Economic Theory and the World of Practice: A Celebration of the ($S, s$) Model

Andrew Caplin and John Leahy

The ($S, s$) model represents in ideal form the interplay between economic theory and the world of application. It was created during a period in which microeconomic theorists and applied mathematicians forged many of the tools that dominate economic analysis to this day, including linear programming, dynamic programming, and game theory. These conceptual advances resulted in large part from the desire to answer important practical questions. One such question of interest in both business and military circles concerned inventory control. For any stored good in more-or-less constant use, there is a need to keep stock on hand. It was the question of how best to balance the costs of ordering and of running out of stock against the costs of holding excess inventory that inspired Arrow, Harris, and Marschak (1951) to introduce the ($S, s$) model. In this celebratory article, we show how this model not only answered important practical questions, but also opened the door to a quite startling range of important and challenging follow-up questions, many of great practical importance and analytic depth. The ($S, s$) model has become one of the touchstone models of economics, opening new vistas of applied economic theory to all who internalize its structure.

We open the paper by providing a general overview of the essential features of the ($S, s$) model, concentrating on the original inventory theoretic setting. On one level, Arrow, Harris, and Marschak (1951) departed in only minor respects from existing models of inventory control, introducing uncertainty and discounting into inventory models with ordering costs. As the follow-up volume of Arrow, Karlin, and Scarf (1958) attests, this basic modeling theme can support a wide

Andrew Caplin and John Leahy are both Professors of Economics, New York University, New York City, New York. Their e-mail addresses are {andrew.caplin@nyu.edu} and {john.leahy@nyu.edu}, respectively.

doi=10.1257/jep.24.1.183
range of variations, allowing an entire field of inventory modeling to develop that captures richer features of reality. Today, the \((S, s)\) model is firmly integrated into the business software with which production and storage processes are run worldwide (Knolmayer, Mertens, and Zeier, 2002; Zipkin, 2000).

The common thread in all \((S, s)\) models is the interaction between fixed costs of adjustment and uncertainty. Both factors are pervasive in economic decision making. Economic agents do not always react immediately or continually to changes in their environment. Consumers buy new houses and cars only on rare occasions. Firms leave prices fixed for months even though information is arriving at a much greater frequency. Fixed costs of adjustment provide a natural explanation for all such forms of inertial behavior. If an agent faces a fixed cost of taking some action and if the loss to nonadjustment is small, then it will pay to leave things be until the benefits of adjustment exceed the costs. Thus, in the second section of the paper, we highlight the role that the \((S, s)\) model plays as a general purpose model of economic decision making in situations with three defining features: 1) a state variable that affects flow payoffs, 2) fixed costs of exerting control over this state variable, and 3) a driving force that causes the state to drift absent control.

We then turn to some of the most exciting recent work on \((S, s)\) modeling, which has concerned aggregation. By definition, macroeconomic fluctuations are intimately linked with microeconomic adjustments to a changing environment. In practice, these microeconomic adjustments are often characterized by long periods of inaction followed by discontinuous jumps. Much macroeconomic analysis is concerned with understanding how these behaviors aggregate up. Up through the early 1980s, the traditional method of modeling such effects involved using “time-dependent” models based on the simplifying assumption that the time between successive changes was independent of the evolution of the economy. Most famous in this regard are the models of monetary nonneutrality based on microeconomic stickiness of nominal prices due to Fischer (1977), Taylor (1980), and Calvo (1983). A student of the \((S, s)\) model is inevitably led to wonder whether or not the resulting stickiness of the price level (and the resulting real effects of monetary policy) are artifacts of this modeling technology. Wouldn’t a higher rate of overall inflation lead firms to change prices more frequently? More broadly, the \((S, s)\) model forces to mind the question of how state-dependent microeconomic inertia can affect macroeconomic outcomes. Does the discrete nature of the decision to sell an owned car and replace it with a new one give rise to fluctuations in pent-up demand, or does demand aggregate up more smoothly? What role does microeconomic inertia play in times of crisis?

The most robust finding to date, and arguably the most disappointing, is an example due to Caplin and Spulber (1987) of how discrete individual adjustments often lead to smooth aggregate adjustment, at least as a reasonable first approximation. Yet there are caveats, with the most intriguing of these relating to information dynamics. In periods of individual inaction, the \((S, s)\) policy not only stores up potential for future action, but also impedes the transmission of information. As a topical example, mortgage default is a costly decision at the individual level.
What combination of falling house prices and policy changes would it take to trigger a large new wave of default? How should mortgage modification strategies be designed given the difficulty of observing yet to be crossed private thresholds? To adapt the immortal phrase of the mathematician David Hilbert, economists must know the answer to these questions, and we will know them. But we do not know them yet.

