Learning Asymmetries in Real Business Cycles

Stijn Van Nieuwerburgh* and Laura Veldkamp†

December 29, 2004‡

Abstract

When a boom ends, the downturn is generally sharp and short. When growth resumes, the boom is more gradual. Our explanation rests on learning about productivity. When agents believe productivity is high, they work, invest, and produce more. More production generates higher precision information. When the boom ends, precise estimates of the slowdown prompt decisive reactions: Investment and labor fall sharply. When growth resumes, low production yields noisy estimates of recovery. Noise impedes learning, slows recovery, and makes booms more gradual than downturns. A calibrated model generates growth rate asymmetry similar to macroeconomic aggregates. Fluctuations in agents' forecast precision match observed countercyclical errors of forecasters.

"There is, however, another characteristic of what we call the trade cycle that our explanation must cover; namely, the phenomenon of the crisis - the fact that the substitution of a downward for an upward tendency often takes place suddenly and violently , whereas there is, as a rule, no such sharp turning point when an upward is substituted for a downward tendency." *J.M. Keynes (1936)*

*New York University, Stern School of Business, Department of Finance

†Corresponding author: New York University, Stern School of Business, Department of Economics. lveldkam@stern.nyu.edu, 44 West 4th St., New York, NY 10012, tel: 212 998-0527, fax: 212 995-4218.

‡First version: May 15, 2001. First submission: July 3, 2003. Many thanks to Tom Sargent, Tim Cogley, Paul Romer, Peter Henry, Bob Hall, Kenneth Judd, Pierre-Olivier Weill, Hanno Lustig, John Shea, Tim Johnson, Robert Rich, Robert King, and attendees of the Stanford Macro Lunch, the 2002 meetings of the Econometric Society in Los Angeles, the conference on learning and bounded rationality at the University of Illinois Urbana Champaign, and the economics seminar at the University of Leuven, Munich and Humbolt University for their helpful advice. Stijn Van Nieuwerburgh thanks SIEPR and the Flanders Fund for Scientific Research for financial support. All remaining errors are our own. JEL classification: E32, D83, E27. Keywords: learning, asymmetry, business cycles, uncertainty.

1 Introduction

When an economic boom ends, the downturn is generally sharp and short, meaning that the growth rate falls far below trend for a short period of time. When a slump ends and growth resumes, the boom is more gradual: longer lasting, with growth rates not as far from trend. This type of asymmetry is referred to as *growth-rate* or *steepness* asymmetry. It is present in many macroeconomic aggregates: output, industrial production, fixed investment and hours worked. In the 1990-91 recession, investment fell by 7% and the unemployment rate rose by 1.8% points over the course of three quarters. The recovery in the first half of the nineties was much more gradual. Likewise, in the first three quarters of the 2001 recession, investment fell by 6.5% and the unemployment rate rose by 1.5% points. In spite of a reduction in the volatility of business cycles over the last two decades, their asymmetry has not diminished. In the decade from 1985 to 1995, changes in monthly industrial production had a skewness of -0.6, compared to -0.2 in the decade before and -0.8 in the period 1945-75.[1]

We explain this asymmetry with an endogenously varying rate of information flow about the aggregate technology. When productivity is high, agents work harder and invest more. More production generates higher-precision information about the level of technology, which is unobservable. When the economy passes the peak of a productivity boom, agents have precise estimates of the extent of the slowdown. Firms abruptly reduce investment projects and labor demand. At the end of a productivity slump, low levels of production yield noisy estimates about the extent of the recovery. This extra uncertainty at the start of an upturn restrains the expansion of investment projects and new hiring. This restraint delays the recovery and makes booms more gradual than crashes.

[1]Negative skewness in the distribution of changes captures the presence of slow increases, a large number of small, positive changes, and sudden decreases, a small number of large negative changes.

The model of asymmetric learning is embedded into a dynamic stochastic general equilibrium model (see Kydland and Prescott (1982), Cooley and Prescott (1995), and Backus, Kehoe and Kydland (1992)). Its distinguishing feature is the additive stochastic term in the output equation: $y_t = z_t f_t + \eta_t$. Because the additive noise term η and the aggregate technology level z are unobserved, agents use observed capital and labor inputs, $f_t = f(k_t, n_t)$, and output y_t to form Bayesian beliefs about technology. Since technology is multiplied by a function of inputs, larger inputs amplify changes in technology. Higher variance in technology relative to noise makes technology more easily observable. Following a change in technology, the *speed of learning* measures how quickly beliefs converge to the truth. When the economy is in recession and inputs are low, filtered estimates of technology are imprecise and learning is slow. In a boom, high capital and labor utilization make learning faster. This variation in the speed of learning over the business cycle produces the asymmetry in growth rates.

Variations in the speed of learning reflect the fact that the signal-to-noise ratio is procyclical. Technology z is amplified by high production f in booms, while the noise η is not. Section 5 argues that this formulation is a natural one. Aggregating production units that have common and idiosyncratic productivity shocks, both of which are multiplicative, generates the additivity. When production units increase, the variance of the sum of their idiosyncratic shocks rises more slowly than the variance of the sum of the common shocks. The more production units there are, the less noise their idiosyncratic shocks add, the higher the signal-to-noise ratio is, and the more precise estimates of aggregate technology are.

A key feature of our explanation for asymmetry is that recessions are times of high uncertainty. Evidence from the Survey of Professional Forecasters supports this prediction (section 3). The median forecast error and the dispersion of the GDP forecast across a panel of forecasters goes up in recessions. This is evidence of more uncertainty about the state of the business cycle in recessions.

A calibration exercise shows that the model is able to replicate the negative skewness of output, investment, employment and consumption (section 6). The model matches not only the magnitude of asymmetry but also its frequency pattern. Skewness in output changes is a high frequency phenomenon. It decreases as the length of the period over which changes are computed increases. It disappears for 3-4 year changes. The model replicates this pattern because short term uncertainty about the current state is resolved after a few periods (section 7).

Understanding the source of business cycle asymmetry improves forecasting of macroeconomic aggregates. Acemoglu and Scott (1997) estimate that allowing for asymmetric booms and crashes captures an additional 12% of US output growth fluctuations. Neftci (1984) shows that Federal Reserve Board estimates of employment perform significantly worse at business cycle turning points because they do not capture the asymmetric nature of booms and crashes. An individual in our model, who behaves as though learning were symmetric, incurs a 4% higher cost of business cycles than agents who understand learning asymmetries. To the extent that there are gains from reducing business cycle fluctuations, this result suggests non-trivial benefits from understanding the source of asymmetry.

2 Literature on Asymmetry and Learning

The literature on business cycle asymmetries measures and models many different types of asymmetry. Most asymmetries fall into three categories: level asymmetries (deepness), growth rate asymmetries (steepness) and delays. All three are distinct features of the data. Level asymmetry refers to the unconditional distribution of detrended output levels, whereas growth rate asymmetry refers to the unconditional distribution of output changes. Growth rate asymmetry means that increases and decreases in output have different distributions. Delay asymmetry results when the output level stagnates at the trough of an otherwise symmetric cycle. Figure 1 illustrates the shape of the cycle produced by level, growth rate and delay asymmetries.

