Volume Title: NBER Macroeconomics Annual 1990, Volume 5

Volume Author/Editor: Olivier Jean Blanchard and Stanley Fischer, editors

Volume Publisher: MIT Press

Volume ISBN: 0-262-02312-1

Volume URL: http://www.nber.org/books/blan90-1

Conference Date: March 9-10, 1990

Publication Date: January 1990

Chapter Title: Kinked Adjustment Costs and Aggregate Dynamics

Chapter Author: Giuseppe Bertola, Ricardo J. Caballero

Chapter URL: http://www.nber.org/chapters/c10976

Chapter pages in book: (p. 237 - 296)

Giuseppe Bertola and Ricardo J. Caballero
PRINCETON UNIVERSITY, AND C.E.P.R./COLUMBIA UNIVERSITY

*Kinked Adjustment Costs and Aggregate Dynamics**

1. *Introduction*

The best-fitting linear representations of relationships among aggregate time series are typically smooth and sluggish. When seeking microfoundations in a representative agent framework, macroeconomists have therefore adopted convex and differentiable—most often quadratic—adjustment cost functions. This implies partial and continuous reactions to all innovations; but even casual observation of, for example, durable purchases and retail prices indicates that microeconomic units make certain adjustments only intermittently and by amounts that are not necessarily small. Obviously, then, real-life individuals are not solving the representative agent's convex adjustment cost problem, and its parameters have no clear "deep structural" interpretation. Resolving the tension between empirical tractability and microeconomic realism is not easy, but the importance of efforts in this direction is becoming more and more evident. Consideration of microeconomic realism is, we believe, essential for macroeconomics to develop sound theoretical foundations.[1]

*For comments on earlier drafts and helpful conversations we are indebted to Roland Benabou, Olivier Blanchard, Alan Blinder, Andrew Caplin, Avinash Dixit, Stanley Fischer, Robert Hall, Eytan Sheshinski, seminar participants at Columbia, Princeton, L.S.E., and M.I.T., and NBER Macroeconomics Annual Conference participants.

1. More careful work on dynamic adjustment may be needed for purely empirical purposes as well. On the one hand, the rich, slow dynamics of aggregate data have proven difficult to rationalize in a representative agent model, whether based on atemporal frictionless optimization or on convex adjustment costs; on the other hand, statistical models of aggregate data often uncover asymmetries and track endogenous variables poorly in the aftermath of large shocks. Hansen and Singleton (1983) and Abel and Blanchard (1986) are examples of careful, structural representative agent models falling short of providing a statistically robust description of aggregate data, which reject the restrictions imposed by optimizing behavior. Neftci (1984), DeLong and Summers

Intermittent large adjustments can be explained in an optimizing framework by the observation that microeconomic adjustment cost functions are often *kinked* at the no-adjustment point. Inaction is costless, but even small changes in endogenous variables may entail finite costs. In this paper we discuss the microeconomics of infrequent adjustment, reviewing well-known qualitative insights along with recent technical advances that make it possible to develop sophisticated and realistic formal models. Because realistic microeconomic adjustment implies much more pronounced sparseness of action than is apparent in aggregate data, notions of *idiosyncratic uncertainty* and *lack of coordination* are essential for macroeconomic applications. We provide a formal framework in which such issues can be addressed, and we discuss the role of microeconomic inertia in shaping the empirical behavior of aggregate data.

The paper is structured as follows: Section 2 discusses the realism of infrequent adjustment in many partial equilibrium problems of macroeconomic interest and solves a stochastic model of adjustment under kinked adjustment costs. The model we propose and the techniques we use for its solution are applicable in a variety of circumstances. As in previous work, we find that in the long run endogenous variables should be well predicted on average by models of costless adjustment, while kinked adjustment costs produce a wide dispersion of possible outcomes at a point in time, and rich, history-dependent dynamics.

Section 3 studies the behavior of a large group of individuals following similar dynamic adjustment policies. Modeling the probabilistic structure of aggregate and idiosyncratic shocks in a variety of simple frameworks, we find that the behavior of aggregate variables depends in an intuitive way on the relative importance of ongoing aggregate and idiosyncratic uncertainty. When the former predominates, the aggregate behaves very much as any one of the individuals would, displaying strong history dependence and sluggishness; but as idiosyncratic shocks become more important, the aggregate behaves more and more as an individual would *in the absence of any obstacle to adjustment*, and therefore quite unlike any one of the actual individuals.

Section 4 proposes an application of these models to U.S. durable goods consumption data. The results are encouraging, and suggest that a good fit of aggregate dynamics can be obtained under realistic assumptions about the dynamics of microeconomic adjustment. Section 5 concludes and outlines directions for further work.