Our final section looks forward to next steps in \((S, s)\) modeling, particularly those relating to household finance. Such questions as when and whether and how to terminate a mortgage, and when to buy and sell assets (including equities and real estate) are characterized by uncertainty and costs of adjustment. It is important to develop \((S, s)\) models of these financial decisions that mirror reality sufficiently closely to allow useful empirical estimation of model parameters. These same models may also serve as decision aids for households facing the corresponding situations in everyday life, decisions for which most of us are woefully underprepared.

The potential for the \((S, s)\) model to assist households in making financial decisions brings us full circle back to the model’s origins. Perhaps the most profound form of model validation involves practitioners skilled in the relevant art incorporating its prescriptions into their decisions. The \((S, s)\) model is exemplary in this respect. To this day, it holds place of honor at the center of inventory theory and ordering policy. If “household operations research” blossoms as we anticipate, in the future we will all rely on variants of the \((S, s)\) model to help us navigate our increasingly complex financial landscape. In turn, such developments will introduce new modeling challenges, continuing the fruitful interaction between economic theory and the world of practice that the \((S, s)\) model epitomizes.

The \((S, s)\) Model and Inventories

Origins of the \((S, s)\) Model

The origins of the \((S, s)\) model lie in efforts to formalize the problem of inventory control dating back to the early twentieth century (Nahmias, 2006). Three costs associated with most practical inventory problems are the holding costs of inventory, costs associated with running short of inventory, and costs associated with receiving and processing orders. The basic question of inventory theory is how best to balance these costs. Rules of thumb for accomplishing this goal were proposed early in the twentieth century (see Whitin, 1953, for a survey). Many years after being introduced in the inventory arena, the resulting “lot size formulae” made their way into popular economic application in the form of the “square root rule of money demand” due to Allais (1947), Baumol (1952), and Tobin (1956).

This money demand model nicely illustrates some key forces at work in inventory problems. The question is how much cash an individual should keep on hand when the alternative is to earn interest that compounds continuously at rate \(r\) per period. Consider the basic case of constant spending, at flow rate of \(y\) dollars per period. In light of the predictability and constancy of spending, cost minimization...
in this continuous-time formulation involves letting cash run down to zero before replenishing to some level \( Q > 0 \). The resulting optimization problem balances the fixed “shoe leather” costs, \( K \), of trips to the bank, against foregone interest. Given that average cash inventory level is \( Q/2 \), the average-level interest foregone can be well approximated by \( Qy/2 \), with the fixed costs being paid every \( y/Q \) periods. Minimizing the sum of these two costs results in a condition expressing the optimal quantity of money to take out as a function of the flow rate of \( y \) dollars per period, the available interest rate \( r \), and the shoe-leather costs \( K \):

\[
Q^* = (2yK/r)^{0.5}.
\]

In the monetary context, the novelty in this “square root rule” was the sharp prediction it suggests for the income and interest elasticity of the demand for money.

**What is an \((S, s)\) Policy?**

The square root rule illustrates the state of the art in inventory theory in the late 1940s. However, the advent of dynamic programming methods revolutionized the field of inventory control, with Arrow, Harris, and Marschak (1951) pulling the ingredients together in definitive style. Three obvious gaps in the earlier models are: that the rate of drawdown of money and other inventories is not completely smooth or predictable; that most inventories cannot be delivered in an instant; and that the appropriate formula for minimizing costs must take account of discounting. Arrow, Harris, and Marschak showed how to allow for all three forces by reformulating the problem of inventory control with fixed ordering costs in the modern language of stochastic dynamic programming. The policies that occurred to Arrow, Harris, and Marschak as natural in their model were those in which an order was to be placed in period \( t \) when inventory fell below a lower trigger level of \( s(t) \), with enough being reordered to bring the inventory to an upper target level of \( S(t) \). In exploring the cost characteristics of these policies, they officially opened up the field of \((S, s)\) modeling. Operation of an \((S, s)\) policy with time-invariant upper and lower barriers is illustrated in Figure 1.

One intriguing twist to the \((S, s)\) story is that while Arrow, Harris, and Marschak (1951) set up the problem of minimizing discounted costs, they did not solve it. Instead, they hewed as closely as possible to the long-run logic of the square root rule, which is to minimize long-run average costs. In the case with the uncertainty of outflows being independent and identically distributed and constant \((S, s)\) bands, they used “renewal theory” to solve for the stationary distribution of inventory levels and then solved for the long-run, average-cost-minimizing \((S, s)\) policies. It is somewhat ironic that the title of the original article is “Optimal Inventory Policy,” since there is no proof that the proposed policies are in fact optimal in the presence of discounting. In fact, the policies need not be optimal if adjustment costs decline over time. It was left to others, starting with Scarf (1959), to prove that the optimal policies for minimizing discounted costs in a large class of inventory problems were indeed of the \((S, s)\) form.
Optimality of \((S, s)\) Policies

The \((S, s)\) model has a fascinating technical sting in its tail, since it is more novel than is immediately apparent. It introduces economies of scale into economic analysis: larger orders are more economical because they allow fixed costs to be divided over more units. It turns out that this fundamentally changes both the nature of micro-economic behavior and the sophistication required to identify optimal policies.