Previous papers have used a variety of mechanisms to produce business cycle asymmetries. Hansen and Prescott (2000) use capacity constraints to prevent booms from being as large a deviation from trend as recessions. In contrast, Kocherlakota (2000) and Acemoglu and Scott (1997) use credit constraints and learning-by-doing to amplify shocks in the trough of a business cycle. Williams (2004) uses large deviation theory to investigate asymmetries in the probability of large recessions and booms. These are examples of level asymmetry.

Most explanations for growth rate asymmetry rely in some way on a learning process. In a partial equilibrium model of lending, Veldkamp (2004) uses the idea of faster learning when more investment projects are undertaken to generate asymmetry in investment and interest rates. Using a varying signal quality rather than a varying number of signals makes more precise calibration in this paper possible. Chakley and Lee (1998) employ noise traders who become a relatively larger fraction of the market in bad times. Boldrin and Levine (2001)'s asymmetric spread of new technologies, Jovanovic (2003)'s asymmetric technology adoption costs, and Zeira (1999)'s learning about market capacity, all have bad signals that are either more informative or more extreme than good signals.

Chamley and Gale (1994) generate delay asymmetry using irreversible investment. Their model shares the feature that investment generates information. At low levels of production, firms wait to produce until they have learned from other firms' investments. This produces the delayed recovery that the authors aim to explain. Since delay is neither an upturn nor a downturn it will not affect the relative speed of booms and crashes.

Our explanation for growth asymmetry relies on the fact that recessions are times when uncertainty is high. This relationship is a recurrent idea in economics. In a model with a GARCH shock to output, a sudden downturn increases expected future output volatility (Ebell (2001)). To guard against consumption disasters, risk-averse agents scale back their inputs, accentuating the downturn when uncertainty is high. This paper generates GARCH-like dynamics endogenously through the learning mechanism. Potter (1999) claims that recessions are caused by uncertainty about the economic environment which makes firms unable to coordinate on a high-output equilibrium. Estimating a Bayesian model where agents are learning about the growth rate, Cogley (2002) finds that consumers may have more difficulty determining the permanent component of their income when output is low.

Work from many corners of economics suggests that learning and output fluctuations are intimately related: learning by doing (Jovanovic and Nyarko (1982)), learning by lending (Lang and Nakamura (1990)), learning about demand (Rob (1991)). Within the business cycle literature, Cagetti, Hansen, Sargent and Williams (2001) have agents who are uncertain about the drift of technology and solve a filtering problem, as in this model. Evans, Honkapohja and Romer (1998) and Kasa (1995) show that business cycles can be produced by learning, rather than by technological fluctuations. In these models, learning is not only related to the business cycle, it is fundamental to the cyclical behavior.

3 Evidence of Business Cycle Learning

Before we show that our model is capable of producing asymmetry, we first ask whether learning is a plausible explanation for business cycle asymmetry. The theory relies on lower precision forecasts in bad times than in good. To test this hypothesis empirically, we measure the precision of analysts' forecasts of future real GDP.

To approximate the amount of uncertainty, we use the median forecast error in a panel of forecasters. For a given quarter, the median forecast error is the log absolute deviation of the median forecast from the final nominal GDP, where the latter is measured two quarters after the end of the quarter. Analysts forecast nominal GDP one through four quarters into the future. The hypothesis is that the median forecast error is negatively correlated with the detrended level of output.

Table 1 contains the results of regressions with detrended GDP as the independent variable and the median forecast error as the dependent variable. Detrending is done with a Hodrick and Prescott (1997) filter. The correlation between uncertainty and output is negative for all forecast horizons (1 to 4 quarters ahead) and increases in magnitude and significance with the forecast horizon. The median forecast error corresponds to the forecast error of the representative agent in the model. This measure supports our main hypothesis: forecast accuracy and output are negatively correlated. Learning is slower when output is low.

Although more loosely connected to the model, we also find a negative relationship between the *dispersion* of analysts' forecasts of real GDP and real GDP deviations from trend (not reported). Dispersion shoots up near business cycle troughs. Median forecast errors and cross-sectional dispersion are related. Using inflation data from the Survey of Professional Forecasters, Rich and Tracy (2003) documents a stable relationship between observed heteroscedasticity in forecast errors and cross-sectional dispersion in forecasts. Veronesi (1999) shows that a similar relationship holds in financial markets. Stock returns are lower when analysts' GDP forecasts are more diffuse. This is consistent with our finding because equity returns are a leading indicator of real GDP.

4 Model

Preferences and Endowments

An infinite-lived representative consumer ranks consumption streams $c = \{c_t\}_{t=0}^{\infty}$ and labor streams $n = \{n_t\}_{t=0}^{\infty}$ according to

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t), \tag{1}$$

where u is in $\mathcal{C}^{2,2}$, nonseparable, strictly increasing and concave in both arguments, and satisfies the Inada conditions:

$$u(c_t, 1 - n_t^s) = \frac{(c_t^{\sigma}(1 - n_t)^{1-\sigma})^{1-\phi}}{1 - \phi}.$$

The parameters ϕ and σ control risk aversion and the intratemporal elasticity of substitution between consumption and labor. The household is endowed with an initial capital stock and one unit of time each period.

Firms and Technology

Competitive firms have access to a risky Cobb-Douglas production technology f that uses capital and labor as inputs. There are two types of production risk. There is an aggregate technology shock, z_t, which is a multiplier on production, and a shock η_t which is additive. The shocks η_t are i.i.d. normal variables with mean zero and constant variance σ_η:

$$y_t = z_t f(k_t, n_t^d) + \eta_t. \tag{2}$$

The technology shock is specified as a two-state Markov switching process $z_t \ \epsilon \ \{z^H, z^L\}$. We assume that the transition matrix $\mathbf{\Pi}$ is symmetric to ensure that all asymmetry in the resulting dynamics is endogenous.

This specification implies that *expected* returns to scale are constant. In this economy, entry of new production units is unprofitable ex-ante.

One feature of the production shock in equation (2) is crucial for learning asymmetry. η must become relatively smaller in magnitude than $z_t f_t$ at the peak of a business cycle, and relatively larger in the trough. In other words, the variance of the output shock σ_η can rise when production increases, but must rise at a rate less than f_t. This condition implies that the *signal-to-noise ratio is procyclical.* For the computation, we will assume that the shock variance σ_η^2 is constant, although this is a more restrictive assumption than necessary for the model to function. Section 5 argues that this assumption is plausible.

Information

Let x^t denote the t-history of any series x. The model is defined on a probability space with filtration $\mathcal{F}_t \equiv \{y^{t-1}, c^{t-1}, d^{t-1}, k^t, i^t, n^{d,t}, n^{s,t}, \theta^t, w^t, p^t\}$. All choice variables in period t are $\mathcal{F}_t$-measurable. Neither the household nor the firm observe the current or past values of technology z^t nor the noise η^t. Output y_t, dividends d_t and consumption c_t are not observed until the end of period t. All other variables are known at the beginning of period t.

