Monetary Policy and Global Equilibria in a Production Economy

Tim Hursey and Alexander L. Wolman

Macroeconomic models that are applied to the study of monetary policy often exhibit multiple equilibria.[1] Prior to the mid-1990s, applied monetary theory typically modeled monetary policy in terms of a rule for the money supply, and it was well understood that multiple equilibria often arose under constant money supply policies. Starting in the mid-1990s, applied work shifted to modeling monetary policy in terms of interest rate rules. This was mainly because of the accumulating observations that central banks in fact operated with interest rate targets rather than money supply targets. A particular class of interest rate rules—so called "active Taylor rules," featuring a strong response of the policy interest rate to inflation—attracted special attention. In linearized models these policy rules were shown to guarantee a locally unique nonexplosive equilibrium. Benhabib, Schmitt-Grohé, and Uribe looked beyond the local dynamics in a series of articles (e.g., 2001a, 2001b, 2002), and showed that active Taylor rules could in fact lead to multiple equilibria. Whereas local analysis ignored the zero bound on nominal interest rates, global analysis showed that the zero bound implied the existence of a second steady-state equilibrium, with low inflation and a low nominal interest rate. This second steady state proved to be the "destination" for paths that had appeared explosive in the local analysis. Benhabib, Schmitt-Grohé, and Uribe's results attracted much attention in the academic literature because the prevailing wisdom had held that active Taylor

[1] Michener and Ravikumar (1998) provide a taxonomy of multiple equilibria in monetary models that predates the recent sticky-price literature.

rules generated a *unique* equilibrium. More recently, the persistence of low inflation and low nominal interest rates has brought attention to Benhabib, Schmitt-Grohé, and Uribe's work in policy circles. Most notably, Bullard (2010) argued that monetary policy in the United States could unintentionally be leading the economy to a steady state in which inflation is below its target.

This article provides an introduction to Benhabib, Schmitt-Grohé, and Uribe's work on multiple equilibria under active Taylor rules, using two simple models. While the type of results presented here is not new, the specific modeling framework—Rotemberg price setting in discrete time—is new, and it fits neatly into the frameworks typically used for applied monetary policy analysis. Furthermore, we provide computer programs in the open source software R to replicate all the results in the article. The programs are available at www.richmondfed.org/research/economists/bios/wolman_bio.cfm.

Section 1 places the topic of this article in historical perspective. Section 2 shows the existence of multiple equilibria in a reduced-form model consisting only of an active Taylor rule and a Fisher equation, assuming that the real interest rate is exogenous and fixed. Section 3 describes the discrete-time Rotemberg pricing model to be used in the remainder of the article. Steady-state equilibria and local dynamics are described in Section 4, and global dynamics are described in Section 5. Section 6 concludes.

1.   HISTORICAL CONTEXT

Multiple equilibria is a common theme in monetary economics, and has been at least since the work of Brock (1975). On the theory side, there has been a steady stream of work on multiple equilibria since the 1970s. In contrast, emphasis on multiple equilibria in applied monetary policy research has fluctuated as new theoretical results have appeared, the tools of analysis have evolved, and economic circumstances have changed. The immediate explanation for why the theoretical results described in this article have attracted attention in policy circles—10 years after those results first appeared—involves economic circumstances, namely the existence of low inflation and near-zero nominal interest rates in the United States. There is a longer history, however, that also involves the ascent of interest rate feedback rules and linearized New Keynesian models, and the accompanying focus on active Taylor rules as a descriptive and prescriptive guide to central bank behavior.

Beginning with Bernanke and Blinder (1992), quantitative research on monetary policy in the United States rapidly shifted from modeling monetary policy as controlling the money supply to modeling monetary policy as controlling interest rates.[2] At around the same time, Henderson and McKibbin

---

[2] Bernanke and Blinder were not the first to suggest modeling monetary policy in terms of interest rates. See for example McCallum (1983).

(1993) and Taylor (1993) influentially proposed particular rules for the conduct of monetary policy. These rules involved the policy rate (federal funds rate in the United States) being set as a linear function of a small number of endogenous variables, typically including inflation and some measure of real activity. Henderson and McKibbin focused on the normative aspects of interest rate rules, whereas Taylor also argued that what would become known as the "Taylor rule" actually provided a reasonable description of short-term interest rates in the United States from 1986–1992.

Just as Taylor rules were attracting more attention, another shift was occurring in the nature of quantitative research on monetary policy. Bernanke and Blinder's 1992 article had used vector autoregressions (VARs) for its empirical analysis and, in their policy analysis, Henderson and McKibbin employed linear rational expectations models with some rule-of-thumb behavior. These two approaches—VARs and linear rational expectations models—had become standard in applied monetary economics for empirical analysis and policy analysis, respectively. Beginning with Yun (1996), King and Wolman (1996), and Woodford (1997), however, the tide shifted toward what Goodfriend and King (1997) called New Neoclassical Synthesis (NNS) models. NNS models represented a melding of real business cycle (RBC) methodology—dynamic general equilibrium—with nominal rigidities and other market imperfections. Nominal rigidities made the NNS models appealing frameworks for studying monetary policy, and the RBC methodology meant that it was straightforward to model the behavior of monetary policy as following a Taylor-style rule.

While NNS models, like RBC models, were fundamentally nonlinear, they were typically studied using linear approximation. In linearized NNS models (as with their predecessors, the linear rational expectations models), the question of existence and uniqueness of equilibrium generally was presumed to be identical to the question of whether the model possessed unique stable local dynamics in the neighborhood of the steady state around which one linearized.[3] In turn, the nature of the local dynamics depended on the properties of the interest rate rule. Although specific conditions can vary across models, the results in Leeper (1991) and Kerr and King (1996) were the basis for a useful rule of thumb in many monetary models: Taylor-style interest rate rules were consistent with unique stable local dynamics only if the coefficient on inflation was greater than one; a coefficient less than one would be consistent with a multiplicity of stable local dynamics. Taylor rules with a coefficient greater than one became known as active Taylor rules, and the rule of thumb

---

[3] For example, see Blanchard and Kahn (1980) or King and Watson (1998). In many economic models, explosive paths for some variables are inconsistent with equilibrium. For example, explosive paths for the capital stock can be inconsistent with a transversality condition (in nontechnical terms, consumers would be leaving money on the table), and explosive paths for real money balances can violate the requirement of a nonnegative price level. See Obstfeld and Rogoff (1983) for a discussion of these issues.

that active Taylor rules guaranteed a unique equilibrium became known as the Taylor principle.[4] Passive Taylor rules, in contrast, are Taylor rules with a coefficient on inflation less than one.

Some intuition for the Taylor principle comes from the much earlier work of Sargent and Wallace (1975) and McCallum (1981). Sargent and Wallace showed that if the nominal interest rate is held fixed by the central bank, then in many models expectations of future inflation will be pinned down, but the current price level is left indeterminate. McCallum followed up by showing that if the nominal interest rate responds to some nominal variable it is also possible to pin down the price level. The Taylor principle states that multiplicity can occur if the nominal interest rate does not respond strongly enough to inflation, consistent with the message of Sargent and Wallace and McCallum.

With widespread understanding of the Taylor principle came empirical applications by Clarida, Gali, and Gertler (2000) and Lubik and Schorfheide (2004). These authors argued that (i) violation of the Taylor principle could help explain the macroeconomic instability of the 1970s, and (ii) a shift in policy so that the Taylor principle did hold could help explain the subsequent stability after 1982. Although this work brought multiple equilibria into the mainstream of applied research on monetary policy, it proceeded under the assumption that the local linear dynamics gave an accurate picture of the nature of equilibrium. These articles also helped to cement the idea that the Taylor principle characterized "good" monetary policy, because the Taylor principle would guarantee that inflation stayed on target.

Beginning with their 2001a article, Benhabib, Schmitt-Grohé, and Uribe (BSU) showed that when there is a lower bound on nominal interest rates, the local dynamics can be misleading about the uniqueness of equilibrium when monetary policy is described by an active Taylor rule. The details of BSU's argument will become clear below. The rough intuition is as follows. Arguments for (local) uniqueness of equilibrium with active Taylor rules posit that without shocks, the model has a unique equilibrium at the inflation rate targeted by the interest rate rule. Any other candidate solutions to the model equations would have the inflation rate exploding to plus or minus infinity, or oscillating explosively. But many of these explosive paths would violate the lower bound on the nominal interest rate. When that bound is imposed and the model is studied nonlinearly, it becomes clear that (i) there is a second steady-state equilibrium at a lower inflation rate, and (ii) there are many

---

[4] Note that Leeper (1991) emphasizes that an active rule guarantees uniqueness only in conjunction with an assumption about fiscal policy, specifically that fiscal policy takes care of balancing the government budget. We maintain that assumption here. Benhabib, Schmitt-Grohé, and Uribe (2002) discuss the implication of alternative assumptions about fiscal policy for multiple equilibria induced by the zero bound on nominal interest rates.

non-steady-state equilibria in which the inflation rate converges to the low-inflation steady state in the long run.

Initially, while the articles by BSU were widely cited, they did not attract much attention in policy circles. This is somewhat surprising because the articles were showing that a policy advocated in large part because it was believed to deliver a unique equilibrium actually delivered multiple equilibria in some models! Furthermore, a rule that violated the Taylor principle—a passive rule—would actually be consistent with keeping inflation close to its targeted value, even though there could be multiple equilibria with this property. Recently however, the results in BSU have attracted substantial attention in policy circles. The simultaneous occurrence of low inflation and low nominal interest rates in the United States is suggestive of some of the equilibria identified by BSU, so it is natural to wonder whether we are experiencing outcomes associated with those global equilibria. Policymakers care about this because the global equilibria involve average inflation below its intended level.

2. A SIMPLE FRAMEWORK WITH ONLY NOMINAL VARIABLES

As a simple framework for communicating some of the key ideas in BSU, this section works through a two-equation model of the nominal interest rate and inflation. That minimal structure is sufficient to illustrate the potential for the local and global dynamics to diverge when monetary policy is given by an active Taylor rule.

Assume that the real interest rate is exogenous and fixed, $r_t = r$, whereas the nominal interest rate (R_t) and the inflation rate (π_t) are endogenous.[5] Expectations are rational. The model consists of a Fisher equation relating the short-term nominal interest rate to the short-term real interest rate and expected inflation,

$$R_t = r E_t \pi_{t+1}, \tag{1}$$

and a rule specifying how the central bank sets the nominal interest rate—in this case as a function only of the current inflation rate, with an inflation target of π^*:

$$R_t = 1 + \left(R^* - 1\right) \left(\pi_t / \pi^*\right)^{\gamma}, \tag{2}$$

where

$$R^* = r\pi^*; \tag{3}$$

---

[5] Throughout the article, interest rates and inflation rates are measured in gross terms—that is, a 4 percent nominal interest rate would be written as $R_t = 1.04$.

that is, the targeted nominal interest rate is the one that is implied by the steady-state Fisher equation when inflation is equal to its target.

The interest rate rule in (2) may look unfamiliar relative to standard linear Taylor rules. We use the nonlinear rule because it will simplify the analysis in the second part of the article.[6] Furthermore, the linear approximation to the rule in (2) around $\{R^*, \pi^*\}$ is

$$R_t - R^* = \gamma \left(\frac{R^* - 1}{\pi^*} \right) \left(\pi_t - \pi^* \right), \tag{4}$$

a simple inflation-only Taylor rule in which the coefficient on inflation is $\gamma \left(R^* - 1 \right) / \pi^*$, and we assume that $\gamma \left(R^* - 1 \right) / \pi^* > r > 1$. The standard local-linear approach around the point $\{R^*, \pi^*\}$ involves combining the linearized Taylor rule (4) with the linearized Fisher equation ($R_t - R^* = (R^*/\pi^*) E_t \left(\pi_{t+1} - \pi^* \right)$), which yields an expectational difference equation in inflation:

$$E_t \left(\pi_{t+1} - \pi^* \right) = \gamma \left(\frac{R^* - 1}{R^*} \right) \left(\pi_t - \pi^* \right).$$

For simplicity, assume perfect foresight—that is, the future is known with certainty, so that $E_t \left(\pi_{t+1} - \pi^* \right)$ can be replaced with $\pi_{t+1} - \pi^*$. Perfect foresight is clearly an unrealistic assumption, but it is a convenient one for illustrating the difference between local and global dynamics. With perfect foresight, we have

$$\left(\pi_{t+1} - \pi^* \right) = \gamma \left(\frac{R^* - 1}{R^*} \right) \left(\pi_t - \pi^* \right). \tag{5}$$

By assumption the coefficient on $\pi_t - \pi^*$ is greater than one—the rule obeys the Taylor principle. Consequently, we can show that there is a unique non-explosive equilibrium. Constant inflation at the targeted steady-state level ($\pi_t = \pi^*$) is clearly an equilibrium because it represents a solution to the difference equation (5). If inflation in period t were equal to any number other than π^*, inflation would have to follow an explosive path going forward because the coefficient on current inflation is greater than one. Any such explosive path would be ruled out as an equilibrium by assumption in the standard local-linear approach.[7]

---

[6] Imposing the zero bound on an otherwise linear rule creates a nondifferentiability, making computation more difficult.

[7] Since the model here is itself ad-hoc, we cannot complain about ruling out explosive paths as equilibria by assumption. Depending on the particular model, explosive paths up or down may or may not be equilibria—see footnote 3. What is important here is that the ad-hoc model we wrote down is nonlinear, and the nonlinear analysis yields different conclusions about equilibrium than the linear analysis.
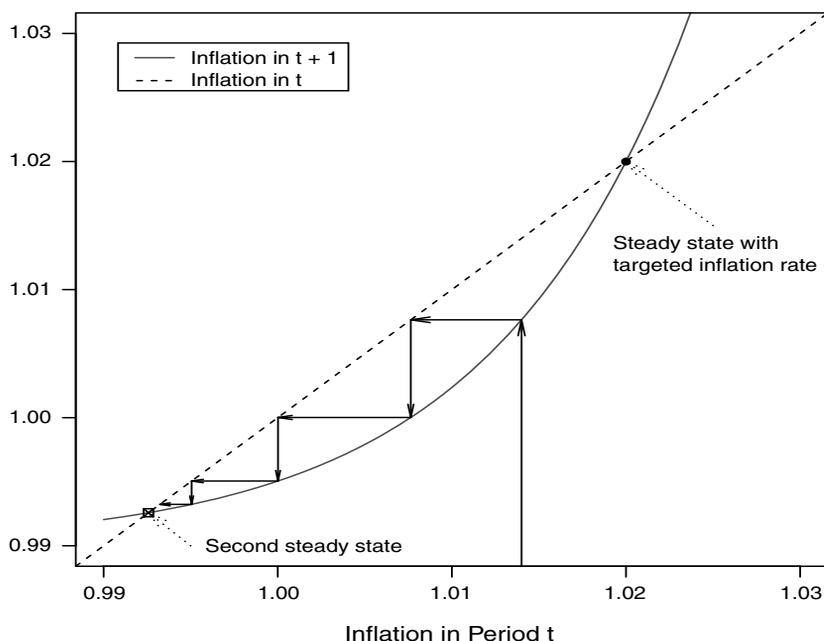
Figure 1 Steady-State Equilibria
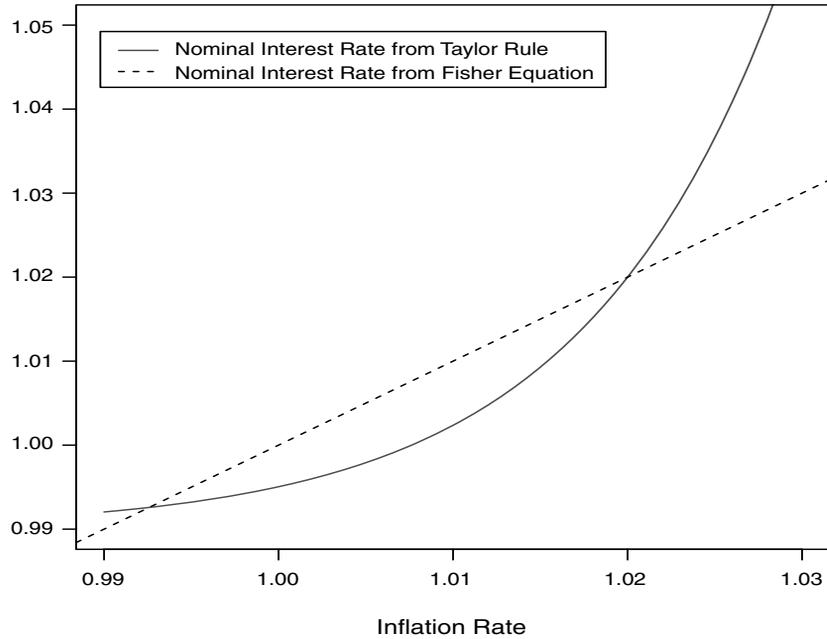
Steady-State Equilibria

It is obvious that $\{R^*, \pi^*\}$ represents a steady-state solution to the Fisher and Taylor equations ([1] and [2]). Less obviously, there is also a second steady-state solution with a lower inflation rate and a lower nominal interest rate. To see this, combine the steady-state Fisher and Taylor equations into a single equation in π:

$$\pi = r^{-1}\left(1 + \left(R^* - 1\right)\left(\pi/\pi^*\right)^{\gamma}\right). \tag{6}$$

Figure 1 displays a plot of the right-hand side of (6) (essentially the Taylor rule) against the 45-degree line—which is also the left-hand side, or the Fisher equation. The two intersections of the right-hand side and left-hand side represent the two steady-state equilibria. The targeted inflation rate is 2 percent, and the other steady state involves slight deflation.

The specific Taylor rule we chose for this example never allows the nominal interest rate to hit the zero bound. Alternatively, if we had chosen a typical linear Taylor rule ($R_t = \max\{R^* + f\left(\pi_t - \pi^*\right), 0\}$), there would be a kink in the steady-state Taylor curve at $\pi = 1/r$, and the second steady state would be at $\pi = \pi^* - (1/f)\,R^*$. BSU (2001a) and Bullard (2010) contain pictures of the analogues to Figure 1 implied by several different interest rate rules that

Figure 2 Example of a Non-Steady-State Equilibrium

all satisfy the Taylor principle at the targeted steady state, and all imply the existence of a second steady state with lower inflation.

Example of a Non-Steady-State Equilibrium

The fact that there are two steady-state equilibria suggests that there may also be equilibria in which inflation and nominal interest rates fluctuate. Returning now to the nonlinear model, by combining the Fisher equation (1) and the interest rate rule (2) and imposing perfect foresight, we have a first-order difference equation for the inflation rate:

$$\pi_{t+1} = r^{-1} \left(1 + \left(R^* - 1 \right) \left(\pi_t / \pi^* \right)^{\gamma} \right). \tag{7}$$

This is the nonlinear analogue of (5). In contrast to the linearized model, we can show that there is a continuum of nonexplosive equilibria.[8] In Figure 2 we plot the right-hand side of (7): It is an identical curve to the solid line in

---

[8] Note the sensitivity of this result to whether current or (expected) future inflation is the argument in the policy rule. If the policy rule responds to π_{t+1} instead of π_t, then the same two steady-state equilibria exist; but the system is entirely static and, under perfect foresight, the two steady-state equilibria are also the only two equilibrium values for inflation in any period. The

Figure 1. The dotted line is the 45-degree line, which is also the left-hand side of (7). The intersections between the two lines are the steady states and, starting with any initial inflation rate below the targeted steady state, we can trace an equilibrium path using the solid line and the 45-degree line. For example, from an initial inflation rate of 1.014, the vertical solid lines with arrows pointing down indicate the successive values of inflation going forward. Generalizing from this example, the figure shows that all perfect foresight equilibria except for the targeted steady state converge to the nontargeted steady state. In contrast, the conventional local linear approach applied to the targeted steady state would conclude that the targeted steady state was the only equilibrium—other solutions are locally explosive and would be ruled out by assumption. Figure 2 conveys the essence of the literature that began with BSU (2001a): Local analysis suggests a unique equilibrium, whereas global analysis reveals that many solutions ruled out as explosive instead lead to a second steady-state equilibrium.

Because the qualitative results involving a second steady state and multiple equilibria will carry over into the model with an endogenous real interest rate and endogenous output, it is interesting to discuss the economics behind these results. In a neighborhood of the targeted steady state, the interest rate rule responds to an upward (downward) deviation of inflation from target by moving the interest rate upward (downward) more than proportionally. This sets off a locally explosive chain: The Fisher equation (1) dictates that an increase in the current nominal interest rate must correspond to a higher future inflation rate, which then is met with a further increase in next period's interest rate, etc. One notable aspect of this process is that there is no sense in which a higher nominal interest rate represents "tighter" monetary policy. The model has only nominal variables, and a higher nominal interest rate *must* correspond to higher expected inflation. In contrast, the Taylor principle is often thought of as ensuring that an increase in inflation is met with a monetary tightening, as represented by a higher nominal interest rate. In models with real effects of monetary policy—such as the one discussed below—an increase in the nominal interest rate does not have to correspond to higher expected inflation. However, we have learned from the two-equation model that this association of higher interest rates with tight monetary policy is not an inherent ingredient in the local uniqueness and global multiplicity associated with the Taylor principle.[9]

"economy" can bounce arbitrarily between those two values in a deterministic way. There may also be rational expectations equilibria with stochastic fluctuations.

[9] See Cochrane (2011) for a similar argument.

3.   A MODEL WITH REAL VARIABLES AND MONETARY NONNEUTRALITY

The model above taught us that the Fisher equation together with a Taylor rule that responds strongly to inflation can lead to multiple steady states and other equilibria because of the lower bound on nominal interest rates. However, the only endogenous variables in that model are nominal variables. One of the simplest ways to endogenize real variables and introduce real effects of monetary policy is with a version of the Rotemberg (1982) model, which has quadratic costs of nominal price adjustment. In this model, there is a representative household that takes all prices and aggregate quantities as given, and chooses how much to consume and how much to work. There is a continuum of monopolistically competitive firms that face convex costs of adjusting their nominal prices, and there is a monetary authority that sets the short-term nominal interest rate according to a time-invariant feedback rule.

The representative household has preferences over consumption (c_t) and (disutility of) labor (h_t) given by

$$\sum_{t=0}^{\infty} \beta^t \left(\ln (c_t) - \chi h_t \right).   \tag{8}$$

There is a competitive labor market in which the real wage is w_t per unit of time. The consumption good is a composite of a continuum of differentiated products ($c_t(z)$), each of which are produced under monopolistic competition:

$$c_t = \left(\int_0^1 c_t(z)^{\frac{\varepsilon-1}{\varepsilon}} \, dz \right)^{\frac{\varepsilon}{\varepsilon-1}}.   \tag{9}$$

Households own the firms. An individual household's budget constraint is

$$c_t + R_t^{-1} B_t / P_t = w_t h_t + B_{t-1}/P_t + \Pi_t/P_t,   \tag{10}$$

where Π_t represents nominal dividends from firms, P_t is the price of the composite good, and B_t is the quantity of one-period nominal discount bonds. As above, R_t is the gross nominal interest rate. The household's intratemporal first-order conditions representing optimal choice of labor input and consumption are given by

$$\lambda_t w_t = \chi,   \tag{11}$$

and

$$\lambda_t = 1/c_t,   \tag{12}$$

and the intertemporal first-order condition representing optimal choice of bondholdings is given by

$$\frac{\lambda_t}{P_t} R_t^{-1} = \beta \cdot \frac{\lambda_{t+1}}{P_{t+1}}.   \tag{13}$$

In these equations, the variable λ_t is the Lagrange multiplier on the budget constraint for period t—it can also be thought of as the marginal utility of an additional unit of consumption at time t. Note that the intertemporal first-order condition (13) corresponds to the Fisher equation from the first model, with the real interest rate now endogenous and given by

$$r_t = \beta^{-1} \frac{c_{t+1}}{c_t}.$$

Firms face a cost (ξ_t) in terms of final goods of changing the nominal price of the good they produce (z):

$$\xi_t(z) = \frac{\theta}{2} \left(\frac{P_t(z)}{P_{t-1}(z)} - 1 \right)^2. \tag{14}$$

Because goods are produced both for consumption and for accomplishing price adjustment, the market-clearing condition is

$$y_t = c_t + \frac{\theta}{2} (\pi_t - 1)^2, \tag{15}$$

where y_t denotes total output of the composite good, π_t denotes the gross inflation rate (P_t/P_{t-1}), and we have imposed symmetry across firms, meaning that all firms choose the same price.

An individual firm chooses its price each period to maximize the expected present value of profits, where profits in any single period are given by revenue minus costs of production minus costs of price adjustment. The demand curve facing each firm is $y_t(z) = (P_t(z)/P_t)^{-\varepsilon} y_t$, so the profit maximization problem for firm z is

$$\max_{P_{t+j}(z)} \sum_{j=0}^{\infty} \beta^j \left(\frac{\lambda_{t+j}}{\lambda_t} \right) \left[\frac{P_{t+j}(z)}{P_{t+j}} \left(\frac{P_{t+j}(z)}{P_{t+j}} \right)^{-\varepsilon} y_{t+j} \right.$$
$$\left. - w_{t+j} \left(\frac{P_{t+j}(z)}{P_{t+j}} \right)^{-\varepsilon} y_{t+j} - \frac{\theta}{2} \left(\frac{P_{t+j}(z)}{P_{t+j-1}(z)} - 1 \right)^2 \right].$$

The first term in the square brackets is the real revenue a firm earns charging a price $P_{t+j}(z)$ in period $t+j$; it sells $\left(P_{t+j}(z)/P_{t+j} \right)^{-\varepsilon} y_{t+j}$ units of goods for relative price $P_{t+j}(z)/P_{t+j}$. The second term in the square brackets (in the second line of the expression) is the real costs the firm incurs in period $t+j$, number of goods sold multiplied by average cost, which is equal to marginal cost and to the real wage because labor productivity is constant and equal to one. Finally, the third term in the square brackets is the real cost of adjusting the nominal price from $P_{t+j-1}(z)$ to $P_{t+j}(z)$. Note that the price chosen in any period shows up only in two periods of the infinite sum. Thus, the part of the objective function relevant for the choice of a price in period t

is

$$\frac{P_t(z)}{P_t}\left(\frac{P_t(z)}{P_t}\right)^{-\varepsilon} y_t - w_t \left(\frac{P_t(z)}{P_t}\right)^{-\varepsilon} y_t$$

$$-\frac{\theta}{2}\left(\frac{P_t(z)}{P_{t-1}(z)} - 1\right)^2 - \beta\left(\frac{\lambda_{t+1}}{\lambda_t}\right)\frac{\theta}{2}\left(\frac{P_{t+1}(z)}{P_t(z)} - 1\right)^2.$$

The first-order condition is

$$(1-\varepsilon)\frac{1}{P_t}\left(\frac{P_t(z)}{P_t}\right)^{-\varepsilon} y_t + \varepsilon w_t \frac{1}{P_t}\left(\frac{P_t(z)}{P_t}\right)^{-\varepsilon-1} y_t$$

$$-\theta\frac{1}{P_{t-1}(z)}\left(\frac{P_t(z)}{P_{t-1}(z)} - 1\right) + \beta\left(\frac{\lambda_{t+1}}{\lambda_t}\right)\theta\frac{P_{t+1}(z)}{P_t(z)^2}\left(\frac{P_{t+1}(z)}{P_t(z)} - 1\right) = 0.$$

If we multiply both sides by P_t and impose symmetry—that is, assume that all firms choose the same price in any given period, the expression simplifies to

$$(1-\varepsilon)\,y_t + \varepsilon w_t y_t$$

$$-\theta\pi_t(\pi_t - 1) + \beta\left(\frac{\lambda_{t+1}}{\lambda_t}\right)\theta\pi_{t+1}(\pi_{t+1} - 1) = 0.$$

Using the goods market clearing condition (15) and the household's optimality conditions, the previous equation simplifies to a form that we will refer to as the New Keynesian Phillips Curve:[10]

$$(\pi_t - 1)\,\pi_t = \left(\frac{c_t}{\theta} + \frac{(\pi_t - 1)^2}{2}\right)(1 - \varepsilon + \chi\varepsilon c_t)$$

$$+\beta E_t\left(\frac{c_t}{c_{t+1}}(\pi_{t+1} - 1)\,\pi_{t+1}\right), \qquad (16)$$

where π_t is the gross inflation rate.

Finally, monetary policy is given by a nominal interest rate rule similar to what was used in the two-equation model, with the one difference that the interest rate responds to expected future inflation instead of to current inflation:

$$R_t = 1 + (\pi^*/\beta - 1)\,(\pi_{t+1}/\pi^*)^\gamma. \qquad (17)$$

Recall that in the two-equation model, using a policy rule identical to (17) would render the model entirely static, whereas the rule that responds to current inflation introduces dynamics. In the current model, optimal pricing already introduces dynamics, so we choose to use the future-inflation version of the policy rule.[11] Combining the policy rule with the household's intertemporal

---

[10] We should note that the term "New Keynesian Phillips Curve" typically refers to the linearized version of (16).