As Arrow, Harris, and Marschak (1951) quickly realized, the problem of minimizing inventory costs could not be set up as a problem with a nicely behaved cost-minimization function in which calculus could be used to identify optimal policies. The critical observation is that fixed costs of adjustment require use of trigger strategies that produce a qualitative change in the future trajectory of costs at whatever inventory level triggers a new order to be placed. Discounted costs change quite differently just above this threshold than just below this threshold. (Formally, as the minimum of two convex cost functions, the low-cost strategy as between placing and not placing an order need not be a convex function of inventory on hand. As a result the present value of costs in the complete dynamic model need not be convex, either.) Following considerable computational exploration of optimal policies by Andrew Clark (as discussed in Scarf, 2002), Scarf found a way around this problem. He introduced a controlled departure from convexity, “\(k\)-convexity,” that characterizes the discounted cost function and that suffices to establish optimality of an \((S, s)\) policy. Scarf’s proof did not call for independence over time and established conditions under which time-varying \((S, s)\) policies would be optimal (for example, the ordering triggers were allowed to vary from period to period as information came in relevant to future sales). To this day, this proof remains one of the most conceptually sophisticated and fruitful applications of the discrete-time principle of optimality, with new variations on the theme still being uncovered (Zipkin, 2008).
In much of economic analysis, the assumption of constant or diminishing returns to scale at the purely microeconomic level has been essential. Indeed, economic theorists have rarely been able to formulate questions, let alone make analytic headway, when returns are increasing. To this day, we have almost no idea how many of the differences between the world that we inhabit and the world dealt with in standard economic models rests on the presence of increasing returns of some form or another. The \((S, s)\) model is by far the best developed case in which increasing returns at the purely microeconomic level have been subject to systematic analysis.

**Variations on the Theme**

One of the main attractions of the formulation of the inventory problem in Arrow, Harris, and Marschak (1951) is its flexibility. Many model refinements have been explored, with some of the key ideas outlined in Arrow, Karlin, and Scarf (1958). In terms of long-run impact, a particularly important variation on the theme involves the consideration of multiple echelons, as when a wholesaler must supply retailers who may themselves face various costs of ordering. The pioneering model in this respect is due to Clark and Scarf (1960).

In current inventory practice, there is much concern with automation of ordering policies, and the coordination of ordering policy into the entire supply chain, from raw materials, through manufacturers, through to wholesalers and retailers. As such, inventory policies are seen as part and parcel of broader “supply-chain management” (Knolmayer, Mertens, and Zeier, 2002; Zipkin, 2000). The entire chain is managed in an automated fashion, and at every stage in which an order is placed, the basic structure of the cost minimization involves \((S, s)\) logic. These programs are used to decide whether or not to pursue a “just-in-time” inventory policy by placing frequent and relatively small orders only when stocks on hand are close to zero, or placing larger orders and keeping a safety stock on hand at all times. The trade-off balancing increased holding costs against increased ordering costs and vulnerability to disruptions in the supply chain is solved based on parameters that are entered by the firm in question and are tailored to its evolving circumstances. In this manner, \((S, s)\) policies are embedded in the policies that are pursued up and down the supply chain, with constant adjustments made in light of changing environmental conditions.

The study of money demand inspired an important new class of \((S, s)\) policy and illustrates how questions feeding back from the world of practice have served to enrich the basic structure of the model. Miller and Orr (1966) noted that businesses both use cash to pay costs and receive cash from sales to customers. Hence, firms, face a two-sided problem of control, for which the natural strategies to explore are “two-sided” \((S, s)\) policies in which there are both upper and lower control boundaries. Two-sided policies of this form are critical in the recent burst of work in economic theory on adjustment costs in continuous-time models with evolving uncertainty. Stokey (2008) provides an invaluable guide both to the underlying models of optimal control of diffusion processes and to economic applications.
The \((S, s)\) Model as a General Purpose Model of Microeconomic Adjustment

Fixed costs of adjustment are pervasive. So is uncertainty. It is therefore no wonder the \((S, s)\) model has migrated from its seemingly specialized origins to take an increasingly central place in many areas of microeconomic analysis. We explore in this section the importance of the \((S, s)\) model across various areas of microeconomic analysis.