Agents use a Bayesian filter to forecast z_t given $\mathcal{F}_t$. This filter is composed of a Bayesian updating formula (3) and an adjustment for the possibility of a state change (4).

$$P(z_{t-1} = z^H|\mathcal{F}_t) = \frac{\phi(y_{t-1}|z^H, \mathcal{F}_{t-1})P(z_{t-1} = z^H|\mathcal{F}_{t-1})}{\phi(y_{t-1}|z^H, \mathcal{F}_{t-1})P(z_{t-1} = z^H|\mathcal{F}_{t-1}) + \phi(y_{t-1}|z^L, \mathcal{F}_{t-1})(1 - P(z_{t-1} = z^H|\mathcal{F}_{t-1}))}, \tag{3}$$

$$\left[P(z_t = z^H|\mathcal{F}_t), P(z_t = z^L|\mathcal{F}_t)\right] = \left[P(z_{t-1} = z^H|\mathcal{F}_t), P(z_{t-1} = z^L|\mathcal{F}_t)\right] \mathbf{\Pi}, \tag{4}$$

$$\tilde{z}_t = \left[P(z_t = z^H|\mathcal{F}_t), P(z_t = z^L|\mathcal{F}_t)\right] \begin{bmatrix} z^H \\ z^L \end{bmatrix}, \tag{5}$$

where ϕ is a normal probability density. Equation (3) applies Bayes' law to compute the posterior probability, at the beginning of time t, that the economy was in the high state in the preceding period. Equation (4) converts this posterior belief about the time $t-1$ state into a prior belief about the time t state. The result of this updating is $\tilde{z}_t$, the expected technology value in period t. It is a convex combination of the high and low true technology states. Equations (3) and (5) form the observation and state equations of a signal extraction problem with time-varying parameters.

Household Problem

Households enter each period with a belief about technology $\tilde{z}_t$ and an ownership share stock portfolio θ_t. Given wages and share prices $\{w_t, p_t\}$, they choose how much labor to supply, n_t^s and what fraction θ_{t+1} of the shares of the firms to purchase. Then the additive shock η_t is realized (but unobserved) and the dividend d_t is received. Finally, the forecast of z is updated.

The household problem is to maximize U in equation (1) subject to a sequence of end-of-period budget constraints and interiority constraints:

$$\begin{aligned} &c_t + p_t\theta_{t+1} \leq w_t n_t^s + p_t)\theta_t + d_t\theta_{t+1},\ \forall t \\ &c_t \geq 0, \qquad 0 \leq n_t^s \leq 1, \qquad \theta_0 \text{ given.} \end{aligned} \tag{6}$$

Households decide on stock purchases and labor supply before the shock is realized. As a result of this timing assumption, consumption is a residual. It absorbs unexpected shocks to output coming from z or η. Stock purchases finance investment projects.

Firm Problem

Competitive firms maximize lifetime expected shareholders' value.

$$S = E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{u_c(c_t, 1-n_t)}{u_c(c_0, 1-n_0)} \right) d_t. \tag{7}$$

Firms choose how much labor to hire and how much to invest at the beginning of the period. Wages and rental rates are set at the beginning of the period. Then the shock η is realized (but not revealed) and output y_t is realized. The profits of the firm (negative or positive) are paid out as dividends d_t to the households.

The firm's problem is to maximize S in equation (7) subject to the dividend determination equation:

$$d_t = z_t f(k_t, n_t^d) + \eta_t - w_t n_t^d - i_t, \ \forall t \tag{8}$$

and the capital accumulation equation:

$$k_{t+1} = (1-\delta)k_t + i_t, \ \forall t \tag{9}$$

$$k_t \geq 0, \qquad k_0 \text{ given.} \tag{10}$$

Dividends, positive or negative, ensure zero profits ex-post. Because η is normally distributed, profits have unbounded support. Households, who own the firms, face unlimited liability.

New investment projects and staffing take time to plan. Hence, the timing assumption of choosing investment and labor supply at the beginning of the period seems consistent with the fact that the National Association of Purchasing Managers index is a leading indicator of economic activity.

Equilibrium

For a given k_0, and an initial distribution of z_0, an equilibrium is a sequence of quantities $\{c_t, i_t, k_t, n_t^d, n_t^s, y_t, \theta_t\}_{t=0}^{\infty}$ and prices $\{w_t, p_t\}_{t=0}^{\infty}$ such that

- Given prices, households solve the household problem,
- Given prices, firms solve the firm problem,
- The markets for goods, labor, and firms' shares clear:

$$y_t = i_t + c_t, \qquad n_t^s = n_t^d, \qquad \theta_t = 1.$$

Optimality Conditions

Firms hire workers until the expected marginal product of labor equals the wage:

$$w_t = E_t \left[z_t f_n(k_t, n_t) \right]. \tag{11}$$

Workers supply labor until the marginal rate of substitution between consumption and labor equals the wage

$$w_t = \frac{E_t \left[u_n(c_t, 1 - n_t) \right]}{E_t \left[u_c(c_t, 1 - n_t) \right]}. \tag{12}$$

Equation (11) differs from the standard condition in that the technology level z_t and hence the marginal product of labor are unobserved. Equation (12) differs from the standard condition in that the consumption level c_t and hence the marginal utilities of consumption and leisure are unobserved at the beginning of the period. We assume that wages are determined at the beginning of the period. In equilibrium they equal the *expected* marginal product of labor, which equals the ratio of the *expected* marginal disutility of labor to the *expected* marginal utility of consumption.

From the firm's first order condition with respect to capital, we obtain the Euler equation:

$$E_t\left[u_c(c_t, 1-n_t)\right] = E_t\left[\beta u_c(c_{t+1}, 1-n_{t+1})(z_{t+1} f_k(k_{t+1}, n_{t+1}) + 1 - \delta)\right]. \tag{13}$$

The Euler equation from the household's problem is:

$$E_t\left[u_c(c_t, 1-n_t)\right] = E_t\left[u_c(c_t, 1-n_t)\left(\frac{d_t}{p_t}\right) + \beta u_c(c_{t+1}, 1-n_{t+1})\left(\frac{p_{t+1}}{p_t}\right)\right]. \tag{14}$$

Equations (13) and (14) relate the cum-dividend return on stocks to the physical return on capital invested in the firm. Learning affects the right hand side of the Euler equations by changing the joint distribution of time $t+1$ consumption and labor.

Social Planner Problem

To solve the model numerically, we work with the planner problem. The planner's problem is to maximize U in equation (1) subject to the capital accumulation equation (9) and the sequence of resource constraints:

$$c_t + i_t = z_t f(k_t, n_t).$$

The planner is subject to the same timing and informational constraints as the agents in the decentralized model. The competitive equilibrium and the Pareto optimum problem are equivalent because they share the same first order conditions, the goods market clearing condition is identical to the planner's resource constraint, technology is convex, and preferences insatiable.

The planner's problem can be formulated in a recursive fashion. The state variables are the expected value of technology and the capital stock: $s_t = (\tilde{z}_t, k_t)$. The planner enters the period with a capital stock k_t and a forecast of the technology shock $\tilde{z}_t$. Then, he chooses n_t and i_t. Next,

z_t and η_t and hence y_t are realized, but only y_t is observed. Finally, the forecast of z is updated and consumption is determined as a residual. The value function $V(s_t)$ solves the following Bellman equation:

$$V(k_t, \tilde{z}_t) = \max_{i_t, n_t} u(c_t, 1 - n_t) + \beta E[V(k_{t+1}, \tilde{z}_{t+1})|\mathcal{F}_t]$$

$$\text{subject to } k_{t+1} = (1 - \delta)k_t + i_t,$$

$$c_t = z_t f(k_t, n_t) + \eta_t - i_t,$$

$$c_t \geq 0,\ 0 \leq n_t \leq 1,\ k_t \geq 0,\ \text{and } k_0 \text{ given},$$

$$\text{taking updating rules (3) (4) and (5) as given.}$$

In choosing optimal policies, we assume that the planner does not take into account the effect of labor and investment choices on the evolution of beliefs. This assumption rules out active experimentation, the costly provision of information. A planner economy with active experimentation has no decentralized counterpart because information externalities invalidate the welfare theorems. In a decentralized economy, information is a public good. Information provision is subject to a free-rider problem. In a large economy, the provision of this public good would fall to zero. Therefore, households and firms do not take into account the effect of their actions on the evolution of aggregate beliefs. To implement passive learning, the planner iterates on the following algorithm. First, given beliefs about $\mathcal{F}_{t+1}$, he chooses policies. Second, given policies, he updates beliefs. The iteration stops when beliefs and policies are consistent.[2]

[2]Quantitatively, using active or passive learning makes a small difference for policy rules in the planner problem.

