentralize the decision about x and Δ two more steps are needed. First, in choosing Δ the firm considers not only its ex-dividend price, but the dividend itself.

And, second, the firm – starting in the representative state (u, A) must choose the pair (x, Δ) that all other firms choose. Let us use bold letters to denote aggregate states so as to distinguish them from the firm's own state. The functional equation for a firm's cum-dividend price is

$$P(u, A, \mathbf{u}, \mathbf{A}) = \max_{x,\Delta} \begin{pmatrix} y(u, \Delta, A) + \\ \beta \int \frac{U'(\delta(\mathbf{u}+\mathbf{x}\varepsilon-\mathbf{\Delta},\mathbf{A}+\mathbf{x}))}{U'(\delta(\mathbf{u},\mathbf{A}))} P(u + x\varepsilon - \Delta, A + x, \mathbf{u} + \mathbf{x}\varepsilon - \mathbf{\Delta}, \mathbf{A} + \mathbf{x})\, dF(\varepsilon) \end{pmatrix} \tag{12}$$

We need to show that

1. At the fixed point for P, the RHS of (12) is maximized by $x = \mathbf{x}$, and by $\Delta = \mathbf{\Delta}$, and that

2. For all (u, A),
$$P(u, A, u, A) = \delta(u, A) + p(u, A). \tag{13}$$

In fact, property 1 implies property 2 as one can deduce by setting $(x, \Delta) = (\mathbf{x}, \mathbf{\Delta})$ so that $(u', A') = (\mathbf{u}', \mathbf{A}')$, in which case substitution from (13) into (12) makes it identical to (11). Thus it suffices to show that property 1 holds. Recall that $U(c) = \ln c$. Then, evaluated at $(x, \Delta) = (\mathbf{x}, \mathbf{\Delta})$, the two first-order conditions for a maximum in (12) are

$$\lambda(u - \Delta) - \theta\Delta - \beta \int \frac{1}{\delta(u + x\varepsilon - \Delta, A + x)} P_1(u + x\varepsilon, A + x, u + x\varepsilon, A + x)\, dF = 0, \tag{14}$$

and

$$\int \frac{1}{\delta(u + x\varepsilon - \Delta, A + x)} (\varepsilon P_1 + P_2)\, dF = 0. \tag{15}$$

The envelope theorem gives us

$$\begin{aligned} P_1 &= -\lambda(u - \Delta)\delta(u, A) + \beta \int \frac{\delta(u, A)}{\delta(u + x\varepsilon - \Delta, A + x)} P_1 dF \\ &= -\theta\Delta\delta(u, A) \end{aligned}$$

where the second equality follows from (14). Thus (14) reads

$$\lambda(u - \Delta) = \theta\Delta - \theta\beta \int \left(\frac{\delta'}{\delta'}\right) \Delta' dF,$$

which, in a symmetric equilibrium is the same as (19).

The envelope theorem also gives us

$$P_2 = \delta(u, A) + \beta \int \frac{\delta(u, A)}{\delta(u + x\varepsilon - \Delta, A + x)} P_2 dF$$

Then let $Q(u, A) \equiv P_2(u, A) / \delta(u, A)$. This is a contraction map that reads

$$Q(u, A) = 1 + \beta \int Q(u + x\varepsilon - \Delta, A + x)\, dF$$

which is the counterpart to (21) and yields the unique solution $Q = 1/(1-\beta)$, which implies that

$$P_2(u, A) = \frac{\delta(u, A)}{1-\beta}.$$

Substituting into (15) in an updated form leads $\delta(u + x\varepsilon - \Delta, A + x)$ to cancel, and leads to $\int \varepsilon\theta\Delta' dF = \frac{1}{1-\beta}$, i.e., which is the same as the equation preceding (22), and which therefore implies (2). Thus the right-hand side of (12) is indeed maximized by $x = \mathbf{x}$, and by $\Delta = \boldsymbol{\Delta}$, which means that every firm will choose the (x, Δ) pair given in (3) and(2).