(1986), and others have focused on the asymmetric cyclical behavior of macroeconomic time series.

2. *Microeconomics*

Individual firms do not continuously adjust their capital stock, prices, and production techniques; consumers do not alter their portfolio composition, labor supply, and consumption habits every hour, day, or even year. This type of behavior cannot be rationalized by strictly convex adjustment costs, which would make it optimal to continuously and partially adjust to all exogenous shocks. In fact, the assumption that average costs of adjustment should be increasing in the speed of adjustment is not generally realistic at the microeconomic level: even as they introduced quadratic adjustment costs and certainty equivalence to the economic literature, Holt et al. (1960) made it clear that such assumptions could only be taken to be reasonable approximations over a range.

Infrequent corrections might be taken to reflect suboptimal behavior at the individual level (Akerlof and Yellen 1985). Alternatively, and to take advantage of optimization-based theory and prediction, inaction can be explained if adjustment costs are specified so as to penalize continuous small reactions—a form of increasing returns to scale. It is not unreasonable to allow for kinked adjustment costs (not differentiable, and possibly discontinuous) at the no-adjustment point. Inaction should be costless, but the cost of even small adjustments may be finite; more generally, if the per-unit cost of reacting to those exogenous changes that typically occur between decisions is large compared to the benefits adjustment would yield, it clearly does not pay to *always* adjust.

While economists have long been aware of the qualitative dynamic effects of adjustment cost nonconvexities (a wide-ranging critique of convex adjustment cost models is in Rothschild 1971), interest in models of infrequent adjustment has recently been rekindled by introduction of techniques providing quantitative insights and exact solutions in realistic applications, adapting stochastic calculus results from engineering, operations research, and finance applications (many relevant results and techniques are usefully summarized and reviewed in Harrison 1985). The assumption of continuous time and state spaces makes it easier to obtain analytical results, as integrals are more readily manipulated than summations. When modeling behavior in continuous time and assuming a differentiable flow benefit function, an adjustment cost function that is not continuously differentiable is sufficient for intermittent adjustment to be optimal. This section reviews recent and less recent contributions to the literature and provides a simple introduction to continuous time models of infrequent dynamic adjustment, highlighting their advantages in terms of realism and analytical tractability.

2.1 LITERATURE REVIEW

The (S, s) two-point rule is the earliest and best-known discontinuous adjustment control policy. It is applicable to cases where adjustment is assumed to be one-directional, and to entail a fixed lump-sum cost per adjustment decision, as may be the case in inventory management at a retail outlet: the cost of ordering nothing is zero, each unit can be purchased at a given unit price once an order has been placed, but a fixed, per-order cost yields a downward sloping unit adjustment cost function. Both the total and unit order cost are then discontinuous at zero, and the optimal ordering strategy calls for infrequent, large orders (see Scarf 1960, and his references for even earlier, less formal work); under simplifying assumptions, cost minimization will call for all orders to be the same size. Money demand has been modeled in a similar framework, assuming the exchange of ready cash for other stores of value to entail lump-sum transaction costs (Baumol 1952, Tobin 1956, Miller and Orr 1966). The controversial—but in some settings realistic—assumption that price changes incur a fixed *menu* cost makes for large, infrequent price adjustments (Barro 1972, Sheshinski and Weiss 1977, 1983), which again can be described by fixed-adjustment-size rules under simplifying assumptions; adjustment will be one-sided if price reductions are never found to be optimal.

The contributions cited above either assume certainty, or provide stylized treatments of simple uncertainty cases. Continuous-time, stochastic models of (S, s) adjustment policies, motivated by realism and tractability, have been studied extensively in the Operations Research literature (an early reference is Bather 1966). Financial economists make extensive use of similar techniques in modeling asset prices, and the first applications to adjustment cost problems other than that of inventory management were, quite naturally, in a financial setting. Constantinides (1986) proposes an approximate solution for the portfolio problem of a consumer-investor in the presence of nondifferentiable portfolio adjustment transaction costs, and Grossman and Laroque (1990) solve a specific model of illiquid durable goods consumption under uncertainty. Frenkel and Jovanovic (1980) provide a rederivation of the Baumol-Tobin model of cash management and money demand in a continuous-time uncertainty framework (Smith 1989 further extends the model to include interest rate variability and proportional transaction costs), and Tsiddon (1988) proposes a continuous-time model of menu cost pricing.

While it is obvious that adjustment should be infrequent and large in size when action entails lump-sum costs, it is perhaps less clear that inaction may be optimal even when adjustment costs are proportional to the size of the correction being undertaken. In fact, when exogenous influences may make adjustment in *both* directions desirable, inaction is optimal if the adjustment cost function, though continuous, fails to be differentiable at zero.