5 Interpreting the Production Shock

An Aggregation Argument

At first glance, the models aggregate additive technology shocks may look unnatural. However, the procyclical signal-to-noise ratio it produces arises from aggregating production units with common and idiosyncratic productivity shocks.

Suppose that the economy is comprised of production units, which are combinations of capital and labor such that $f(k_t^i, n_t^i) = 1$. Higher aggregate production is achieved when more of these production units are operating. Let F_t denote the number of production units in operation at time t. The output of each production unit is multiplied by its productivity, which has a common component z_t and an idiosyncratic component η_t^i, so that $y_t^i = (z_t + \eta_t^i) f(k_t^i, n_t^i) = (z_t + \eta_t^i)$. Then, aggregate output is the sum of output of all the production units:

$$y_t = \sum_{i=1}^{F_t} (z_t + \eta_t^i) = z_t F_t + \sum_{i=1}^{F_t} \eta_t^i \tag{15}$$

Define the signal-to-noise ratio to be the variance of $z_t F_t$, divided by the variance of the noise $(\sum_{i=1}^{F_t} \eta_t^i)$.

Proposition 1. *If $Corr(\eta_t^i, \eta_t^j) < 1$, $\forall\ (i, j)$ and the signal-to-noise ratio of an individual production unit is greater than zero and constant over the business cycle, then the aggregate signal-to-noise ratio is increasing in F_t.*

Proof: If the idiosyncratic shocks are uncorrelated across production units, the signal-to-noise ratio is $(F_t^2 \sigma_z^2)/(F_t \sigma_\eta^2) = F_t \sigma_z^2 / \sigma_\eta^2$; Since σ_z^2 and σ_η^2 are positive, it is increasing in F_t. Independence is not required. As long as the correlation of shocks across production units is less than one, $\mathrm{var}(\sum_{i=1}^{F_t} \eta_t^i) < F_t^2 \sigma_\eta^2$, and the signal-to-noise ratio is still increasing in F_t (procyclical). This happens because the effect of each production units idiosyncratic shock on the signal is lessened when many production units are aggregated.□

What might these idiosyncratic shocks be? Like the "iceberg costs" in international trade these could represent shipping accidents, or they could be machine failures, or labor strikes. As long as there is some transitory, heterogeneous component to productivity, an additive noise term, less procyclical than output, appears in aggregate production.

This aggregation argument could be reinterpreted as a story about measurement error. If having more production units in operation gives the econometrician more observations about the relationship between inputs and output, then his estimate of aggregate productivity will become more precise in booms. We do not formulate the model as an estimation error model because the representative agent can infer output from the utility he gains through consumption. However, in a decentralized model, agents might not observe the consumption of other agents.

Finally, it is possible that there is a constant number of production units, but that each one has additive output shocks. These shocks would represent events at the firm level whose effects on production are stochastic but do not scale up with production (at least not at business cycle frequency). For example, a change in the fixed cost of a firms production (overhead such as advertising or accounting) would generate a (non-multiplicative) change in a firms value-added, its measured contribution to GDP. Industrial economics has long emphasized the importance of non-convexities in production at the firm level. Appendix A shows how firm-level non-convexities aggregate to our representative agent production function.

It is not necessary that every production unit, or every sector has additive production shocks. We just need to relax the stringent assumption that every productivity shock, in every sector, is multiplicative. A small deviation, in a few types of business activities, can generate the required asymmetry in learning.

An Economy with Growth

Even though the model and data are analyzed in a stationary environment, learning asymmetry should not disappear when growth is added. As the economy grows, learning will become trivial unless the standard deviation of the production shock grows at the same long-run rate as output.

Our formulation of the additive noise term can be reconciled with asymmetry that does not vanish in a growing economy. In a stationary environment, the model assumes that the variance of η is constant over the business cycle. In an economy with growth rate γ, we must replace η with $\gamma\eta$. The variance of the additive shock $\gamma^2\sigma_\eta^2$ does not vary with the business cycle. This keeps the long-run signal-to-noise ratio constant. There are constant returns to scale in the long-run.[3]

If this assumption were not made, the signal-to-noise ratio would increase over time, the signal extraction problem would become easier and asymmetry would eventually disappear. The skewness of monthly and quarterly changes in industrial production does not support the hypothesis that asymmetry is disappearing over time.[4]

6 Data, Calibration and Computation

If crashes are more sudden than booms, then the most extreme downward changes in output will be larger in magnitude than the largest upward movements. These changes are measured as differences in the log of output. If the negative changes are more extreme, then the unconditional distribution of log differences will exhibit negative skewness.

[3] In the decentralization above, multiplying the idiosyncratic production unit shock η_t^i by γ keeps the signal-to-noise ratio from exploding in an economy with growth.

[4] Skewness statistics for the last six decades are available upon request.

Data

We begin the calibration exercise by computing second and third moments of macroeconomic aggregates for the last 50 years. Panel A of table 2 lists moments of U.S. output (y), capital(k), employment (n), and consumption (c). All data are quarterly from 1952:1 until 2002:1. Output data is from U.S. Department of Commerce, Bureau of Economic Analysis (BEA). Quarterly capital data is obtained by combining an annual capital stock series from the fixed asset tables of the BEA and quarterly fixed private investment from the Bureau of Economic Analysis used to interpolate between the years. Similarly, quarterly hours worked is constructed using quarterly employment (BLS data: non-farm, seasonally adjusted) to interpolate between yearly hours worked (from the BEA). Nominal GDP, consumption and investment are deflated with their respective GDP deflators (all from BEA). All variables are per-capita, where the total population data is from the Census Bureau.

Methods of Detrending

The choice of detrending procedure is not innocuous for the asymmetry measure. Equally critical is the time horizon of changes in the data. To get a full picture of the frequencies at which growth asymmetry in output appears, we examine various choices of filters and horizons.

The most common filters are the Hodrick and Prescott (1997) filter and the Baxter and King (1995) bandpass filter. Both are two-sided, meaning that they use information both from the past and the future to estimate the filtered value. The use of a two-sided filter raises two issues. First, in a non-stationary version of the model, agents would not have access to the future information necessary to detrend. Second, if the trend line starts to fall because a sudden crash is ahead, this will reduce the magnitude of the crash in the detrended data. The skewness in detrended real per capita GDP (1952:1-2002:4) is .06 for the Hodrick-Prescott and -.44 for the bandpass filter. A one-sided version of the bandpass filter does not have these drawbacks. The skewness in detrended GDP with the one-sided filter is -.22. Finally, we use geometric detrending with two break points to account for the productivity slowdown in the 1970's (1969:4 and 1980:3). The skewness is -.29. Simple log differencing yields a skewness estimate of -.40.