Why is equilibrium symmetric?—A representative-firm equilibrium exists because the opportunity to differentiate one's dividend stream is quite limited: A firm can make its dividend less correlated with the aggregate dividend only by reducing its x, or by raising its Δ. Each deviation would reduce its expected dividends by too much. When all firms choose the same technology, a shareholder's lifetime utility in the state u, A is the same as Crusoe's.

Cross-section evidence on skewness.—The cross-section evidence on skewness is mixed. On the one hand, the distribution of firms' efficiencies has a longer left tail – see Caves and Barton (1990) and Figure 2 of Baily *et al* (1992). On the other hand, evidence on small start-up ventures and the value of patents show a long *right* tail – most start-ups fail and most patents are worth next to nothing, but a few yield very high returns. It may at first seem like this evidence is irrelevant here because in the decentralized model firms are identical. But minor adjustments to the model would have implications for cross sections as well as for time series. In particular, if each firm had its own technology ladder so that the ε's were firm-specific, the model would imply a cross-section distribution of productivity with a longer left tail. The long-run distribution of u for a given firm would still be negatively skewed, but it now would also be the cross-section distribution of u's among firms at each date; aggregate output would grow smoothly. The model thus seems to fit better the experience of established firms than that of of start-up firms, but it is the established firms that produce most of the economy's output.

5 The issue of technological commitment

Technological commitment is central to the results of the model; output sometimes declines because of commitment to technology before an unfavorable shock ε is realized, and skewness properties depend on this as well. Can the results be saved if, instead of being full, commitment were only partial?

If a firm could revert quickly and costlessly to technologies it used before, its output would never fall – it would always use the best technology that it had ever sampled to date. Costless recall of past technologies, indeed, is what Jovanovic and Rob (1990) assume. In their model output never declines, and instead of having a left tail, the distribution of growth rates acquires a spike at zero and a *right* tail. Therefore we need at least partial commitment

Suppose, then, that Crusoe has to use a technology for at least one period after he first adopts it, but that later he can revert to the technology he used in the previous period. Denote the sufficient statistics of that previous technology by (u_{-1}, A_{-1}). Were he to revert to it in the next period, Crusoe would get the value of $W(u_{-1}, A_{-1}, u, A)$. Then instead of (1), his Bellman equation would read

$$W(u, A, u_{-1}, A_{-1}) = \max_{\Delta} \left\{ \begin{array}{c} U[y(u, \Delta, A)] + \\ \beta \max \left[W(u, A, u_{-1}, A_{-1}), \max_x \int W(u + x\varepsilon - \Delta, A + x, u, A) \, dF(\varepsilon) \right] \end{array} \right\}$$

The subset of R^4

$$\left\{ (u_{-1}, A_{-1}, u, A) \mid W(u, A, u_{-1}, A_{-1}) > \max_x \int W(u + x\varepsilon - \Delta, A + x, u, A) \, dF(\varepsilon) \right\}$$

is the set of states on which recall is optimal. In contrast to how u enters V in (5), u_{-1} and u will no longer enter W in the form of a square: Crusoe is now better off if u and u_{-1} are of the same sign, because then he can choose h' so as to reduce both of them simultaneously. When u and u_{-1} are opposite in sign, reducing one necessarily raises the other.

While this version is harder to analyze, one can easily guess one property of the solution: As long as he chooses a positive x, Crusoe can not prevent output from falling, at least for one period. As the period of commitment becomes short (period length affects β), Crusoe would definitely want a positive x. Thus it appears that allowing recall in this sense preserves one key implication of the model, namely that TFP can decline. And if $F(\varepsilon)$ had a sufficiently pronounced mode at $\varepsilon = 0$, negative skewness might survive too, but what other conditions would also be needed is unclear.

6 Conclusion

This paper has modeled the business cycle as the result of luck in adoption of technology. We saw how declines in output and TFP and their negatively skewed distributions can emerge when technologies have unpredictable skill needs. The estimates suggested that adjustment costs amount to about 14 percent of output, and technological inefficiency to about 28 percent. It seems that firms abandon technologies long before they are perfected – current-practice TFP is 20 percent below its maximal level.