Much of the focus has been on investment. A common theme in microeconomic studies of business behavior is that factor adjustment often exhibits long periods of relative inactivity punctuated by infrequent and large, or lumpy, changes in factor use. Doms and Dunne (1998) study a large collection of manufacturing plants in the United States and conclude that a significant fraction of investment activity at the plant is associated with large variations in the capital stock and that fluctuations in aggregate investment are associated with variations in the frequency of these investment spikes. For example, 25 percent of total investment derives from firms that adjust their capital stock in a given year by more than 30 percent. Cooper, Haltiwanger, and Power (1999) show further that the conditional probability of a large investment episode rises in the time since the last such episode, as would be predicted by the \((S, s)\) model.

Caballero, Engel, and Haltiwanger (1995) argue that investment can be characterized by a "generalized \((S, s)\) policy" in which the probability of a large investment episode is increasing in the deviation between the firm's actual capital stock and its desired capital stock. On the theoretical side, Abel and Eberly (1994) provide a canonical model of investment allowing for adjustment costs of all forms.

Adjustment policies of the \((S, s)\) variety have also been used to characterize labor hiring and firing. Hamermesh (1989) considers data from a small set of plants and argues that fixed costs of adjustment are needed to explain the sluggishness of employment in response to small shocks and the dramatic jumps in response to large shocks. Caballero, Engel, and Haltiwanger (1997) consider a large collection of manufacturing plants and argue that a generalized \((S, s)\) model fits the data well. Other \((S, s)\) models of firm behavior relate to the dynamic pattern of trade flows when it is costly to set up and dismantle supply channels, and the path of diffusion of a new technology when there are costs of adoption.

There are also important applications of the \((S, s)\) model to household behavior. The early applications involved purchases of real estate and large durables such as automobiles. Consumers do not move to a new house every time their circumstances change, presumably because large fixed costs arise in finding, purchasing, and moving into a new home. Grossman and Laroque (1990) construct an \((S, s)\) model of housing demand with fixed costs and show how real estate holdings interact with portfolio choice. Bertola and Caballero (1990) similarly model the demand for durables in an \((S, s)\) framework. Portfolio choice in the face of fixed costs of adjustment and stock ownership has been studied by Vissing-Jorgensen (2002).
The \((S,s)\) approach can also be fruitfully applied to pricing behavior, which typically involves large adjustments taken at certain threshold moments, rather than continual adjustments. Up through the early 1980s, the modeling of price and wage stickiness typically involved the simplifying assumption that the time between price changes was independent of the evolution of the economy (for example, Fischer, 1977; Taylor, 1980; Calvo, 1983). While time per se is important in some pricing decisions made at a standardized time in the calendar year, other pricing decisions are made on a more or less continuous basis. In the latter case, one would expect the frequency of price changes to increase the higher is the rate of overall inflation.

To introduce flexibility in the time of price changes, Barro (1972) and Sheshinski and Weiss (1977) developed pioneering models of optimal pricing with fixed “menu costs” associated with price changes. Uncertainty concerning the evolution of underlying inflation was introduced by Sheshinski and Weiss (1983), with Danziger (1999) providing an elegant alternative formulation. Benabou (1989) introduced another form of uncertainty deriving from the need to keep durable good price changes unpredictable in a strategic setting. It is these models of menu cost pricing in the face of uncertainty that provided the largest impetus to the migration of the \((S,s)\) model into macroeconomics.

**The \((S,s)\) Model and Macroeconomic Dynamics**

A natural economic question is how the system as a whole behaves when decisionmakers face adjustment costs and hence pursue some form of \((S,s)\) policy. Are the separate \((S,s)\) decisions of firms generally timed in such a way that their aggregate effects are essentially nil, or might widespread application of such policies lead to fluctuations at the macroeconomic level? We begin by outlining early work on this question in the inventory arena, then move to the paradigmatic setting of state-dependent pricing, and then to demand for durables. We close by considering informational and other spillovers that have the potential to increase alignment in decisions to overcome barriers to adjustment, and to amplify underlying economic fluctuations.

Before outlining recent advances in \((S,s)\) analysis, we need to dismiss one intuitive line of argument against the larger significance of models with microeconomic inertia. Intuition suggests that microeconomic adjustment costs need themselves to be significant before they can cause any larger macroeconomic phenomena. In responding to this claim in relation to the seemingly small menu costs of nominal price changes, Akerlof and Yellen (1985) and Mankiw (1985) made a point of general interest: agents facing fixed costs may display economically significant inertia even when costs of adjustment are small, since losses to small deviations from the optimum are second order. Dixit (1991) followed up on this point with particular force. In principle, small adjustment costs at the level of individual agents may have important effects at the macroeconomic level. Understanding how important
these effects are in practice has become an important focus of macroeconomic research over the last 20 or so years.

**Early Insights**

Perhaps the most basic question in the field of \((S, s)\) aggregation concerns how the correlations in shocks faced by distinct agents translate into correlations among their adjustment decisions. Early examples due to Blinder (1981) suggested that much of the observed volatility of aggregate inventory investment might be related to pursuit of \((S, s)\) policies by retailers. With correlations in inventory levels, waves of new orders can be triggered as different retailers respond to a common shock to sales.