Multi-period Changes To capture the idea that slow booms and big crashes may not only occur at one quarter horizons, we examine skewness of N-period changes in output, varying N between 1 and 16 quarters. In spite of the level differences, all four detrending methods show negative skewness in output in our sample period (1956:1 to 1998:1). Negative skewness diminishes, and eventually vanishes at horizons of around 3-4 years. Figure 2 shows the skewness of N-quarter log differences of real GDP (line with squares). Thus, negative skewness is a subtle but robust feature of the data that mainly arises as a 1 to 12 quarter phenomenon.

Calibration

Following Kydland and Prescott (1982), and Cooley and Prescott (1995), we calibrate the model at quarterly frequency. The depreciation rate of $\delta = 0.0186$ matches the investment to capital ratio and the discount factor $\beta = 0.98$ is chosen to match the capital to output ratio. Since the data is quarterly, δ and $(1 - \beta)$ are one-fourth of their annual rates. Preference parameter $\sigma = 0.386$ is chosen such that steady state labor supply is $1/3^{rd}$. Production is Cobb-Douglas with exponent $\alpha = 0.34$ chosen to match the labor share of income. Both labor hours and the fraction of labor income are fairly constant in the post-war U.S. data. Finally, we choose the coefficient of relative risk aversion to be $\phi = 4$ and check the robustness of our results with low risk aversion $\phi = 2$.

Three parameters are specific to our learning model. These are the probability of a change in the technology state, the relative values of those states, and the variance of the production shock σ_η^2.

The probability of a technology change is chosen so that the implied autocorrelation of technology matches Cooley and Prescott (1995)'s estimate of 0.95. The absolute level of the technology state is unimportant since results compare percentage deviations from trend. However, the distance between z^H and z^L will determine the volatility of technology. On the basis of estimates of Solow residuals, Cooley and Prescott (1995) set the standard deviation of the technology shock at 0.007 in an AR(1) specification for $ln(z)$. In our formulation, this corresponds to a technology process with standard deviation of 0.032. Setting z^H and z^L to $(1 + 0.032, 1 - 0.032)$ produces the same standard deviation.

The shock variance σ_η^2 is important because it determines how easy or difficult it is to learn z. If the shocks are large, estimates will be bad and agents will have difficulty learning that transitions have occurred. If the shocks are very small, learning becomes trivial. The standard deviation of the shock $\sigma_\eta = 0.02$ is chosen to match the observed correlation between the median forecast error and observed real GDP.

Computational Details

To estimate the value function and policy rules for the social planner, we use value iteration on a grid, with linear interpolation between nodes. There are two continuous state variables: the capital stock, k_t, and the believed technology level, $\tilde{z}_t$. Because capital leaves the grid in some states of the world, we need an extrapolation method. We use the function $V_z(k) = -\psi k^{-\kappa}$ and determine the two unknown coefficients by matching the level and slope of the interpolated function at the grid boundaries. The value of infinite capital is zero, and zero capital has negative infinite value.

We simulate the model for 50,000 periods and compute statistics on 250 sub-samples of 200 periods each. This is the length of the data. Skewness of the first-differenced simulated data is our measure of growth asymmetry.

Learning and Output Volatility

A significant challenge in comparing the model to data lies in a fundamental tension between learning and output volatility. If the (i.i.d.) shock η is to disguise the true technology state z, then it has to be big enough to make a boom look like a recession. Any process that would make learning non-trivial would also imply an unrealistically high volatility and low autocorrelation in output.

To resolve this conflict, the time series that we compare to observed GDP data is not y, but rather the estimate of the persistent component of output at the end of the period, $\hat{y}$. The reason a data collecting agency might want to report $\hat{y}$ is because a measure of the persistent component of output is useful for predicting future outcomes. To be precise, the data collecting agency would like to report $y_t - \eta_t$, but does not observe η. Instead, it constructs and reports a filtered GDP series, given public information available at the end of the period: $\hat{y}_t = E[z_t|\mathcal{F}_{t+1}]f(k_t, n_t)$. The filtered time series $\hat{y}$ can be interpreted as revised data.[5] The interpretation of η as the contribution of intangibles to output is consistent with this goal of revisions since intangibles are likely to be a component of output that is poorly-measured and hard to predict. Lastly, for the national income accounts to balance, consumption must be filtered in a manner identical to output. Consumption is a residual determined by $\hat{c}_t = \hat{y}_t - i_t$.

Three pieces of empirical evidence support the interpretation of observed GDP as filtered estimates. First, revisions to GDP are often quite large. Mean absolute revision from preliminary to final estimates of annual US real GDP growth data are 1.3 percentage points for the period 1983-2000. Data collecting agencies report revisions to previous estimates that make use of data

[5]An alternative view of GDP revisions is that they correct measurement error. However, Mankiw and Shapiro (1986) and Grimm and Parker (1998) find empirical evidence that subsequent revisions to US GNP/GDP growth data display increasing variances. Increasing variances is consistent with updating filtered estimates, but inconsistent with successive reduction of measurement error. Sargent (1989) combines both views in an economy where a data collecting agency observes error-ridden data, but reports a filtered version. The filtering also generates a reported time series that is smoother than the true output.

available only after the period to which the estimate pertains, i.e. they filter (see Fixler and Grimm (2002)). Second, Faust, Rogers and Wright (2001) study GDP announcements for different countries and show that successive revisions move GDP estimates closer to trend. They conclude that revisions reflect the removal of idiosyncratic "noise" from the series. Third, the Bank of England, one of the only banks that disclose their estimation procedure, explicitly states that a goal of GDP revisions is to reduce "variation around a relatively fixed level" (ESA95 Inventory, Chapter 2.1, p.33, January 2001).

7 Results

No-Learning Model

To isolate the effect of learning on business cycles, we compute results for a calibrated version of the model with no learning. The model is identical to the learning model except that the technology level is revealed at the start of each period: $z^t \in \mathcal{F}^t$. This no-learning model differs from the standard real business cycle model in the literature in two respects. First, technology follows a Markov switching, rather than an AR(1) process. Second, and more substantively, there are additive *i.i.d.* shocks to output. Since labor and investment are chosen at the beginning of the period, the shocks move consumption around. Agents insure against a bad shock to consumption by investing less and working more.

Moments of the simulated no-learning model are summarized in panel B of table 2. Following Cooley and Prescott (1995), we HP-filter all simulation results before comparing second moments in model and in the data. Because of the distortionary effect of the two-sided HP filter on skewness, we compute skewness of changes in the unfiltered series.

In matching the second moments, this model suffers from the same problems as the original real business cycle models. One example is the low volatility of labor. Changing from CES preferences to Hansen (1985) preferences could remedy this.

Even without learning, the model generates some asymmetry. This asymmetry comes from the assumption that consumption is a residual that absorbs output shocks. When productivity has been high and falls, agents slash investment in order to prevent a large fall in consumption. When productivity has been low and rises, a small capital stock keeps the output level low. An aggressive investment policy with small planned consumption could result in an unexpected consumption disaster. Instead, risk-averse agents increase investment slowly as output rises, and the "boom" is more gradual. This asymmetry can be seen in the negative skewness of output estimates, labor, and investment. This is one source of asymmetry in the model. However, it can reproduce only half the asymmetry in the output and it does not produce any skewness statistics that are significantly different from zero.