An extreme example is that of finite, proportional adjustment cost in one direction, and prohibitive costs in the other, as may be realistic in models of investment (Arrow 1968, Nickell 1974, Pindyck 1988, Bertola 1989a).[2] More generally, inaction is optimal when either the adjustment technology has increasing returns to scale, like in the lump-sum cost models discussed above; or, adjustment has constant returns to scale (proportional adjustment costs), but it is costly (not necessarily impossible) to retrace one's steps, and exogenous variables may return to their original values after an innovation.

For example, labor turnover costs may not be strictly convex even when no lump-sum adjustment costs are present. New employees need to be screened and trained, and the cost of doing so may well be proportional (if not less than proportional) to the total number of hirees; and employment contracts penalize—explicitly or implicitly—a firm's firing decisions. It is quite clearly not optimal to hire a new worker just before desired employment ceases rising, and fire her as it starts falling: the hiring and firing costs would have to be paid at essentially the same moment to almost no avail, since the marginal worker would not have time to produce flow revenues in her short tenure. These insights are modeled by Kemp and Wan (1974), Nickell (1978), and Bertola (1989b) in a certainty framework, while the effects of uncertainty have been studied by Caplin and Krishna (1986), Gavin (1986), and Bentolila (1988) in discrete-time models, and by Bentolila and Bertola (1990) in a more general continuous-time framework.

Other dynamic adjustment problems have been studied along these lines. Inasmuch as durable goods can be seen as factors of utility production, the problem of a consumer faced by transaction costs in the purchase and sale of durable goods is similar to that of a producer choosing an optimal capital accumulation policy (the budget constraint introduces additional complications, however). Bar-Ilan and Blinder (1987) and Grossman and Laroque (1990) propose models incorporating these features. Finally, marketing of a product may also entail lumpy and, more generally, nondifferentiable costs. Baldwin and Krugman (1989) apply this insight to the responsiveness of prices and quantities of internation-

2. If desirability of downward adjustment were ruled out (as it is in (S, s) models of inventories with positive net sales at all times), irreversibility of adjustment would be completely irrelevant. In general, however, irreversibility has important consequences on optimal (necessarily infrequent) dynamic adjustment.

ally traded goods to exchange rate fluctuations: exporting and import-competing firms should be wary of reacting to such fluctuations, if doing so is costly and there is a possibility of exchange rates reverting back to their earlier values. When adjustment costs are kinked, short-lived exchange rate swings may have long-lived effects on international trade, and, more generally, the dynamic relationship between exchange rates, activity levels, and trade balances is highly nonlinear. Continuous-time stochastic models of such phenomena have been developed by Dixit (1989a,b,c) and Dumas (1988).

This brief review suggests that issues of infrequent adjustment have been addressed early and often in the economic literature. Models, however, that are so realistic as to make applied work possible have only recently begun to be developed in economics. We proceed to illustrate the new techniques with a relatively simple model.

2.2 A PROBLEM OF ADJUSTMENT UNDER KINKED COSTS

Let the flow benefits accruing to a microeconomic unit be described by a function $\Pi(x,y)$ of x, controllable, and y, exogenous; y denotes a collection of variables describing the environment of the microeconomic unit and the character of its stochastic evolution in time. Let $\Pi(.,.)$ be twice differentiable and strictly concave in x, with a well-defined unrestricted maximum

$$x^*(y) \equiv \arg\max_x \Pi(x,y).$$

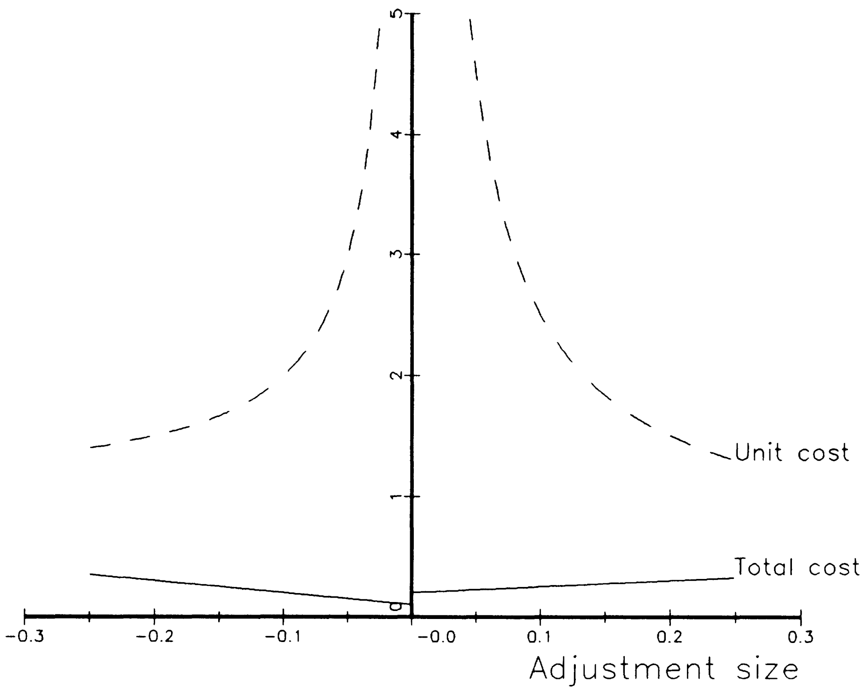
In the presence of adjustment costs and uncertainty, we write the dynamic problem of a risk-neutral optimizer with discount rate ρ as