The mechanism focuses not on labor-market and capital-market frictions, but on costs of adjusting the firm's organization capital. The model treats these costs as being internal to firms. It is possible, however, that, as in some other contexts, it does not matter for aggregates whether costs of adjustment are internal or external.

Finally, the growth process is Pareto optimal, and so the model does not call for policy intervention.

References

[1] Acemoglu, Daron, and Andrew Scott. "Asymmetric Business Cycles: Theory and Time Series Evidence." *Journal of Monetary Economics* 40 (December 1997): 501-522.

[2] Baily, Martin Neil, Charles Hulten and David Campbell. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity. Microeconomics* (1992): 187-267.

[3] Caves, Richard, and Barton, D. R., *Efficiency in U.S. Manufacturing Industries*. The MIT Press, Cambridge, Massachusetts (1990).

[4] Chalkley, Martin, and In-Ho Lee. "Learning and Asymmetric Business Cycles." *Review of Economic Dynamics* 1 (1998): 623-45.

[5] Comin, Diego. "An Uncertainty-Driven Theory of the Productivity Slowdown: Manufacturing." 2000.

[6] De Long and Lawrence Summers. "Are Business Cycles Symmetrical?" in *The American Business Cycle: Continuity and Change*. Robert J. Gordon, ed. Chicago: University of Chicago Press, (1986): 166-178. Also NBER w1444, November 1986.

[7] Greene, William H., *Econometric Analysis*, 4th edition, Upper Saddle River, NJ: Prentice Hall (2000).

[8] Jovanovic, Boyan and Rafael Rob. "Long Waves and Short Waves: Growth Through Intensive and Extensive Search." *Econometrica* 58, no. 6 (November 1990): 1391-1409.

[9] King, Robert and Sergio Rebelo. "Resuscitating Real Business Cycles", in John Taylor and Michael Woodford (eds.) *Handbook of Macroeconomics* 1B, Elsevier Science (1999): 927-1007.

[10] Klenow, Peter. "Learning Curves and the Cyclical Behavior of Manufacturing Industries." *Review of Economic Dynamics* 1, no. 2 (April 1998): 531-550.

[11] Kontolemis, Zenon. "Does Growth Vary over the Business Cycle? Some Evidence from the G7 Countries." *Economica* N.S. 64, no. 255. (August 1997): 441-460.

[12] Lippi; Marco, and Lucrezia Reichlin. "Diffusion of Technical Change and the Decomposition of Output into Trend and Cycle." *Review of Economic Studies* 61, no. 1 (Jan., 1994): 19-30.

[13] Lucas, Robert E. Jr. "Asset Prices in an Exchange Economy." *Econometrica* 46, no. 6 (November 1978): 1429-1445

[14] McQueen, Grant and Steven Thorley. "Asymmetric Business Cycle Turning Points." *Journal of Monetary Economics* 31 (1993): 341-362.

[15] Prescott Edward and Michael Visscher. "Organization Capital." *Journal of Political Economy* 88, no. 3 (June 1980): 446-61.

[16] Ramey, Garey, and Valerie Ramey. "Technology Commitment and the Cost of Economic Fluctuations." NBER W3755, June 1991.

[17] Sepulveda, Facundo. "Training and Business Cycles." Michigan State University, 2002. http://www.economics.adelaide.edu.au/workshops/doc/sepulveda.pdf

[18] Sichel, Daniel. "Business Cycle Asymmetry: A Deeper Look." Economic Inquiry 31 (April 1993): 224-36.

[19] Veldkamp, Laura. "Slow Boom, Sudden Crash." *Journal of Economic Theory*, forthcoming.