Learning Model

To match the third moments of the data, the learning mechanism is needed. The learning model (table 2 panel C) matches the asymmetry in output, labor and capital. All of the aggregates, except capital, have skewnesses that are significantly different from zero. For investment, the model skewness is too low, but the point estimate in the data is within the 95 percent confidence interval of the model's skewness.

Due to the smoothing effect of estimating trend GDP, the volatility of output is slightly too low. This tells us that, if governments are releasing smoothed estimates as GDP numbers, the volatility of productivity shocks may be higher than previously estimated. Lowering the risk aversion parameter to $\phi = 2$ remedies the problem somewhat, while leaving the third moments largely unchanged.[6]

[6] Results available upon request.

The model can also reproduce the pattern of negative but diminishing skewness, as the horizon increases. The reason skewness diminishes is that 80 percent of the change in output from the boom or crash occurs within 8 periods, on average. Changes over a period that includes the entire boom or crash will not measure the speed of these movements. We compute skewness in multi-period output changes ($y_t - y_{t-N}$) for the learning model, using different detrending methods. The model starts off with short-horizon asymmetry below zero. The asymmetry approaches zero as the horizon lengthens (figure 2, line with circles). Finally, we compare this to the N-period skewnesses for the no-learning model (line with crosses). Asymmetry in the no-learning model disappears after 9 quarters. In the learning model, asymmetry diminishes, but stays negative at all horizons, similar to the log-differenced data.

Another way to see the effect learning has on asymmetry is to vary the difficulty of learning. The theory predicts that both high-noise (high σ_η) and low-noise (low σ_η) environments should have little asymmetry. Asymmetry should be strongest where learning is neither impossible nor trivial. Figure 3 shows that output skewness is most negative when the variance of η is close to half the variance in output. As the ratio of noise to output variance approaches zero and one, skewness approaches zero.

Finally, we ask what is the effect of our assumption that agents are all passive, rather than active learners? We solve a model in which there is no separation between the filtering and the control problem. The resulting moments are almost indistinguishable from the results of the passive learning model.

Welfare Analysis

Accounting for asymmetries can have important implications for predicting the future fluctuations of macroeconomic aggregates. To isolate the cost of ignoring learning asymmetry, we consider two agents. The first agent updates beliefs, using the model specified in this paper. The second agent is unaware of the learning asymmetry in the economy. He believes that the standard deviation of the additive shock is proportional to production f_t, making learning equally difficult throughout the business cycle. Placing both of these agents in the environment of our model, we ask how the one with the misspecified model fares, relative to the agent who understands learning asymmetry.

To compare the welfare of our two agents, we compute the total cost of business cycles for each agent. Lucas (1987) defines the total cost of business cycles as the compensation in consumption, χ, needed to equalize lifetime utility with a world without business cycles.

$$E_0 \sum_{t=1}^{\infty} \beta^t u((1+\chi)c_t, n_t) = \sum_{t=1}^{\infty} \beta^t u(E_0(c_t, n_t)).$$

The sources of risk in the model are technology shocks z and the additive shocks η.

We find that the cost of business cycles for the agent with the misspecified model is 3.8% higher. The consumption multipliers, χ, are 0.16% and 0.15% for the agents with the asymmetric and symmetric models. Given the low volatility of consumption and hours worked in the standard RBC model, the small costs are not surprising. However, to the extent we think there are gains from reducing business cycle fluctuations, this result suggests some benefits from understanding the source of asymmetry.

8 Conclusion

Learning about the technology level over a business cycle can generate asymmetry in booms and crashes. The reason for the asymmetry is that when agents believe that the level of technology is high, they invest more and work harder. Because technology movements are amplified by the production function, more production makes the technology process is more transparent. It becomes easier to forecast the future accurately. Because of the mean-reversion in technology, large downturns are most likely when technology is high and large upturns when technology is low. Therefore, when large downturns happen, agents quickly develop accurate estimates of the decline. The high speed of learning at the peak of a cycle causes the response to a bad technology shock to be sudden and the high precision of the estimates causes that reaction to be large. Booms happen when production is low and it is very difficult to observe changes in technology. More gradual changes in beliefs and more uncertainty about productivity cause booms to be more gradual. We embed this mechanism into a dynamic stochastic general equilibrium model.

While suffering from the same second moment problems as a standard model, the learning model is a substantial improvement in matching the asymmetry in output, investment and employment. As in the data, the model predicts that the degree of asymmetry diminishes at lower frequencies.

The crucial assumption for asymmetric learning is the assumption that the standard deviation of the idiosyncratic shocks does not rise proportionately with production over the business cycle. If noise and signal grew at the same rate, then the filtering problem would be just as difficult in good times and bad. Data on analyst forecasts support the main premise of the model, that uncertainty varies counter-cyclically. This work suggests that there are significant asymmetries in the data that can be explained by uncertainty and learning about productivity over the business cycle.

Learning asymmetry also arises when the amount of information available about productivity is constant, but its transmission is cyclical. VanNieuwerburgh and Veldkamp (2004) show that when the value of an asset rises, information about that asset becomes more valuable. In a business cycle setting with costly information acquisition, agents would obtain more precise signals in booms, when the value of capital is high. More precise information in booms implies a procyclical signal-to-noise ratio. As this paper has shown, that generates business cycle asymmetry.

The results may shed some light on related asset-pricing puzzles. Asymmetric learning generates counter-cyclical uncertainty that may account for counter-cyclical movements in asset price volatility (Bekaert and Wu (2000)), counter-cyclical equity premia (Campbell and Cochrane (1999)) and return asymmetry (Chen, Hong and Stein (2001)). To generate quantitatively meaningful asymmetry predictions for asset pricing, this model would need to be augmented with habit persistence (Boldrin, Christiano and Fisher (2001)) or frictions such as endogenous borrowing constraints (Lustig and VanNieuwerburgh (2004)).

A Appendix: Aggregation over Productive Units

This appendix shows that the main feature of the model, that the signal-to-noise ratio is procyclical, is preserved under aggregation. Model predictions are agnostic about the scale of operation and the number of production units (industries, firms or plants). We show the following property. Holding the number of production units fixed, if the scale of production units varies over the business cycle, then the aggregate signal-to-noise ratio is procyclical.

The output of each production unit is given by $y_t^i = z_t f(k_t^i, n_t^i) + \eta_t^i$. Let there be N productive units in the economy each operating at a scale $f_t^i = f(k_t^i, n_t^i)$. The technology shock is aggregate and continues to follow a 2-state Markov chain with variance σ_z^2. The function f is increasing, concave, and twice continuously differentiable in both arguments and is homogenous of degree one. The additive productivity shock η_t^i is normally distributed with mean zero and constant variance σ_η^2.

We investigate two (natural) cases for the cross-correlation of additive shocks:

- *(i)* $Corr(\eta_t^i, \eta_t^j) = 0, \ \forall\ (i,j)$ and
- *(ii)* $Corr(\eta_t^i, \eta_t^j) = 1, \ \forall\ (i,j)$.

The aggregate output equation $y_t = z_t f(k_t, n_t) + \eta_t$ is obtained by summing over the outputs of individual production units. By definition $\eta_t = \sum_{i=1}^N \eta_t^i$ and $f(k_t, n_t) = \sum_{i=1}^N f(k_t^i, n_t^i)$.

Because the variance of the additive shock η_t^i is constant over the business cycle and production f_t^i is procyclical, the individual production unit's signal-to-noise ratio is procyclical and equals $(f_t^i)\frac{\sigma_z}{\sigma_\eta}$.