[11] Note that with current inflation in the policy rule, the steady states do not change and it would be possible to study dynamic equilibria in the same way we do here—tentative results

first-order condition (13), using the definition of the inflation rate to eliminate the price level, and using the household's intratemporal first-order condition (12) to eliminate λ, we have

$$\left(\frac{c_t}{\pi_{t+1}c_{t+1}}\right)^{-1} = \beta \left(1 + \left(\pi^*/\beta - 1\right)\left(\pi_{t+1}/\pi^*\right)^\gamma\right). \tag{18}$$

The model has now been reduced to two nonlinear difference equations (16) and (18) in the variables c_t, π_t, c_{t+1}, and π_{t+1}.

4. LOCAL DYNAMICS AROUND STEADY-STATE EQUILIBRIA

As with the ad-hoc model in Section 2, there are two steady-state equilibria. That there are two steady-state equilibrium inflation rates is immediately apparent from (18)—in a steady state it is identical to (6). One of the steady states has inflation equal to the targeted inflation rate π^*, and the other steady state has a lower inflation rate.[12] The steady-state levels of consumption are determined by (16).

To study dynamic equilibria, we follow the same steps as in the two-equation model, beginning with the linearized model and then moving on to the exact nonlinear model. The two dynamic equations (16) and (18) can be represented as

$$\begin{bmatrix} F\left(c_t, c_{t+1}, \pi_t, \pi_{t+1}\right) \\ G\left(c_t, c_{t+1}, \pi_t, \pi_{t+1}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where

$$F\left(c_t, c_{t+1}, \pi_t, \pi_{t+1}\right) =$$
$$\left(\pi_t - 1\right)\pi_t - \left(\frac{c_t}{\theta} + \frac{\left(\pi_t - 1\right)^2}{2}\right)\left(1 - \varepsilon + \chi\varepsilon c_t\right)$$
$$-\beta\left(\frac{c_t}{c_{t+1}}\left(\pi_{t+1} - 1\right)\pi_{t+1}\right)$$

$$G\left(c_t, c_{t+1}, \pi_{t+1}\right) = \pi_{t+1}c_{t+1} - \beta c_t\left(1 + \left(\pi^*/\beta - 1\right)\left(\pi_{t+1}/\pi^*\right)^\gamma\right).$$

---

suggest that qualitatively similar results apply with current inflation in the policy rule. Our approach in this article is positive rather than normative. For a policymaker choosing a rule, whether multiple equilibria arise would be one important consideration in that choice.

[12] This statement relies again on γ being sufficiently large. In contrast, for low enough γ such that $R'\left(\pi^*\right) < 1$, the second steady state will involve inflation higher than π^*.

Table 1 Parameter Values

β	0.99
ε	6
θ	17.5
χ	5
γ	90
π^*	1.005

Linearizing around the steady state with the targeted inflation rate (denoted $[c^*, \pi^*]$) yields

$$
\begin{bmatrix} F_2\,(c^*, c^*, \pi^*, \pi^*) & F_4\,(c^*, c^*, \pi^*, \pi^*) \\ G_2\,(c^*, c^*, \pi^*) & G_3\,(c^*, c^*, \pi^*) \end{bmatrix} \begin{pmatrix} c_{t+1} - c \\ \pi_{t+1} - \pi \end{pmatrix} \approx
$$
$$
- \begin{bmatrix} F_1\,(c^*, c^*, \pi^*, \pi^*) & F_3\,(c^*, c^*, \pi^*, \pi^*) \\ G_1\,(c^*, c^*, \pi^*) & 0 \end{bmatrix} \begin{pmatrix} c_t - c \\ \pi_t - \pi \end{pmatrix}, \qquad (19)
$$

where $H_j\,(\mathbf{s})$ denotes the jth partial derivative of the generic function $H\,()$, evaluated at $\mathbf{s}$.

The existence and uniqueness of a nonexplosive equilibrium in the linearized model depends on the eigenvalues of the Jacobian matrix J, given by

$$
J = - \begin{bmatrix} F_2\,(.) & F_4\,(.) \\ G_2\,(.) & G_3\,(.) \end{bmatrix}^{-1} \begin{bmatrix} F_1\,(.) & F_3\,(.) \\ G_1\,(.) & 0 \end{bmatrix}.
$$

Neither c_t nor π_t are predetermined variables, so the condition for a unique nonexplosive equilibrium is that both eigenvalues of J be less than one in absolute value. Because we are not able to provide a general proof of the parameter conditions under which equilibrium exists and is unique, we turn to a numerical example, which we will stay with for the rest of the article.[13] Table 1 contains the parameters for that example; they are chosen to be consistent with a 2 percent annual inflation target (the model is a quarterly model), a 4 percent real interest rate, a markup of 20 percent, and a coefficient in the Taylor rule of 1.33 when the Taylor rule is linearized around the targeted steady state. In addition, our choice of θ implies that price adjustment costs are less than $\frac{2}{10}$ percent of output.

At the targeted steady state, the local (nonexplosive) dynamics are unique, in a trivial sense. The Jacobian's eigenvalues are $0.99771321 \pm 0.12791602i$, which means that both eigenvalues have absolute value 1.0059. Local to the

---

[13] If the targeted inflation rate were zero ($\pi^* = 1$) then it would be straightforward to characterize uniqueness conditions analytically—this is the standard New Keynesian Phillips Curve. With a nonzero inflation target there are price-adjustment costs incurred in steady state, and the analysis is less straightforward.

targeted steady state, the fact that both eigenvalues have absolute value greater than one and are imaginary means that any solution to the difference equation system (19) other than the steady state itself oscillates explosively. In the linearized model the local dynamics *are* the global dynamics, so the only nonexplosive solution is the targeted steady state itself.

Suppose instead that we linearize around the low-inflation steady state. There the Jacobian's eigenvalues are 1.1291231 and 0.89509305. This eigenvalue configuration, with one explosive root and one stable root (less than one), means that there is a saddlepath: Given an initial value for c (or an initial value for π), there is a unique initial value for π (or for c) such that the economy will converge from that point to the steady state with low inflation. If either inflation or consumption were predetermined variables, then this saddlepath would describe the unique equilibrium at any point in time. Because neither variable is predetermined, the saddlepath represents one dimension of equilibrium indeterminacy at any point in time. That is, any value of c (or π) is consistent with equilibrium in period t, but as was stated above, once that value of c (or π) has been selected, the associated value of π (or c) is pinned down, as is the entire subsequent equilibrium path.[14]

The conventional linearization approach to studying NNS models, as followed, for example, by King and Wolman (1996), involves implicitly ignoring the steady state with low inflation. In that approach it is presumed that the only relevant steady state is the targeted one. From the same kind of reasoning used in the discussion following (5), the explosiveness of paths local to the targeted steady state means there is a unique nonexplosive equilibrium, the steady state itself. One can then proceed to study the properties of the model when subjected to shocks, for example to productivity or monetary policy. However, the fact that there are two steady states suggests that it may be revealing to investigate the global dynamics. Furthermore, if one extrapolates the local dynamics around the two steady states, it leads to the conjecture that paths that explode locally from the targeted steady state may in fact end up as stable paths converging at the low-inflation steady state. This is indeed what we will find in studying the global dynamics.

5. GLOBAL DYNAMICS

Studying the model's global dynamics means analyzing the nonlinear equations ([18] and [16]). We will combine the nonlinear equations with information about the local dynamics to trace out the global stable manifold of the low-inflation steady state. The global stable manifold is the set of inflation and

---

[14] Because we are dealing here with perfect foresight paths, the discussion of period t really should apply only to an *initial* period, prior to which the perfect foresight assumption does not apply. After that initial period the equilibrium outcomes are unique.

consumption combinations such that if inflation and consumption begin in that set, there is an equilibrium path that leads in the long run to the low-inflation steady state. While this approach may not yield a comprehensive description of the perfect foresight equilibria, it will provide a coherent picture of how the two steady states relate to the dynamic behavior of consumption and inflation.[15] We will find that the local saddlepath can be understood as part of a path (the global stable manifold) that begins arbitrarily close to the targeted steady state and cycles around that steady state with greater and greater amplitude before converging monotonically to the low-inflation steady state.

From Local to Global

Before plunging into the global dynamics, it may be helpful to take stock of our knowledge. There are two steady-state equilibria, one with the targeted inflation rate (π^*) and one with a lower inflation rate (π_l). The levels of consumption in the two steady states are c^* and c_l. Local to the targeted steady state, all dynamic paths oscillate explosively. Local to the low inflation steady state many paths explode and one path converges to that steady state. To go further, we will combine the forward dynamics local to the low inflation steady state with the nonlinear *backward dynamics*. This approach will allow us to compute the global stable manifold of the low-inflation steady state. Since all paths diverge around the targeted steady state, no analogous approach can be applied there.

As described above, the local dynamics around $\{c_l, \pi_l\}$ involve a unique path in $\{c, \pi\}$ space that converges to the steady state. If we begin with a point on that path, very close to the low-inflation steady state, and then iterate the nonlinear system *backward*, we can trace out the global dynamics associated with the saddlepath—the global stable manifold. We now describe this process algorithmically.

1. To find a point on the local saddlepath of the low-inflation steady state, follow the approach described in Blanchard and Kahn (1980). First, decompose the Jacobian matrix J into its Jordan form: $J = P\Lambda P^{-1}$, where Λ is a diagonal 2×2 matrix whose diagonal elements are the eigenvalues of J, and where P is a 2×2 matrix whose columns are the eigenvectors of J. Next, rewrite the system in terms of canonical variables $x_{1,t}$ and $x_{2,t}$, which are linear combinations of c_t and π_t: $\begin{bmatrix} x_{1,t} & x_{2,t} \end{bmatrix}' = P \begin{bmatrix} c_t - c_l & \pi_t - \pi_l \end{bmatrix}'$. The system is

$$\begin{bmatrix} x_{1,t+1} \\ x_{2,t+1} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix}. \tag{20}$$

---

[15] While we have not proved that the global stable manifold contains all perfect foresight equilibria, we conjecture this to be the case.

Note that at the steady state c_l, π_l, we have $x_{1,l} = x_{2,l} = 0$. Recall that one of the roots (λ_1, λ_2) is greater than one. Without loss of generality, assume that $\lambda_1 > 1$. Any point on the local saddlepath must have $x_{1,t} = 0$, because $x_{1,t+j} = \lambda_1 x_{1,t+j-1}$, and if $x_{1,t} \neq 0$ then $x_{1,t+j}$ could not approach 0 as $j \to \infty$. Select one such point within an ε ball of the low-inflation steady state and call that point $\{c_T, \pi_T\}$. Set $t = T$.

2. From (18) we have

$$c_{t-1} = \frac{c_t}{\beta} \left(\frac{\pi_t}{1 + (\pi^*/\beta - 1)(\pi_t/\pi^*)^\gamma} \right).$$

3. Compute π_{t-1} by solving (16):

$$\left(1 - \frac{1}{2}(1 - \varepsilon(1 - \chi c_{t-1}))\right)\pi_{t-1}^2 - \varepsilon(1 - \chi c_{t-1})\pi_{t-1} -$$
$$\left(\frac{c_{t-1}}{\theta}(1 - \varepsilon(1 - \chi c_{t-1})) + \beta\left(\frac{c_{t-1}}{c_t}(\pi_t - 1)\pi_t\right)\right) = 0. \quad (21)$$

With c_{t-1}, c_t, and π_t all known, (21) is a quadratic equation in π_{t-1}. The presence of two solutions is rooted in the properties of the firm's profit-maximization problem—while there is a unique profit-maximizing price, there are multiple solutions to the first-order condition. Only the positive root of the quadratic is consistent with the firm maximizing profits—the negative root typically implies a negative gross inflation rate, which would imply a negative price level.

4. Set $t = t - 1$, return to step 2.

Figure 3 describes the results of iterating backward for 450 periods in steps 2 through 4. The figure is in c, π space. It plots the two steady states and the global stable manifold of the low-inflation steady state, constructed as just described. The arrows represent forward movement in time, as opposed to the backward movement that characterizes the algorithm. The algorithm starts at a point close to the low-inflation steady state and goes backward in time. The figure shows that the only path that converges to a steady-state equilibrium initially involves spirals around the targeted steady state and ends with monotonic convergence to the low-inflation steady state. The figure provides us with a unified understanding of the local results around the two steady states. From the local dynamics we learn that all paths local to the targeted steady state oscillate explosively. From Figure 3, we see that *one* of those paths is not globally explosive, instead converging at the low-inflation steady state. This path is what we refer to as the global stable manifold.

Figure 3 Global Stable Manifold of Low-Inflation Steady State
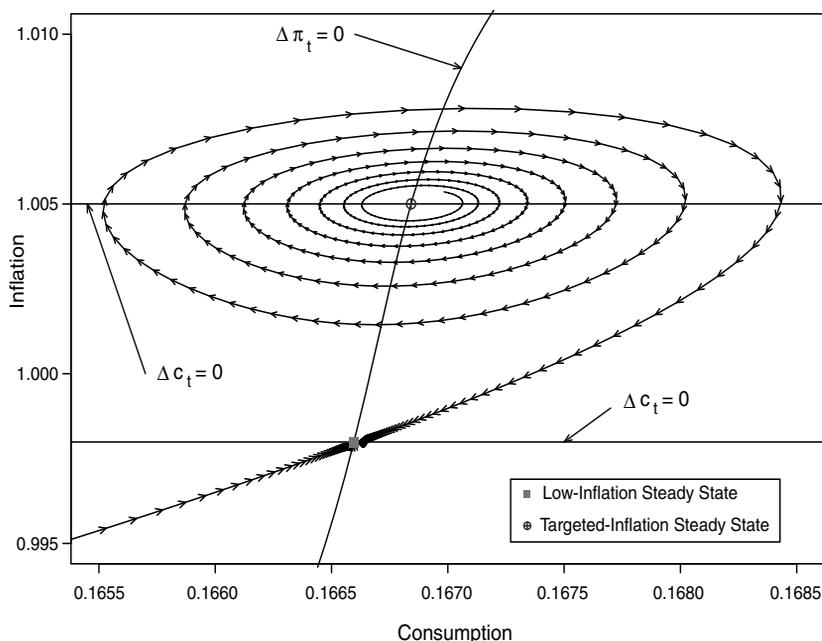
6. CONCLUSION

Since late 2008, both inflation and nominal interest rates have been extremely low in the United States. These facts have focused attention on ideas motivated by the theory in BSU (2001a, 2001b, 2002): An active Taylor rule, together with a moderate inflation target, could have the unintended consequence of leading the economy to undesirably low inflation with a near-zero nominal interest rate. The article by St. Louis Federal Reserve Bank President James Bullard (2010) represents the leading example of this attention.

The aim of this article was to provide an accessible introduction to the ideas in BSU (2001a). Much of the literature in this area uses models that are either set in continuous time or that assume prices are flexible. In contrast, the model in this article is set in discrete time and has sticky prices. Discrete time reduces mathematical tractability, but makes it easy to compute specific solutions; in addition, the quantitative literature on monetary policy overwhelmingly uses discrete time models. Sticky prices are also a central element in the applied monetary policy literature. In adapting BSU's analysis to a discrete-time framework with sticky prices, we have seen that the general conclusions of their work also apply to the specific example we have analyzed. First, with an active Taylor rule, the presence of a lower bound on the nominal interest

rate leads to the presence of two steady states, one at the targeted inflation rate and one at a lower inflation rate. Second, the targeted steady state, which is a unique equilibrium according to the conventional local analysis, instead is the source for a global stable manifold of the low-inflation steady-state equilibrium.

In closing we will offer some caveats regarding using the kind of analysis in this article to interpret current economic outcomes. It is tempting to conclude from Figure 3 that the low-inflation steady state is "more likely" because it does possess a stable manifold while the targeted steady state does not. However, the model only tells us what equilibria exist, not how likely they are to occur. It is also tempting to conclude from this work that policy may be unwittingly leading the economy to the unintended steady state. However, the theoretical analysis is based on perfect information about the model and the equilibrium by all agents. It is interesting to think about situations where policymakers and private decisionmakers do not understand the structure of the economy, but that is not the situation analyzed here. Finally, we should stress that before using this kind of framework for quantitative analysis, it would be desirable to enrich the model to incorporate capital accumulation. The behavior of the capital stock plays a key role in interest rate determination, and at this point it is an open question whether the kind of dynamics described here carry over to models with capital accumulation.

REFERENCES

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2001a. "The Perils of Taylor Rules." *Journal of Economic Theory* 96 (January): 40–69.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2001b. "Monetary Policy and Multiple Equilibria." *American Economic Review* 91 (March): 167–86.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2002. "Avoiding Liquidity Traps." *Journal of Political Economy* 110 (June): 535–63.

Bernanke, Ben S., and Alan S. Blinder. 1992. "The Federal Funds Rate and the Channels of Monetary Transmission." *American Economic Review* 82 (September): 901–21.

Blanchard, Olivier, and Charles M. Kahn. 1980. "The Solution of Linear Difference Models Under Rational Expectations." *Econometrica* 48 (July): 1,305–11.

Brock, William A. 1975. "A Simple Perfect Foresight Monetary Model." *Journal of Monetary Economics* 1 (April): 133–50.

Bullard, James. 2010. "Seven Faces of the Peril." Federal Reserve Bank of St. Louis *Review* 92 (September): 339–52.

Clarida, Richard, Jordi Gali, and Mark Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics* 115 (February): 147–80.

Cochrane, John H. 2011. "Determinacy and Identification with Taylor Rules." http://faculty.chicagobooth.edu/john.cochrane/research/Papers/taylor_rule_jpe_revision.pdf

Goodfriend, Marvin S., and Robert G. King. 1997. "The New Neoclassical Synthesis and the Role of Monetary Policy." In *NBER Macroeconomics Annual 1997*, Vol. 12, edited by Ben Bernanke and Julio Rotemberg. Cambridge, Mass.: MIT Press, 231–96.

Henderson, Dale, and Warwick J. McKibbin. 1993. "A Comparison of Some Basic Monetary Policy Regimes for Open Economies: Implications of Different Degrees of Instrument Adjustment and Wage Persistence." *Carnegie-Rochester Conference Series on Public Policy* 39 (December): 221–317.

Kerr, William, and Robert G. King. 1996. "Limits on Interest Rate Rules in the IS Model." Federal Reserve Bank of Richmond *Economic Quarterly* 82 (Spring): 47–75.

King, Robert G., and Alexander L. Wolman. 1996. "Inflation Targeting in a St. Louis Model of the 21st Century." Federal Reserve Bank of St. Louis *Review* 78 (May): 83–107.

King, Robert G., and Mark W. Watson. 1998. "The Solution of Singular Linear Difference Systems Under Rational Expectations." *International Economic Review* 34 (November): 1,015–26.

Leeper, Eric. 1991. "Equilibria Under 'Active' and 'Passive' Monetary and Fiscal Policies." *Journal of Monetary Economics* 27 (February): 129–47.

Lubik, Thomas A., and Frank Schorfheide. 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *American Economic Review* 94 (March): 190–217.

McCallum, Bennett T. 1981. "Price Level Determinacy with an Interest Rate Policy Rule and Rational Expectations." *Journal of Monetary Economics* 8: 319–29.

McCallum, Bennett T. 1983. "A Reconsideration of Sims' Evidence Concerning Monetarism." *Economics Letters* 13: 167–71.

Michener, Ronald, and B. Ravikumar. 1998. "Chaotic Dynamics in a Cash-in-Advance Economy." *Journal of Economic Dynamics and Control* 22 (May): 1,117–37.

Obstfeld, Maurice, and Kenneth Rogoff. 1983. "Speculative Hyperinflations in Maximizing Models: Can We Rule Them Out?" *Journal of Political Economy* 91 (August): 675–87.

Rotemberg, Julio J. 1982. "Sticky Prices in the United States." *Journal of Political Economy* 90 (December): 1,187–211.

Sargent, Thomas J., and Neil Wallace. 1975. "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule." *Journal of Political Economy* 83 (April): 241–54.

Taylor, John B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39 (December): 195–214.

Woodford, Michael. 1997. "Control of the Public Debt: A Requirement for Price Stability?" In *The Debt Burden and Monetary Policy*, edited by G. Calvo and M. King. London: Macmillan (published version is excerpted from NBER Working Paper No. 5684 [July]).

Yun, Tack. 1996. "Nominal Price Rigidity, Money Supply Endogeneity and Business Cycles." *Journal of Monetary Economics* 37 (April): 345–70.

Hidden Effort, Learning by Doing, and Wage Dynamics

Arantxa Jarque

M any occupations are subject to learning by doing: Effort at the workplace early in the career of a worker results in higher productivity later on.[1] In such occupations, if effort at work is unobservable, a moral hazard problem arises as well. The combination of these two characteristics of effort implies that employers need to provide incentives for the employee to work hard, possibly in the form of pay-for-performance,[2] while taking into account at the same time the optimal path of human capital accumulation over the duration of the contract.

The recent crisis had a big impact on the labor market with high job-destruction rates. If firm-specific human capital accumulation is important, the effect of these separations on welfare may come from several channels. A direct channel is through the loss of human capital prompted by the exogenous separation, as well as the loss in welfare from the decrease in wealth because of unemployment spells of workers. A less direct channel, but potentially an important one, is the change in the cost of providing incentives when the (exogenous to the incentive provision) separation rate increases. However, we are far from being able to understand and measure the importance of this

[1] See Arrow (1962), Lucas (1988), and Heckman, Lochner, and Taber (1998) for a complete discussion of this issue, as well as alternative specifications of learning by doing.

[2] Lemieux, MacLeod, and Parent (2009) report that, for a Panel Study of Income Dynamics sample of male household heads aged 18–65 working in private sector wage and salary jobs, the incidence of pay-for-performance jobs was about 38 percent in the late 1970s and increased to about 45 percent in the 1990s. They define pay-for-performance jobs as employment relationships in which part of the worker's total compensation includes a variable pay component (bonus, commission, piece rate). Any worker who reports overtime pay is considered to be in a non-pay-for-performance job. See also MacLeod and Parent (1999).

cost, since little is known so far about the structure of incentive provision in the presence of learning by doing.[3] This article constitutes a modest first step in this direction: Abstracting from separations and in a partial equilibrium setting, this article studies the time allocation of incentives and human capital accumulation in the optimal contract. This simplified analysis should be a helpful benchmark in future studies of the fully fledged model with separations and general equilibrium.

We modify the standard repeated moral hazard (RMH) framework from Rogerson (1985a) to include learning by doing. In the standard framework, a risk-neutral employer, the principal, designs a contract to provide incentives for a risk-averse employee, the agent, to exert effort in running the technology of the firm. Both the principal and the agent commit to a long-term contract. The agent's effort is private information and it affects the results of the firm stochastically: The probability distribution over the results of the firm (the agent's "productivity") in a given period is determined by the effort choice of the agent *in that same period only*. We introduce the following modification to this standard framework: We specify learning by doing by assuming that the probability distribution over the results of the firm in each period is determined by *the sum of past undepreciated efforts* of the agent, as opposed to his current effort only. In other words, the agent's productivity is determined by his "accumulated human capital." More human capital implies higher expected output, although all possible output levels may realize under any level of human capital. In this specification, the agent determines his human capital deterministically by choosing effort each period. Lower depreciation of past effort is interpreted as "more persistence" of effort.

We present a model of two periods. The first period represents the junior years, when the worker has just been hired and has little experience. The second period represents the mature worker years, when human capital has been potentially accumulated and there are no more years ahead in which to exploit the productivity of the worker. A contract contingent on the observed performance of the agent is designed by the principal to implement the path of human capital accumulation that maximizes the principal's expected profit (expected output minus expected payments to the agent).

In our analysis, we find the following two main implications of the presence of learning by doing. First, the principal does not find it optimal to require a high level of human capital in the last period of the contract, since there is not much time left to exploit the productivity of the worker. Hence, the more experienced workers are not the most productive ones, since they optimally are asked to let their human capital depreciate. This implies that workers exert

---

[3] The only articles dealing with effort persistence in a repeated moral hazard problem are, to our knowledge, Fernandes and Phelan (2000), Mukoyama and Şahin (2005), Kwon (2006), and Jarque (2010).

the most effort in their junior years, and the least in their pre-retirement years. In a comparison with the standard RMH problem, we find that the frontloading of effort, as well as the low requirement at the end of the worker's career, differ markedly from the optimal path of effort in a context without learning by doing. Second, and in spite of this difference in effort requirements over the contract length, we find that learning by doing does not imply a change in the properties of consumption paths; hence, the properties of consumption paths found by previous studies, such as Phelan (1994), remain true in this context (see also Ales and Maziero [2009]).

It is worth noting that in our analysis we assume perfect commitment to the contract both from the employer and the employee, and we do not allow for separations to be part of the contract. This means we need to abstract from the usual career concerns that have been explored in the literature (see Gibbons and Murphy [1992]). The implications of the hidden human capital accumulation that we model here should be viewed as complementary to the implications of career concerns.

As pointed out above, the problem studied here differs from the standard RMH in that the contingent contract needs to take into account the persistent effects of effort on productivity. On the technical side, this highly complicates solving for the optimal contract. The fact that both past and current effort choices are not observable means that, at the start of every period, the principal does not know the preferences of the agent over continuation contracts (that is, the principal does not know the true productivity of the agent for a given choice of effort today). Jarque (2010) deals with this difficulty and presents a class of problems with persistence for which a simple solution can be found. The article studies a general framework in which past effort choices affect current output, as opposed to other forms of persistence that one may consider, such as through output autocorrelation (see, for example, Kapička [2008]). The learning-by-doing problem that we are interested in, hence, constitutes a fitting application of the results in Jarque (2010). We adapt the assumptions in Jarque (2010) to a finite horizon and we show how this specification of learning by doing greatly simplifies the analysis of the optimal contract.

In Section 1 we introduce the common assumptions throughout the article. Section 2 presents, as a benchmark, the case in which the principal can directly observe the level of effort chosen by the agent every period, and hence can control his human capital at all times. For reference, we also discuss the case in which the effort of the agent does not have a persistent effect in time. The analytical properties of the problem are discussed in both cases. Then we analyze the main case of interest of this article, in which effort is unobservable and contracts that specify payments contingent on the observable performance of the agent are needed to implement the desired sequence of human capital accumulation. In Section 3, we discuss the case without persistence—a standard two-period repeated moral hazard problem. In Section 4 we discuss the

technical difficulties of allowing for effort persistence in problems of repeated moral hazard, and the solutions provided in the literature. Section 5 presents the framework of hidden human capital accumulation, a particular case of effort persistence. As the main result, we provide conditions under which the problem with hidden human capital can be analyzed by studying a related auxiliary problem that is formally a standard repeated moral hazard problem. Hence, the discussion of the properties of the standard case in Section 3 becomes useful when deriving the properties of the case with persistence. The numerical solution to an example is presented in Section 6, together with a comparison to the standard RMH without learning by doing, and a discussion of the main lessons about the effects of hidden human capital accumulation on wage dynamics. Section 7 concludes.

1.   DESCRIPTION OF THE ENVIRONMENT

The results in this article apply to contracts of finite length T; however, in order to keep the exposition and the notation as simple as possible, we discuss here the case of a two-period contract, $T = 2$. We assume that both parties commit to staying in the contract for the two periods. For tractability, we assume that the principal has perfect control over the savings of the agent. They both discount the future at a rate β. We assume that the principal is risk neutral and the agent is risk averse, with additively separable utility that is linear in effort.

Assumption 1 *The agent's utility is given by $U(c_t, e_t) = u(c_t) - v e_t$, where u is twice continuously differentiable and strictly concave and c_t and e_t denote consumption and effort at time t, respectively.*

There is a finite set of possible outcomes in each period, $Y = \{y_L, y_H\}$. Histories of outcomes are assumed to be observable to both the principal and the agent. We assume both consumption and effort lie in a compact set: $c_t \in [0, y_t]$ and $e_t \in E = [\underline{e}, \overline{e}]$ for all t.

We model the hidden accumulation of human capital by assuming that the effect of effort is "persistent" over time, in a learning-by-doing fashion. That is, we depart from the standard RMH framework, which assumes that the probability distribution over possible outcomes realizations at t depends only on e_t. In our human capital accumulation framework, the probability distribution at t depends on *all* past efforts up to time t. Assumption 2 states this formally for the two-period problem.