$$V(x_t,y_t) \equiv \max_{\{x_\tau\}} E_t \left\{ \int_t^\infty e^{-\rho(\tau-t)} \left(\Pi(x_\tau,y_\tau)d\tau - [\text{adjustment costs}] \right) \right\}.$$

Ideally, the optimizer would like to choose a stochastic process for x_τ such that $x_\tau = x^*(y_\tau)$ at all τ. If altering the level of x is costly, however, these costs have to be traded off against the benefits of tracking $x^*(y_\tau)$, the *frictionless* optimum, more closely. Specifically, let every upward adjustment of x by some amount $k > 0$ cost $C_u + c_u k$; and, similarly, let downward adjustments by $-k < 0$ cost $C_l + c_l k$.[3] This is the piecewise

3. The adjustment cost parameters $C_i, c_i, i = l, u$ need not all be positive: if some adjustment decisions bring on rewards, rather than costs, the optimization problem may still we well defined. For the existence of a well-defined solution to the problem, however, C_l, C_u, c_l, c_u must be such that $C_u + C_l + c_l k + c_u k \geq 0$ for all $k \geq 0$; otherwise, adjustment up and down by k would produce revenues rather than costs, and nothing in the problem would prevent this from happening infinitely often, yielding an unbounded value for the program.

Figure 1



linear adjustment cost function plotted in Figure 1. The function is *kinked* at zero, the no-adjustment point; it is discontinuous whenever $C_u \neq 0$ and/or $C_l \neq 0$, and nondifferentiable whenever $c_u \neq -c_l$.

To separate the essentially static problem of choosing $x^*(y)$ from the dynamic adjustment aspects, let us define the value of the program in the absence of any adjustment cost,

$$V^*(y_t) = E_t \left\{ \int_t^\infty e^{-\rho(\tau - t)} \Pi(x^*(y_\tau), y_\tau) dt \right\},$$

and write

$$V(x_t, y_t) = V^*(y_t) + \tilde{v}(x_t, y_t).$$

Thus, $\tilde{v}(x_t, y_t)$ denotes the loss in value terms due to the presence of adjustment costs, as a function of the current state of endogenous and exogenous state variables. The adjustment policy should minimize this loss, trading off the costs incurred at the times when adjustment is undertaken against the expected present value of flow costs due to deviations from the frictionless optimum.

We now introduce assumptions that yield a simple form for the adjustment policy. First, we assume that the $\{x^*(y_t)\}$ stochastic process is well described by an arithmetic Brownian-motion process, and the endogenous stock x_t depreciates linearly when the optimizer chooses not to act:

$$dx_t^* = \vartheta^* dt + \sigma dW_t, \quad dx_t = -\delta dt + \text{adjustment}.$$

Second, we assume that

$$\tilde{v}(x_t, y_t) = E_t\left\{ \int_t^\infty e^{-\rho(\tau - t)} \left((-\tfrac{1}{2} b)\,(x_\tau)^2 d\tau - [\text{adjustment costs}] \right) \right\}.$$

Defining $z_t \equiv x_t - x^*(y_t)$, we note that in the absence of adjustment z_t follows a Brownian-motion process with constant drift $-\vartheta$ (with $\vartheta \equiv \vartheta^* + \delta$) and standard deviation σ, and that the loss $\tilde{v}(.,.)$ is a function of z_t only:

$$\tilde{v}(x_t, y_t) \equiv v(z_t) \equiv -\min_{\{z_\tau\}} E_t\left\{ \int_t^\infty e^{-\rho(\tau - t)} \left((\tfrac{b}{2} z_\tau^2)\, d\tau + [\text{adjustment costs}] \right) \right\}. \tag{2.1}$$