7 Appendix 1: Proof of Proposition 1

The first-order condition for the optimality of Δ is

$$\lambda (u - \Delta) - \theta\Delta - \beta \int V_1 dF = 0, \tag{16}$$

and the first-order condition for the optimality of x is

$$\int (\varepsilon V_1 + V_2)\, dF = 0. \tag{17}$$

Solving for Δ.—The envelope theorem gives

$$V_1 = -\lambda (u - \Delta) + \beta \int V_1 dF = -\theta\Delta, \tag{18}$$

where the second equality uses (16). Substituting into (16) we have

$$\lambda (u - \Delta) = \theta \Delta - \beta \theta \int \Delta' dF(\varepsilon). \tag{19}$$

We seek a solution of the form (3) where α is a constant to be solved for. If (3) holds, (19) reads

$$\begin{aligned} \lambda \alpha u &= \theta (1-\alpha) u - \beta \theta \int (1-\alpha)(\alpha u + x\varepsilon) dF(\varepsilon) \\ &= \theta (1-\alpha) u - \beta \theta (1-\alpha) \alpha u, \end{aligned}$$

which, after cancellation of u leaves a quadratic in α, namely $\theta (1-\alpha) - \beta \theta (1-\alpha) \alpha - \lambda \alpha = 0$, or

$$\beta \alpha^2 - \left(1 + \beta + \frac{\lambda}{\theta}\right) \alpha + 1 = 0. \tag{20}$$

This implicit function has the solution for α given in (4).

Solving for x.—The envelope theorem also gives

$$V_2 = 1 + \beta \int V_2 dF = \frac{1}{1-\beta}. \tag{21}$$

The second equality follows because the right-hand side of (21) is a contraction map with at most one solution for V_2. Substituting from (3) into (18) and from there (in an updated form) into (17) gives

$$\begin{aligned} 0 &= -\int \varepsilon \theta \Delta' dF + \frac{1}{1-\beta} \\ &= -\int \varepsilon \theta (1-\alpha)(\alpha u + x\varepsilon) dF + \frac{1}{1-\beta} \end{aligned} \tag{22}$$

because $\Delta' = (1-\alpha) u' = (1-\alpha)(\alpha u + x\varepsilon)$. But $E(\varepsilon u) = 0$, which leads to (2).

We must show that this function solves (1). Let us proceed with the method of undetermined coefficients. Since $V = aA - bu^2 + c$,

$$\begin{aligned} aA - bu^2 + c &= A - \psi u^2 + \beta \int \left(a(A+x) - b(u + x\varepsilon - \Delta)^2 + c\right) dF(\varepsilon) \\ &= A - \psi u^2 + \beta \int \left(a(A+x) - b(\alpha u + x\varepsilon)^2 + c\right) dF(\varepsilon) \\ &= A - \psi u^2 + \beta (a(A+x) + c) - \beta b \int (\alpha u + x\varepsilon)^2 dF(\varepsilon) \\ &= A - \psi u^2 + \beta (a(A+x) + c) - \beta b \alpha^2 u^2 - \beta b x^2 \sigma^2 \end{aligned}$$

Equating coefficients: $a = 1 + a\beta$, $b = \psi + b\beta\alpha^2$, and $c = \beta (ax + c - bx^2\sigma^2)$, so that

$$a = \frac{1}{1-\beta}, \quad b = \frac{\psi}{1-\beta\alpha^2}, \text{ and } c = \frac{\beta}{1-\beta}\left(ax - bx^2\sigma^2\right).$$

This leads to

$$V(u,A)=\frac{A}{1-\beta}-\frac{\psi}{1-\beta\alpha^2}u^2+\frac{\beta\left(\frac{x}{1-\beta}-\left[\frac{\psi}{1-\beta\alpha^2}\right]x^2\sigma^2\right)}{1-\beta}$$

where $x=\frac{1}{\theta\sigma^2(1-\beta)(1-\alpha)}$. Then $V_2(u,A)=1/(1-\beta)$ which is consistent with (21). It remains to be shown that $V_1(u,A)$ agrees with (18) and (3). Now, since $\psi=\frac{1}{2}\left(\lambda\alpha^2+\theta(1-\alpha)^2\right)$, they agree only if

$$\frac{\psi}{1-\beta\alpha^2}=\frac{1}{2}\theta(1-\alpha)$$

i.e., if

$$\left(\lambda\alpha^2+\theta(1-\alpha)^2\right)=\theta(1-\alpha)\left(1-\beta\alpha^2\right)$$

i.e., if

$$\left(\frac{\lambda}{\theta}\alpha^2+(1-\alpha)^2\right)=(1-\alpha)\left(1-\beta\alpha^2\right) \tag{23}$$