Assumption 2 *The agent affects the probability distribution over outcomes according to the following function:*

$$\Pr(y_t = y_H | s_t) \equiv \pi(s_t),$$

where

$$s_1 = e_1, \tag{1}$$

$$s_2 = \rho s_1 + e_2, \tag{2}$$

and $\pi (s)$ is continuous, differentiable, concave, and $\rho \in (0, 1)$.

In the human capital accumulation language, we could equivalently write the law of motion for human capital as

$$s_1 = e_1$$

$$s_2 = (1 - \delta) s_1 + e_2,$$

where $\delta = 1 - \rho$ would represent the depreciation rate. Then,

$$f(s_t) = \begin{cases} y_H & \text{with probability } \pi(s_t) \\ y_L & \text{with probability } 1 - \pi(s_t) \end{cases}$$

could be interpreted as the production function or technology of the firm.

In the rest of the article, we loosely refer to Assumption 2 as effort being "persistent," we refer to s_t as the accumulated human capital at time t, and we refer to ρ as the persistence rate.

The strategy of the principal consists of a sequence of consumption transfers to the agent contingent on the history of outcome realizations, $\mathbf{c} = \{c_i, c_{ij}\}_{i,j=L,H}$, to which the principal commits when offering the contract at time 0. The agent's strategy is a sequence of period best-response effort choices that maximize his expected utility from t on, given the past history of output: $\mathbf{e} = \{e_1, e_{2i}\}_{i=L,H}$. At the beginning of each period, the agent chooses the level of current effort, e_t. Then output y_t is realized according to the distribution determined by all effort choices up to time t. Finally, the corresponding amount of consumption is given to the agent.

A contract is a pair of contingent sequences $\mathbf{c}$ and $\mathbf{e}$. For the analysis in the rest of the article, it will be useful to follow Grossman and Hart (1983) in using utility levels $u_i = u(c_i)$ and $u_{ij} = u(c_{ij})$ as choice variables.[4] To denote the domain for this new choice variable, we need to introduce the following set notation:

$$\mathcal{U}_i = \{u | u = u(c_i) \text{ for some } c_i \in [0, y_i], i = L, H\}$$

$$\mathcal{U}_{ij} = \{u | u = u(c_{ij}) \text{ for some } c_{ij} \in [0, y_j] i, j = L, H\}.$$

---

[4] If the reader is knowledgeable about contract theory, he or she may notice that this is not a simple change of notation. In fact, when computing the solution to numerical examples (see Section 6), we will follow the two-step procedure proposed in Grossman and Hart (1983). This procedure consists of splitting the expected profit-maximization problem of the principal in two steps: (1) cost minimization of implementing a given effort level (on a grid of efforts), and (2) choosing the effort on the grid that implies the highest expected profit for the principal. Using utility as the choice variable, it is easy to show that under the assumptions of this article there will exist a unique minimum in the cost minimization problem.

The contingent sequence of utility is then denoted $\mathbf{u} = \{u_i, u_{ij}\}_{i,j=L,H}$, and we assume that $u_i \in \mathcal{U}_i$ and $u_{ij} \in \mathcal{U}_{ij}$.

In order to keep the expressions in the article as simple as possible, and abusing notation slightly, we also introduce some notation shortcuts. We denote $c_i = u^{-1}(u_i)$ for all i. We also write $\Pr(y_t = y_H | s_t)$ as $\pi_H(s_t)$ and $\Pr(y_t = y_L | s_t)$ as $\pi_L(s_t)$.

The expected profit of the principal, denoted by $V(\mathbf{u}, \mathbf{e})$, depends on the contract as follows:

$$V(\mathbf{u}, \mathbf{e}) \equiv \sum_{i=L,H} \left\{ \pi_i(s_1) \left[y_i - c_i + \beta \sum_{j=L,H} \pi_j(s_{2i}) \left(y_j - c_{ij} \right) \right] \right\},$$

where s_t changes with e_t as detailed in (1). In the same way, we can write the agent's expected utility of accepting to participate in the contract as

$$W_0(\mathbf{u}, \mathbf{e}) = \sum_{i=L,H} \pi_i(s_1) \left[u_i + \beta \sum_{j=L,H} \pi_i(s_{2i}) u_{ij} - v e_{2i} \right] - v e_1. \quad (3)$$

Within this environment we are now ready to set up the problem of finding the optimal contract that will provide the right incentives for human capital accumulation at the least expected cost. Before analyzing the hidden human capital accumulation case, however, we go through a series of related and simpler cases that will serve in clarifying the main case of interest.

2.   OBSERVABLE EFFORT

The case of observable effort is often referred to in the literature as first-best (FB) since it represents the maximum joint utility achievable in the contractual relationship between the principal and the agent. This is because, if effort is observable, the principal can directly control the choice of effort of the agent and, hence, there is no need for incentives. This implies that there is no need to impose risk on the agent, which results in lower expected transfers from the principal to the agent. Although we are interested in the case of unobservable effort, it is useful to also analyze this simpler benchmark to learn about the differences between the problem with effort persistence (human capital accumulation) and the standard RMH problem (in which human capital fully depreciates every period).

We will refer to the problem of the principal when effort is observable as problem FB:

$$\max_{(\mathbf{u}, \mathbf{e})} V(\mathbf{u}, \mathbf{e})$$

$$s.to$$

$$\mathbf{e} \quad \in \quad \left[\underline{e}, \bar{e}\right]^3 \tag{ED}$$

$$u_i \quad \in \quad \mathcal{U}_i, \ u_{ij} \in \mathcal{U}_{ij} \ \forall i \ \forall i, j \tag{CD}$$

$$w_0 \leq W_0\left(\mathbf{u}, \mathbf{e}\right). \tag{PC}$$

The solution to problem FB is a contract that consists of a pair of contingent sequences of utility and effort that maximize the expected profit of the principal subject to the participation constraint (PC)—which assures that the agent expects as much utility from accepting the contract than staying out—and the domain constraints for consumption (CD) and effort (ED). Characterizing the solution to this problem when considering all the possible combinations of (ED) and (CD) binding constraints is very lengthy and tedious. In the interest of space, we choose to discuss here only the case in which neither of the constraints in (CD) or (ED) bind.

What are the properties of consumption and effort in the optimal contract? We learn them from looking at the first-order conditions of the problem. Let $\lambda \geq 0$ be the multiplier of the (PC).[5] We have:

$$
\begin{aligned}
(u_i) \quad &: \quad \frac{1}{u'\left(c_i\right)} = \lambda, \ \text{ for } i = L, H \\
\left(u_{ij}\right) \quad &: \quad \frac{1}{u'\left(c_{ij}\right)} = \lambda, \ \text{ for } i, j = L, H \\
(e_1) \quad &: \quad \left[\pi'\left(s_1\right) + \beta\rho\pi'\left(s_{2i}\right)\right]\left(y_H - y_L\right) = v\lambda \\
\left(e_{2_i}\right) \quad &: \quad \pi'\left(s_{2i}\right)\left(y_H - y_L\right) = v\lambda, \ \text{ for } i = L, H.
\end{aligned}
\tag{4}
$$

We analyze in turn the case with and without persistence.

Full Depreciation

First we analyze the observable effort version of a standard two-period RMH problem (see, for example, Rogerson [1985a]). This case is nested in the common framework presented above, for a value of the persistence parameter $\rho = 0$. In this case, effort does not have a persistent effect on the output distribution, that is, there is no learning by doing. Hence, we can say that the human capital of the agent fully depreciates every period.

Here and throughout the rest of the paper, we use stars to denote the solutions to the problems. When necessary, we index the solutions by two arguments: the first one takes a value P if $\rho > 0$ (persistence) and a value NP if $\rho = 0$ (no persistence). The second one takes a value FB if we are in the case of observable effort and a value SB if we are in the case of

---

[5] Standard arguments for $\lambda > 0$ hold in this setup with persistence. The basic intuition is that $V^*\left(\mathbf{c}, \mathbf{e}; w_0\right)$ is strictly decreasing in w_0.

unobservable effort. Hence, here we denote the solution to problem FB when $\rho = 0$ as $\mathbf{u}^* (NP, FB)$ and $\mathbf{e}^* (NP, FB)$. Note that, whenever it does not lead to confusion, we do not include these arguments to keep the notation light.

Since the right-hand sides of all the first-order conditions for utility are equal to λ, we conclude that the level of utility, and hence consumption, should be the same independent of the output realizations and the period: $u_i^* = u_{ij}^* = u^*$ for all i, j. The first-order conditions of effort, in turn, imply that effort requirements are independent of output realizations and the period: $e_1^* = e_{2i}^* = e^*$ for all i. It is easy to see that, given these properties of consumption and effort, the (PC) in problem FB simplifies to

$$w_0 = (1 + \beta) \left(u^* - v e^* \right).$$

Hence, we can solve for the level of utility in the solution to the FB problem:

$$u^* \equiv \frac{w_0 + v (1 + \beta) e^*}{1 + \beta}. \tag{5}$$

Let $c^* \equiv u^{-1} (u^*)$. Let $\pi_j' (e_2)$ denote the derivative of $\pi_j (e_2)$. Noting that $\pi_H' (e) = -\pi_L' (e)$, we can combine the first-order conditions for consumption and effort to get

$$u' \left(c^* \right) \pi_H' \left(e^* \right) (y_H - y_L) = v \ \forall t. \tag{6}$$

That is, the optimal effort level is such that the marginal benefit from increased effort (the marginal increase in expected output times the marginal utility of output) equals the marginal utility cost of effort.

The following properties summarize our conclusions about the FB problem with nonpersistent effort:

1A. We have that $c_1^* = c_2^* = c^*$.

2A. We have $e_1^* = e_2^* = e^*$.

The main property of the optimal *consumption* sequence of the FB contract in the standard RMH problem is that the contract insures the agent completely against consumption fluctuations whenever feasible. The intuition for this result is straightforward: Since the agent has concave utility in consumption, this is the cheapest way of providing the agent with his outside utility. The main property of the optimal *effort* sequence of the FB contract in the standard RMH problem is a constant effort requirement over time. The tradeoff between increasing the disutility suffered by the agent and increasing the expected output is exactly the same in each period, and hence the solution is the same each time.

It is worth noting that the solution in the observable-effort case coincides with that of a repeated static problem ("spot" contract) in which neither the agent nor the principal commit to the two-period contract, and the outside

Table 1 Parameters of the Numerical Example

υ	Marginal effort disutility	5.00
β	Discount factor	0.65
y_H	Output realization, high state	30.00
y_L	Output realization, low state	20.00
w_0	Outside utility	6.55

utility of the agent is $\frac{w_0}{2}$ each period. Hence, commitment has no value in the case of observable effort and no persistence.

An example

Throughout this article, we illustrate the properties of each particular case of the environment presented by solving a particular numerical example. This makes it easy to compare across the different cases presented. The common parameters of the example are listed in Table 1.

We also assume $u(c) = 2\sqrt{c}$ and a probability function

$$\pi(s) = \sqrt{s}, \tag{7}$$

as well as $\underline{e} = 0.01$ and $\bar{e} = 0.99$.

We now solve for c^* and e^*. Since we are in the case of full depreciation of human capital, we use $\rho = 0$ and the formulas derived above. For our example, we have that (6) becomes

$$\frac{1}{2\sqrt{e^*}}(30 - 20) = 5\sqrt{c^*}$$

$$\frac{1}{2\sqrt{e^*}} = \sqrt{c^*}$$

$$c^* = \frac{0.25}{e^*}.$$

Together with (5) this gives us the solutions listed in Table 2.

Observable Human Capital Accumulation

We now turn to analyzing the case in which the effects of effort are persistent in time, with $\rho > 0$. That is, we analyze the optimal contract in the presence of human capital accumulation, or learning by doing.

We established above that the main property of the optimal consumption sequence of the FB contract in the standard RMH problem is that the contract insures the agent completely against consumption fluctuations. Here we will learn that this property remains true in the case with effort persistence. The main property of the optimal effort sequence of the FB contract in the standard

RMH problem is also a constant effort requirement over time. We will learn that when effort is persistent this property no longer holds: Effort requirements will vary over time even in the observable effort benchmark.

We now proceed to derive these results by formally analyzing the problem of the principal FB for the case of $\rho > 0$. She chooses an optimal contract: a pair of contingent sequences $\mathbf{u}^*(P, FB)$ and $\mathbf{e}^*(P, FB)$ that solve problem FB, i.e., they maximize the expected profit of the principal subject to (PC) and the domain constraints (CD) and (ED). We initially discuss the case in which neither the (CD) nor the (ED) constraint bind. However, the lower (ED) constraint (the non-negativity constraint on effort) may bind, with persistence, in not-so-trivial cases. Because of its relevance, the case of this constraint binding will be discussed in turn.

We can derive the properties of the solution by analyzing the first-order conditions in (4) for the case of $\rho > 0$. The first thing to note is that, as in the case without persistence, neither consumption nor effort are contingent on output realizations. However, effort recommendations will depend on the time period. We can use the (PC) here as well to derive the optimal level of utility:

$$u^* \equiv \frac{w_0 + v\left(e_1^* + \beta e_2^*\right)}{1 + \beta}.$$

The optimal level of consumption will be $c^* \equiv u^{-1}(u^*)$. We can substitute the first-order condition for effort e_2 into that for e_1, as well as the expression of λ from the consumption first-order conditions, to get an expression for the tradeoff determining the choice of e_1:

$$u'\left(c^*\right) \pi'_H\left(s_1^*\right)(y_H - y_L) = v\left(1 - \beta\rho\right). \tag{8}$$

Comparing this to the tradeoff determining the choice of e_2,

$$u'\left(c^*\right) \pi'_H\left(s_{2i}^*\right)(y_H - y_L) = v, \tag{9}$$

we learn that the marginal cost of increasing effort in the first period is different (smaller) than that in the second period. The optimal choice takes into account that any effort e_1 exerted in the first period persists into the second one, i.e., it "saves" the agent the equivalent of the discounted disutility of effort of exerting ρe_1 in the second period. This difference in the effective cost of effort that appears because of persistence implies that the principal sets the effort requirements in a way that implies a higher probability of observing y_H in the first period than in the second. We can see exactly how this difference is determined by using the first-order conditions of effort to get the following relationship:

$$\frac{\pi'_H\left(s_1^*\right)}{1 - \beta\rho} = \pi'_H\left(s_2^*\right). \tag{10}$$

This implies $s_1 > s_2$ since $1 - \beta\rho$ is always between 0 and 1. From the accumulation of human capital in (1) we have that

$$
\begin{aligned}
e_1^* &= s_1^*, \\
e_2^* &= s_2^* - \rho s_1^*,
\end{aligned}
\tag{11}
$$

which implies a higher effort in the first period than in the second, $e_1^* > e_2^*$.

The following properties summarize our conclusions about the case with persistence and observable effort:

1B. We have that $c_1^* = c_2^* = c^*$.

2B. We have that $e_1^* > e_2^*$.

That is, whenever c^* is feasible in both states, the principal provides complete consumption smoothing, both across states and across time. As for effort requirements, the principal decreases the requirement from the first to the second period. We repeat the intuition for this result: In the first period, the effort disutility incurred by the agent is a sort of "investment," since it improves the conditional distribution not only in the current period but also in the following one. At $t = 2$, however, there is no period to follow, so the marginal benefits of effort are not as high, while the marginal cost is the same as in the first period.[6]

An example

We now solve for the optimal contract with persistence and observable effort. For this case with accumulation of human capital, we use $\rho = 0.2$ and the formulas derived above. We list the solution in Table 2. Note that the level of $s_2^*(P, FB)$ in this case is 0.16, smaller than that of the second-period effort in the no-persistence case of the previous section, which was $e_2^*(NP, FB) = 0.17$. Comparing the equations that determine each ([6] for $e_2^*(NP, FB)$ and [9] for $s_2^*(P, FB)$), we can see that $c^*(P, FB) < c^*(NP, FB)$ implies $1/u'(c^*(P, FB)) > 1/u'(c^*(NP, FB))$, and hence $\pi_H'(s_2^*(P, FB)) > \pi_H'(e_2^*(NP, FB))$. Given the concavity of $\pi(\cdot)$, it follows that $s_2^*(P, FB) < e_2^*(NP, FB)$.

The Nonnegativity Constraint on Effort

In light of this solution we can discuss the case of the lower constraint in (ED) binding. As an introduction to why this case is of particular relevance to the

---

[6] In a $T > 2$ framework with $s_0 = 0$, we would have that $e_1 \geq e_t$ for $t < T$, that $e_t = e_2$ for $t = 2, ..., T - 1$, and $e_T \leq e_2$. Again, the intuition is that in all $t < T$, effort improves the conditional distribution not only in the current period, but also in the periods that follow. At $t = 1$, since $s_0 = 0$, effort is higher than in any other period.

Table 2 Solutions for the Numerical Example, FB Problem

FB Solutions	c_1^*	c_2^*	e_1^*	e_2^*	s_1^*	s_2^*
NP	5.82	5.82	0.17	0.17	0.17	0.17
P	4.95	4.95	0.22	0.12	0.22	0.16

problem with persistence, it is useful to consider the effect of changes in the persistence parameter, ρ, on the effort solution just presented. For a value of persistence $\rho = 0$, effort equals accumulated effort trivially, and its level is constant across periods. On the other hand, if we instead substitute a value of persistence $\rho = 1$, $(1 - \beta\rho)$ takes its minimum value in (10) and the solution implies the maximum difference between the level of s_1 and s_2, with s_1 much higher than s_2. However, carefully inspecting (11), we can already see that such high level of persistence cannot be compatible with an interior solution for effort in period 2: The principal would choose $e_2^* = 0$. Since $s_1^* > s_2^*$ for all values of $\rho > 0$, effort e_2^* may not be interior for other high enough values of ρ. In other words, persistence implies that, in many interesting cases, the lower domain constraint on effort (ED) cannot be safely ignored.

Constraint (ED) is represented by the following set of inequalities:

$$s_{2i} \leq \rho s_1 + \bar{e} + \underline{e}, \tag{12}$$

and

$$s_{2i} \geq \rho s_1 + \underline{e}. \tag{13}$$

Constraint (12) may be binding for some parametrizations. However, we choose not to discuss this case explicitly here because it is easy to impose ex ante conditions on the parameters that preclude it from binding; for example, for the specification of the probability in (7), it is easy to see that $s \geq \bar{e}$ is never chosen in the optimal contract. The lower bound on s represented in (13), however, is endogenous, and equation (13) cannot be checked without having the solution s_1^* in hand. Fortunately, in the case of observable effort that we are analyzing here, we are able to include constraint (13) explicitly in the maximization problem FB. This allows us to study how the solution properties differ from those in 1B and 2B discussed above when this constraint binds.

Let $\gamma_i \geq 0$ be the multiplier associated with constraint (13) in the version of problem FB for the case $\rho > 0$. We have that the first-order condition for e_{2i} is modified as follows:

$$(e_{2i}) : \pi'(s_{2i})(y_H - y_L) = v\lambda - \gamma_i, \quad \text{for } i = L, H. \tag{14}$$

Note that, again, the choice for effort in the second period is not contingent on the first-period outcome, so we have $\gamma_L = \gamma_H = \gamma$. Then we can substitute

(14) into the unmodified first-order condition for first-period effort, (e_1), to get a general version of equation (8) that allows for the lower domain constraint of effort to be binding:

$$\pi'(s_1)(y_H - y_L) = v\lambda(1 - \beta\rho) + \beta\rho\gamma. \tag{15}$$

From the Kuhn-Tucker conditions, we know that whenever $\gamma > 0$ we have $e_2^* = 0$ and, hence, $s_2^* = \rho s_1^*$.

An example

In some special cases, we can check ex ante whether $\gamma = 0$ is a feasible solution to the FB problem, and hence we can restrict ourselves to the simpler analysis without domain constraints. In particular, with the specification for the probability function in (7) that we are using for our example, equation (10) becomes

$$\frac{1}{(1 - \beta\rho)2\sqrt{s_1^*}} = \frac{1}{2\sqrt{s_2^*}},$$

or, rewriting,

$$s_2^* = (1 - \beta\rho)^2 s_1^*. \tag{16}$$

This is the relationship that should hold between the level of s_1^* and s_2^* whenever $\gamma = 0$. Hence, the domain condition $e_2 \geq 0$ is satisfied whenever $s_2^* \geq \rho s_1^*$, or, substituting s_2^* from (16), whenever

$$(1 - \beta\rho)^2 \geq \rho. \tag{17}$$

A closer inspection of condition (17) shows that, for $\beta \leq 0.5$, it is always satisfied. For higher β values, however, the condition is satisfied only for low enough ρ values, i.e., when effort is not "too persistent." In our example, for $\beta = 0.65$, we need to check whether (17) is satisfied: The left-hand side is equal to 0.76, which is clearly greater than the right-hand side, 0.2.

To summarize the findings of our analysis, we have shown that for the numerical example presented here, we can provide ex ante conditions (a functional form for the probability as in equations [7] and [17]) on the parameters of the problem that assure us that the domain constraints on (ED) do not bind. Under such restrictions, the characteristics of the solution to the first-best problem 1B and 2B presented earlier in this section are valid.

In relation to those characteristics, it is worth pointing out that the properties of effort requirements depend strongly on our assumption that the utility of the agent is linear in effort. Linearity implies that there is no tradeoff between the efficient accumulation path of human capital and smoothing effort disutility over time. In other words, the smoothing of effort requirements over the duration of the contract does not increase the overall utility of the agent, as is the case with consumption smoothing; hence, the principal only takes

into account the effects that different accumulation paths have on the utility of the agent and his own profit through the changes in expected output over time. In the numerical example in Section 6, we will revisit the solution to the observable-effort case discussed here, and we will see the direct consequence of this: It is optimal to ask the agent to exert effort earlier rather than later in the contract, since effort that is done early improves the distribution over future output, holding constant the level of future effort.

3. UNOBSERVABLE EFFORT WITH FULL DEPRECIATION

When effort is not directly observable, the principal must rely on observed output realizations, which are imperfect signals about the effort level of the agent, in order to implement the desired sequence of human capital. Contrary to the case of observable effort, here consumption in a given period will need to vary with the output realization in order to provide incentives for the worker to choose the recommended level of effort.

Formally, the problem of the principal, which we will refer to as the second-best (SB), is:

$$\max_{(\mathbf{u},\mathbf{e})} V(\mathbf{u},\mathbf{e})$$
$$s.to$$

$$\mathbf{e} \in \left[\underline{e},\overline{e}\right]^3 \tag{ED}$$
$$u_i \in \mathcal{U}_i, \ u_{ij} \in \mathcal{U}_{ij} \ \forall i \ \forall i,j \tag{CD}$$

$$U \leq W_0(\mathbf{u},\mathbf{e}) \tag{PC}$$

$$W_0(\mathbf{u},\mathbf{e}) \geq W_0(\mathbf{u},\widehat{\mathbf{e}}) \ \ \forall \widehat{\mathbf{e}} \neq \mathbf{e}. \tag{IC}$$

The incentive constraint (IC) ensures that the expected utility that the agent gets from following the principal's recommendation is at least as large as that of any other effort sequence.

In order to illustrate clearly the differences that derive from the presence of effort persistence in this two-period problem, we analyze first the version without persistence ($\rho = 0$), that is, with full depreciation of human capital every period, or no learning by doing. Moreover, because the main result that we will derive when we study the case with $\rho > 0$ is that, in some cases, the properties for consumption in the optimal contract will be the same as those of the optimal contract in a framework without persistence, it is useful to analyze in detail the properties of the solution with observable effort.

Without persistence, the structure of the incentive constraints simplifies considerably. This influences the solution, but also the ways in which the problem can be studied. In particular, the standard RMH problem has a simple recursive formulation that is not available with persistence. In this section we

provide an illustration of this difference. Then, we discuss the difficulties of introducing persistence, along with some potential solutions, in Section 4. In Section 5 we discuss our example with human capital accumulation, a particularly simple case with effort persistence for which a solution can easily be found.

A Simplified Incentive Compatibility Constraint

In the case without persistence the structure of the incentive constraints simplifies considerably. In particular, the expected utility of the agent in the second period is independent of the first-period effort choice. Define

$$W_{1i}(\mathbf{u}, \mathbf{e}) = \sum_{j=L,H} \pi_j(s_{2i}) u_{ij} - v e_{2i}, \quad \text{for } i = L, H, \tag{18}$$

as the expected utilities for the second period, contingent on the first-period realization. This expression for the continuation utility simplifies, when $\rho = 0$, to

$$W'_{1i}(\mathbf{u}, e_{2i}) \equiv \sum_{j=L,H} \pi_j(e_{2i}) u_{ij} - v e_{2i}, \quad \text{for } i = L, H. \tag{19}$$

(Note that, to distinguish the notation for continuation utilities here from those of the general case that allows for persistence in (18), we denote them here with a prime and we make explicit the independence of e_1.)

What is the simplification of the incentive constraints that follows from this independence? As it turns out, all the sequences that have the same choice of effort in the second period, regardless of the first-period effort choice, provide the agent with the same expected utility in the second period, conditional on the first-period output realization being the same. In other words, the deviations of the agent in the second period can be evaluated independently of the first-period effort choice, and also independently at each node following the first-period output realization. As a consequence, the number of relevant incentive constraints for the agent is drastically decreased.

To see this formally, denote by $w_{1i} \equiv W'_{1i}(\mathbf{u}, e_{2i})$ the continuation utilities evaluated at the effort requirement of the principal. Then all the incentive constraints that involve deviations *only* in the second period, or that have the same effort choice for the first period, simplify to

$$w_{1i} \geq W'_{1i}(\mathbf{u}, \widehat{e}_{2i}) \quad \forall \widehat{e}_{2i} \neq e_{2i} \text{ for } i = L, H. \tag{20}$$

We refer to equation (20) as the "second-period incentive constraints."[7]

---

[7] For a more concrete illustration, consider the case with discrete effort and $E = [e_L, e_H]$. Then the initial number of IC constraints would be seven, and they would simplify to three: one first-period constraint and two second-period constraints.

Now note that the independence of $W'_{1i}(\mathbf{u}, \widehat{e}_{2i})$ on e_1 also implies the following: Imposing the second-period incentive constraints in (20) serves to assure that all potential deviations $(\widehat{e}_1, \widehat{e}_{2L}, \widehat{e}_{2H})$ that consider effort choices in the second period that are not e_{2H} and e_{2L} are dominated by a strategy $(\widehat{e}_1, e_{2L}, e_{2H})$ that considers the same deviation in period 1 and none in the second period. Formally, what we are saying is that

$$\sum_{i=L,H} \pi_i(\widehat{e}_1)[u_i + \beta w_{1i}] - v\widehat{e}_1 \geq \sum_{i=L,H} \pi_i(\widehat{e}_1)\left[u_i + \beta W'_{1i}(\mathbf{u}, \widehat{e}_{2i})\right] - v\widehat{e}_1$$

trivially simplifies to the second-period incentive constraint in (20). This is useful because it means that when we are evaluating deviations in the first period we forget about potential deviations in the second period as well, and simply substitute w_{1i} into the second period utility:

$$\sum_{i=L,H} \pi_i(e_1)[u_i + \beta w_{1i}] - ve_1 \geq \sum_{i=L,H} \pi_i(\widehat{e}_1)[u_i + \beta w_{1i}] - v\widehat{e}_1. \quad (21)$$

We refer to these constraints as the "first-period incentive constraints."

The independence of second-period expected utility on first-period effort choice not only decreases the number of IC constraints that we need to consider, but also allows the problem of the principal to be analyzed period by period. This is precisely because all future period payoffs can be summarized through the promised utility w_{1i} without specifying the particular consumption transfers or effort recommendations that will deliver w_{1i} in the future. From a practical point of view, it is important to note that the range of values that w_{1i} can take is independent of the agent's action in the first period, and hence can be calculated by simply using the domain restrictions for consumption and second-period effort, together with the second-period IC in (20). This is a very useful feature when we want to compute the solution for a particular numerical example, as we will do in Section 6.

To summarize, the simplifications we just discussed are the reason why the recursive formulation first introduced by Spear and Srivastava (1987) is possible. In a finite two-period problem like the one presented here, this also means that we can solve the problem backward and characterize the properties of the solution. We proceed to do that now.

A Backward Induction Solution to the Optimal Contract

As a first step, we use the fact that incentives in the second period are independent of choices and utilities in the first period. This allows us to split the problem of the agent in the IC into two problems: a first-period problem and a second-period problem. The second-period problem, PIC_2, is

$$\max_{e_{2i} \in [\underline{e}, \overline{e}]} \sum_{j=L,H} \pi_j(e_{2i})u_{ij} - ve_{2i},$$

and the first-period problem, PIC_1, is

$$\max_{e_1 \in [\underline{e}, \overline{e}]} \sum_{i=L,H} \pi_i (e_1) (u_i + \beta w_{1i}) - v e_1,$$

where w_i is the expected utility for the second period in equilibrium.