Under regularity conditions, $v(z)$ is less than zero and is bounded below (see the Appendix). These simplifying assumptions ensure that the state space of the optimizer's problem is continuous and Markovian in terms of a single-state variable, z. The optimizer always has the option to alter the current level z_t instantaneously, not necessarily by infinitesimal amounts, and the optimal policy can be expressed in terms of *fixed* trigger points and adjustment steps in this state space. The simple form of the solution is exact under the assumptions above, which are not dissimilar from those made in earlier macroeconomic applications; we discuss in *Section* 2.4 the restrictions they impose on the underlying structure. The solution may also be taken as an approximation to that of more general problems with kinked adjustment costs leading to some form of inaction.[4]

We describe the adjustment policy in terms of four, not necessarily distinct, parameters (L, l, u, U). Specifically, adjustment only occurs when

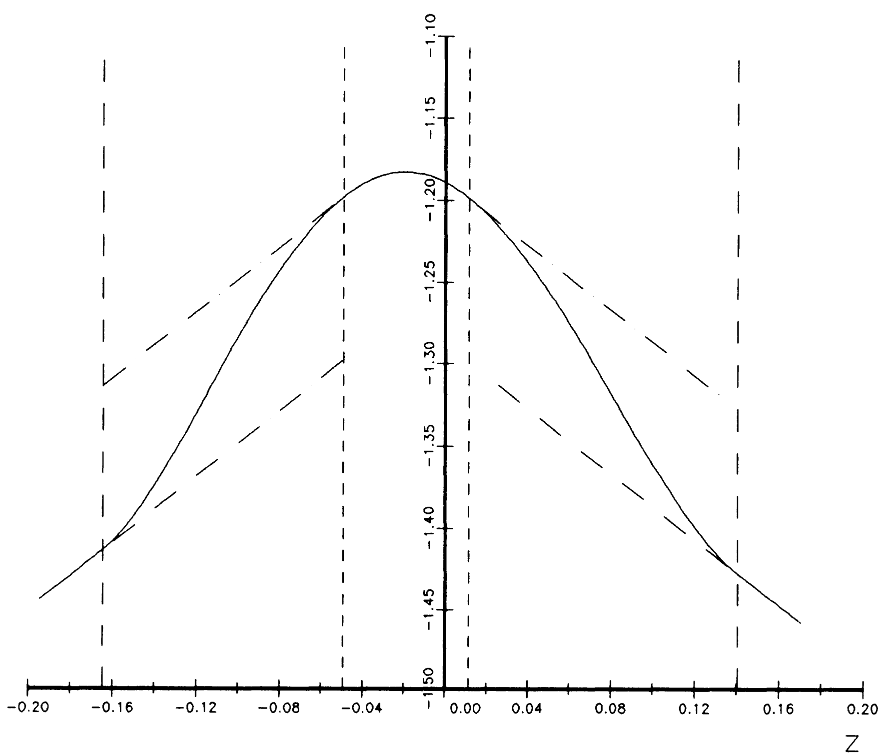
4. Since $\partial_x \Pi(x^*(y), y) \equiv 0$ by assumption, this expression could be justified in terms of a second-order approximation around the *moving* point $x^*(y_\tau)$, as long as $\partial_{x^2} \Pi(x^*(y_\tau), y_\tau)$ is constant (and equals b). $\partial_{x^j} f(\bar{x}, y)$ denotes the j^{th} partial derivative of a function $f(.)$ with respect to x, evaluated at $x = \bar{x}$.

z is at points L or U, $L \leq U$; when z reaches L, control moves it instantaneously to l, with $L \leq l < U$; and when z reaches U, control moves it back to a point u, with $L < u \leq U$. We proceed to characterize these four points in terms of the value function $v(.)$ defined by (2.1). If the optimal control policy is unique, it is necessary and sufficient for optimality of a candidate policy that costs and benefits of any action undertaken by the optimizer be equal along the optimal path, on the one hand, and costs of potential actions be weakly larger than their benefits when the optimizer is inactive, on the other hand. Formally, $v(.)$ and (L,l,u,U) must be such that

$$\begin{array}{ll} v(l) - v(L) = C_l + c_l(l - L), & v(x) - v(y) \leq C_l + c_l(x - y)\ \forall x > y \\ v(u) - v(U) = C_u + c_u(U - u), & v(x) - v(y) \leq C_u + c_u(y - x)\ \forall x < y. \end{array} \quad (2.2)$$

These relationships are illustrated in Figure 2 (similar diagrams appear in Constantinides and Richard 1978, and in Caplin and Krishna 1986 for the

Figure 2: $v(z)$



case of proportional costs only). Moreover, we note in the Appendix that (2.2) and differentiability of $v(.)$ imply:

$$\begin{aligned} v'(l) &= v'(L) = c_l \\ v'(u) &= v'(U) = -c_u \end{aligned} \tag{2.3}$$

These conditions on $v'(z)$ can be intuitively interpreted in terms of optimality of adjustment size. Once the optimizer has decided to take action, the lump-sum cost C_u (or C_l) is sunk, and the size of the jump must be such that the marginal return to adjustment exactly offsets the proportional cost at the return point. Considering that the optimizer might have decided to initiate adjustment at points different from the candidate triggers, the same reasoning applies to L and U.