But from (20),

$$\frac{\lambda}{\theta}=\frac{1}{\alpha}+\beta\alpha-1-\beta.$$

Substitute for λ/θ into (23) to conclude that $V_1(u,A)$ is consistent with (18) and (3) if and only if

$$\left(\alpha+\beta\alpha^3-\alpha^2-\beta\alpha^2+(1-\alpha)^2\right)=(1-\alpha)\left(1-\beta\alpha^2\right) \tag{24}$$

But expanding the left-hand side of (24) yields

$$\left(\alpha+\beta\alpha^3-\alpha^2-\beta\alpha^2+1+\alpha^2-2\alpha\right)=\beta\alpha^3-\beta\alpha^2+1-\alpha.$$

Conversely, expanding the right-hand side of (24) yields

$$(1-\alpha)\left(1-\beta\alpha^2\right)=1-\alpha-\beta\alpha^2+\beta\alpha^3$$

Therefore (24) holds, and V is therefore given by (5).

8 Appendix 2: Estimation details

The parameters are β, σ, λ, and θ. They determine α, x, and ψ via (4), (2) and (6). The main regression equation is

$$\ln y_t=A_0+xt-\psi u_t^2. \tag{25}$$

Secondly, multiplying both sides of (9) by ψ we have

$$\psi u_{t+1}^2=\alpha^2\left(\psi u_t^2\right)+\psi x^2\varepsilon_{t+1}^2+\psi 2\alpha x\varepsilon_{t+1}u_t. \tag{26}$$

Now let the OLS regression in (25) be denoted by $\ln y_t = \hat{a} + \hat{x}t + U_t$. Note that

$$E\left(\psi u_t^2\right) = x^2 \frac{\psi\sigma^2}{1-\alpha^2} \equiv v. \tag{27}$$

If the estimates are consistent, the residual $U_t = \ln y_t - \hat{a} - \hat{x}t$ converges to $\psi u_t^2 - v$ and $\hat{a}$ converges to $A_0 - v$. Now substitute for $\psi u_t^2 = -\hat{U}_t + v$ into (26) to get

$$-U_{t+1} + v = \alpha^2\left(-U_t + v\right) + \psi x^2 \varepsilon_{t+1}^2 + \psi 2\alpha x \varepsilon_{t+1} u_t.$$

Rearranging, $-U_{t+1} = -\alpha^2 U_t - \left(1-\alpha^2\right) v + \psi x^2 \varepsilon_{t+1}^2 + \psi 2\alpha x \varepsilon_{t+1} u_t$. But $E\left(\psi x^2 \varepsilon_{t+1}^2\right) = \psi x^2 \sigma^2 = \left(1-\alpha^2\right) v$ (using [27]), so that

$$U_{t+1} = \alpha^2 U_t + \eta_t \tag{28}$$

where

$$\eta_{t+1} = -\psi 2\alpha x \varepsilon_{t+1} u_t - \psi x^2\left(\varepsilon_{t+1}^2 - \sigma^2\right)$$

is a zero-mean disturbance that is uncorrelated with U_t.

We cannot identify σ, λ, and θ separately, only $(\lambda\sigma^2, \theta\sigma^2)$. Therefore we set σ to unity in this exposition and proceed as follows:

1. Get $\hat{x}$ and $\hat{U}$ using OLS in (25),
2. get $\hat{\alpha}$ using from OLS in (28) as described,
3. use (4) and $\beta = 0.95$ to solve for $\frac{\lambda}{\theta} \equiv \hat{\rho}$,
4. use (2) to get
$$\hat{\theta} = \frac{1}{\hat{x}\left(1-0.95\right)\left(1-\hat{\alpha}\right)},$$
5. get $\hat{\lambda} \equiv \hat{\rho}\hat{\theta}$.