If we want to characterize the optimal contract, first we need to transform these maximization problems into an equality constraint that we can include in the problem of the principal. Following the spirit of the first-order approach (see Rogerson [1985b]), we establish concavity of the maximization problems in PIC_1 and PIC_2. Then we can substitute them by their first-order conditions, which are necessary and sufficient for a maximum. In our two-outcome example, this concavity is fairly straightforward to guarantee. It is easy to see that, for any positive first-period effort recommendation to satisfy the original first-period IC in (21), we need $u_H + \beta w_H > u_L + \beta w_L$. Also, for any second-period positive effort recommendation to satisfy the second-period IC in (20), we need $u_{iH} > u_{iL}$. Since we have assumed that $\pi_H (\cdot)$ is a concave function of effort, concavity of the expected utility of the agent in effort follows.[8] Hence, we can substitute PIC_1 for its first-order condition,

$$(e_1): \quad \sum_{i=L,H} \pi'_i (e_1) (u_i + \beta w_i) - v = 0, \tag{22}$$

and we can substitute PIC_2 by its corresponding first-order condition,

$$(e_{2i}): \quad \sum_{j=L,H} \pi'_j (e_{2i}) u_{ij} - v = 0. \tag{23}$$

Using these in place of the original IC allows us to derive some properties for the optimal contract.

As a second step in characterizing the optimal contract, we appeal to the same logic that we spelled out to show the independence of second-period utility of the agent on his first-period actions, to argue that the same independence holds for the expected profit of the principal. The objective function in problem SB can be written as

$$V (\mathbf{u}, \mathbf{e}) = \sum_{i=L,H} \pi_i (e_1) [y_i - c_i + \beta V_{1i} (w_{1i})],$$

where

$$V_{1i} (w_{1i}) = \sum_{j=L,H} \pi_j (e_{2i}) (y_j - c_{ij}).$$

Hence, to solve problem SB subject to (PC) and (22) and (23)—assuming the domain constraints are not binding—we can simply split the problem across

---

[8] For a higher number of output levels, the conditions on the probability function that would assure concavity have not been determined (see Rogerson [1985b] and Jewitt [1988] for a discussion of these conditions in the context of a static contract).

the two periods and solve it backward using subgame perfection. First, we solve the second-period problem, P_{2i}, for an unspecified value of w_{1i}:

$$\max_{u_{iL}, u_{iH}, e_{2i}} V_{1i}(w_{1i})$$

$$s.t.\ (23)\ \text{and}$$

$$w_{1i} = \sum_{j=L,H} \pi_j(e_{2i}) u_{ij} - v e_{2i}.$$

Let μ_i and λ_i be the multipliers of the first and second constraints, respectively. For each $i = L, H$, the first-order conditions with respect to utility are

$$(u_{ij}): \frac{1}{u'(c_{ij})} = \lambda_i + \mu_i \frac{\pi'_j(e_{2i})}{\pi_j(e_{2i})}, \quad j = L, H. \tag{24}$$

This condition will be familiar to the reader acquainted with basic contract theory: Since the second-period problem is, in fact, a static moral hazard, we find that this first-order condition links consumption to likelihood ratios in the same way as in a static contract (see Prescott [1999] for a review of this textbook case). The likelihood ratios capture the informational value of each possible output realization. The same static intuition prevails in the case for effort. The first-order conditions are

$$(e_{2i}): \sum_{j=L,H} \pi'_j(e_{2i})(y_j - c_{ij}) + \mu_i \sum_{j=L,H} \pi''_j(e_{2i}) u_{ij} = 0. \tag{25}$$

It is easier to see the intuition when we substitute $\pi'_L(e) = -\pi'_H(e)$ in the expression above and get

$$(e_{2i}): \pi'_H(e_{2i})[y_H - y_L - (c_{iH} - c_{iL})] + \mu_i \pi''_H(e_{2i})(u_{iH} - u_{iL}) = 0.$$

We see that the principal equates the marginal increase in the expected net profit that comes from a higher probability of y_H with the change in the marginal increase in expected compensation associated with it, given that $u_{iH} > u_{iL}$.

Note, however, that the solution for the second period is contingent on the value of w_{1i} (which plays the role of the period outside utility in a static problem). With the solution to the second-period problem in hand, we can calculate the value to the principal of promising a level of utility of w_{1i} to the agent for the second period. Hence, we know the value of $V_{1i}(w_{1i})$ and we can substitute it in the first-period problem, P_1:

$$\max_{\substack{u_L, u_H, e_1, \\ w_{1L}, w_{1H}}} \sum_{i=L,H} \pi(e_1)[y_H - c_H + \beta V_{1i}(w_{1i})]$$

$$s.t.\ (22)\ \text{and}$$

$$w_0 \leq \sum_{i=L,H} \pi_i(e_1)(u_i + \beta w_{1i}) - v e_1.$$

Let μ and λ be the multipliers of the first and second constraints, respectively. The first-order conditions for consumption are

$$(u_i) : \frac{1}{u'(c_i)} = \lambda + \mu \frac{\pi_i'(e_1)}{\pi_i(e_1)}, \quad i = L, H. \tag{26}$$

These mirror the conditions in (24) for the second period: The ranking of consumption is again determined by the likelihood ratios, although the dispersion is potentially different and depends on the multiplier of the first-period incentive constraint, μ. The values of μ and μ_i, as well as λ and λ_i, are difficult to get for generic utility functions. (To see this, note that the first-order conditions give us information about $u'(c)$, while the constraints of the problems P_1 and P_{2i} are written in terms of $u(c)$; this makes for a highly nonlinear system of equations that seldom has an explicit solution.) This is why computing numerically the solution to particular problems is a popular strategy in dynamic contract theory.[9]

Recall that in this first period the principal has an extra choice variable with respect to problem P_{2i}: the contingent levels of expected utility of the agent in the second period, w_{1i}. The importance of the value of w_{1i} relative to that of u_i in the optimal contract is at the heart of dynamic incentives. We can explore the optimal tradeoff between the two variables by looking at the first-order condition for the continuation utility:

$$(w_{1i}) : V_{1i}'(w_{1i}) + \lambda + \mu \frac{\pi_i'(e_1)}{\pi_i(e_1)} = 0, \quad i = L, H. \tag{27}$$

To interpret this condition we need to figure out the derivative of the value function of the principal, $V_{1i}'(w_{1i})$. We do this by using the envelope theorem and the second-period problem P_{2i} that determines $V(w_{1i})$:

$$V_{1i}'(w_{1i}) = -\lambda_i.$$

Substituting this derivative into (27) we get

$$\lambda_i = \lambda + \mu \frac{\pi_i'(e_1)}{\pi_i(e_1)}, \quad i = L, H.$$

Note that this, combined with (26), implies $\lambda_i = \frac{1}{u'(c_i)}$. What does the λ_i multiplier represent in the second period? It is the shadow value of relaxing the "promise keeping" constraint of the principal in the second period. The principal has committed to deliver a level of expected utility of w_{1i}. How costly this is for him depends on the necessary spread of utilities in order to satisfy incentives in the second period. This can be seen formally by multiplying the first-order conditions for u_{ij} in (24) for each j times $\pi_j(e_{2j})$, and then

---

[9] For details on these computations see, for example, Phelan and Townsend (1991) or Wang (1997).

summing the resulting equations for $j = L$ and $j = H$; doing this we get that

$$\frac{\pi\left(e_{2_i}\right)}{u'\left(c_{iH}\right)} + \frac{1 - \pi\left(e_{2i}\right)}{u'\left(c_{iL}\right)} = \lambda_i.$$

The shadow value depends on the expected tradeoff between the marginal value to the principal of increasing consumption, -1, and the marginal increase in utility of spending this extra unit of consumption, $u'(c)$. Now we take this condition further: Since we had established that $\lambda_i = \frac{1}{u'(c_i)}$, we get the following relationship of the inverse of the marginal utility of consumption:

$$\frac{\pi\left(e_{2_i}\right)}{u'\left(c_{iH}\right)} + \frac{1 - \pi\left(e_{2i}\right)}{u'\left(c_{iL}\right)} = \frac{1}{u'\left(c_i\right)}. \tag{28}$$

This is the so called "Rogerson condition," first derived in Rogerson (1985a). It summarizes how the optimal dynamic contract with commitment allocates incentives over time and histories. We now discuss its implications for the choices of effort and consumption.

Effort and Consumption Choices Over Time

To illustrate the implications of the Rogerson condition, consider, for the sake of comparison, a slightly different model to the one presented here: Everything else equal, assume no commitment to long-term contracts for both the principal and the agent. This is often referred to as "spot contracting." For the purpose of our comparison, set the per period outside utility for the agent to $\frac{w_0}{2}$ in both periods. It is easy to see that the solution to this problem without commitment is the repetition of the one-period optimal contract. This implies that the second-period consumptions would be independent of the first-period realizations, and hence identical to those in the first: $c_H = c_{LH} = c_{HH}$, as well as $c_L = c_{HL} = c_{LL}$. It is immediate that this solution to the spot contract violates (28).

How is the contract with commitment different than the repetition of the static contract? The main difference is that with commitment the contract exhibits memory, i.e., the level of consumption in the second period, contingent on a second-period realization, is different depending on the first-period realization. Why is it optimal for the contract with commitment to be different than the repetition of the static contract? Because it allows incentives to be provided in a more efficient way. The reason becomes clear if we consider how the principal can improve on the repetition of the static contract once he has commitment to a two-period contract. If the agent gets a y_H realization in the first period, his overall expected utility increases if he trades off some of the consumption that the static contract assigns him in the first period with some expected consumption in the second. Because c_H was high to start with, the decrease in his first-period utility from postponing some consumption translates into a bigger increase in expected utility in the second period, where he

has positive probability of facing low consumption whenever y_L realizes. This means the principal can, with this deviation from the spot contract solution, keep some of the consumption for himself while leaving constant the expected utility following the high realization node in the first period, i.e., $u_H + \beta w_H$. In the same way, if the agent gets a y_L realization in the first period, he is better off by trading some expected utility in the second period for some consumption in the first, and this again saves resources for the principal. Hence, in the optimal contract, we have that $w_{1H} > w_{1L}$. It is worth noting that these optimal tradeoffs result in a violation of the Euler equation of the agent, which is incompatible with (28).[10]

The last first-order condition of problem P_1 left for analyzing is that of effort in the first period:

$$(e_1): \sum_{i=L,H} \pi_i'(e_1)(y_i - c_i + \beta V_{1i}(w_{1i})) + \mu \sum_{i=L,H} \pi_i''(e_1)(u_i + \beta w_{1i}) = 0.$$

This condition captures the same tradeoff discussed after deriving the second-period effort first-order conditions in (25). Of course, the values of the variables and multipliers will typically be different than in the second period, implying a different solution across periods. To gain some important insight in the properties of effort requirements over time, it is again useful to compare the effort solution here to that of the spot contract without commitment. It is easy to see that the repetition of the static contract would imply $e_1 = e_{2H} = e_{2L}$.[11] Here, instead, this is not the case. If we recall that the optimal contract implies $w_{1H} > w_{1L}$, a simple inspection of the second-period problem P_2 tells us that, for the principal, effort incentives are more expensive following a y_H realization than a y_L realization. The continuation utility w_{1i} plays the role of the outside utility in a static contract. It is immediate from the risk aversion of the agent that, for the same spread of utility that would satisfy the IC in (23), a higher level of outside utility translates into more consumption. Hence, the principal will optimally choose $e_{2H} < e_{2L}$. Moreover, in the second period the principal cannot provide incentives for effort as efficiently as in the first period, since the intertemporal tradeoff of consumption that we described above is not available (there are no future periods after $t = 2$). This will typically imply a lower effort requirement in the second period than in the first. We conclude that, in contrast with the first-best property summarized in 2A, effort requirements will fluctuate over time and across histories in the unobservable effort case in order to provide incentives more efficiently.

The solution to this version of our numerical example is presented in Table 3 and Figures 1 and 2. We defer the discussion of this solution example until

----

[10] This follows from Jensen's inequality and the convexity of $1/u'(c)$. For details, see Rogerson (1985a).

[11] Simply set $w_{1H} = w_{1L}$ and note that $\pi_H'(e_1) = -\pi_L'(e_1)$.

Section 6, where we compare the solution to the unobserved effort case both with full depreciation and without.

4.   DEALING WITH PERSISTENCE

The simplifications outlined in the previous section, when effort is not persistent, do not hold for the general case of $\rho > 0$. Before we go on to analyze a particular case of human capital accumulation in Section 5 and illustrate the differences, we discuss here the main particularities that persistence of effort introduces in the analysis of the optimal contract.

Two main differences with respect to the standard framework appear when effort is persistent. First, it is no longer the case that a given choice for effort in the second period provides the agent with the same expected utility w_{1i} regardless of his first-period effort choice e_1. It follows that the number of relevant incentive constraints is much higher in the problem with persistence. Second, the problem of the principal cannot, in general, be written in the usual recursive form in which the promised utility w_{1i} summarizes all relevant information about past periods. The relevant summary variable is the original W_{1i} (**u**, **e**), which depends on both the first- and the second-period effort choices. The dependence of W_{1i} (**u**, **e**) on e_1 complicates the calculation of its possible values. In particular, this state variable is not a number (like w_{1i} was) but a function: The principal needs to take into account all possible choices for e_1, including those off the equilibrium path. Finally, the conditions for concavity of the agent's problem in the IC are difficult to establish, even in the two-outcome case presented here.

These issues have so far been addressed in the literature with two main strategies. The first strategy limits the effort choices to a two-point set, and includes explicitly in the problem of the principal the complete list of relevant incentive constraints for all possible combinations of effort choices. The second strategy allows for a continuum of effort choices, but puts restrictions on the functional form of π (e_1, e_2) in order to simplify the set of constraints. These approaches are now discussed in some detail.

A Hands-On Analysis of the Joint Deviations Problem

Within the first approach, the main contribution is Fernandes and Phelan (2000). They provide a tractable setup in which an augmented recursive formulation of the problem of the principal is possible. Intuitively, this formulation has an increased number of state variables with respect to the recursive formulation of the moral hazard problem without persistence first presented in Spear and Srivastava (1987). The simplified framework that allows for the recursive formulation limits the effort choices and the output realizations to two. Also, the contract lasts for an infinite number of periods but persistence

lasts only for one period; that is, effort at time t affects only the probability distribution over outcomes at time t and $t + 1$. The recursive formulation of the problem of the principal has three state variables, one of which is the standard promised utility on Spear and Srivastava's formulation. The two extra states allow the principal to keep track of the marginal disutility of effort for the agent across periods, as well as the set of utilities achievable by the agent off the equilibrium path.

Still within the first approach, Mukoyama and Şahin (2005) limit the effort choices and the output values to two and analyze a two-period problem. They assume that high effort is optimal every period. They are able to provide analytical conditions on the conditional probability function under which the implications of persistence are drastically different than those of no persistence: When the first-period effort affects the second-period probability in a sufficiently stronger way than the second-period effort, the optimal contract exhibits perfect insurance in the initial period. Using a recursive formulation in the spirit of Fernandes and Phelan (2000), Mukoyama and Şahin also analyze a three-period problem numerically.

Kwon (2006) uses a very similar framework with discrete effort choices (0 or 1), also assuming that high effort is implemented every period. He imposes concavity of $\pi(\cdot)$ on the sum of past effort choices, so past effort is more effective than current effort. These assumptions allow him to analyze a $T > 2$ period problem that shares the same perfect insurance characteristic as in Mukoyama and Şahin (2005).

A Particularly Simple Case of Persistence

The second approach, presented in Jarque (2010), allows for a continuum of effort choices but assumes that the conditional probability depends on past effort choices only through the sum of undepreciated effort in the same manner as stated in Assumption 2. Note that, even for a concave probability function $\pi(s)$, Assumption 2 implies that past effort is less effective than current effort in contrast to what was assumed in Mukoyama and Şahin (2005) or Kwon (2006). The article shows that, for a subset of problems with this particular form of persistence, the computation of the optimal contract simplifies considerably. For these problems, an auxiliary standard repeated moral hazard problem without persistence can be used to recover the solution to the optimal contract. The linearity in effort of both variable s (which determines the probability distribution) and the utility of the agent dramatically simplifies the structure of the joint deviations across periods; in practice, we can think of s as the choice variable, and the structure of the resulting transformed problem is (under some conditions) equivalent to that of a standard repeated moral hazard.

In the next section, a finite version of the model in Jarque (2010) is presented and this result is explained in detail. The finite version allows for the numerical computation of the optimal contract in an example in which the stochastic structure is interpreted as unobservable human capital accumulation.

5. HIDDEN HUMAN CAPITAL ACCUMULATION

The problem of the principal is again as in problem SB, but now we consider the case $\rho > 0$. We argued in Section 4 that this case is more complicated because of the dependence of second-period utility and optimal actions of the agent on first-period choices. In order to go around some of these difficulties, here we adapt to our two-period finite example the strategy presented in Jarque (2010) for solving problems with persistence. Following this work we will show that, under our assumptions, the structure of the problem simplifies to that of the standard repeated moral hazard presented above, *provided the domain constraints in (ED) do not bind*. This is an important qualification since, as we learned when analyzing the case of observable human capital accumulation in Section 2, in the presence of persistence the effort domain constraints in (ED) will sometimes bind, especially for high values of the persistence parameter ρ. To deal with this issue, we follow the approach in Jarque (2010): First, we find a candidate solution assuming that the constraint in (ED) does not bind. Then we need to check numerically that this constraint is indeed satisfied to be sure that we have found a true solution. Unfortunately, a general analysis of the optimization problem of the principal including the inequality constraints for effort (again, as in Section 2) is more difficult with unobserved effort. Hence, finding the properties of the general case when constraint (ED) binds remains a question for future research.

Rewriting the Problem

Jarque (2010) shows that, whenever the effort domain constraint (ED) is not binding, we can find the solution to the problem with persistence using a related RMH problem without persistence as an auxiliary problem. The key observation for that result is that we can write the expected utility of the agent, $W_0(\mathbf{u}, \mathbf{e})$, as a function of the s variable only. This is convenient because s is the variable that effectively determines the probability distribution over outcomes each period; different combinations of effort choices that give rise to the same s are equivalent both for the principal and for the agent. Hence, once we rewrite the problem with s as the choice variable, there is no need to consider joint deviations across periods, the recursive structure is recovered, and we can solve for the optimal contract as we do with a standard repeated moral hazard.

Let $\widetilde{W}_0\left(\mathbf{u}, \mathbf{s}\right) = W_0\left(\mathbf{u}, \mathbf{e}\right)$ for all the pairs of $\mathbf{s}$ and $\mathbf{e}$ sequences such that $\mathbf{s}$ results from effort choices in $\mathbf{e}$ according to the law of accumulation of human capital in (1). Writing the effort in the second period as

$$e_{2i} = s_{2i} - \rho s_1,$$

we have

$$\widetilde{W}_0\left(\mathbf{u}, \mathbf{s}\right) \;=\; \sum_{i=L,H} \pi_i\left(s_1\right) u_i - v s_1$$

$$+\beta \sum_{i=L,H} \pi_i\left(s_1\right) \left[\sum_{j=L,H} \pi_j\left(s_{2i}\right) u_{ij} - v\left(s_{2i} - \rho s_1\right) \right].$$

Note that we have explicitly written the utility accrued in the first period in the first row of this expression, and that of the second period in the second row. With utility spelled out this way it is easy to see that, although s_1 is all accumulated in the first period, it appears both in the first- and second-period utility. Also, since s_1 is not contingent on any realization, it appears in the second period both after observing a first-period y_H and a first-period y_L. Hence, we can group the s_1 terms of the second period together with those of the first, to get an expression of the form

$$\widetilde{W}_0\left(\mathbf{u}, \mathbf{s}\right) \;=\; \sum_{i=L,H} \pi_i\left(s_1\right) u_i - v\left(1 - \beta\rho\right) s_1$$

$$+\beta \sum_{i=L,H} \pi_i\left(s_1\right) \left[\sum_{j=L,H} \pi_j\left(s_{2i}\right) u_{ij} - v s_{2i} \right]. \quad (29)$$

This allows us to interpret s as the variable being chosen by the agent. In the first period, we can interpret $v\left(1 - \beta\rho\right)$ as the "marginal disutility of exerting s_1." In the second period, the "marginal disutility of exerting s_2" is instead v.

This rearrangement of terms and thinking about s as the choice variable is a useful trick. Note that in the second row the expression inside the square brackets is independent of s_1. Interpreting s_{2i} as the choice variable, we can see that we can do here as we did in the case of no persistence and write the continuation utility of the agent independently of the first period's choice for s_1:

$$\sum_{j=L,H} \left[\pi_j\left(s_{2i}\right) u_{ij} - v s_{2i}\right] = \widetilde{W}'_{1i}\left(\mathbf{u}, s_{2i}\right).$$

Hence, we obtain expressions that parallel those of the standard RMH formulation in (19). The expression in (29) can then simply be rewritten as

$$\widetilde{W}_0\left(\mathbf{u}, \mathbf{s}\right) = \sum_{i=L,H} \pi_i\left(s_1\right) \left[u_i + \beta\widetilde{W}'_{1i}\left(\mathbf{u}, s_{2i}\right)\right] - v\left(1 - \beta\rho\right) s_1.$$

Note also that the structure of the incentive constraints simplifies as it did in the case of the RMH; in the second period, the first-period choice of s drops

out:

$$\sum_{j=L,H} \pi_j \left(s_{2i} \right) u_{ij} - v s_{2i} - \rho v \widehat{s_1} \geq \sum_{j=L,H} \pi_j \left(\widehat{s_{2i}} \right) u_{ij} - v \widehat{s_{2i}} - v \rho \widehat{s_1}, \quad \forall \widehat{s_1}, \widehat{s_{2i}}.$$

Again, all these changes of notation are simply aimed at pointing to the following fact: The problem in which effort is persistent has a similar structure to that of a standard RMH problem in which s is interpreted as effort that is not persistent, but has marginal disutility of $v (1 - \beta \rho)$ at $t = 1$ and of v at $t = 2$. To make this explicit, using the intertemporal regrouping of s_1, the problem of the principal in SB can be written as problem SB':

$$\max_{\mathbf{u}, \mathbf{s}} V \left(\mathbf{u}, \mathbf{s} \right)$$
$$s.to$$
$$w_0 \leq \widetilde{W}_0 \left(\mathbf{u}, \mathbf{s} \right)$$
$$\widetilde{W}_0 \left(\mathbf{u}, \mathbf{s} \right) \geq \widetilde{W}_0 \left(\mathbf{u}, \widehat{\mathbf{s}} \right) \quad \forall \widehat{\mathbf{s}} \neq \mathbf{s}$$
$$u_i \in \mathcal{U}_i \quad \forall i$$
$$s_1 \in S_1$$
$$s_{2i} \in S_2 \quad i = L, H,$$

with $S_1 = \left[\underline{e}, \overline{e} \right]$ and $S_2 = \left[\rho s_1 + \underline{e}, \rho s_1 + \underline{e} + \overline{e} \right]$. This rewriting leads to the following observation: If problem SB' were in fact formally equivalent to a standard RMH problem (with the modified structure of the marginal disutility), this would help us enormously to find and characterize the solution to SB, since we would know how to solve it (or at least compute it numerically). However, a close inspection of SB' points to a small but potentially important difference with a standard RMH problem: In problem SB', the domain S_2 depends on the choice of s_1, while in a standard RMH problem this domain would be exogenously given.

Using a Related RMH Problem without Persistence as an Auxiliary Problem

Following Jarque (2010), we now show that, in some instances, we can work around the difficulty that an endogenous domain S_2 poses by using a related auxiliary problem for our purposes instead of SB'. Consider a problem SB$_{aux}$ that is equal to SB' except for the domain S_2, which is substituted by an auxiliary domain $\widetilde{S}_2 = \left[\underline{e}, \overline{e} \right]$. Note that $\widetilde{S}_2$ is exogenous so, interpreting s as effort, problem SB$_{aux}$ is a standard RMH. We will now argue that, under some conditions, the solution to SB' coincides with the solution to SB$_{aux}$, and hence we can easily obtain a solution to our problem with persistence.

The solution to problems SB' and SB$_{aux}$ coincides when two conditions are satisfied: (i) $W_0 \left(\cdot \right)$ is concave in $\mathbf{s}$, and (ii) the resulting optimal choices for effort are interior. This is a set of sufficient conditions because if the

expected utility of the agent is concave in his choice of s, then the relevant effort deviations are those close to the optimal (interior) s, and not those at the limits of the domain. This implies that using an auxiliary domain that does not exactly overlap with the true domain is not changing the solution to the problem, as long as this true solution is contained in the auxiliary domain. Are each of these conditions satisfied in our framework?

(i) **Concavity of** $W_0(\cdot)$ **in s.** In our particular example, it is easy to argue that the problem of the agent is concave in s_t for all t. In fact, the argument is the same that we used earlier to argue that problems PIC_1 and PIC_2 were concave: There are only two outcomes, the probability of observing y_H is concave in s_t, and current and future utility assigned to y_H is always higher than current and future utility assigned to y_L.

(ii) **Effort is interior**. This is not satisfied trivially. Constraint (ED) implies that two restrictions need to be checked to establish that the true solution is contained in the proposed auxiliary domain:

$$s_{2i} \quad < \quad \rho s_1 + \underline{e} + \overline{e}, \ i = L, H, \tag{30}$$

$$s_{2i} \quad > \quad \rho s_1 + \underline{e}, \ i = L, H. \tag{31}$$

Under the probability specification in (7), equation (30) is always satisfied. Other specifications are easy to find for which the upper bound of effort in (30) is not binding. The lower bound, however, is endogenous, and equation (31) cannot be checked without having the solution for s in hand. We conclude that the interiority cannot easily be guaranteed ex ante. The strategy to go around this problem that is proposed in Jarque (2010) is the following: Solve the problem assuming that the domain constraint can be substituted—and hence the equivalence to the RMH can be used—and then, with a candidate solution for s in hand, check the constraint ex post. We follow this route in the numerical computation of an example presented next. As it turns out, it is easy to find parametrizations for which the ex post check on the nonnegativity of effort is satisfied.

The Optimal Contract for Hidden Human Capital Accumulation

What do we conclude about the properties of the optimal contract in the presence of hidden human capital accumulation? Denote as $\widetilde{\mathbf{c}}^*$ and $\widetilde{\mathbf{e}}^*$ the solution to problem SB_{aux}. Whenever the sufficient conditions discussed above are satisfied, we have that, in the optimal contract:

1. The optimal consumption sequence in problem SB, $\mathbf{c}^*(P, SB)$, is equal to $\widetilde{\mathbf{c}}^*$.

2. The optimal human capital sequence in SB, $\mathbf{s}^*(P, SB)$, is equal to $\widetilde{\mathbf{e}}^*$.

3. The optimal effort sequence in SB, $\mathbf{e}^*(P, SB)$, can be recovered from the effort solution to problem SB_{aux} using

$$
\begin{aligned}
e_1^*(P, SB) &= \widetilde{e}_1^* \\
e_2^*(P, SB) &= \widetilde{e}_2^* - \rho \widetilde{e}_1^*.
\end{aligned}
$$

Importantly, the optimal consumption sequence has the same properties as in the solution to a standard RMH problem without persistence. Also, the optimal *human capital sequence* has the same properties as the *effort sequence* in a standard RMH problem. These properties were discussed at length in Section 3. Using these properties, we can reflect on the economic meaning of the ex post check implied by equation (31).

Whenever the ex post check in (31) is satisfied, the optimal contract asks the agent to increase human capital *in every period*. That is, the remaining level of human capital from the previous period, after depreciation, ρs_1, is never sufficient to cover the requirement of human capital for the current period, s_{2i} for $i = L, H$. In light of the properties of effort in a standard RMH problem, it is easy to see that this condition may not be satisfied in some examples since a decrease in the level of human capital from one period to the next could be part of the optimal solution for the principal. In particular, we learned in Section 3 that in an interior solution we will typically have $e_{2H} < e_1$, since the smoothing of incentives that is present in the first period is not available in the second, making effort in the second period relatively more expensive. Given the results we just established for the case with persistence, this means that we will typically have $s_{2H} < s_1$ in the optimal contract with hidden human capital accumulation. How does this lead to a violation of the ex post check in equation (31)? For certain parameters, we may have that s_{2H} is *so much smaller* than s_1 that, in fact, we have $s_{2H} < \rho s_1 + \underline{e}$, violating the interiority of effort choices. That is, if it were feasible, the principal would choose to have s_2 lower than $\rho s_1 + \underline{e}$. However, in the true problem with human capital accumulation (problem SB), effort needs to stay within its domain in each period, i.e., $e_{2i} > \underline{e}$ for all i, which rules out the possibility of decreasing s_2 below $\rho s_1 + \underline{e}$. Any adjustment should be made in the first period, when the principal anticipates the added cost of future incentives. That is, the solution for s_1 should differ from the one that was just presented. Unfortunately, characterizing how exactly the solution for s_1 changes is not easy. Solving for the optimal contract in this case becomes more complicated. As we argued, the independence of second-period choices from first-period choices breaks down, both for the principal and for the agent. In practice, even the numerical computation of examples is more involved, since all feasible combinations of effort across the two periods (and choices contingent on realizations of output) need to be tested for incentive compatibility. The simple recursive structure with w_{2i} as a state variable is no longer valid, and the dimensionality of the

computational problem is similar to that of the strategy proposed in Fernandes and Phelan (2000).