Thus, optimal action and return points must then be such that $v'(x)$ equals the marginal cost of action whenever action is undertaken ("smooth pasting"), and the value function at the trigger and return points must differ by the total cost of adjusting between the two points ("value matching"). By differentiability of $v(.)$, we can write

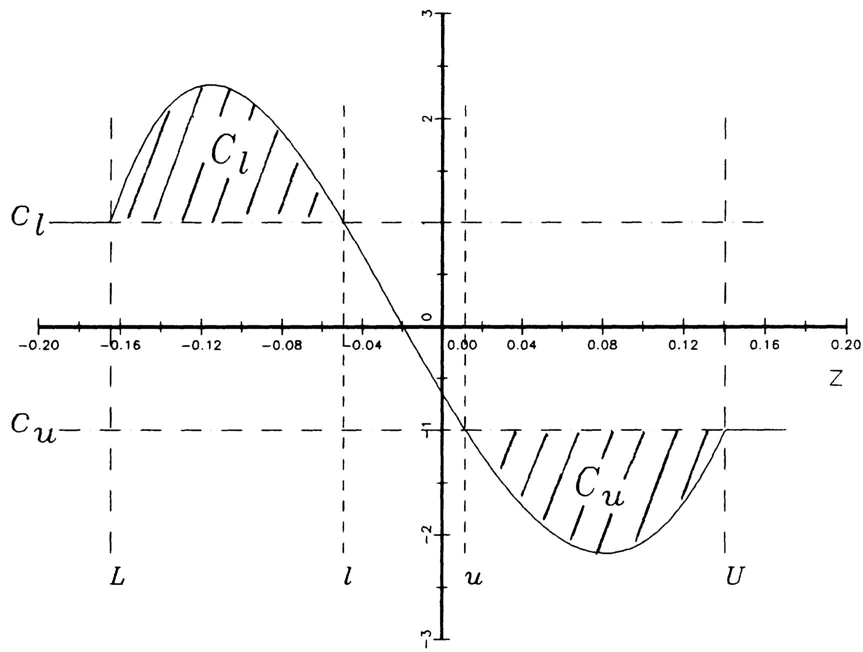
$$v(l) = v(L) + \int_L^l v'(z)dz, \quad v(U) = v(u) + \int_u^U v'(z)dz,$$

and the solution of the dynamic optimization problem can be represented as in Figure 3 (similar diagrams appear in Constantinides and Richard 1978, Harrison, Selke, and Taylor 1983, Dixit 1989d). In the Figure, smooth pasting constrains the level of the S-shaped $v'(z)$ function at the action and return points, and value matching requires that the shaded areas be equal to the lump-sum costs of adjusting in that direction.

When there is no lump-sum component ($C_u = C_l = 0$), there is never any reason for adjustment to be larger than infinitesimal in an ongoing optimization program, as the path of $\{z_t\}$ is continuous in the absence of regulation. Hence, $U = u$ and $L = l$, and the common value of these parameters needs to be determined, by the conditions in (2.3) alone, at the points where the S-shaped curve of Figure 3 is horizontal. In the case of nonzero lump-sum adjustment costs, adjustment must have finite size; given differentiability of the value function, infinitesimal changes in z would not yield benefits large enough to match a finite cost. Then, the value of the four points defining the optimal policy will be derived from joint consideration of (2.2) and (2.3).

To make use of the optimality conditions, we need a functional form for $v(z)$. Since $v(z)$ is flat around its maximum, it is certainly suboptimal

Figure 3: $v'(z)$



to correct small z_t deviations, as doing so entails first-order costs. The optimal adjustment policy then allows z_t to wander some finite distance (L, U) from zero before taking correction action. In this range, $\{z_t\}$ behaves as a Brownian motion, and we show in the Appendix that this makes it possible to characterize the value function's behavior in the absence of control, and to obtain an explicit functional form for $v(.)$ up to integration constants to be determined at the boundaries of the inaction region. These boundaries and the integration constants are jointly determined by conditions (2.2) and (2.3), which are not difficult to solve numerically.