Proposition 2. *For any* $Corr(\eta_t^i, \eta_t^j)$, $\forall\ (i,j)$, *if the individual signal-to-noise ratio is procyclical but the number of firms is constant over the business cycle, then the signal-to-noise ratio is procyclical at the aggregate level.*

With N production units, the aggregate signal-to-noise ratio is: *(i)* $\sqrt{N}(f_t^i)\frac{\sigma_z}{\sigma_\eta}$ *(ii)* $(f_t^i)\frac{\sigma_z}{\sigma_\eta}$. In each case (and any intermediate case) and for any given level of N, the aggregate signal-to-noise ratio is at least as procyclical at the aggregate as at the production unit level.

Section 5 in the main text shows a related property: Holding the scale of a productive unit fixed, if the number of production units varies over the business cycle, then the aggregate signal-to-noise ratio is also procyclical. The difference between the above setups is that in the first one, there are N signals observed but the quality of the individual signals varies procyclically. In the second formulation, there is a constant signal quality at the individual level, but there is a procyclical number of signals. Either mechanism supports an aggregate production function with a procyclical signal-to-noise ratio.

References

Acemoglu, D., and A. Scott 1997, Asymmetric business cycles: Theory and time-series evidence, Journal of Monetary Economics 40, 501–33.

Backus, D. P. Kehoe and F. Kydland 1992, International real business cycles, Journal of Political Economy 100(4), 745–75.

Baxter, M., and R. King 1995, Measuring business cycles: Approximate band-pass filters for economic time series, NBER Working Paper no 5022.

Bekaert, G., and G. Wu 2000, Asymmetric volatility and risk in equity markets, The Review of Financial Studies 13(1), 1–42.

Boldrin, M., and D. Levine 2001, Growth cycles and market crashes, Journal of Economic Theory (96), 13–39.

Boldrin, M. L. Christiano and J. Fisher 2001, Habit persistence, asset returns, and the business cycle, American Economic Review 91(1), 149–66.

Cagetti, M. L. Hansen T. Sargent and N. Williams 2001, Robustness and pricing with uncertain growth, Review of Financial Studies.

Campbell, J., and J. Cochrane 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior., Journal of Political Economy 107(2), 205–251.

Chakley, M., and I.-H. Lee 1998, Learning and asymmetric business cycles, Review of Economic Dynamics 1, 623–45.

Chamley, C., and D. Gale 1994, Information revelation and strategic delay in a model of investment, Econometrica 62(5), 1065–85.

Chen, J. H. Hong and J. Stein 2001, Forecasting crashes : Trading volume, past returns, and conditional skewness in stock prices, Journal of Financial Economics 61(3), 345–81.

Cogley, T. 2002, Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey, Journal of Monetary Economics 49, 309–334.

Cooley, T., and E. Prescott 1995, Frontiers of Business Cycle Research (Princeton University Press, Princeton, NJ).

Ebell, M. 2001, Why are asset returns more volatile during recessions? a theoretical explanation, Working Paper Studienzentrum Gerzensee Workin Paper No. 01.01.

Evans, G. S. Honkapohja and P. Romer 1998, Growth cycles, American Economic Review 88(June), 495–515.

Faust, J. J. Rogers and J. Wright 2001, News and noise in g-7 gdp announcements, Mimeo Federal Reserve Board , International Finance Division.

Fixler, D., and B. Grimm 2002, Reliability of gdp and related nipa estimates, Survey of Current Business 9–27.

Grimm, B., and R. Parker 1998, Reliability of the quarterly and annual estimates of gdp and gross domestic income, Survey of Current Business 12–21.

Hansen, G. 1985, Indivisible labor and the business cycle, Journal of Monetary Economics 16, 309–27.

Hansen, G., and E. Prescott 2000, Capacity constraints, asymmetries and the business cycle, mimeo.

Hodrick, R., and E. Prescott 1997, Post-war u.s. business cycles: An empirical investigation, Journal of Money, Credit and Banking 29, 1–16.

Jovanovic, B. 2003, Asymmetric cycles, Mimeo New York University.

Jovanovic, B., and Y. Nyarko 1982, Learning by doing and the choice of technology, Econometrica 50(6), 1345–70.

Kasa, K. 1995, Signal extraction and the propagation of business cycles, Federal Reserve Bank of San Francisco Working Papers in Applied Economic Theory, no.95-14.

Kocherlakota, N. 2000, Creating business cycles though credit constraints, Federal Reserve Bank of Minneapolis Quarterly Review Summer, 2–10.

Kydland, F., and E. Prescott 1982, Time to build and aggregate fluctuations, Econometrica 50(6), 1345–70.

Lang, W., and L. Nakamura 1990, The dynamics of credit markets in a model with learning, Journal of Monetary Economics 26, 305–18.

Lucas, R. 1987, Models of Business Cycles Yrjo Jahnsson Lectures).

Lustig, H., and S. VanNieuwerburgh 2004, A theory of housing collateral, consumption insurance and risk premia, Working Paper NYU Stern and University of Chicago.

Mankiw, G., and M. Shapiro 1986, News or noise? an analysis of gnp revisions, Survey of Current Business 66, 20–25.

Neftci, S. 1984, Are economic times series asymmetric over the business cycle?, Journal of Political Economy 92(2), 307–28.

Potter, S. 1999, Fluctuations in confidence and asymmetric business cycles, Mimeo Federal Reserve Bank of New York.

Rich, R., and J. Tracy 2003, Modeling uncertainty: Predictive accuracy as a proxy for predictive confidence, Federal Reserve Bank of New York Staff Reports no 161.

Rob, R. 1991, Learning and capacity expansion under demand uncertainty, Review of Economic Studies 58(4), 655–75.

Sargent, T. 1989, Two models of measurement and the investment accelerator, Journal of Political Economy 97(2), 251–87.

VanNieuwerburgh, S., and L. Veldkamp 2004, Information acquisition and portfolio underdiversification, NYU Working Paper.

Veldkamp, L. 2004, Slow boom, sudden crash, forthcoming Journal of Economic Theory.

Veronesi, P. 1999, Stock market overreaction to bad news in good times: A rational expectations equilibium model, The Review of Financial Studies 12(5), 975–1007.

Williams, N. 2004, Small noise asymptotics for a stochastic growth model, Journal of Economic Theory 119, 271–298.

Zeira, J. 1999, Informational overshooting, booms and crashes, Journal of Monetary Economics 43, 237–57.