The next section presents an example for which the ex post check in (31) is satisfied, and hence solving for the optimal contract is simple. Using the numerical solution, we discuss the implications of persistence for consumption and effort paths by comparing the solution to that of the case without persistence ($\rho = 0$).

6. NUMERICAL EXAMPLE WITH UNOBSERVED EFFORT: A COMPARISON

For cases in which the equivalence to a RMH is valid, we can find the solution to our problem with persistence using the usual numerical methods to solve standard RMH problems without persistence.

Figures 1 and 2 illustrate the implications for effort and consumption in the solution to an example with the parameter values listed in Table 1. The example without persistence has $\rho = 0$, while the example with persistence has $\rho = 0.2$. For the numerical examples we use the functional forms $u(c) = 2\sqrt{c}$ and the probability specification in (7). We also set $\underline{e} = 0.01$ and $\overline{e} = 0.99$ in order to restrict to cases with full support.

In Figure 1, the solution for s and e in the SB problem with persistence is plotted with a solid line. As we can see in the top panels, the level of s_1 in problem SB is always higher with persistence than without persistence (dashed line). Since $s_1 = e_1$, a higher level of s_1 with persistence reflects the fact that human capital is accumulated in the first period with the same cost as nonpersistent effort, but it lasts (partially) until the following period.[12]

The solutions for the paths of optimal s in the FB model are also represented in Figure 1 (dotted and dash-dotted line respectively for the persistent and nonpersistent case). The comparison clearly shows that human capital accumulation makes frontloading of s optimal. (This also translates into frontloading of effort as shown clearly in the bottom panels of Figure 1.) The main difference with the solutions to the respective SB problem is the level (higher in the FB problem). A second difference is that, even without persistence, in the second period the requirement for s may decrease in the SB problem, for incentive reasons, following both realizations (although the decrease may be more pronounced after y_H), and hence we have $s_1 > s_{2i}$ for all i.

As we can see in the bottom panels of the solution to the SB problem, both with persistence and without, effort is higher in the initial period than in

---

[12] The level of s_{2i} in this example coincides with and without persistence for all i. This is particular to this example and is violated if, for example, the level of w_0 is modified. Although human capital in the second period is equivalent to nonpersistent effort (because there are no further periods to exploit the persistence of human capital), the optimal choice for w_{2i} will typically be different across the two models.

**Figure 1 Contingent Paths for Human Capital and Effort in the
        Optimal Contract with and without Effort Persistence, both
        for the First-Best and the Second-Best Models**
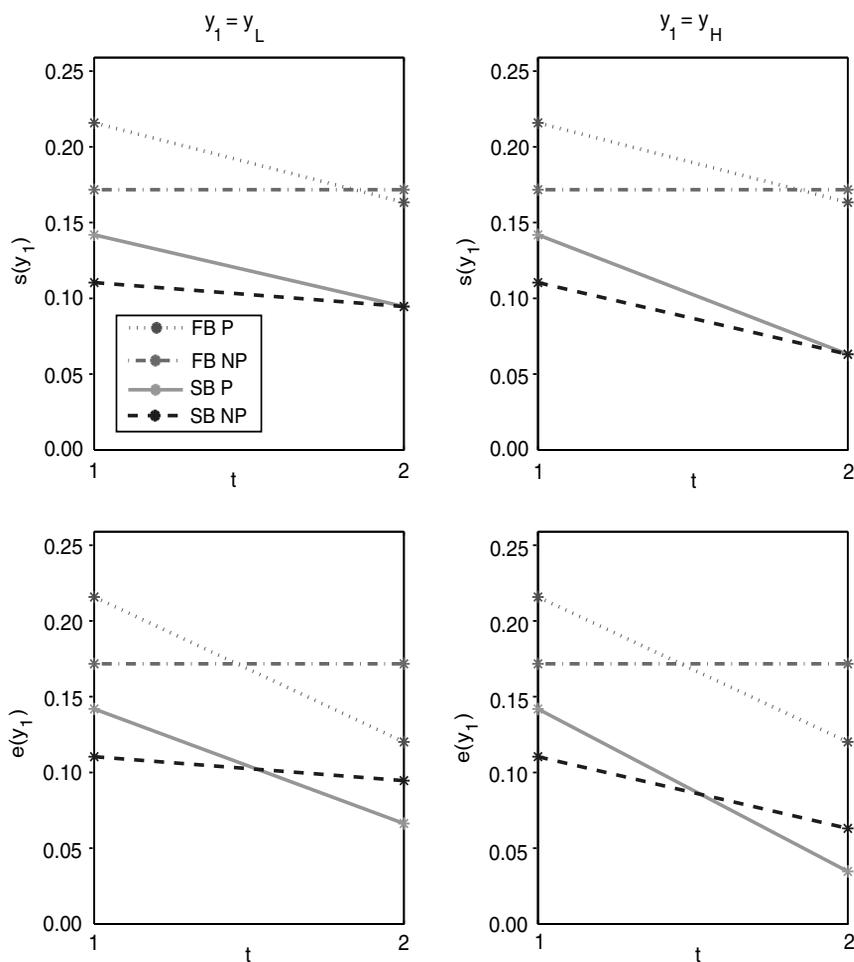
Notes: $s(y_1)$: human capital contingent on history y_1; $e(y_1)$: effort contingent on history y_1; P: see Table 3, $\rho = 0.2$; NP: see Table 3, $\rho = 0$.

the second. However, the frontloading of effort is much more pronounced with persistence. This is also true when comparing the solutions for the FB problem: While effort stays constant from one period to the next in the case without persistence, with persistence it is frontloaded, as discussed in Section 2.

Figure 2 Contingent Paths for Consumption in the Optimal Contract with and without Effort Persistence, both for the First-Best and the Second-Best Models
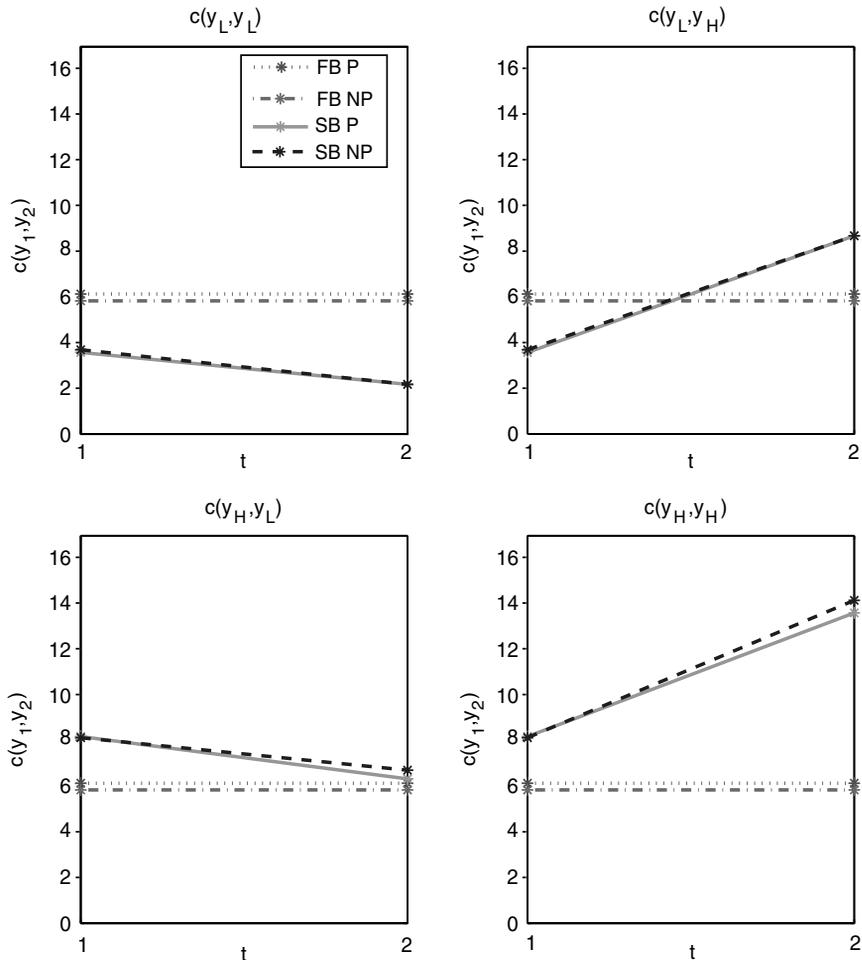
Notes: $c(y_1, y_2)$: consumption contingent on history (y_1, y_2); P: see Table 3, $\rho = 0.2$; NP: see Table 3, $\rho = 0$.

Consumption, depicted in Figure 2, is, in the SB case, virtually the same with and without persistence. It simply increases when the realization is y_H and decreases when it is y_L for the standard incentive provision reasons discussed in the earlier sections. However, we can see in the FB case that consumption is slightly lower in the case with persistence. Since the FB case

Table 3 Summary Statistics

	$\rho = 0.2$ **(FB)**		$\rho = 0.0$ **(FB)**		$\rho = 0.2$ **(SB)**		$\rho = 0.0$ **(SB)**	
	$t = 1$	$t = 2$	$t = 1$	$t = 2$	$t = 1$	$t = 2$	$t = 1$	$t = 2$
$E\left[c_t^*\right]$	6.12	6.12	5.82	5.82	5.30	5.47	5.16	5.30
$E\left[u\left(c_t^*\right)\right]$	4.95	4.95	4.83	4.83	8.26	11.32	7.74	10.90
$Var\left[c_t^*\right]$	0	0	0	0	4.96	13.69	4.34	13.96
$E\left[e_t^*\right]$	0.22	0.16	0.17	0.17	0.14	0.05	0.11	0.084
$Var\left[e_t^*\right]$	0	0	0	0	0	0.00023	0	0.00022
$E\left[s_t^*\right]$	0.22	0.16	0.17	0.17	0.14	0.0828	0.11	0.0842
$Var\left[s_t^*\right]$	0	0	0	0	0	0.00023	0	0.00022

is calculated numerically but without using a grid, we conclude that most likely consumption is also slightly lower with persistence in the true solution to the unobservable effort case.

Table 3 reports the value of some simple statistics of the comparison across the two models presented in Figures 1 and 2. The FB model statistics are included for reference, since they correspond to the solutions reported already in Sections 1 and 2. All expectations in the first period are conditional on s_1^*, and those in the second are conditional on s_{2i}^*. When comparing the statistics for the SB problem, we see that persistence implies a higher level of expected consumption, expected utility, and a slightly higher variance of consumption in the first period. When looking at these three moments across periods we see that persistence implies a steeper increase of expected consumption in time. Again, the statistics on consumption need to be interpreted with care since they are likely influenced by the use of a grid.

As for expected effort, we see that the level is higher with persistence in the initial period, but it drops below the no persistence case in the second period (a much steeper decrease than without persistence). The comparison of the expected accumulated human capital explains this: The expected level of s_1 with persistence is much higher than the level of e_1 without persistence, but the solution for s_2 with persistence is similar (in this particular example, identical) to the solution for e_2 without persistence.

7. CONCLUSION

When learning by doing is an important factor in a repeated agency relationship, solving for the optimal contract is generally very difficult. In the framework studied here, with linear disutility of effort and the productivity of the agent being a distributed lag of past efforts, we provide an example with a simple solution. This allows us to numerically establish some properties of

the optimal contract. On one hand, the human capital of the agent in equilibrium and, hence, his productivity tend to be higher with learning by doing than without. Moreover, the optimal contract offered to the employee implies a lower productivity in the final years of the contract. The human capital of the agent is left to depreciate since, close to the end of the contract, the cost of incentives of requiring a higher productivity is not justified by the benefit of future productivity. This implies that, over the contractual relationship, effort is frontloaded and follows a steeper decreasing pattern than in the case without learning by doing. On the other hand, we find that the properties of wage dynamics remain unchanged with respect to those of the optimal contract without learning by doing.

REFERENCES

Ales, Laurence, and Pricila Maziero. 2009. "Accounting for Private Information." Mimeo.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning by Doing." *The Review of Economic Studies* 29 (June): 155–73.

Fernandes, Ana, and Christopher Phelan. 2000. "A Recursive Formulation for Repeated Agency with History Dependence." *Journal of Economic Theory* 91 (April): 223–47.

Gibbons, Robert, and Kevin J. Murphy. 1992. "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." *Journal of Political Economy* 100 (June): 468–505.

Grossman, Sanford J., and Oliver D. Hart. 1983. "An Analysis of the Principal-Agent Problem." *Econometrica* 51 (January): 7–45.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (January): 1–58.

Jarque, Arantxa. 2010. "Repeated Moral Hazard with Effort Persistence." *Journal of Economic Theory* 145 (November): 2,412–23.

Jewitt, Ian. 1988. "Justifying the First-Order Approach to Principal-Agent Problems." *Econometrica* 56 (September): 1,177–90.

Kapička, Marek. 2008. "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach." Mimeo, University of California, Santa Barbara.

Kwon, Illoong. 2006. "Incentives, Wages, and Promotions: Theory and Evidence." *RAND Journal of Economics* 37 (Spring): 100–20.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent. 2009. "Performance Pay and Wage Inequality." *Quarterly Journal of Economics* 124 (February): 1–49.

Lucas, Robert E., Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July): 3–42.

MacLeod, W. Bentley, and Daniel Parent. 1999. "Job Characteristics, Wages, and the Employment Contract." Federal Reserve Bank of St. Louis *Review* (May): 13–27.

Mukoyama, Toshihiko, and Ayşegül Şahin. 2005. "Repeated Moral Hazard with Persistence." *Economic Theory* 25: 831–54.

Phelan, Christopher. 1994. "Incentives, Insurance, and the Variability of Consumption and Leisure." *Journal of Economic Dynamics Control* 18: 581–99.

Phelan, Christopher, and Robert M. Townsend. 1991. "Computing Multi-Period, Information-Constrained Optima." *Review of Economic Studies* 58 (October): 853–81.

Prescott, Edward S. 1999. "A Primer on Moral-Hazard Models." Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Winter): 47–77.

Rogerson, William P. 1985a. "Repeated Moral Hazard." *Econometrica* 53 (January): 69–76.

Rogerson, William P. 1985b. "The First-Order Approach to Principal-Agent Problems." *Econometrica* 53 (November): 1,357–67.

Spear, Stephen E., and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies* 54 (October): 599–617.

Wang, Cheng. 1997. "Incentives, CEO Compensation and Shareholder Wealth in a Dynamic Agency Model." *Journal of Economic Theory* 76 (September): 72–105.

News Shocks and Business Cycles

Per Krusell and Alisdair McKay

T he discussion surrounding the recent deep recession seems to have shifted the focus from currently used business cycle models to the standard Keynesian model (by which we mean the "old Keynesian," as opposed to the new Keynesian, model). In the Keynesian model, pessimism among consumers and investors about the economy will simultaneously lower aggregate consumption and aggregate investment, as well as aggregate output, through an increase in the rate of unemployment, and more generally through lower capacity utilization. Moreover, in the Keynesian model, pessimism and optimism are not determined within the model—they appear exogenously and they disappear exogenously. The analysis is then about how the economy reacts to these exogenous events. Undoubtedly, there are many indications that consumers and investors seemed pessimistic about their prospects during the recession, but does such pessimism necessitate the reversion back to the Keynesian model? The present article reviews and contributes to a recent strand of the "modern" business cycle literature, i.e., the literature that insists on building a model of the economy that is explicit about its microeconomic foundations and that addresses a related question: Can news shocks generate positive co-movement among our macroeconomic aggregates? An example of a negative news shock would be the sudden arrival of information indicating that future productivity will not be as high as previously thought. Thus, such a shock would generate current pessimism, and yet be grounded in real and fundamental developments. Another kind of news shock would be a government announcement about a policy change to be implemented on a future date (say, that taxes will be raised beginning next year). In this recent literature, thus, optimism and pessimism are examined as determinants of business

cycle fluctuations, but as add-ons to otherwise microfounded macroeconomic models, and moreover they are tied in a systematic way to anticipated changes in the economy's fundamentals.

Models of business cycles that rely on microeconomic foundations generate fluctuations in economic activity in response to fluctuations in fundamentals, such as preferences, technology, or government policy. The first generations of these models (Kydland and Prescott 1982) relied on technology shocks, i.e., shocks to aggregate productivity; such a shock, if positive and persistent, would raise output directly, via an increase in aggregate employment, and as a consequence raise both consumption and investment, thus generating the kind of co-movement we observe in aggregate time series. Shocks to government expenditures have been considered as well, as have preference shocks (for consumption now versus consumption in the future), though these shocks alone do not easily generate co-movement in the remaining aggregate variables. For example, when government spending rises there is strong pressure on either consumption or investment to fall, unless hours worked (or perhaps capital utilization) rises significantly; hours worked might increase if there is a significant wealth effect in labor supply, but in standard parameterizations the wealth effects are not strong enough.

The new literature begins with Beaudry and Portier (2006, 2007), who analyze time-series data and conclude that news about future productivity may be an important driver of business cycles and then go on to discuss in what model economies news can generate co-movement. We briefly review the data analysis in Section 1. In Section 2, we explain why news shocks, like some other shocks, do not readily generate co-movement in standard neoclassical settings. Beaudry and Portier suggest their own setting, wherein news shocks have the desired effect, but there are other frameworks that generate co-movement in response to news shocks as well. Section 3 describes a very simple setting that we think has most, if not all, of the necessary qualitative effects: the Pissarides (1985) model. This model is a general-equilibrium description of labor markets with search/matching frictions in which unemployment is an equilibrium phenomenon. Capital does not play a major role in the simplest version of the model, though the number of firms, which is endogenous and depends on labor market conditions and on (current and future) productivity, can be given the interpretation of capital, and the creation of new firms can be interpreted as investment. We show that in this model, news about, say, a decline in future productivity—pessimism—will lead fewer firms to enter on impact. Thus, investment falls. Moreover, there is a rise in unemployment, along with a stock market bust, which we measure as the value of the firms in the market. If, in addition, the economy has access to a storage technology, or the economy is open, a fall in consumption can result as well. Thus, the model can generate co-movement in all macroeconomic variables. We then

review, in Section 4, other settings proposed in the literature that achieve the same goals, and in Section 5 we offer conclusions.

1.   EVIDENCE OF NEWS SHOCKS AND THEIR EFFECTS

Typical Business Cycle Co-Movements

What features of the business cycle might one expect models to capture? Perhaps the key characteristic of the business cycle is the co-movement of broad measures of economic activity. A business cycle expansion typically involves rapid growth of output, consumption, and investment and high levels of employment and hours worked. Another distinguishing feature of business cycles is the frequency of expansions and contractions. Business cycle fluctuations are typically thought to have a frequency of longer than one year but shorter than one decade. Finally, one might ask a model to match the magnitude of business cycle fluctuations in both absolute as well as relative terms. While an ideal model of the business cycle would be accurate along all these dimensions, the focus of the discussion here is on matching co-movements.

VARs and Other Evidence

Much of the interest in news shocks stems from the empirical work of Beaudry and Portier (2006, 2007), who present evidence that news of productivity shocks arrives in advance of actual changes to productivity. Their evidence is based on two structural vector autoregressions (VARs). The VARs use the same two variables, stock prices and total factor productivity (TFP), but they differ in their structural identification schemes. In the first VAR, the authors identify a shock to stock prices that is orthogonal to the current TFP shock. In the second VAR, they use a long-run restriction to identify shocks to long-run TFP. The authors find that the stock price shock from the first VAR and the long-run TFP shock from the second VAR are highly correlated, which suggests that stock market participants are able to predict future innovations to TFP. Information about future economic conditions should be reflected by many forward-looking variables beyond stock prices. Beaudry and Portier (2006) introduce consumption and hours into their VAR system and obtain similar results to their baseline bivariate VAR. Moreover, the authors show that these "news" shocks explain a substantial fraction of movements in consumption, investment, and hours worked at business cycle frequencies.

The empirical relevance of news and other informational shocks for business cycle analysis is an active area of research. Barsky and Sims (2008) consider another forward-looking variable: consumer confidence as measured by the Michigan Survey of Consumers. One of the questions in the Michigan survey asks respondents for their expectations of national economic conditions for the next five years. Barsky and Sims show that consumer confidence

is a useful predictor of changes in macroeconomic variables. They consider two interpretations of this finding, which they term the "animal spirits" view and the "superior information" view. The animal spirits view is that consumer confidence, or confidence more broadly, directly causes an expansion of economic activity. The superior information view is that consumer confidence reflects early knowledge of future economic conditions. The authors use a VAR analysis to distinguish between these two possibilities. The key findings are that innovations to confidence are highly correlated with innovations to long-run output and not correlated with transitory innovations to output. These results suggest that the superior information channel is the operative one because output growth that is not associated with increases in potential output, as in the animal spirits view, should be short-lived. These results support the finding of Beaudry and Portier that agents receive signals about productivity changes ahead of the actual change in productivity.

Sims (2009) proposes a method for identifying news shocks that is an alternative to the Beaudry and Portier approach. He estimates a VAR with data on TFP (corrected for capacity utilization), output, consumption, hours, stock prices, inflation, and consumer confidence. The latter two variables are intended to augment the information about future productivity provided by stock prices. After estimating the reduced-form VAR, Sims identifies the unanticipated shock to TFP with the reduced-form innovation to TFP and then identifies the news shock as the linear combination of the reduced-form innovations that best explains the remaining movements in future TFP. The response of the economy to news shocks under Sims's identification is quite different from its response to news shocks under the Beaudry and Portier identification. Sims finds that a favorable news shock leads to an increase in consumption but declines in hours, investment, and output on impact. As we discuss in Section 2, these are the co-movements that the standard real business cycle (RBC) model would predict for a news shock.

Blanchard, L'Huillier, and Lorenzoni (2009) investigate news shocks in a context in which agents are unsure about the exact nature of the innovation to productivity. Their model features permanent shocks to productivity that build up gradually over time as well as transitory shocks to productivity. Agents are not able to observe the two components of productivity separately, but instead observe the level of productivity and a noisy signal about the permanent component of productivity. The noisy signal fluctuates for two reasons: news and noise. Here news shocks are the permanent productivity shocks that because of their gradual effect on productivity, are largely information about future productivity rather than changes in current productivity. Noise shocks, by contrast, are shocks to the signal that are unrelated to changes in productivity. Ideally agents would ignore the noise shocks, but they are unable to fully distinguish between noise and news. The authors assume that agents smooth consumption completely in the sense that they set consumption equal

to their estimate of the permanent component of productivity. In response to a permanent productivity shock, consumption responds only gradually because the agents are unsure if the productivity shock is permanent, and over time they revise their estimates in favor of the shock being permanent. In response to a transitory shock or a news shock, consumption responds initially, but over time agents learn that the shock is transitory or nonexistent and consumption returns to its initial level. Importantly, the authors demonstrate that a VAR applied to data on productivity, consumption, and the productivity signal cannot produce impulse responses that match the true ones implied by the model. The reason is that the model posits that consumption is a random walk, and so the VAR, which makes use of current and past observations, cannot identify a shock that has a transitory impact on consumption. If it could identify such a shock, then the agents in the model, who have at least as much information as the econometrician, also would see the transitory dynamics in consumption and would adjust their consumption to eliminate them. Therefore, the consumption response to any shock the econometrician can identify must be flat. Moreover, it is not enough to allow the econometrician to use observations from the future. The problem that arises is related to the invertibility problems discussed by Fernández-Villaverde et al. (2007). When some state variables are hidden from the econometrician, an innovation in the statistical model may either be the result of an economic shock or the result of a discrepancy between the econometrician's beliefs about the state variables and the true state. Only if the econometrician can infer the value of the state with certainty can he or she be certain about what is a shock and what is a "mistake" about the state. Blanchard, L'Huillier, and Lorenzoni show numerically that even with a large amount of data from the future, the econometrician is still uncertain about the state and therefore still uncertain about the shocks that generated the data. While news and noise shocks cannot be identified using VAR analysis, the model can be estimated structurally and information about the shocks can be recovered using the Kalman smoother. By imposing more structure on the data, the authors are able to summarize, but not completely eliminate, the uncertainty about the state variables and the economic shocks. The resulting structural estimates imply that noise shocks are an important source of short-run volatility, accounting for 50 percent of the variance in consumption at a four-quarter horizon. The remaining 50 percent of the variance in consumption is attributable to permanent and transitory productivity shocks in roughly equal measures. The results suggest that the manner by which information about changes in productivity disseminates is an important part of business cycle analysis. An interesting avenue for further research would be to see how the importance of noise shocks holds up in a richer model.

Additional evidence that noise shocks might be factors in aggregate fluctuations comes from the work of Rodríguez Mora and Schulstad (2007). These authors observe that official estimates of gross national product (GNP) are

revised over time, and the revisions are often quite substantial. They treat the final estimate of GNP as the true level of activity in a given quarter and the initial estimate as the perception of that level at the time. Their main finding comes from a regression of the true growth in GNP on the true growth and the perception of growth in the preceding quarter. They find that perceptions of growth in the previous quarter are useful in predicting future growth, but the true growth in the previous quarter is not. Moreover, they show that perceptions of growth in the previous quarter affect GNP growth through investment spending rather than consumption or government spending. These results suggest that the evolution of macroeconomic aggregates depends in part on perceptions of economic fundamentals that may not always be correct.

Finally, Schmitt-Grohé and Uribe (2009) investigate the importance of news shocks using a structural estimation approach. These authors estimate an RBC model that incorporates a number of real rigidities and structural shocks. Specifically, they include permanent and transitory shocks to TFP, investment-specific productivity shocks, and government spending shocks. Each of the shocks is composed of innovations that are anticipated at different horizons ranging from zero quarters (unanticipated) to three quarters. Their posterior mode attributes about 70 percent of the variance of output growth to anticipated shocks and the posterior probability that this share is less than 50 percent is essentially zero. Moreover, they find that output, hours, consumption, and investment all increase in response to a positive anticipated transitory shock. However, hours fall in response to a positive anticipated permanent shock. The results in this article strongly support anticipated technology shocks as sources of business cycle fluctuations.

All in all, much of the literature points to news and other informational shocks as potentially important drivers of aggregate fluctuations. However, it is far from clear yet how to best model and identify these disturbances. Relatedly, if one wanted operational measures of news shocks that could be fed into a model and used to predict aggregate economic variables over the near term, how would these shocks be constructed in practice (perhaps based on current events)? The empirical studies discussed above define the shocks as residuals based on an empirical (structural or semi-structural) specification; direct measurement is hard, and estimates via, say, surveys regarding "consumer confidence" would tend to mix news shocks with other shocks. This empirical problem, of course, is shared with, and arguably less severe than in, traditional Keynesian methods.

2.   THEORETICAL CHALLENGES

In light of the evidence that changes in TFP can be anticipated to a significant extent, a natural question is to ask how such news shocks play out in the standard real business cycle model. The standard one-sector RBC model has

time-additive preferences for consumption and leisure of the form

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t u \left(C_t, \bar{H} - H_t \right) \right], \tag{1}$$

where β is the discount factor, $u(\cdot, \cdot)$ is the period utility function, C_t is the flow of consumption, and H_t represents hours worked out of a maximum $\bar{H}$. In the standard model, output is produced according to a constant returns-to-scale production function that combines capital and labor. The stochastic disturbances that drive the business cycle enter through the production function in the form of technology shocks. The most commonly assumed functional form for the production function is Cobb-Douglas, which leads to

$$Y_t = F\left(K_t, H_t, z_t \right) = z_t K_t^{\alpha} H_t^{1-\alpha},$$

with K_t being the stock of capital at the beginning of the period and z_t being the level of technology in period t. Resources evolve according to

$$K_{t+1} = F\left(K_t, H_t, z_t \right) + (1 - \delta) K_t - C_t. \tag{2}$$

This resource constraint implies that there is a single homogeneous good that is freely used for consumption or as capital.