2.3 DYNAMICS, LONG-RUN DISTRIBUTION, AND AVERAGES

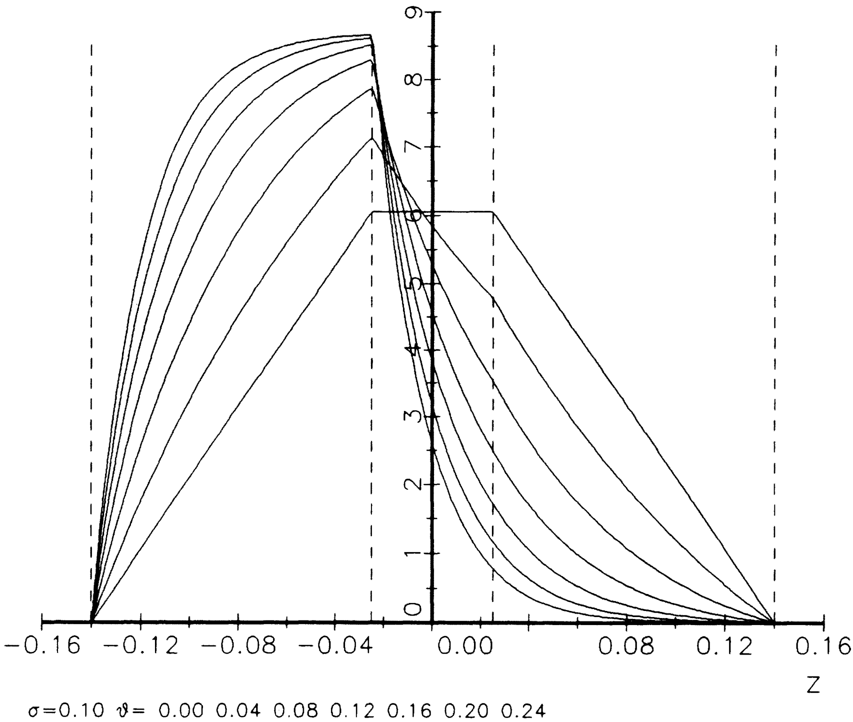
In the absence of obstacles to continuous and complete adjustment, the economics of the optimization problem would provide us with a tight relationship $x^*(y)$ between the exogenous state variables, y, and the endogenous one, x. This relationship would typically be used to draw positive implications on the position and dynamics of x from knowledge of y. Adjustment costs make such inferences more difficult and less precise. At a point in time, the actual value x_t can deviate from $x^*(y_t)$ by the potentially large amount z_t, depending on the past history of the exogenous variables and on the resulting path of adjustment. As to dynamic reactions, they also depend on the past history (as summarized by the current value of z): x may fail to respond to a small change in y if no adjustment is triggered, or may react disproportionately if lump-sum costs of adjustment are present and the y change triggers a jump in x.

When we are asked to interpret the likely evolution of x in the face of exogenous shocks, we cannot always have complete information about initial conditions and the history of exogenous variables. It is interesting, then, to examine the implications of the model at the other extreme: suppose we have no information as to the current position of x, though we know the parameters of the individual's dynamic problem, and consider the long-run behavior of $\{z_t\}$ deviations in the simple optimization program above. Since the $\{z_t\}$ process never leaves $[L,U]$, and reaches any point in that interval with probability one over the infinite time horizon we consider, it possesses an invariant, ergodic distribution—and if we literally know nothing about the past history of the optimizer, our inferences about x_t from knowledge of y_t should be probabilistic, based on the ergodic distribution over the $[x^*(y_t) + L, x^*(y_t) + U]$ interval.

In the Appendix, we derive the ergodic distribution exploiting its invariance property. Under the assumptions above, the stable density $f(z)$ is piecewise linear if $\vartheta = 0$, piecewise exponential otherwise. The shape of the ergodic distribution depends on ξ, the ratio of the drift ϑ to uncertainty per unit time σ^2. A positive x^* drift (or a large δ) tends to concentrate the z distribution toward the lower boundary of a given inaction interval (see Figure 4), but the extent to which this occurs is decreasing in the degree of uncertainty about the fluctuations of the regulated process $\{z\}$. Intuitively, we expect z to be low if it usually drifts downward, but are less and less sure about this inference the larger the uncertainty. In the limit, the distribution tends to uniformity over the relevant action range when the ratio ϑ/σ^2 tends to plus or minus infinity; in this sense, one-sided rules of the (S,s) type, and their uniform ergodic distribution, emerge as a limit of the more general four-points rules examined here (see Section 3.3 below for a further discussion of this point).