The practical importance of Blinder’s insight depends on the extent of the underlying correlation in inventory levels. Caplin (1985) advanced understanding by characterizing the long-run probability distribution of stock levels across a finite number of firms pursuing constant \((S, s)\) policies. He solved the model based on a natural analogy between the operation of \((S, s)\) policies in continuous time and the workings of “modulo” arithmetic, originally created by the great mathematician Gauss. Figure 2 illustrates the inventory state space of a single firm in such circular fashion. Placing \(S\) at 12 o’clock and little \(s\) at its immediate left, demands out of inventory translate into clockwise angular movements of equivalent size around this circular state space, with a new order joining the lower to upper adjustment barrier. To give a convenient example, with \(S = 15\) and \(s = 3\), each order placed in the continuous-time version of the model is precisely of size \(S - s = 12\), and each individual sale moves the clock forward one hour. In a discrete time period with sales of three, the clock moves three hours forward. If the initial inventory was eleven it is reduced to eight; while if initial inventory was five, it increases to 14 as the net result of sales of three and a new order of size twelve. In geometric terms, the effect of individual sales on inventories correspond to equidistant movements around this circle, just as the passage of one hour rotates the hour hand an equal amount wherever it starts on the clock.
Modulo arithmetic gives rise to uniformity. To understand why, think about trying to predict what the second hand of a clock will say when you have completed five jobs each of which take roughly, but not exactly, five minutes. Even if a coworker were simultaneously to undertake five similar jobs that were correlated with yours in terms of time taken, the position of the second hand at the end of your work effort would contain essentially no information concerning its position at the end of your coworker’s. The analogy with respect to inventory levels is that, over time, \((S, s)\) policies drive inventory levels of distinct firms to mutual independence even when sales are correlated. This observation allowed for simple characterization of the stepped-up volatility of orders placed by retailers in comparison with final sales.

An obvious gap in the earliest aggregation models is that demand shocks are treated as exogenous. Further progress requires feedback effects between the policies being pursued and the stochastic processes driving the model. The challenges involved in placing \((S, s)\) models in a full general equilibrium context remain severe. Fisher and Hornstein (2000) specify a general equilibrium model that endogenously yields time-invariant \((S, s)\) adjustment rules and a constant order size per adjusting firm. But the greatest headway in taking account of feedback effects has involved models of state-dependent pricing.

### State-dependent Pricing and Monetary Policy

Consider a central bank that controls the supply of money (or any policymaker with ability to control nominal aggregate demand). Given the path that policymakers set, to what extent will pursuit of \((S, s)\) pricing policies affect the aggregate dynamics of prices and therefore of real output? The intricacy is that the driving forces to which firms are responding, real output and the price level, affect the nominal pricing decision in different ways. At the same time, these pricing decisions are what must be aggregated to identify the evolution of the endogenous variables that are driving prices. So complex is this interaction that progress involves shining narrow points of light and analyzing how far they spread.

An early point of light was provided by Caplin and Spulber (1987), which takes as its starting point the reinterpretation of the \((S, s)\) policy in terms of the modulo arithmetic. Consider adding not a finite number of firms but rather a continuum of firms to this circle. Now think of the analogy of the independence result for inventories that suggests a tendency for the real prices chosen by distinct firms to become independent over time. After this has played out, one would find firms uniformly distributed around the circle. From that point on, the distribution would be unchanging, whatever the nature of the shocks to the system. Upon identification of the state space as prices relative to the monetary base (in logarithms), one concludes that the distribution of such real prices will in steady state remain unchanged in the face of monetary shocks. If each price setter increases prices by 10 percent a time, then a calendar quarter in which the money supply increases by 2.5 percent will trigger price changes to be made by 25 percent of firms. All firms move clockwise 90-degrees in
The Caplin and Spulber (1987) model presents the apparent paradox of frictionless aggregate adjustment in the face of lumpy individual adjustment. Money does not have real effects in this model, despite the fact that all prices are adjusted infrequently. Once the distribution of relative prices is uniform, it will not change; nor will output, in a simple quantity theory model.

Several aspects of this original formulation seem restrictive and potentially fragile. One of the first papers to explore the limitations of the neutrality result was ours (Caplin and Leahy, 1991). The Caplin and Spulber (1987) model applied one-sided ($S, s$) reasoning, based on implicit assumptions that the money supply never falls and that common shocks are dominant. In practice, many nominal prices decline, especially in periods of low inflation. Thus, in Caplin and Leahy (1991), we allow for price adjustment in both directions, analyzing the extreme case of complete symmetry of shocks and of the resulting ($S, s$) strategy. When aggregate shocks dominate idiosyncratic shocks, we show that nominal prices still move to uniformity over a range that corresponds to each price adjustment, as in the one-sided case. However, in the two-sided case this entire uniform distribution can shift in position relative to the money supply. The average level of output tracks the middle of this range precisely. In economic terms, this implies that the money supply affects output over a certain range, but after a while of heading in one direction, this effect ceases; at that point, a reversal in the trajectory of money will affect output, while a further continuation will not. In other words, monetary shocks can have a significant effect on output when they reverse a prior pattern but have no effect when they continue in one direction for a significant time.