As the standard model is frictionless, the equilibrium behavior of the economy can be found through solving a planner's problem. The planner chooses stochastic processes for C, H, and K to maximize expected utility according to equation (1) subject to equation (2), the stochastic process for z, and the initial condition for K_0. The first-order conditions of this problem can be expressed as the usual Euler equation

$$u_C \left(C_t, \bar{H} - H_t \right) = \beta E_t \left[R_{t+1} u_C \left(C_{t+1}, \bar{H} - H_{t+1} \right) \right], \tag{3}$$

where R_{t+1} is the marginal product (in equilibrium, the rental rate) of capital in period $t + 1$:

$$R_{t+1} = F_K \left(K_{t+1}, H_{t+1}, z_{t+1} \right) + 1 - \delta, \tag{4}$$

and the efficiency condition for the labor-leisure tradeoff:

$$u_C \left(C_t, \bar{H} - H_t \right) w_t = u_H \left(C_t, \bar{H} - H_t \right), \tag{5}$$

where w_t is the marginal product of labor (in equilibrium, the wage rate) in period t:

$$w_t = F_H \left(K_t, H_t, z_t \right). \tag{6}$$

Though a full account of the effects of shocks requires a full solution of the stochastic general-equilibrium model and examination of its simulated time-series properties, one can obtain significant insight by looking at "unexpected shocks to steady states." That is, assume that an economy is in steady state and will stay there until there is an actual change in the technological opportunities that occurs with probability zero. The question at hand here is how knowledge

of a future change in technology will affect the economy in the intervening periods before the changes actually occur. While the Beaudry-Portier evidence suggests that positive news about future productivity should lead to something of a business cycle expansion, the standard one-sector RBC model cannot generate such a response. To see why the standard model has trouble generating a business cycle expansion in response to a positive news shock, consider what is required of the four main variables in the model: output, consumption, investment, and hours. An expansion is marked by an increase in both consumption and investment. In the standard model there are no imports or exports and no government spending, so the aggregate resource constraint requires that output must rise to allow consumption and investment to rise simultaneously. The only option for an increase in output is for hours worked to rise as the technological opportunities are initially unchanged and the capital stock is predetermined by what was installed in the previous period. However, consumption and leisure are normal goods under standard preferences, so that at a given wage (marginal product of labor) a household will choose to adjust consumption and leisure in the same direction, i.e., consumption and hours in opposite directions. To see this mathematically, equation (5) can be used to implicitly differentiate H with respect to C. Doing so yields

$$H'(C) = -\frac{u_{CC}w - u_{CH}}{u_{HH} - u_{CH}w}, \qquad (7)$$

and decreasing marginal utility ($u_{CC}, u_{HH} < 0$) together with the weak complementarity of consumption and hours ($u_{CH} \geq 0$) imply this expression is negative. So hours and consumption must move in opposite directions when wages are held constant. The only hope for the model is that in equilibrium wages increase so that the substitution effect raises hours, but, as was already noted, the capital stock and technology have not changed so increased hours will lead to lower wages in equilibrium. The implication is that the equilibrium response of the standard one-sector RBC model to a positive news shock does not look like a business cycle expansion.

If the model does not generate a boom in response to a news shock, what happens instead? If the preferences exhibit a strong wealth effect then positive news about future productivity will lead to an increase in consumption. This increase in consumption is associated with a decline in hours worked as before, which in turn implies a reduction in output and the aggregate resource constraint implies a reduction in investment. In contrast, with a weak wealth effect all of these implications can be reversed.[1]

It will be useful to consider an extreme case for preferences, both for the sake of understanding the workings of the basic model and for the sake of

---

[1] Using a particular set of functional forms, Beaudry and Portier (2004) show that consumption and investment respond in opposite directions for any set of parameter values.

understanding the behavior of the model that is presented in Section 3, which is based on the Diamond-Mortensen-Pissarides framework. Consider a utility function that is just linear in consumption $u(C, H) = C$, so that leisure is not valued and labor supply is fixed exogenously. In this case, the return on capital is pinned down by the discount rate in all periods as shown by the Euler equation:

$$1 = \beta E_t \left[F_K \left(K_{t+1}, \bar{H}, z_{t+1} \right) + 1 - \delta \right]. \tag{8}$$

This Euler equation implies that in an experiment with perfect foresight, the capital stock will perfectly track the level of technology—K_t is only a function of z_t and parameters of the model. The result is that in response to a news shock the capital stock remains unchanged until the period before the change in productivity takes place when (for a positive news shock) consumption is reduced to raise the capital stock to its new steady-state level. While this case yields a simple transition to the new steady state, the dynamics it does generate have consumption and investment moving in opposite directions and with a delay.

   An important element of the Beaudry and Portier analysis is the response of the stock market or, in terms of the model, the relative price of capital. In the standard one-sector model there is in essence a single good that is used for both consumption and capital. Therefore, the relative price of capital is fixed at one unit of the consumption good at all times. A truly satisfactory explanation of the Beaudry and Portier results would be able to replicate the behavior of the stock market as well as the usual macroeconomic aggregate quantities. Christiano et al. (2007), reviewed below, do discuss stock prices within their model.

3. A SEARCH MODEL

The overall question we discuss in this article is what kinds of theoretical settings can deliver co-movement in response to news shocks. In Section 4, we survey the recent literature and the range of models discussed there. Here, mostly for the purpose of illustration, we look at a specific, and very simple, model: one based on the Diamond-Mortensen-Pissarides search-and-matching model. What we present here is related to Den Haan and Kaltenbrunner (2009), who study a similar setting. The setting with search frictions offers something that the standard neoclassical model does not have: "free resources," namely, a set of unemployed agents who would gladly work if they could just find an employer. Therefore, it is at least imaginable that the frictions are such that when a news shock arrives, employment responds relatively quickly, provided that frictions are endogenous and respond to the news. The response of frictions in this model is governed by flows of firms in and out of the market for workers. The idea is, in principle, very simple: If there is

positive news, firms flow in immediately and look for workers, which makes it easier for workers to find employment, leading to an increase in employment and higher production, so that the overall resources available are increased. Firms begin posting vacancies immediately upon learning the positive news because an employed worker is immediately more valuable since, with some probability, that worker will still be employed by the firm when productivity rises.

Model Framework

The model framework is the standard continuous-time Diamond-Mortensen-Pissarides search-and-matching model. A more detailed discussion of the model framework and the determination of steady-state values can be found in Pissarides (2000) or Hornstein, Krusell, and Violante (2005).

The model economy is populated by a unit continuum of workers. Workers have linear utility for consumption discounted at the rate r, which implies they are risk-neutral. The workers each supply one unit of labor inelastically. Workers can be either employed or unemployed. Employed workers receive a wage income of w and unemployed workers receive an unemployment benefit of b, which also can be interpreted as the value of home production during unemployment. The wage is an endogenous variable that will depend on, among other things, the tightness of the labor market. The unemployment benefit is an exogenous feature of the economic environment. Workers cannot save and consume their income flows immediately.

The economy is also populated by an endogenous number of firms that are also risk-neutral and discount future profits at rate r. Firms all have access to the same production technology so there are no productivity differences across firms. Firms are free to enter the labor market, but posting a vacancy involves a flow cost in the amount c. Production requires a single worker and a single firm and the amount of output produced by such a pair, $p(t)$, varies through time. It is assumed that production is always efficient in the sense that $p(t) > b$.

There is a search friction in the labor market so that, at any point in time, there will be a fraction $u(t)$ of workers who are unemployed and looking for firms and there will be a measure $v(t)$ of firms with vacant jobs looking for workers. These two groups meet at a rate, $m(t)$, that is determined by a constant-returns-to-scale matching function $M(u(t), v(t))$. We use a Cobb-Douglas matching function, $M(u, v) = Au^\alpha v^{1-\alpha}$. Given the rate at which new matches occur, the rate at which an unemployed worker finds a firm is $\lambda_w(t) = m(t)/u(t)$ and the rate at which a vacant firm finds a worker is similarly $\lambda_f(t) = m(t)/v(t)$. The gains from forming a productive worker-firm pair are divided between the worker and the firm by Nash bargaining, with β going to the worker and $1 - \beta$ going to the firm. Existing worker-firm pairs separate at the exogenous rate σ.

Steady State

To determine the steady-state values of unemployment and wages, we begin by writing the conditions that must be satisfied by the values for the employed worker, unemployed worker, matched firm, and vacant firm. Respectively, these are:

$$rW(t) = w(t) + \sigma[U(t) - W(t)] + \dot{W}(t) \tag{9}$$

$$rU(t) = b + \lambda_w(t)[W(t) - U(t)] + \dot{U}(t) \tag{10}$$

$$rJ(t) = p(t) - w(t) + \sigma[V(t) - J(t)] + \dot{J}(t) \tag{11}$$

$$rV(t) = -c + \lambda_f(t)[J(t) - V(t)] + \dot{V}(t), \tag{12}$$

where a dot over a variable represents the derivative with respect to time. Each of these equations can be interpreted in terms of the relationship between the flow value and the capital value of a state. For example, equation (9) states that the flow value of being an employed worker is equal to the income flow plus the expected value of the capital loss that occurs upon separation when the worker becomes unemployed and the change in value over time, possibly stemming from a changing environment.[2]

The total surplus of a worker-firm match is the sum of the worker's gain and the firm's gain, $S \equiv (W - U) + (J - V)$. The Nash-bargaining determination of wages implies that the total surplus is divided between workers and firms according to their bargaining powers:

$$W - U = \beta S \tag{13}$$

$$J - V = (1 - \beta)S. \tag{14}$$

A useful expression for S can be found by adding and subtracting equations (9)–(12) and using equations (13) and (14):

$$rS = p - b + c - \sigma S - \lambda_f(1 - \beta)S - \lambda_w\beta S + \dot{S}. \tag{15}$$

This can be viewed as an "asset-pricing" equation: The value of the match—the worker and the employer—equals a current payoff plus future payoffs, which are captured by the $\dot{S}$ term; they can, in principle, be successively substituted in so that the price of the asset equals the present value of all payoffs, present and future. The equation can be rearranged to yield

$$S = \frac{p - b + c + \dot{S}}{r + \sigma + \lambda_f(1 - \beta) + \lambda_w\beta}. \tag{16}$$

Now use the fact that firms are free to enter (and exit) the labor market, so the value of a vacant firm must be zero. Setting V equal to zero in equations (12)

---

[2] See footnote 12 in Hornstein, Krusell, and Violante (2005) for a detailed derivation of these conditions.

Table 1 Model Parameter Values

Symbol	Description	Value
p	Productivity	1.000
b	Unemployment benefit	0.950
α	Elasticity of the matching function	0.720
A	Matching function efficiency	1.350
β	Worker's bargaining share	0.050
r	Interest rate	0.012
σ	Separation rate	0.100
c	Vacancy posting cost	0.357

Notes: One unit of time is equal to one quarter. See Hornstein, Krusell, and Violante (2005) for additional details.

and (14) and combining the results yields

$$S = \frac{c}{\lambda_f \, (1 - \beta)}. \tag{17}$$

Combining equations (16) and (17) yields one equation in the two meeting rates:

$$\frac{p - b + \dot{S}}{r + \sigma + \lambda_w \beta} = \frac{c}{\lambda_f \, (1 - \beta)}. \tag{18}$$

As the matching function is constant returns to scale, the meeting rates can be expressed in terms of a single variable that represents labor market tightness:

$$\begin{aligned}
\theta &\equiv v/u \\
\lambda_w &= M(u, v)/u = M(1, \theta) = A\theta^{1-\alpha} \\
\lambda_f &= M(u, v)/v = M(1/\theta, 1) = A\theta^{-\alpha}. \tag{19}
\end{aligned}$$

In steady state the total surplus is constant, $\dot{S} = 0$, so equation (18) is one equation in the unknown θ. Once θ has been found, the λs and values follow immediately from the equations above and equations (17), (14), and (11) can be used to find the wage as a function of θ and p.

The unemployment rate evolves slowly as workers gradually flow into and out of unemployment. The evolution of the unemployment rate follows

$$\dot{u}(t) = \sigma \, [1 - u(t)] - \lambda_w(t)u(t), \tag{20}$$

and in steady state, unemployment is simply equal to $\sigma \, / \, (\lambda_w + \sigma)$.

Solving for the steady state of the model requires solving a nonlinear equation in θ (equation [18], with $\dot{S} = 0$). We do this numerically after calibrating the model following Hagedorn and Manovskii (2008). This calibration leads to a steady-state unemployment rate of 6.9 percent and features stronger effects on firm entry of productivity shocks than in alternative calibrations

such as Shimer (2005); for a discussion, see Hornstein, Krusell, and Violante (2005). The parameter values used in our calibration appear in Table 1.

Transition to Steady State

Before considering the effects of a news shock it is necessary to consider how the economy transitions to the steady state in a stationary environment. The key result for the transition dynamics is that labor market tightness immediately reaches its steady-state value regardless of the initial conditions for the economy. As unemployment is a predetermined variable, the response of the labor market is driven by a jump in posted vacancies. To see that this must be the case, rewrite equation (18) into

$$\dot{\theta} = \left[\frac{r + \sigma}{\alpha} + \frac{\beta A}{\alpha} \theta^{1-\alpha} - \frac{(p - b)(1 - \beta)A}{c\alpha} \theta^{-\alpha} \right] \theta, \qquad (21)$$

so that dynamics are expressed in terms of θ (thus including its time derivative $\dot{\theta}$).[3] Notice that the term inside the brackets is increasing in θ, so for θ below the steady-state value the time derivative is negative and for θ above the steady-state value the time derivative is positive; therefore, the steady state is unstable, and the only nonexplosive solution to the problem is for θ to jump immediately to the steady state.[4] It then follows that the λs also must jump to their new steady-state values and so then must the values W, U, J, V, and S. Given the new, constant level of λ_w, one can use equation (20) to trace out the evolution of the unemployment rate to its new steady-state value, and vacancies are then determined by the relationship $v = \theta u$. In the end, there are very limited transition dynamics resulting from an unexpected productivity shock, and if productivity is expected to remain constant in the future, then θ must be at its steady-state value.

News Shock (Recession)

We now consider how the model responds to a negative news shock. In particular we perform the following experiment: Before $t = 0$, the economy is in steady state and expected to remain there in perpetuity. At $t = 0$, news arrives that at time $T = 5$ productivity, p, will drop by 1 percent. The arrival of this news is a zero-probability event, which implies that agents put no weight

---

[3] Equation (17) and its time derivative imply $\dot{S}\lambda_f (1 - \beta) = -c\dot{\lambda}_f / \lambda_f$, and equation (19) can be used to relate $\dot{\lambda}_f$ to $\dot{\theta}$.

[4] Pissarides (2000) considers the system of differential equations formed by equations (20) and (21). The boundary conditions for this system are the initial condition on u and the requirement that the system converge. These conditions can only be met if θ immediately assumes its steady-state level.

on the event in forming their expectations, but does not imply that it cannot occur.

To calculate the equilibrium response of the economy to this news, we use the fact that θ must be at its new steady-state value at time T when the change in productivity occurs because after that point the environment is expected to be stationary. We use this as a terminal condition and solve the ordinary differential equation (21) from time $t = 0$ to T. Having done so, we are able to calculate the λs, trace the evolution of the unemployment rate, and solve for all the other equilibrium quantities in the model. Interestingly, our version of the Pissarides model has nontrivial dynamics, whereas the standard model does not; in the standard model, there is always an immediate jump in θ in response to a change in productivity, since this change is known as it is realized. The slow-moving θ we look at, thus, comes from knowing that productivity will change at a known future date.

The results appear in Figure 1, and we begin by comparing the two steady states. The lower level of productivity results in a decrease in the total surplus of a match, one implication of which is that the value of a productive firm is lower. This induces fewer firms to enter the labor market until market tightness falls sufficiently and the probability of finding a new worker rises to keep the value of a vacant firm at zero. The weaker labor market leads to a lower job-finding rate for unemployed workers, which leads to a higher steady-state unemployment rate. In equilibrium the wage decreases, but by less than productivity, so profits also decrease eventually, though profits first rise since wages, which are forward-looking, fall and productivity has not yet fallen. Total resources fall smoothly, which is the effect sought: Firms leave in anticipation of future falls in profit, which creates additional unemployment— there are now even more "free resources" in the form of workers who are not working. The fact that fewer firms are posting vacancies means that fewer resources are spent on vacancy posting, which we interpret as investment. Resources net of investment costs rise somewhat during transition but then drop and are lower in the long run.[5] During the transition to the new steady state from time $t = 0$ to time T, the value of a productive firm drops initially and then smoothly falls toward the new steady-state value. Labor market tightness follows the same pattern, which is achieved by an initial jump and then decreasing path for vacancies. The weaker labor market decreases the speed at which workers flow out of unemployment and results in a gradual rise in the unemployment rate. Unlike the other variables, vacancies overshoot their steady-state level. This overshooting stems from the fact that unemployment

---

[5] A model in which there is endogenous separation—say, because workers or matches are heterogeneous so that only some firm-worker contacts lead to lasting matches—might generate another channel through which more resources are left idle since then some existing matches could also break up in reaction to negative news.

Figure 1 News about a Coming Fall in Productivity
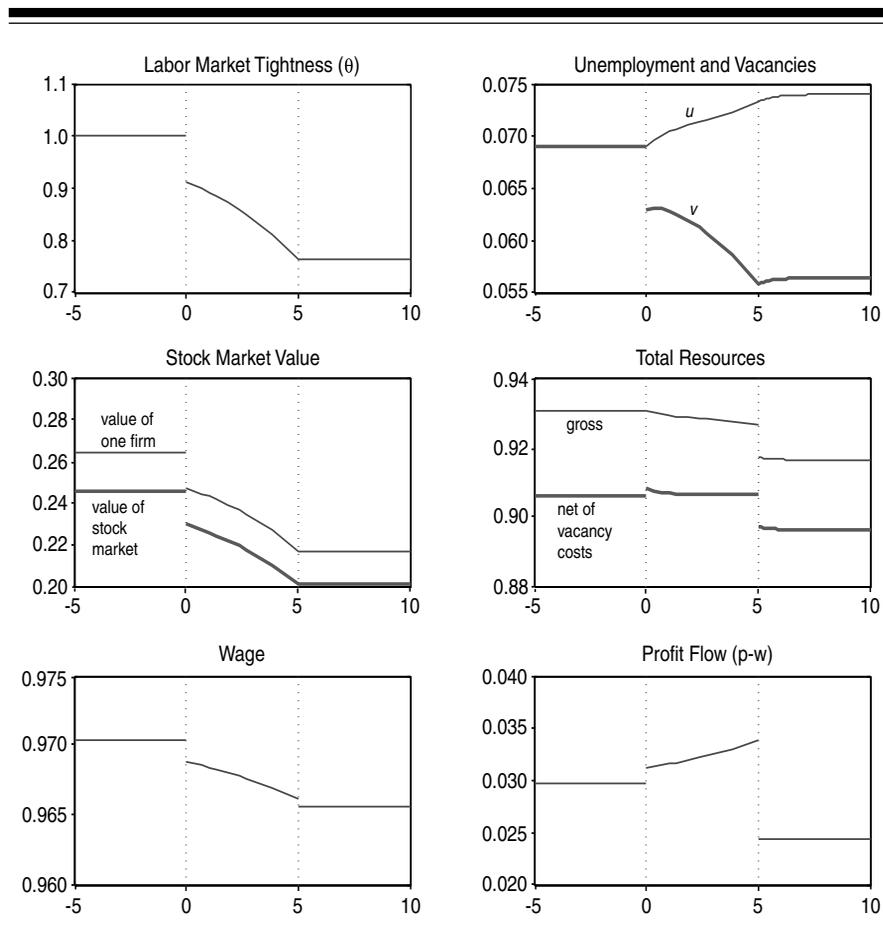
has not reached its steady state at time T, but is still below that level. Vacancies must then also be below their steady-state level at T so that labor market tightness can remain at its steady-state level from T onward. The increase in the unemployment rate mechanically leads to a decrease in output, and the level of output jumps when all employed workers become less productive at T.

The model is successful in generating a decline in employment, output, and the stock market. What about investment and consumption? If we interpret firm vacancy-posting costs as investment, then the model also generates a fall in investment. Consumption, however, must rise on impact if the economy is closed: No existing matches are broken up endogenously, so on impact no resources are lost, but investment falls, and thus consumption must rise. An

open-economy version of the model with decreasing marginal utility would reverse this result, as consumers would then want to smooth consumption over time and use intertemporal international trade to achieve a smoothly declining path for consumption.

As Figure 1 shows, labor market tightness, θ, drops initially when the news is received and then converges to its new steady-state level at date T. This pattern will hold for any choice of parameter values. Quantitatively, however, the initial impact of the news on labor market tightness depends on the way the model is calibrated, and there are two ways that the parameters can affect this initial impact. First, different calibrations lead to different steady-state responses of θ to changes in p. This sensitivity is the focus of the literature that studies the implications of search-and-matching models for unemployment fluctuations in response to unanticipated productivity shocks (Shimer 2005; Hagedorn and Manovskii 2008; Pissarides 2009). The more θ must have changed by date T, the more it must jump initially. The second consideration is the speed with which the market tightness adjusts to its new steady-state level. If the model dynamics are such that θ moves rapidly when it is out of steady state, then a small drop in θ is needed at date 0 to achieve the same level of θ at date T. What then determines the speed of convergence and therefore the size of the initial impact of the news? Mathematically, if the right-hand side of equation (21) is increasing more quickly in θ, then the speed of convergence will be higher, and the initial impact of the news will be smaller. For example, differentiation of equation (21) shows that an increase in the interest rate, r, leads to a faster speed of convergence. This result is intuitive as an increase in the interest rate leads firms to discount the future more heavily and so the value of a firm depends more on the immediate future and less on the distant future. As the productivity change does not happen for some time after the news arrives, firms with high discount rates do not respond as much as firms with low discount rates. Similar logic holds when the separation rate is high. In this case, firms discount the future because the match is likely to be destroyed before the change in productivity occurs.

Differentiation of equation (21) also shows that the speed of convergence is increasing in the worker's bargaining share, β. Therefore, when workers have more bargaining power, the initial impact of the news is smaller. To see the importance of the worker's bargaining share, consider the case when β is set to zero. In that case, the worker's wage is always equal to the value of leisure, b, and the firm's flow profit is $p - b$, which is unchanged until date T when it jumps down. Now consider a positive bargaining share, $\beta > 0$. As shown in Figure 1, the worker's wage falls at date 0 and remains below its initial level thereafter. With a lower wage, flow profits actually rise between dates 0 and T. So with a positive β, firms are partially compensated for the future reduction in productivity by a short-term increase in profits. This

Figure 2 Misleading News about a Coming Decline in Productivity
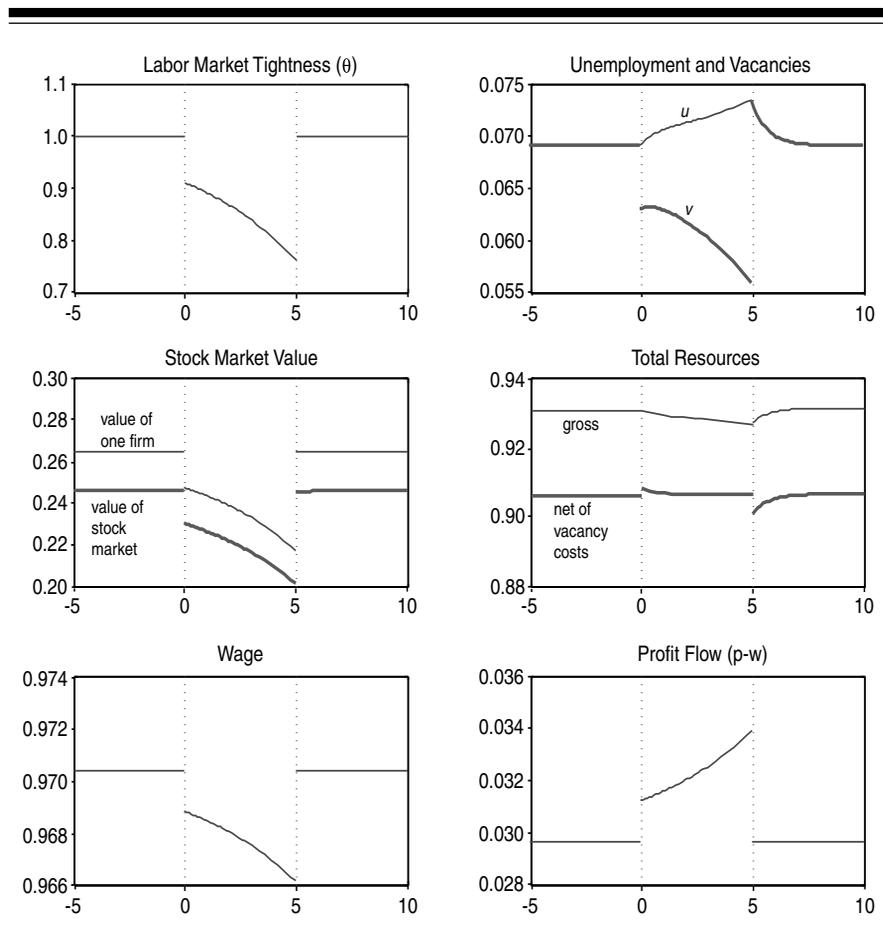
short-term increase in profits motivates firms to post vacancies just after the news arrives, and this force reduces the initial drop in θ.

The News Shock Turns Out to Be Wrong

The second experiment that we consider is to ask what happens if the expected lower productivity is not realized at time T, but instead productivity remains at its initial level both before and after T. Specifically, we assume that after the news shock arrives, there remains a possibility that productivity fails to decline at time T, although this possibility has zero probability; thus, we consider what happens when that zero-probability event occurs. The experiment is displayed in Figure 2, with $T = 5$ again. Before time T the economy behaves exactly

as in the case when the productivity shock is realized because agents fully expect that it will be realized. At time T, however, the productivity shock does not materialize and labor market tightness and the value of a productive job immediately return to their initial steady states. These developments imply a stock market boom and an immediate increase in posted vacancies. The new tightness in the labor market increases the rate at which unemployed workers find jobs and leads to a gradual fall in the unemployment rate. As employment rises output also rises, but, as before, the increase in vacancy posting costs is large enough that it offsets the rise in output so the resources available for consumption actually decrease.

Looking at this experiment, one might label the shock whose effects are displayed in Figure 2 "misleading." More generally, realizing that all shocks containing "news" do not necessarily always lead the economy in the right direction, one can speak of "noise" perhaps: shocks that are believed to have relevance for productivity but in the end do not. For example, the Internet technology bubble during the last years of the last millenium could have reflected beliefs that eventually turned out to be too optimistic (but may well have been rational). Thus, the literature on news shocks should be viewed as closely related to ideas about noise as well. The very recent literature (e.g., Angeletos and La'O [2009], or Blanchard, L'Huillier, and Lorenzoni [2009]) takes an explicit signal extraction approach and thus formalizes news and noise, as shocks driving business cycles, in a slightly different way.

4.   OTHER APPROACHES IN THE LITERATURE

We now briefly discuss the main features of the different models, all with neoclassical underpinnings, that have been proposed as a way of generating co-movement in response to news shocks. In this discussion, we omit the very recent contributions to this literature that build on signal processing and "noise shocks."

Other Approaches to Labor Market Frictions

Den Haan and Kaltenbrunner (2009) present a version of the RBC model with a search friction in the labor market. Specifically, production occurs within "projects" that require an entrepreneur and a worker. Creating a new project is a time-consuming process as entrepreneurs and workers must search for one another. In response to a news shock, entrepreneurs and workers begin preparing for the future productivity increase by entering the labor market to begin the process of establishing relationships through which they can exploit the higher future productivity when it arrives, just as in Section 3. There are two main differences between the model in Section 3 and Den Haan and Kaltenbrunner's work. First, in Section 3 the labor supply is inelastic, while

Den Haan and Kaltenbrunner allow it to be elastic. With elastic labor supply, one of the effects of a news shock is an increase in the demand for leisure through the wealth effect, which might reverse the result that employment increases in response to the news shock. Den Haan and Kaltenbrunner show that this effect is sufficiently weak to be overcome by the household's motivation to enter the labor market to find a job in anticipation of higher productivity in the future. Therefore, the result that employment increases in response to a news shock is not an artifact of the inelastic labor supply. Second, the standard Diamond-Mortensen-Pissarides model considered in Section 3 does not include capital, so there are no predictions for the response of investment to a news shock. Den Haan and Kaltenbrunner show that investment does respond positively to a news shock except in the first period after the shock. Production is fixed in the first period because the capital stock and employment are predetermined, so it is impossible for consumption and investment to rise simultaneously in that period. However, the increase in employment that occurs in response to the news shock quickly increases output to finance higher investment as well as higher consumption in subsequent periods.