B Tables and Figures

Forecast Horizon	+1	+2	+3	+4
Median Error	-0.00	-0.11	-0.25^{**}	-0.33^{**}
(t-stat)	(0.07)	(1.58)	(3.65)	(4.98)

Table 1: Dependent variable is the absolute value of the median analyst forecast error. Independent variable is real GDP as a percentage deviation from tend. T-statistics, based on HAC Newey-West corrected standard errors, are in parentheses. A * denotes significance at the 5% level, ** denotes significance at the 1% level. Analyst forecast data is from the survey of professional forecasters, available at *www.phil.frb.org/econ/liv/index.html*. Data is quarterly from 1968:4 to 2003:1 (137 observations). The number of analysts varies between 9 and 76 with an average of 36 per quarter.

	standard deviation	relative std deviation	first-order autocorrelation	correlation with y	skewness
	Panel A: Data				
y	1.69	1.00	0.85	1.00	-0.40
inv	7.41	4.39	0.79	0.89	-0.72
n	1.55	0.92	0.89	0.85	-0.16
k	0.27	0.16	0.96	0.31	-0.11
c	1.26	0.74	0.84	0.89	-0.33
	Panel B: No Learning Model				
$\hat{y}$	1.63	1.00	0.71	1.00	-0.23
	(0.028)	(0.000)	(0.004)	(0.000)	(0.177)
inv	6.34	3.88	0.70	0.97	-0.24
	(0.109)	(0.012)	(0.004)	(0.013)	(0.177)
n	0.57	0.35	0.70	0.97	-0.16
	(0.010)	(0.002)	(0.004)	(0.009)	(0.176)
k	0.41	0.30	0.95	0.04	-0.01
	(0.007)	(0.026)	(0.001)	(0.009)	(0.052)
$\hat{c}$	0.80	0.50	0.72	0.99	-0.18
	(0.014)	(0.008)	(0.004)	(0.000)	(0.167)
	Panel C: Learning Model				
$\hat{y}$	1.52	1.00	0.79	1.00	-0.41
	(0.022)	(0.000)	(0.004)	(0.000)	(0.084)
inv	6.15	3.98	0.70	0.82	-0.42
	(0.088)	(0.007)	(0.006)	(0.004)	(0.096)
n	0.55	0.36	0.70	0.82	-0.33
	(0.008)	(0.001)	(0.006)	(0.004)	(0.097)
k	0.40	0.26	0.96	-0.15	-0.04
	(0.006)	(0.001)	(0.001)	(0.004)	(0.055)
$\hat{c}$	1.09	0.72	0.29	0.82	-0.22
	(0.014)	(0.003)	(0.009)	(0.001)	(0.079)

Table 2: Key Moments of Aggregate Macroeconomic Series. Panel A reports macroeconomic aggregates from 1952:3-2002:2 (200 observations). All series are real, per-capita. Second moments are computed for percentage deviations from Hodrick-Prescott trend. Skewness is of the first-differenced log series. Panel B reports moments from the no-learning model (50,000 simulations). Benchmark calibration; the coefficient of relative risk aversion is $\phi = 4$. Second moments are computed for percentage deviations from HP trend. Skewness is of first-differenced series. $\hat{y}, \hat{c}$ are the true persistent components of output and consumption. Standard errors (in parentheses) are computed from 250 sample moments of series, each 200 observations long. Panel C reports the moments from the learning model (50,000 simulations). $\hat{y}, \hat{c}$ are agents' estimated persistent components of output and consumption.

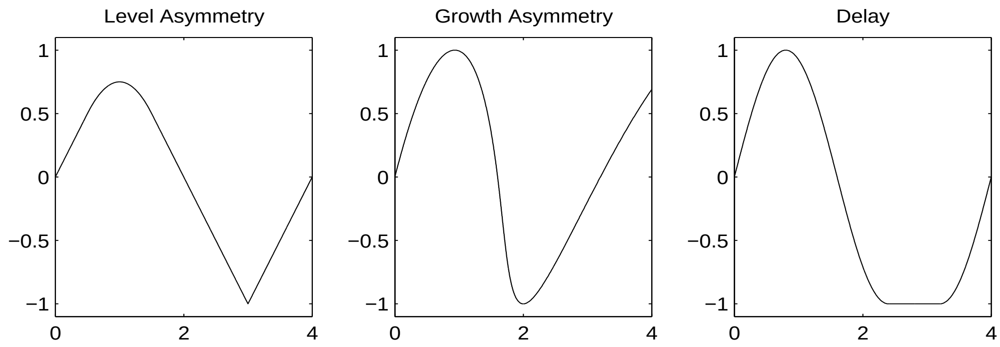

Figure 1: Types of Asymmetry.

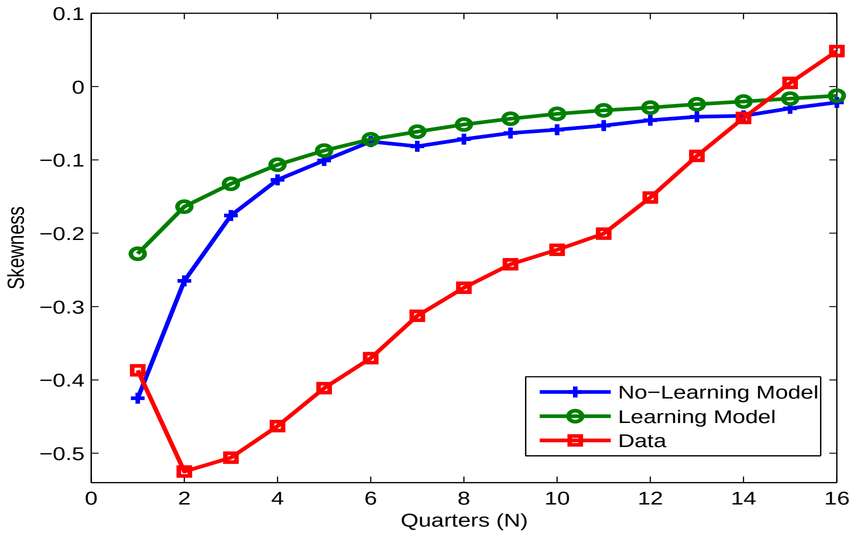

Figure 2: Skewness of N-Period Changes in Learning Model, No-Learning Model, and Data. Skewness of $(y_t - y_{t-N})$, where y_t is log of real GDP per capita. Data are for the sample period 1956:1-1998:1. Model is for benchmark parametrization.

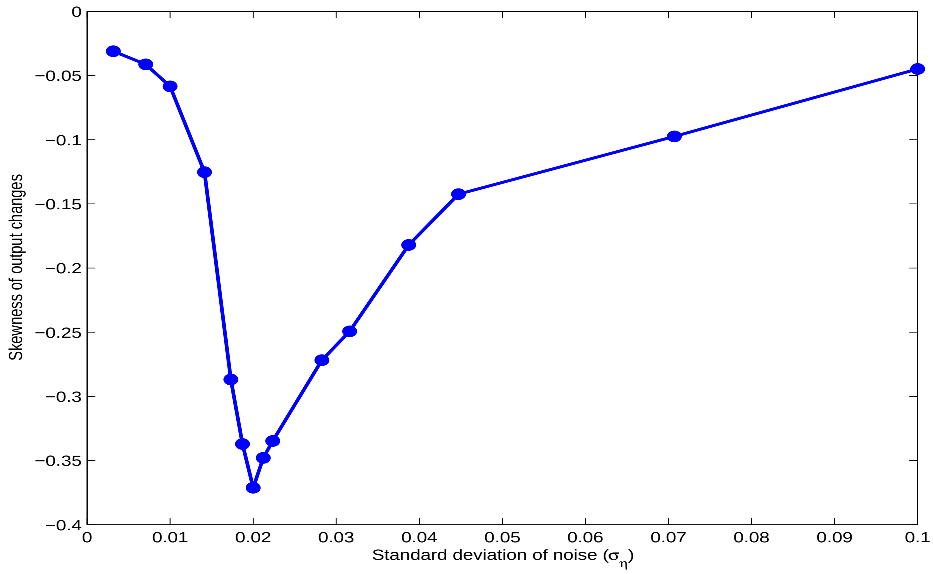

Figure 3: Skewness of output changes ($\Delta\hat{y}$) as the standard deviation of noise (σ_η) varies.