Our Caplin and Leahy (1991) model has some attractive features. For example, it implies a “Keynesian” impact of money on output over some range, and then a sharp stop when the end of the range is reached. Internally consistent versions of the model have been developed in which the strategies are optimal (Caplin and Leahy, 1997; Damjanovic and Nolan, 2006; Stokey, 2008). Yet the model’s reliance on the dominance of common shocks is questionable. Most estimates suggest that idiosyncratic shocks are typically dominant, in which case the cross-sectional distribution of prices relative to the money supply will change little unless there are very large changes in the money supply. As a result, neutrality of money will hold to a reasonable first approximation.

Current advances in understanding of the real effects of monetary policy with state-dependent pricing are occurring on two fronts. First, the empirical possibilities for such research have been hugely enhanced by the data on nominal price changes gathered by Bils and Klenow (2004), which in turn followed earlier work by Lach and Tsiddon (1992) and Kashyap (1995). At the same time, and in a highly complementary manner, there are increasing efforts to build up general equilibrium formulations that match moments from the pricing data and the aggregate economy. An early general equilibrium treatment of state-dependent pricing was provided by Dotsey, King, and Wolman (1999). The
recent research of Golosov and Lucas (2007), Gertler and Leahy (2006), and Midrigan (2009) has increasingly refined the ability of these models to match moments derived from the microeconomic studies of price changes. If any broad theme has emerged from this research, it is that the mechanism illustrated in the simple example of Caplin and Spulber (1987)—that is, an outcome in which money is neutral even when prices are adjusted infrequently—is less fragile than it initially appeared.

**Demand for Durable Goods**

A particularly natural area of macroeconomic application concerns aggregate demand for consumer durables. The critical questions are how much “pent-up” demand there is, and what changes in the environment will trigger a wave of purchases. For example, to what extent does a temporary program of incentives destroy future demand (Adda and Cooper, 2000)? This question is highly relevant to the success of demand-shifting programs that have featured heavily in recent years, from offering temporary incentives to trade in old cars for new, to programs that temporarily subsidize first-time home purchases.

Bertola and Caballero (1990) study in depth the nature of pent-up demand for durables in a partial equilibrium model built on an \((S, s)\) foundation. Caballero (1993) applies this framework to show that an aggregate \((S, s)\) model may look like a partial adjustment model, while Attanasio (2000) provides a detailed application to automobile purchases. As was the case in early inventory aggregation models, the pioneering models of durables aggregation limit the feedback from demand to pricing. This is a major limitation. If \((S, s)\) policies create some bunching up in orders for durables, prices will have to adjust as a result, radically changing the actual pattern of sales, requiring solution of a full equilibrium model with feedback from the policy to the structure of demand.

Just as in the inventory case, it has proven difficult to develop general models that incorporate feedbacks from demand to pricing in a general equilibrium setting. Yet there are some points of light. In the model of Thomas (2002), endogenous price movements greatly smooth out surges in aggregate demand. For example, if an unusually large number of people are about to purchase a car, the price of cars will rise, thereby dissuading some agents from making purchases. Similarly, in Caplin and Leahy (2004), we provide an example in which \((S, s)\) behavior at the individual level produces aggregate dynamics qualitatively similar to those of a representative agent model, suggesting that fixed costs at the individual level may have relatively little discernible effect on aggregate outcomes.

\((S, s)\) Policies and Information Externalities

In some respects, the flow of the literature on \((S, s)\) aggregation is a little disappointing. The most robust answer to date is that microeconomic indivisibilities do not have first-order implications for macroeconomic dynamics. We who intuitively believe that fixed costs are fundamentally significant for macroeconomic fluctuations remain far from having established our case. It is our hypothesis that if and when we
do, it will be due in the first place to increased understanding of how \((S, s)\) policies affect the flow of information. The essential insight is that pursuit of an \((S, s)\) policy can make inference difficult for outside observers. In periods of inaction, \((S, s)\) policies are effectively stores of private information about the conditions that would trigger future action, and what action would then be triggered. To take a prosaic example, how can a car dealer know when to anticipate orders from potential customers who have choice in terms of precisely when to replace their current cars, and what to replace them with? Is a temporary reduction in demand for one class of vehicle a positive sign for future demand given the further deterioration in the outstanding stock, or is it a sign that this type of vehicle is no longer of the same interest to buyers?