Multiple Sectors

The standard one-sector RBC model has a tight link between consumption and investment decisions: Investment directly reduces the resources available for consumption. Beaudry and Portier (2004) present a three-sector model with final goods, nondurable intermediate goods, and capital produced in different sectors. The latter two sectors use labor and a sector-specific fixed factor of production. In this model the link between consumption and investment is much weaker because output from the capital goods sector cannot be used for consumption and the presence of the fixed factors limits the extent to which the planner is willing to alter the amount of labor in the sectors. This uncoupling of the consumption and investment decisions allows consumption and investment to both increase in response to a positive news shock. Specifically, Beaudry and Portier assume the news concerns the future productivity of the nondurable goods sector, and the crucial assumption is that nondurable goods and capital are complementary in the production of final goods. Under these assumptions, the planner chooses to build up the capital stock in response to positive news about future nondurable goods productivity because the complementarity implies that capital will be more productive in the future because nondurables will be cheaper. The accumulation of capital, however, makes nondurables more valuable, which leads the planner to expand their production as well. In the end, the production of capital and nondurables increases, which is achieved through an expansion of hours worked in each sector and therefore in total. More capital and nondurables directly translate into more final output for which the only use is consumption. In this way the

model delivers an expansion of consumption, investment, hours, and output in response to positive news about nondurables productivity.

Other Model Features

An alternative approach taken in the literature is to keep the single-sector framework, but modify the standard RBC model along several other dimensions. Jaimovich and Rebelo (2009) present a model with three key modifications. They assume a functional form for preferences that has extremely weak short-run wealth effects on labor supply. In fact, the preferences used nest those of Greenwood, Hercowitz, and Huffman (1988) in which there is no wealth effect on labor supply. The calibration used by Jaimovich and Rebelo is extremely close to this case. Since these preferences imply a zero wealth effect, they allow the model to generate an increase in hours despite a substantial increase in consumption. The second modification introduced by Jaimovich and Rebelo is an adjustment cost for the rate of investment, which serves to produce an investment boom in response to a positive news shock as the planner wishes to minimize adjustment costs by smoothing investment over time. Finally, the authors add variable capacity utilization to the model, which allows the amount of resources to be expanded in the initial periods in order to finance simultaneous consumption and investment booms. The resulting model succeeds in generating a sizable boom in response to news of a future increase in TFP and in response to news of future investment-specific technical change.

Christiano et al. (2007) make similar modifications to the standard model in order to generate a boom in response to a positive news shock. Their key modifications are to introduce habits in consumption and the adjustment cost to the flow of investment. Jaimovich and Rebelo also have non-time-separable preferences, but the calibration is such that the habit persistence is very weak. The habits and adjustment costs in Christiano et al.'s work motivate the planner to engineer a smooth transition to the new steady state and begin consuming and investing in advance of the change in productivity. Hours are able to increase to provide resources for the consumption and investment booms because there is no longer a tight link between current hours and current consumption in the presence of habit persistence.

A troubling feature of these models is the response of the price of capital to a news shock, which is a decline. As investment is raised to reduce adjustment costs in anticipation of higher investment in the future, there is, in a sense, an excess of capital before the shock occurs. The result is that the relative price of capital falls during the boom. Walentin (2009) presents a model that is close to that of Christiano et al. (2007), with the modification that there is limited enforcement of financial contracts. With limited enforcement, there is a wedge between the value of the firm and the cost of its capital and, in

Walentin's model, this wedge increases in response to a news shock so that the value of the firm increases despite the fall in the cost of capital.

Investment-Specific Technical Change

In a model with adjustment costs, the planner chooses to start investing early in order to minimize the cost of building up the capital stock in response to a sector-neutral productivity shock. If, however, productivity shocks are investment-specific, then the only way to take advantage of them is through investment. Flodén (2007) uses a vintage capital model to argue that the news that next period's vintage of capital will be very productive leads to a boom in the current period. The mechanism draws on the model elements presented by Greenwood, Hercowitz, and Krusell (2000), which are shocks to the relative price of capital and variable capacity utilization. The cost of more intensive utilization of the capital stock is typically modeled as faster depreciation. When the relative price of capital declines, the replacement cost of the depreciated capital stock falls. As a result, an investment-specific technology shock leads to more intensive utilization in the current period, which raises the marginal product of labor and elicits higher labor supply. The additional resources produced through the increases in utilization and labor supply allow consumption to increase at the same time as investment.

Flodén only considers news shocks at a horizon of one period. That is, the economy learns that the capital being installed in the current period will be more productive in the next period and thereafter. This short horizon makes the expectations-driven boom somewhat short-lived, but it may be possible to extend the boom by extending the period between the receipt of the news and the technological change.

There is some ambiguity about the timing of the technology shock in that investment-specific technical change relates to the evolution of resources between periods rather than the productivity within a period. For example, Greenwood, Hercowitz, and Krusell adopt the timing convention that the shock relates to the productivity of investment this period and is therefore a shock in the current period, while Flodén considers the same shock to be a shock to the productivity of the capital when it is used in the future, which is then a shock that arrives in the future but is learned about in the current period through the news shock. Both interpretations are valid, but an important consideration is the interpretation used in the construction of the National Income and Product Accounts (NIPA). In principle the NIPA investment data are adjusted for quality, and if the vintage of capital that is being installed is going to be more productive in the future, this may be accounted for in the measurement of current investment and current TFP. However, if the shock raises current TFP, it would not be classified as a news shock by Beaudry and Portier (2006) because news shocks are orthogonal to current TFP shocks.

Financial Frictions

Another way of modifying the model to generate expectations-driven business cycles is to introduce financial frictions. Chen and Song (2008) consider a model with two sectors, only one of which requires the use of working capital. In their model, entrepreneurs have the ability to divert working capital, and the optimal contract in response to this limited debt enforcement leaves the sector financially constrained. When a positive news shock arrives, the entrepreneurs' continuation value rises because future profits will be higher, which relaxes the financial constraint. By reducing financial frictions, the news of higher TFP in the future triggers a reallocation of capital between the two sectors and raises current TFP. The increase in current TFP leads to more output that can be used for both more consumption and more investment. The more efficient use of capital, as well as the accumulation of more capital, raises the marginal product of labor, which leads to an increase in hours under Greenwood-Hercowitz-Huffman preferences.

If financial frictions like the ones Chen and Song have proposed are important features of the macroeconomy, then there are implications for other issues besides expectations-driven business cycles. In particular, there would be a need for government policy to alleviate the financial constraints of firms. This could be achieved in a variety of ways; for the same reasons as future profit increases would improve the current allocation of capital, any policy that increases future profits would have a desirable effect (production subsidies would suffice for this purpose).[6] Whether the economy is subject to this strong inefficiency is perhaps questionable. If there is already government policy in place designed to correct the inefficiency, no reallocation of capital in response to news shocks will take place.

5.   CONCLUSION

The news shocks literature has generated some interesting new insights about macroeconomic dynamics that seem relevant for understanding co-movement of macroeconomic aggregates. The above-discussed settings, including the simple Pissarides (1985) search/matching model used for illustration, do admit some channels that are promising ways forward. Some of these settings have more nonstandard features than others, and it is an open question whether they will survive more microeconomic scrutiny. It is also, as discussed in Section 1, still an open question how to identify news shocks and whether they really do lead to co-movements. All in all, this new literature does offer a challenge to

---

[6] Such policies might involve time inconsistency, since it is only by support of *future* policy that the desired effect is attained.

existing macroeconomic settings that do not admit co-movement in response to news shocks, and, as such, they should perhaps move our priors.

As also briefly mentioned above, a very recent strand of articles is now exploring explicit signal extraction channels by which news as well as noise can drive fluctuations. The focus here is on asymmetric information and, even though Lucas (1972) certainly sparked interest in the importance of this phenomenon for macroeconomics, there is no quantitatively oriented model available off the shelf to evaluate. A central reason for this is the theoretical difficulty of aggregation across agents with different information sets. Therefore, we may have to wait for a closer comparison between models relying on these ideas and existing representative-agent macroeconomic models.

Finally, the underlying notion in our discussion here is to examine whether co-movement is possible, in response to the arrival of information, in settings that are fully microfounded. It should be noted that none of these settings build on, or admit, coordination failures, which would seem to more easily admit strong effects of news or noise. With multiple equilibria, however, it is not clear how the movement across equilibria is supposed to occur, and there is nothing inherently more attractive about productivity-related shocks as coordination devices than other shocks, so it would seem that an approach based on coordination failures would have to be augmented with a theory of what triggers changes across equilibria. The earlier literature on sunspots (see Cass and Shell [1983] and later studies) offers an answer, but sunspots are just coordination devices, and it might be hard, in a reduced-form sense, to distinguish sunspots from true news shocks. If a news shock indicates high future productivity of capital, investment likely will go up today. Alternatively, in a model with multiple equilibria because of some form of increasing returns to capital, say, as an externality of capital use across firms, a sunspot would trigger either high or low investment, which would both be self-enforcing under the assumption of increasing returns. So this latter model would indeed justify later movements in productivity, not because of changes in technology, but through increasing returns and aggregate activity. Ultimately, these two "stories" could only be told apart by more detailed empirical scrutiny. One route is through better productivity measurements, perhaps finding ways of establishing what the returns to scale are on different levels of aggregation. Alternatively, a more detailed structural description of the model and examination of how the two kinds of economies respond to other shocks could help identification.

REFERENCES

Angeletos, George-Marios, and Jennifer La'O. 2009. "Noisy Business Cycles." Working Paper 14982. Cambridge, Mass.: National Bureau of Economic Research (May).

Barsky, Robert B., and Eric R. Sims. 2008. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." Mimeo, University of Michigan.

Beaudry, Paul, and Franck Portier. 2004. "An Exploration Into Pigou's Theory of Cycles." *Journal of Monetary Economics* 51 (September): 1,183–216.

Beaudry, Paul, and Franck Portier. 2006. "Stock Prices, News, and Economic Fluctuations." *American Economic Review* 96 (September): 1,293–307.

Beaudry, Paul, and Franck Portier. 2007. "When Can Changes in Expectations Cause Business Cycle Fluctuations in Neo-Classical Settings?" *Journal of Economic Theory* 135 (July): 458–77.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni. 2009. "News, Noise, and Fluctuations: An Empirical Exploration." Manuscript.

Cass, David, and Karl Shell. 1983. "Do Sunspots Matter?" *Journal of Political Economy* 91 (April): 193–227.

Chen, Kaiji, and Zheng Song. 2008. "Financial Frictions on Capital Allocation: The Engine of TFP Fluctuations." Unpublished manuscript, University of Oslo and Fudan University.

Christiano, Lawrence, Cosmin Ilut, Roberto Motto, and Massimo Rostagno. 2007. "Monetary Policy and Stock Market Boom-Bust Cycles." Manuscript, Northwestern University and the European Central Bank.

Den Haan, Wouter J., and Georg Kaltenbrunner. 2009. "Anticipated Growth and Business Cycles in Matching Models." *Journal of Monetary Economics* 56 (April): 309–27.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson. 2007. "ABCs (and Ds) for Understanding VARs." *American Economic Review* 97 (June): 1,021–6.

Flodén, Martin. 2007. "Vintage Capital and Expectations Driven Business Cycles." CEPR Discussion Paper 6113.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman. 1988. "Investment, Capacity Utilization, and the Real Business Cycle." *American Economic Review* 78 (June): 402–17.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman. 2000. "The Role of Investment-Specific Technological Change in the Business Cycle." *Europoean Economic Review* 44 (January): 91–115.

Hagedorn, Marcus, and Iourii Manovskii. 2008. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited." *American Economic Review* 98 (September): 1,692–706.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante. 2005. "Unemployment and Vacancy Fluctuations in the Matching Model: Inspecting the Mechanism." Federal Reserve Bank of Richmond *Economic Quarterly* 91 (Summer): 19–50.

Jaimovich, Nir, and Sergio Rebelo. 2009. "Can News About the Future Drive the Business Cycle?" *American Economic Review* 99 (September): 1,097–118.

Kydland, Finn E., and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica* 50 (November): 1,345–70.

Lucas, Robert E., Jr. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (April): 103–24.

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages." *American Economic Review* 75 (September): 676–90.

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory*. Cambridge, Mass.: MIT Press.

Pissarides, Christopher A. 2009. "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" *Econometrica* 77 (September): 1,339–69.

Rodríguez Mora, José V., and Paul Schulstad. 2007. "The Effect of GNP Announcements on Fluctuations of GNP Growth." *European Economic Review* 51 (November): 1,922–40.

Schmitt-Grohé, Stephanie, and Martín Uribe. 2009. "What's News in Business Cycles." Manuscript, Duke University.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review* 95 (March): 25–49.

Sims, Eric R. 2009. "Expectations Driven Business Cycles: An Empirical Evaluation." Mimeo, University of Michigan.

Walentin, Karl. 2009. "Expectation Driven Business Cycles with Limited Enforcement." Sveriges Riksbank Working Paper Series 229 (April).

Risk Sharing, Investment, and Incentives in the Neoclassical Growth Model

Emilio Espino and Juan M. Sánchez

T he amount of risk sharing among households, regions, or countries is crucial in determining aggregate welfare. For example, pooling resources at the national level can help regions better deal with natural disasters like floods. Similarly, pooling resources with an insurance company can help individuals deal with shocks like a house fire or a car accident.

Capital accumulation and economic growth also are crucial in determining aggregate welfare. In particular, they determine the stock of wealth available for consumption and investment. Importantly, wealthier households, regions, or countries possess a buffer stock of precautionary assets, a form of self-insurance.

These two important factors in determining welfare have interesting interactions with one another. An important one is how insurance and savings substitute for each other. For example, individuals may want to save more when they do not have access to insurance than when they do because the extra savings can protect against the consequences of an uninsured shock. Therefore, capital accumulation and growth would be faster in an economy without perfect insurance than in one with perfect insurance.

This article explores the tradeoffs between insurance and growth in the neoclassical growth model with two agents and preference shocks. Most of the analysis reviews the full information version of the model, where there are no limits on insurance between the two agents, though there is still aggregate uncertainty that affects aggregate savings behavior. Private information is

then added to the model to limit the ability to insure the two agents. This is a much harder problem, as has been observed in the literature, and only a partial characterization is provided.

Literature Review

Our article relates to the voluminous consumption/savings/capital accumulation literature on two levels. On one hand, there has been a growing literature focusing on the accumulation effects of demand side shocks in dynamic stochastic general equilibrium models, following the pioneering work of Baxter and King (1991) and Hall (1997). In general equilibrium models, demand side shocks (such as preference shocks to consumption demand) have a strong tendency to crowd out investment.[1] On the other hand, there is literature on the impact of inequality on capital accumulation. If preferences aggregate in the Gorman sense, the distribution of wealth does not affect the evolution of aggregate variables—see Chatterjee (1994) and Caselli and Ventura (2000). In our setting, preferences do not aggregate in that strong sense. Thus, distribution matters for aggregate savings and the corresponding dynamics of the aggregate stock of capital.[2]

The literature analyzing economic growth and private information is not as large, and the valuable contributions have relied on different simplifying assumptions to make the analysis tractable. This article is related to those articles because we are interested in understanding when information is (more) important to implement the full information allocation. However, we solve the full information model to obtain the full information allocation and characterize only the incentives to misreport the shocks under that allocation.

Pioneering contributions in the literature on constrained efficient allocations with private information abstracted from capital accumulation, as the main goal was to study wealth distribution. In a pure exchange economy setting, Green (1987) and Atkeson and Lucas (1992) show that (constrained) efficient allocations, independent of the feasibility technologies, display extreme levels of "immiserization": The expected utility level of (almost) every agent in the economy converges to the lower bound with probability one. This result is also present in Thomas and Worrall (1990). Then, in an early contribution that includes capital accumulation, Marcet and Marimon (1992) examine a two-agent model where a *risk-neutral* investor with *unlimited resources* invests in the technology of a risk-averse producer whose output is subject to privately observed productivity shocks. They show that the full information investment policy can be implemented in the private information

---

[1] See Wen (2006) for an overview and references therein.

[2] See Lucas and Stokey (1984) for a general early discussion and, more recently, Sorger (2002).

environment. That is, in their setting, a risk-neutral investor can make the risk-averse entrepreneur follow the full information investment policy and allocate his consumption conditional on output realizations. Thus, they find that growth levels are as high as with perfect information. The key simplification in this article is that the second agent in the economy is risk-neutral with unlimited resources.

Khan and Ravikumar (2001) extend Marcet and Marimon (1992) to impose a period-by-period feasibility constraint and endogenous growth. In particular, they examine the impact of incomplete risk sharing on growth and welfare in the context of the AK model. The source of market incompleteness is private information since household technologies are subject to idiosyncratic productivity shocks not observable by others. Risk sharing between households occurs through contracts with intermediaries. This sort of incomplete risk sharing tends to reduce the rate of growth relative to the complete risk-sharing benchmark. However, "numerical examples indicate that, on average, the growth and welfare effects on incomplete risk sharing are likely to be small." One key simplification in this case is that the allocation solved is not necessarily the best incentive-compatible allocation.

Recently, Greenwood, Sanchez, and Wang (2010a) embedded the costly state verification framework into the standard growth model.[3] The relationship between the firm and lender is modeled as a static contract. In the economy in which information is too costly, undeserving firms are overfinanced and deserving ones are underfunded. A reduction in the cost of information leads to more capital accumulation and a redirection of funds away from unproductive firms toward productive ones. Greenwood, Sanchez, and Wang (2010b) show that this mechanism has quantitative significance to explain cross-country differences in capital-to-income ratios and total factor productivity.

Other studies use similar models for other purposes. Espino (2005) studies a neoclassical growth model that includes a discrete number of agents, like the one presented in this article. However, he uses the economy with private information about the preference shock to analyze the validity of Ramsey's conjecture about the long-run allocation of an economy in which agents are heterogeneous in their discount factor. Clementi, Cooley, and Giannatale (2010) study a repeated bilateral exchange model with hidden action, along the lines of Spear and Srivastava (1987) and Wang (1997), that includes capital accumulation. The two agents in the economy are a risk-neutral investor and a risk-averse entrepreneur. They show that the incentive scheme chosen by the investor provides a rationale for firm decline.

This article is organized as follows: Section 1 presents the physical environment and the planner's problem, and derives the optimal allocation. Section

---

[3] See also Khan (2001) and Chakraborty and Lahiri (2007).

2 describes the calibration and the numerical solution of the full information allocation. Section 3 studies in which cases the full information allocation would be incentive compatible in an economy with private information. Section 4 offers concluding remarks.

1. MODEL

Environment

There is a constant returns to scale aggregate technology to produce the unique consumption good that is represented by a standard neoclassical production function, $F(K, L)$, where K is the current stock of capital and L denotes units of labor. There are two agents in the economy, $h = 1, 2$. Agent h is endowed with one unit of time each period and does not value leisure, i.e., the time endowment is supplied inelastically in the labor market. The initial stock of capital at date 0 is denoted by $K_0 > 0$. Capital depreciates at the rate $\delta \in (0, 1)$.

At the beginning of date t, agent 1 faces an idiosyncratic preference shock $s_t \in S_t = \{s_L, s_H\}$, where $s_H > s_L$. This shock is assumed to be i.i.d. across time, where $\pi_i > 0$ is the probability of s_i, $i = L, H$. Notice that s_t is also the aggregate preference shock at date t. The aggregate history of shocks from date 0 to date t, denoted $s^t = (s_0, ..., s_t)$, has probability at date 0 given by $\pi(s^t) = \pi(s_0)...\pi(s_t)$.

Given a consumption plan $\{c_{1,t}\}_{t=0}^{\infty}$ such that $c_{1,t} : S^t \to \mathbb{R}_+$, agent 1's state-dependent preferences are represented by

$$\mathcal{U}_1(c_1) = E \left\{ \sum_{t=0}^{\infty} \beta^t \, u_1(s_{1,t}, c_{1,t}) \right\} = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi(s^t) \, u_1(s_t, c_1(s^t)),$$

where $u_1 : \mathbb{R}_+ \to \mathbb{R}$ is strictly increasing, strictly concave, and twice differentiable, $\lim_{c_t \to 0} u'(c_t) = +\infty$, and $\beta \in (0, 1)$. Similarly, given $\{c_{2,t}\}_{t=0}^{\infty}$, agent 2's preferences are represented by

$$\mathcal{U}_2(c_2) = E \left\{ \sum_{t=0}^{\infty} \beta^t u_2(c_{2,t}) \right\} = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi(s^t) u_2(c_2(s^t)).$$

Planner's Problem

Consider the problem of a fictitious planner choosing the best feasible allocation. Let $K' = \{K_{t+1}\}_{t=0}^{\infty}$ be an *investment plan* that every period allocates next period's capital for all t. Similarly, let $C = \{C_t\}_{t=0}^{\infty}$ be a *consumption plan* where $C_t = (c_{1t}, c_{2t})$. Given K_0, a sequential allocation (C, K') is feasible if, for all s^t,

$$K_t(s^t) + c_{1t}(s^t) + c_{2t}(s^t) \leq F(K_t(s^{t-1}), 1) + (1 - \delta)K_t(s^{t-1}).$$

We will assume throughout the article that the production function F is Cobb-Douglas with exponent γ.

The Pareto-optimal allocation in this economy is a feasible allocation such that there is no other feasible allocation that provides all the agents the same or more lifetime utility. One reason to be interested in these allocations is that, under certain conditions, they are equivalent to competitive equilibrium allocations. Under our assumptions, Pareto-optimal allocations can be obtained by solving the following problem:

$$\max_{(C,K')} \alpha \mathcal{U}_1(c_1) + (1-\alpha)\mathcal{U}_2(c_2)$$

subject to

$$K(s^t) + c_1(s^t) + c_2(s^t) \leq F(K(s^{t-1}), 1) + (1-\delta)K(s^{t-1}), \forall s^t,$$

where K_0 is given and $\alpha \in [0, 1]$ is the weight that the planner assigned to agent 1—referred to hereafter as Pareto weight. Notice that different values of α characterize different points in the Pareto frontier. Later, we will consider a different allocation varying the value of α.

To characterize the problem further, it is simpler to consider the methods developed by Lucas and Stokey (1984) to solve for Pareto-optimal allocations in growing economies populated with many consumers. It is actually simple to adapt their method to analyze this economy. The idea is to make next period welfare weights conditional on the current shock.[4]

The planner's recursive problem is a fixed point, V, of the function equation

$$V(k, \overline{\alpha}) = \max_{c,k',w} \alpha \{\pi_L [u_1(s_L, c_{1L}) + \beta w_{1L}] + \pi_H [u_1(s_H, c_{1H}) + \beta w_{1H}]\} +$$
$$(1-\alpha) \{\pi_L [u_2(c_{2L}) + \beta w_{2L}] + \pi_H [u_2(c_{2H}) + \beta w_{2H}]\} \quad (1)$$

subject to

$$f(k) + (1-\delta)k \geq k'_L + c_{1L} + c_{2L}, \quad (2)$$
$$f(k) + (1-\delta)k \geq k'_H + c_{1H} + c_{2H}, \quad (3)$$
$$\min_{\overline{\alpha}'_L} V(k'_L, \overline{\alpha}'_L) - \alpha'_L w_{1L} - (1-\alpha'_L)w_{2L} \geq 0, \quad (4)$$
$$\min_{\overline{\alpha}'_H} V(k'_H, \overline{\alpha}'_H) - \alpha'_H w_{1H} - (1-\alpha'_H)w_{2H} \geq 0, \quad (5)$$

where $\overline{\alpha} = \{\alpha, 1-\alpha\}$ and w are the from-tomorrow-on utilities. The idea in (1)–(5) is to represent the problem of choosing an optimal allocation for a given stock of capital k and a vector of Pareto weights $(\alpha, 1-\alpha)$ as one of choosing a feasible current period allocation of consumption $c = \{c_{1L}, c_{1H}, c_{2L}, c_{2H}\}$ and capital goods $k' = \{k'_L, k'_H\}$, and a vector of from-tomorrow-on utilities $w =$

---

[4] See Beker and Espino (2011) for a discussion about the implementation and the corresponding technical details.

$\{w_{1L}, w_{1H}, w_{2L}, w_{2H}\}$, subject to the constraint that these utilities be attainable given the capital accumulation decision, as guaranteed by constraints (4)–(5). As in Lucas and Stokey (1984), the weights $\{\overline{\alpha}'_L, \overline{\alpha}'_H\}$ that attain the minimum in (4) and (5) will be the new weights used in selecting tomorrow's allocation, and so on, ad infinitum.

Characterization

Assume preferences are represented by

$$u_1(s, c) = s \frac{c^{1-\sigma}}{1 - \sigma} \text{ and } u_2(c) = \frac{c^{1-\sigma}}{1 - \sigma}.$$

The first-order conditions (FOC) for consumption are

$$\begin{aligned}
\alpha \pi_L s_L (c_{1L})^{-\sigma} &= \lambda_L, \\
\alpha \pi_H s_H (c_{1H})^{-\sigma} &= \lambda_H, \\
(1 - \alpha)\pi_L (c_{2L})^{-\sigma} &= \lambda_L, \\
(1 - \alpha)\pi_H (c_{2H})^{-\sigma} &= \lambda_H,
\end{aligned}$$

where λ_i is the Lagrange multiplier in the resource constraints in state $i = L, H$. From these equations it is simple to obtain that the consumption of each agent will be a share of the aggregate consumption, $\mathbf{C}_i$,

$$\begin{aligned}
c_{1L} &= \frac{(\alpha s_L)^{1/\sigma}}{(\alpha s_L)^{1/\sigma} + (1 - \alpha)^{1/\sigma}} \mathbf{C}_L, \\
c_{1H} &= \frac{(\alpha s_H)^{1/\sigma}}{(\alpha s_H)^{1/\sigma} + (1 - \alpha)^{1/\sigma}} \mathbf{C}_H, \\
c_{2L} &= \frac{(1 - \alpha)^{1/\sigma}}{(\alpha s_L)^{1/\sigma} + (1 - \alpha)^{1/\sigma}} \mathbf{C}_L, \\
c_{2H} &= \frac{(1 - \alpha)^{1/\sigma}}{(\alpha s_H)^{1/\sigma} + (1 - \alpha)^{1/\sigma}} \mathbf{C}_H.
\end{aligned} \tag{6}$$

The FOC with respect to w are

$$\begin{aligned}
\alpha \pi_L \beta &= \mu_L \alpha'_L, \\
\alpha \pi_H \beta &= \mu_H \alpha'_H, \\
(1 - \alpha)\pi_L \beta &= \mu_L (1 - \alpha'_L), \\
(1 - \alpha)\pi_H \beta &= \mu_H (1 - \alpha'_H).
\end{aligned}$$

These imply that

$$\alpha \pi_L \beta + (1 - \alpha)\pi_L \beta = \mu_L \alpha'_L + \mu_L (1 - \alpha'_L),$$

and therefore $\pi_L \beta = \mu_L$ and $\pi_H \beta = \mu_H$. Using the FOC with respect to w again, these two conditions imply $\alpha = \alpha'_L = \alpha'_H$. Thus, the Pareto weights will be constant in this problem.

Using the fact that individual consumption is a share of aggregate consumption and that Pareto weights are constant, this problem can be rewritten as one solving for the consumption (or capital accumulation) of a representative consumer with aggregate preference shocks. In that case, the state-dependent utility of the representative consumer, u_R, would be

$$u_R(s, \mathbf{C}) = \left((s\alpha)^{1/\sigma} + (1-\alpha)^{1/\sigma} \right)^\sigma \frac{\mathbf{C}^{1-\sigma}}{1-\sigma}.$$

Notice here that the level of the shock depends not just on the size of s, but also on α. This representation is useful to understand that the optimal investment decision is affected by the realization of the preference shock and the distributional parameter α. When s is larger, the representative agent prefers to increase consumption today and decrease investment. Given the same shock, the size of the drop in investment depends on the Pareto weight of the agent that received the shock.

The FOC with respect to capital accumulation are

$$\lambda_L = \mu_L \frac{\partial V(k'_L, \overline{\alpha}'_L)}{\partial k'_L},$$

$$\lambda_H = \mu_H \frac{\partial V(k'_H, \overline{\alpha}'_H)}{\partial k'_H}.$$

An application of the envelope conditions makes these conditions imply the standard Euler equations determining capital accumulation,

$$1 = (F'(k'_L) + (1-\delta))\beta \frac{\left(\pi_L s_L (c'_{1L})^{-\sigma} + \pi_H s_H (c'_{1H})^{-\sigma} \right)}{s_L (c_{1L})^{-\sigma}},$$

$$1 = (F'(k'_H) + (1-\delta))\beta \frac{\left(\pi_L s_L (c'_{1L})^{-\sigma} + \pi_H s_H (c'_{1H})^{-\sigma} \right)}{s_H (c_{1H})^{-\sigma}}.$$

2. NUMERICAL SOLUTION

This model can be solved in the computer once the values of the parameters are determined. Most of the parameters are standard in the neoclassical growth model and take standard values. Others, such as the size of the preference shock and the probability of occurrence, were chosen only to illustrate the behavior of the model. In particular, a high preference shock happens on average every 6.7 years, but it is large enough to demand a significant amount of resources. Think, for example, that a country in an economic union requires help or assistance on average every 6.7 years. Table 1 presents the values for all the parameters of the model.