The standard errors for λ, θ and λ/θ were obtained with the Delta method (Greene, Theorem 4.16 pag.118) which I report here

Theorem 1 *If z_n is a $K \times 1$ sequence of vector valued random variables such that $\sqrt{n}\left(z_n - \mu\right) \rightarrow N\left(0, \Sigma\right)$ and if $c\left(z_n\right)$ is a set of J continuous functions of z_n not involving n, then*

$$\sqrt{n}\left(c\left(z_n\right) - c\left(\mu\right)\right) \rightarrow N\left(0, C\left(\mu\right)\Sigma C\left(\mu\right)'\right)$$

where $C\left(\mu\right)$ is the $J \times K$ matrix $\partial c\left(\mu\right)/\partial\mu'$. The j'th row of $C\left(\mu\right)$ is the vector of partial derivatives of the j'th function with respect to μ'.

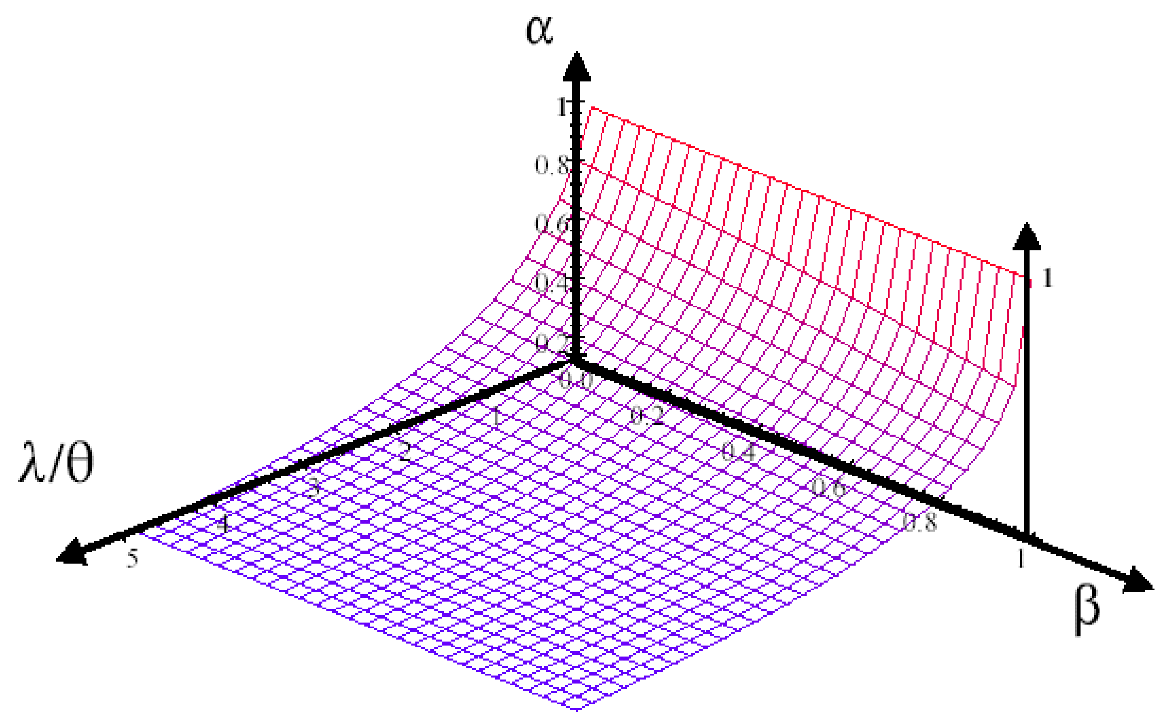


Figure 6: Three-dimensional plot of α from (4).