The broad problem the above example illustrates is that inertial behavior in periods of non-adjustment keeps information of common interest in private hands. When adjustment thresholds are finally crossed, the resulting actions may release information to the market, causing something of a cascade. In the above example, seeing signs that previous owners of a given vehicle were trading them in for an entirely new type of vehicle would resolve the problem one way, while a burst of direct replacement would resolve it the other way. Zeira (1987) constructed the first model in which an informational channel of this form associated with fixed costs resulted in a boom–bust cycle at the aggregate level, with some small change in behavior revealing to all that a run of market growth has stopped, with this common recognition causing a market crash. This basic mechanism was expanded on by Rob (1991) and by us in Caplin and Leahy (1994). It is also related to models of information externalities and herd behavior (Chamley, 2004).

We believe these theoretical concepts have important parallels in the world of practice. Think how much more is known about the quality of mortgage originations now that the housing market has crashed. The low and ever-diminishing quality of these originations was known to many market insiders for more than a decade (Caplin, Chan, Freeman, and Tracy, 1997). However this information failed to become widely internalized by investors, due to the relatively low default rates that were visible in the period of rising house prices. Once house prices fell, defaults rocketed, and the nagging doubts of many concerning the underlying quality of the mortgages were confirmed, with ongoing fallout effects.

To this point, few models have been developed that capture this pattern of boom and bust induced by informational barriers. One impediment to this research is technical: no one has solved anything other than highly restrictive models, which are of necessity fragile. The other is institutional: economists with roots in the academy have only very limited understanding of the actual flow of information in modern markets. Thereby hangs another tale.

### The \((S, s)\) Model and Household Finance

Household financial decision making provides many settings in which fixed costs of adjustment are important, in which the \((S, s)\) model can play a role in
capturing the associated phenomena. We will argue that adapting the \((S, s)\) model to household finance presents both new challenges and new opportunities. One important challenge is that adjustment costs in household financial decisions are not purely technological, but may also have psychological and social components. The central opportunity lies in the development of “household operations research” to aid households in making hugely complex decisions for which almost all of us are so patently underprepared.

**Adjustment Costs and Household Debt**

Consider a homebuyer who used a mortgage to finance a past purchase transaction. Assume that default is not an issue and that the only question facing the household is whether and when to refinance the mortgage as the interest rate on a newly originated alternative mortgage drifts lower. This decision shares classical features of the \((S, s)\) model. The state variable that drifts is the payment on the current mortgage relative to the payment that would be available by refinancing to the currently best available rate. There are costs of terminating the current mortgage and applying for the replacement. There will be future possibilities for refinancing if interest rates drift down further, so there can be repeated downward adjustments in the monthly payment. All such future possibilities must be taken into account in the current decision. While modeling this decision is technically tricky due to the need to model the dynamic path of future interest rates realistically and to keep track of changes in mortgage application costs, the resulting optimization problems is clearly of the \((S, s)\) variety.

When and whether to refinance a mortgage is far from the only financial decision households make in which there are fixed costs of adjustment and a state variable that drifts absent control. In the financial crisis and its aftermath, many households are facing a decision with regard to when and how to terminate a mortgage. For many, default is under active consideration. Here the path of house values and projected income is important. The effect of uncertainty concerning future house value on default incentives was first noted more than 20 years ago in the context of the options-based models of mortgage default (Foster and Van Order, 1984). Transactions costs associated with the act of default, in the form of future market-based penalties, separation from the home, and possible social stigma, were incorporated by Crawford and Rosenblatt (1995). Putting these forces together, the decisionmaker is again facing a problem of the \((S, s)\) variety in terms of when and whether to default.

The two decisions above are part of the larger complex of financial decisions that many households face. The importance of these decisions to the health of advanced credit-based economies should be clear to all in light of the ongoing market crisis. Yet modeling in this area is very underdeveloped. Hazard rate models (as in Foote, Gerardi, and Willen, 2008) that statistically summarize past default provide a crucial first step in understanding default. Yet a structural model is needed to infer the extent of the transactions costs of default and how they differ between homeowners and investors. At present, policymakers designing
policies to affect future mortgage defaults are essentially flying blind. Additional research on the \((S, s)\) model as applied to household finance is urgently needed.

**What Are the Household Adjustment Costs?**

The cases of household debt decisions with explicit financial costs of adjustment appear to be merely the tip of the inertial iceberg in household finance. Ameriks and Zeldes (2004) have intriguing findings concerning the low frequency with which households adjust their asset portfolios. Equally shocking are the results of Madrian and Shea (2000) and Choi, Laibson, Madrian, and Metrick (2002) relating to the low frequency with which households change the proportion of their employment income that they set aside to provide funds for retirement. More broadly, the evidence suggests that individuals resist change in many respects, a subject that was opened up by studies of status quo bias in individual decision making (Samuelson and Zeckhauser, 1988). Much of this may fruitfully be conceived of as resulting from psychological inertia that can only be overcome by the weight of accumulated evidence—which is precisely the mechanism that the \((S, s)\) model describes.