The right-hand side of (1)–(5) defines a contraction. The computation is based on value function iteration as follows. Guess a function V. Then solve for $\max_{c, w', k'}$ using V, the FOC described above, and numerical maximization.

Table 1 Parameter Values

	Parameter	Value
γ	Exponent of capital in production function	0.30
δ	Depreciation rate of capital	0.07
β	Discount factor	0.97
σ	Relative risk aversion	0.50
s_L	Low value of the preference shock	0.95
s_H	High value of the preference shock	1.05 and 2.00
π_L	Probability of low value of the preference shock	0.85
π_H	Probability of high value of the preference shock	0.15

With this solution, construct a new function V' and restart the maximization unless V' is sufficiently close to V. Now we discuss the results using the parameters in Table 1 with $s_H = 2$ and Pareto weights $\{0.75, 0.25\}$. Figure 1 presents time series for aggregate consumption and capital accumulation in the steady state of this economy. On the top panel that aggregate consumption jumps after a preference shock and then returns slowly to a relatively constant value until a new shock hits. As a consequence, capital accumulation drops after a high preference shock to accommodate larger aggregate consumption, as shown on the top panel. The effect of this change on the incentives to misreport a shock—if it would be unobservable—is discussed in the next section. The distribution of consumption among agents is determined by equations (6), i.e., agent 1's share of aggregate consumption increases with the value of the shock. More on this later.

Figure 2 depicts the stationary distribution of the main variables for the same example analyzed in Figure 1. The top left panel shows that 15 percent of the time there is a large preference shock equal to 2 and most of the time (85 percent) a low shock equal to 0.95. The top right panel presents the stationary distribution of capital. It is somehow surprising that very different values (e.g., 3 and 6) are reached with positive probability. Most of its mass is accumulated on the higher values, however. Those correspond to periods with low preference shocks. The lowest values of capital correspond to periods of several consecutive high preference shocks. Something similar happens with c_2, on the bottom right panel. A priori, these properties could have been expected since k' and c_2 are the two sources to finance transfers to agent 1 after a high preference shock. The distribution of c_1, presented on the bottom left panel, has most of the mass around lower values and some mass at higher values. The highest values correspond to a high preference shock hitting the economy after a long period of low shocks.

Figure 1 Consumption and Capital Paths in the Stationary Distribution
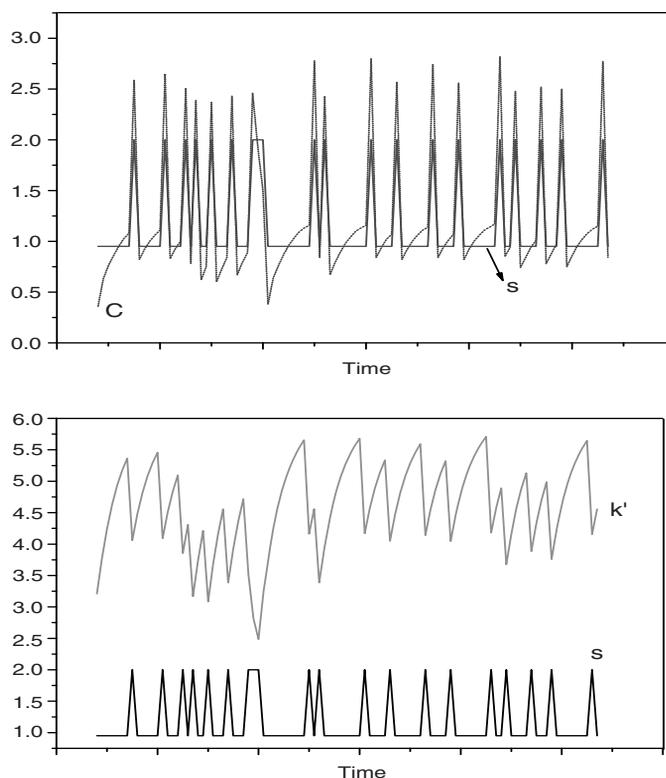
Notes: These histograms were computed from time series data of these variables for 5,000 periods after deleting the first 500 realizations.

3.   THE ROLE OF INFORMATION

This section investigates the incentives to misreport preference shocks by agent 1 whenever the full information allocation described above is the target to be implemented. To do so, consider the value of the following (implicit) incentive compatibility constraints:

$$icc_{HL} \;=\; s_H u\,(c_{1H}) + \beta w_{1H} - [s_H u\,(c_{1L}) + \beta w_{1L}], \qquad (7)$$

$$icc_{LH} \;=\; s_L u\,(c_{1L}) + \beta w_{1L} - [s_L u\,(c_{1H}) + \beta w_{1H}]. \qquad (8)$$

The interpretation of these variables is very important for the analysis hereafter. If the variable icc_{HL} is positive, it means that when the state H realizes, agent 1 would prefer truthfully reporting a high preference shock and obtaining $\{c_{1H}, w_{1H}\}$ instead of misreporting it and receiving $\{c_{1L}, w_{1L}\}$. Similarly, a

Figure 2 Stationary Distribution, Main Variables
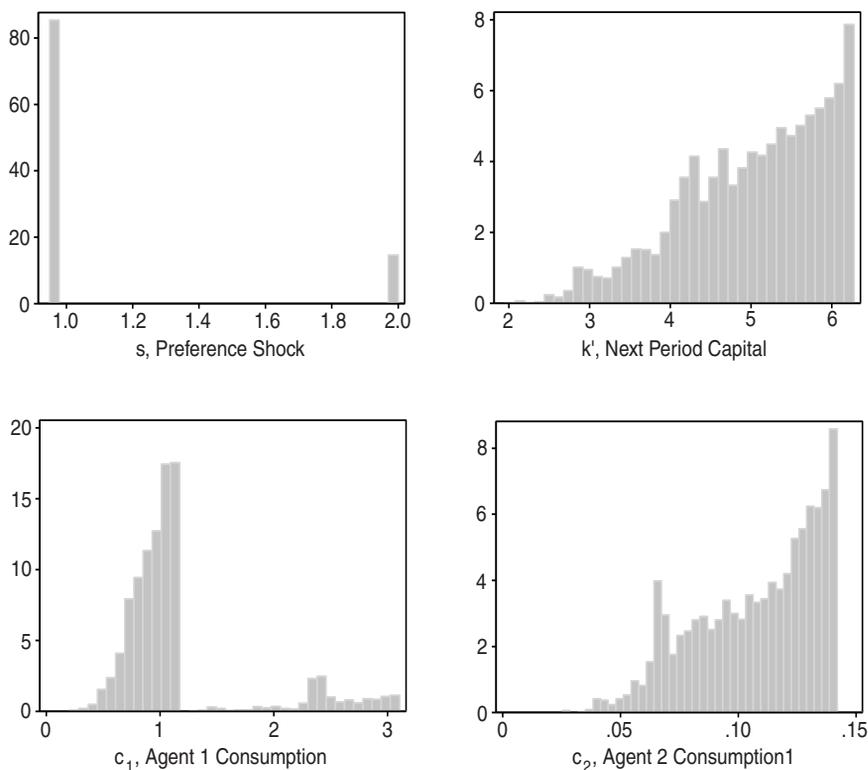
Notes: These histograms were computed from time series data of these variables for 5,000 periods after deleting the first 500 realizations.

negative value of icc_{LH} means that agent 1 would prefer misreporting a high preference shock and obtaining $\{c_{1H}, w_{1H}\}$ to truthfully reporting a low shock and receiving $\{c_{1L}, w_{1L}\}$.

Since $c_{1H} > c_{1L}$, one may expect that there is no incentive to report the low shock when the high shock was actually realized, i.e., a positive value of icc_{HL}. This is actually what happens in the stationary distribution, as shown on the top panel of Figure 3. In contrast, agent 1 may be tempted to misreport a high preference shock to obtain higher consumption. Remember that this would imply that $icc_{LH} < 0$. This does not need to always be the case, however. Since k' is lower after a high preference shock, agent 1's prospects worsen after a high preference shock. Thus, it will be a race between more consumption today, $c_{1H} > c_{1L}$, and less future consumption, $w_{1L} > w_{1H}$. The results for

Figure 3 Incentive Compatibility in the Stationary Distribution
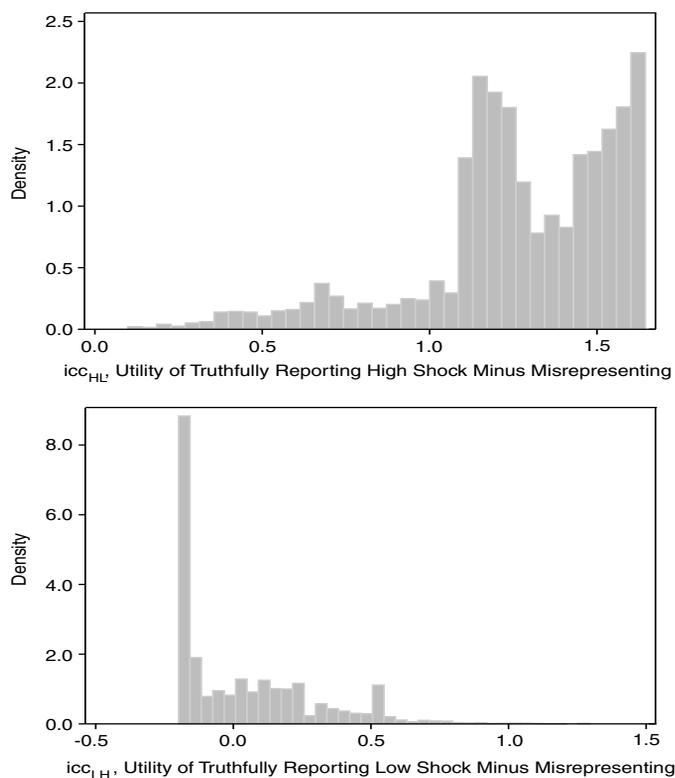
Notes: These histograms were computed from time series data of these variables for 5,000 periods after deleting the first 500 realizations.

the example described above are presented in the bottom panel of Figure 3. There, icc_{LH} is negative more than 80 percent of the time but positive in some instances. This means that in all such instances, the drop in from-tomorrow-on utilities caused by reporting a high preference shock is enough to compensate for the difference in current consumption. What determines whether icc_{LH} is negative or positive will be studied next by analyzing different examples.

The next two examples capture the role of the size of redistribution versus disinvestment. The first example is presented in Figure 4. This is the same example in all the previous figures, but the difference is that the Pareto weight of agent 1 is only 0.33 (instead of 0.75) and the weight of agent 2 is 0.67. This implies that agent 2's consumption is larger, as shown in the top left panel. The top right panel presents the behavior of capital accumulation. Notice that

Figure 4 Paths with Large Redistribution of Aggregate Consumption
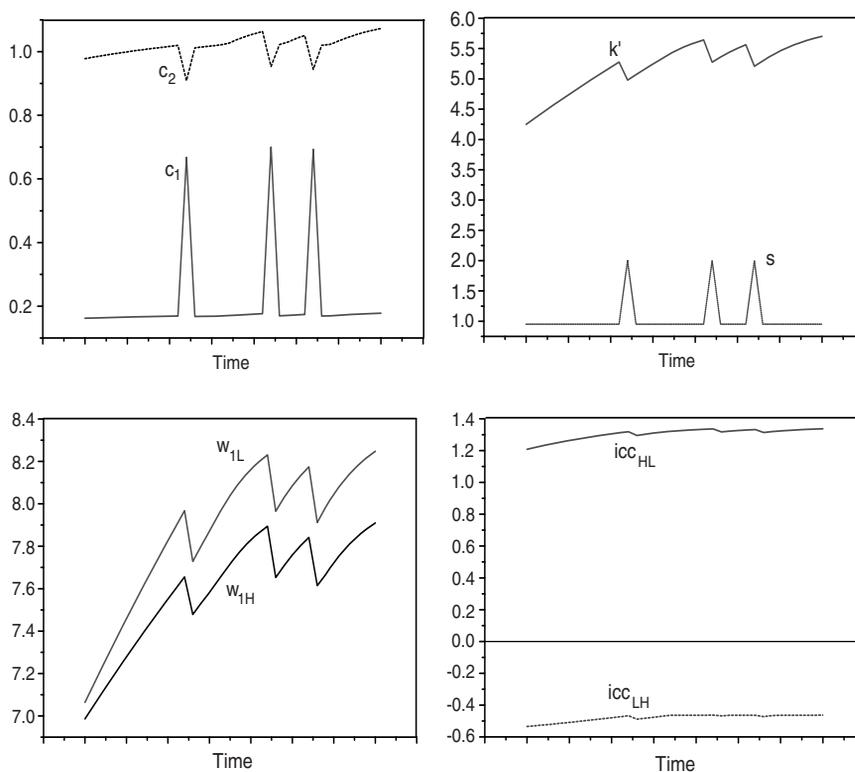
Notes: In this economy, the weights on agents 1 and 2 are 0.33 and 0.67, respectively. The time series data in all four graphs correspond to the initial 35 periods after the economy is started with a stock of capital smaller than the steady-state level.
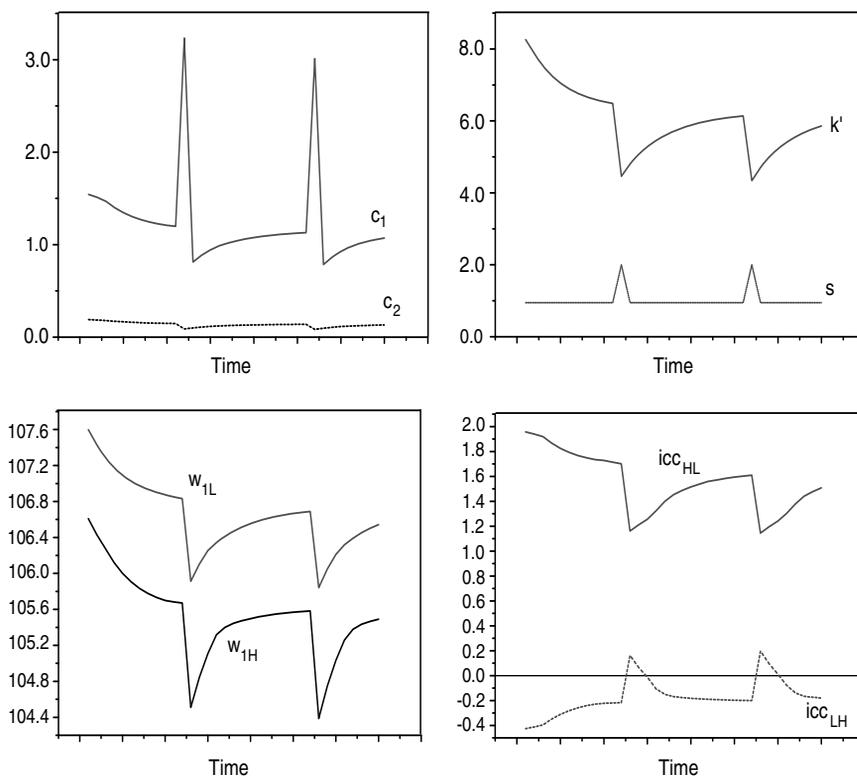
the time series in the graphs correspond to the transition toward a higher level of capital. From these two figures it is clear that a nontrivial part of the rise in agent 1's consumption after a high preference shock comes from redistribution of consumption across agents. As a consequence, the promised utilities from next period on are not that different after a report of high or low preference shock, as shown in the bottom left panel. In turn, this implies that icc_{LH} is always negative, as presented in the right bottom panel. Thus, this is an example in which the full information allocation would not be implementable under private information: After a low preference shock, agent 1 would prefer to falsely report a high preference shock.

Figure 5 Paths with Large Variation in Investment
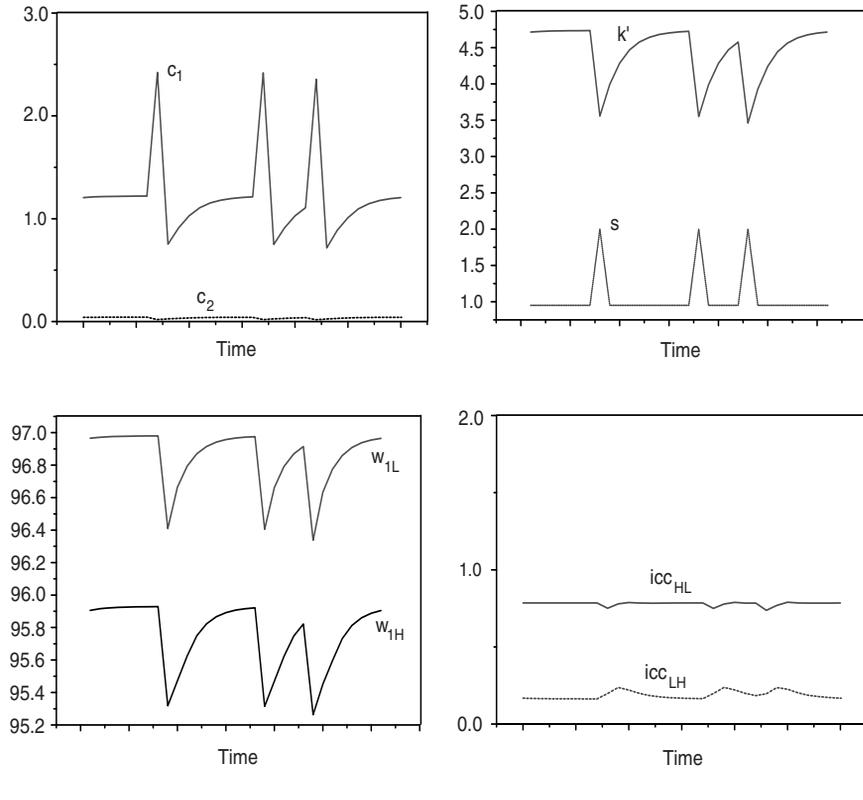
Notes: In this economy, the weights on agents 1 and 2 are 0.85 and 0.15, respectively. The time series artificial data in all four graphs correspond to 30 periods created after the steady-state level of capital is reached.

Now consider the example presented in Figure 5. Here, the behavior of the same series is presented for an economy in which the Pareto weight of agent 1 is 0.85 and the steady-state distribution of capital is reached. This implies that agent 1's consumption is much larger than that of agent 2, as shown in the top left panel. As a consequence, capital accumulation must vary significantly to provide more consumption to agent 1 after the realization of a high preference shock. This is shown in the top right panel. Therefore, as presented in the bottom left panel, the difference in from-tomorrow-on utilities associated with low and high preference shocks is large. Thus, both incentive compatibility constraints are positive in the stationary distribution

Figure 6 Incentive Compatibility and Capital Accumulation

Notes: In this economy, the weights on agents 1 and 2 are 0.75 and 0.25, respectively. The time series data in all four graphs correspond to the initial 35 periods after the economy is started with a stock of capital larger than the steady-state level.

of this economy (see bottom right panel), and the full information allocation would be implementable under private information.

The previous two examples are useful to understand that the relative importance of the agent who privately observes the shock matters for the role of private information. When this agent is more important, her share of aggregate consumption is larger, and the rise of that agent's consumption after a shock comes mainly from disinvestment. This makes misreporting a high preference shock too costly in terms of her own future consumption, and hence the full information allocation is implementable under private information. Thus, the *size* of disinvestment, determined by the importance of the agent with the preference shock, matters for the provision of incentives under private

information. This *suggests* that in a fully specified model with private information, the planner would like to increase the Pareto weight of the agent with private information to reduce the incidence of this friction.

The next example illustrates the role of the outlook for economic growth at the time of disinvestment in preventing misrepresentation of preference shocks. First, consider the example in Figure 6. It displays the transition to the steady state from a larger stock of capital. The weights of agents 1 and 2 are 0.75 and 0.25, respectively. Initially, consumption, capital, and from-tomorrow-on utilities decrease. During this initial phase, while capital is large and decreasing, icc_{LH} is negative and increasing. This means that when there is extra capital in the economy, as compared to the stationary distribution, the optimal drop in capital that a high preference shock would require (and its corresponding drop in promised utility) is not large enough to provide incentives to make the report of that shock incentive compatible. Eventually, a high preference shock hits the economy, the consumption of agent 1 jumps, and capital drops significantly. Now, the economy is expected to grow in the coming years, which implies that another high preference shock would hurt both agents more. Therefore, reporting a high preference shock becomes incentive compatible for a few years, until the stock of capital reaches a higher level. The same story occurs again in a few years, when a high preference shock hits the economy again. Thus, this example illustrates the interaction of growth and information. Misrepresentation of preference shocks is more costly if the economy is expected to grow. This finding suggests that a planner solving for the best incentive-compatible allocation would delay growth to facilitate the provision of incentives.

The last example confirms the importance of the size of disinvestment and the outlook for economic growth. Consider the time series artificial data presented in Figure 7. The Pareto weight for agent 1 is larger than in previous examples, 0.85, but the value of the high preference shock is smaller, $s_H = 1.05$. First, notice that this example confirms the result in the previous figure: It is easier to provide incentives (icc_{LH} is larger) when the economy is expected to grow. However, in this case, icc_{LH} is never greater than zero. Notice that this happens despite agent 1's weight being larger than in all other examples. The key difference is that the shock is not that large. Thus, the size of the drop in capital accumulation is not very relevant, and therefore the difference between w_{1L} and w_{2L} is small.

4. CONCLUSIONS

This article studies the interaction between growth and risk sharing. First, it answers how investment is affected by insurance needs. A stochastic growth model with two agents and preference shocks is used to answer this question. Only one of the agents (or groups, regions, countries) is affected by this shock,

Figure 7 Paths for the Model with Small Shocks
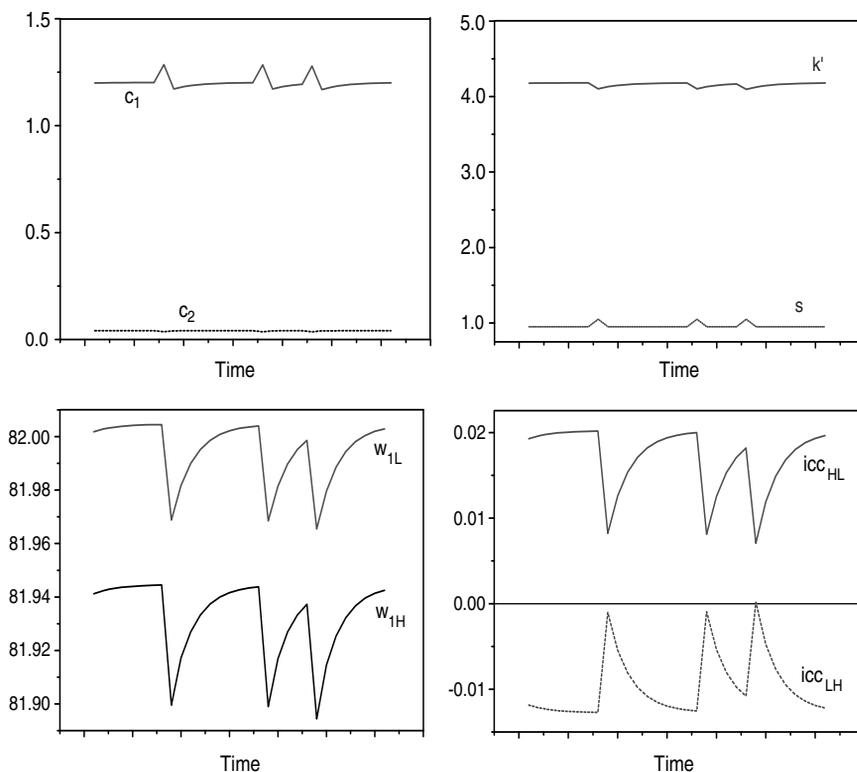
Notes: In this economy, the weights on agents 1 and 2 are 0.85 and 0.15, respectively. The time series artificial data in all four graphs correspond to 30 periods created after the steady-state level of capital is reached.

which basically increases the need of consumption for this agent. When both agents are risk-averse, the socially optimal response to this shock requires both decreasing the consumption of other agents and decreasing capital accumulation. Thus, the occurrence of this shock slows down the convergence toward the stationary distribution of capital.

Then, we analyze if the best path of capital accumulation and consumption allocation is implementable if needs are privately observed by the agents. That is, if the shocks are privately observed by individuals, do they have incentive to misrepresent? The value of the incentive compatibility constraints implied by the full information allocation is used to answer this question. Because investment drops when an agent reports a high preference shock, the prospects

of all agents deteriorate after such a report. This may be enough to prevent misreporting. The size of disinvestment after the report of a high preference shock and the outlook for economic growth at the time of disinvestment are important to induce individuals to report a low realization of the preference shock truthfully. This analysis *suggests* that in a fully specified model with private information, the best incentive compatible allocation would tend to hurt growth, by decreasing investment, and increase inequality, by augmenting the share of consumption of the agent with private information. Of course, this is only a conjecture. Solving for the constrained-efficient allocation in this environment is necessary to verify the validity of this conjecture. This is left for future research.

REFERENCES

Atkeson, Andrew, and Robert E. Lucas, Jr. 1992. "On Efficient Distribution with Private Information." *Review of Economics Studies* 59 (July): 427–53.

Baxter, Marianne, and Robert G. King. 1991. "Productive Externalities and Business Cycles." Institute for Empirical Macroeconomics at Federal Reserve Bank of Minneapolis Discussion Paper 53.

Beker, Pablo F., and Emilio Espino. 2011. "The Dynamics of Efficient Asset Trading with Heterogeneous Beliefs." *Journal of Economic Theory* 146 (January): 189–229.

Caselli, Francesco, and Jaume Ventura. 2000. "A Representative Consumer Theory of Distribution." *American Economic Review* 90 (September): 906–26.

Chakraborty, Shankha, and Amartya Lahiri. 2007. "Costly Intermediation and the Poverty of Nations." *International Economic Review* 48 (1): 155–83.

Chatterjee, Satyajit. 1994. "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model." *Journal of Public Economics* 54 (May): 97–119.

Clementi, Gian Luca, Thomas Cooley, and Soni Di Giannatale. 2010. "A Theory of Firm Decline." *Review of Economic Dynamics* 13 (October): 861–85.

Espino, Emilio. 2005. "On Ramsey's Conjecture: Efficient Allocations in the Neoclassical Growth Model with Private Information." *Journal of Economic Theory* 121 (April): 192–213.

Green, Edward J. 1987. "Lending and the Smoothing of Uninsurable Income." In *Contractual Arrangements for Intertemporal Trade*, edited by Edward C. Prescott and Neil Wallace. Minneapolis: University of Minnesota Press, 3–25.

Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2010a. "Financing Development: The Role of Information Costs." *American Economic Review* 100 (September): 1,875–91.

Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2010b. "Quantifying the Impact of Financial Development on Economic Development." Economie d'Avant Garde Research Report 17.

Hall, Robert E. 1997. "Macroeconomic Fluctuations and the Allocation of Time." *Journal of Labor Economics* 15 (January): S223–50.

Khan, Aubhik. 2001. "Financial Development and Economic Growth." *Macroeconomic Dynamics* 5 (June): 413–33.

Khan, Aubhik, and B. Ravikumar. 2001. "Growth and Risk-Sharing with Private Information." *Journal of Monetary Economics* 47 (June): 499–521.

Lucas, Robert Jr., and Nancy L. Stokey. 1984. "Optimal Growth with Many Consumers." *Journal of Economic Theory* 32 (February): 139–71.

Marcet, Albert, and Ramon Marimon. 1992. "Communication, Commitment, and Growth." *Journal of Economic Theory* 58 (December): 219–49.

Sorger, Gerhard. 2002. "On the Long-Run Distribution of Capital in the Ramsey Model." *Journal of Economic Theory* 105 (July): 226–43.

Spear, Stephen E., and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies* 54 (October): 599–617.

Thomas, Jonathan, and Tim Worrall. 1990. "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem." *Journal of Economic Theory* 51 (August): 367–90.

Wang, Cheng. 1997. "Incentives, CEO Compensation, and Shareholder Wealth in a Dynamic Agency Model." *Journal of Economic Theory* 76 (September): 72–105.

Wen, Yi. 2006. "Demand Shocks and Economic Fluctuations." *Economics Letters* 90 (March): 378–83.