These examples raise the question of how one can identify empirical counterparts to the fixed costs of adjustment that lie at the heart of the \((S, s)\) model. This is relatively simple in the case of inventories, since most firms have a technologically-driven understanding of delivery costs. In most other applications, the fixed costs of adjustment are far more subtle. In one fascinating study, Levy, Bergen, Dutta, and Venable (1997) examine the process of changing nominal prices in large organizations. For producers of intermediate goods, pricing is part of an intricate web connecting the firm to its clients, and every thought is given to keeping this connection as strong as possible in the face of price changes. Balancing the need to respond to cost and market changes against the needs of long-standing customers requires a concentrated effort at communication across different divisions of the firm. What may seem like trivial “menu costs” of adjusting prices turn out to be profound, with the entire organization needing to coalesce around and defend alterations in its pricing structure.

The challenge of how best to identify adjustment costs is particularly severe in the case of household finance. It is also of the utmost practical importance. To take a topical example, might increased levels of mortgage default lower the “stigma” involved in so defaulting, effectively lowering the associated costs and triggering yet more default? More broadly, selection by policymakers of a status quo option that is costly for the household to change and is regarded as socially beneficial (like default participation in a pension scheme) is increasingly being promoted as a form of “benign paternalism” (Thaler and Sunnstein, 2008). The efficacy and policy benefits of the “nudge” that a particular status quo option provides depends on why it is left in place, and by whom. It is one matter if those who care enough change while those who care little leave the decision unchanged. It is quite another if households who desperately need money in hand feel that they have no real option of changing behavior without facing social stigma and private humiliation,
and instead run up higher credit card debts that blight their financial future. Research is urgently needed on why inaction is so prevalent before policymakers can comfortably use private inertia as a policy instrument.

**Household Operations Research**

There is widespread interest in improving the quality of household financial decision making. We believe that a field of household operations research will arise to help fill in this gap. Just as the original development of \((S, s)\) inventory models was driven by real economic needs of businesses, so we believe that the next generation of modeling will be driven in large part by the desire of many households to seek decision aids given the various costs of adjustment that affect their financial decisions. The first stirrings of this form of “household” operations research are already evident in software that has been developed and designed to aid households in their work, savings, and portfolio decisions (for example, Bernheim, Forni, Gokhale, and Kotlikoff, 2000; Kotlikoff and Rapson, 2006). The chief goal of household operations research will be to develop workable algorithms for complex household financial decisions.

Households appear to make many suboptimal financial decisions throughout the lifecycle. For example, many households have a large number of sources of debt finance, including credit card debts, home equity lines of credit, auto loans, students loans, and others. Given the multiple forms of debt open to households, the common state variable that is affected by debt-related decisions may be the credit record. After all, this record summarizes the market penalty in terms of future access to credit associated with individual acts of default. Yet few households even understand how their decisions about debt affect their credit score and how this credit score impacts future opportunities, let alone how to optimize in this dimension. Thus, a potentially productive approach might be to develop realistic \((S, s)\) models in which households make costly financial decisions based on a desire to keep the credit record in an acceptable range. The ideal for household operations research would be to build such a well-developed model that personal parameters could be entered, with recommendations and rationalizations then being produced. Such models could serve both prescriptive as descriptive roles, as was true with the original \((S, s)\) model.

**Concluding Remarks**

All model builders strive to eliminate inessential elements to get to the heart of whatever matter they are studying. The key to evaluating the success of an economic model is the extent to which those choices open up new vistas of theoretical insight and application. By this standard, the \((S, s)\) model has continually proven its worth as one of the touchstone models of economics.

At a theoretical level, because of the focus on indivisibilities, those who develop the \((S, s)\) model often find themselves grappling with technical issues
quite distinct from those faced elsewhere in the profession. But the \((S, s)\) model reveals its true significance in its relationship with the world of application. To this day, the \((S, s)\) model holds place of honor at the center of inventory theory and ordering policy, having most recently been incorporated into the business software with which production and storage processes are run worldwide (Knolmayer, Mertens, and Zeier, 2002; Zipkin, 2000). As ever-more realistic \((S, s)\) models are developed to capture household financial decisions, they too will be encapsulated in algorithms that will guide practice in the field. The unfolding of this process will continue the cycle between economic theory and the world of practice that the \((S, s)\) model so wonderfully epitomizes.

We thank Herbert Scarf for his contributions to the spirit and the substance of this article, as well as Nancy Stokey and Paul Zipkin for important guidance. We thank also the editors of this journal for their encouragement, their structural guidance, and their substantive contributions.

References


Economic Activity, no. 2, pp. 443–505.