Figure 4 displays ergodic densities for different drift-variance ratios (ξ) over a given inaction range, highlighting effects on the shape of $f(z)$ but disregarding the impact of ξ on optimal policies for given adjustment costs. The latter issue is illustrated in Figure 5. There, we plot the stable distributions of z for different values of ϑ, setting the action and return points at their optimal levels (given σ and the other parameters). In a smaller panel, we also plot against ϑ the action and return points, and the mean of the ergodic distribution.

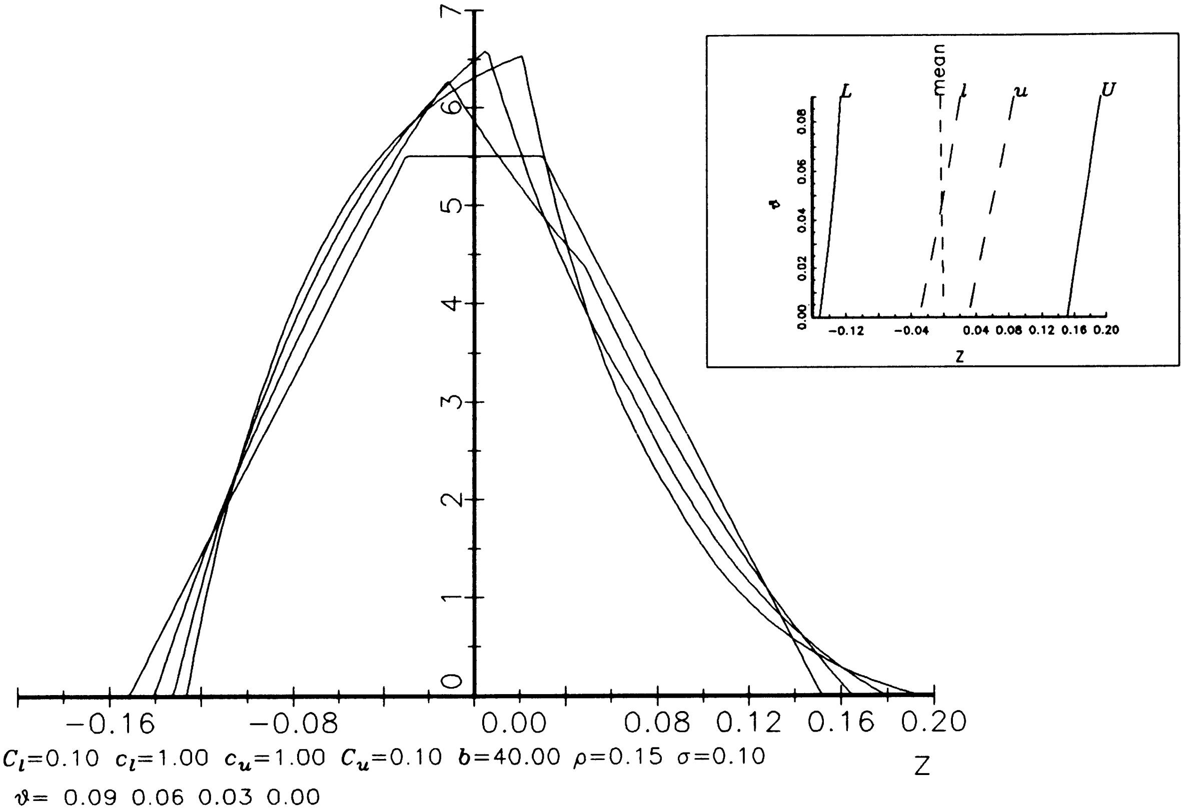
Figure 4: $f(z)$



It is apparent that, while the *shape* of the distribution is strongly influenced by the ratio of φ to σ, different drifts do not have a noticeable effect on the *mean* value of z (different σ would also, while affecting the shape of $f(z)$, have quite minor effects on its mean). A positive drift implies a tendency toward positive deviations, but induces the optimizer to correct those deviations sooner and by a larger amount; when the expected change of the instantaneously optimal level x^* is strongly positive (so that the drift of z is strongly negative), the difference between x and x^* is not allowed to become very negative—because such deviations would be expected not only to persist but to become larger in the absence of corrective action. As a consequence, although the ergodic distribution tilts heavily in the direction of L with a large negative drift in z (Fig. 4), the average deviation of x from x^* is hardly affected by the size of the drift, precisely because behavior is altered in ways that by and large tend to maintain z quite close to zero on average (Fig. 5).

In partial equilibrium, if the path of x_t^* can be taken as exogenous, this insight is quite general; although "small" adjustment costs and "small"

Figure 5: $f(z)$



7
6
5
4
3
2
1
0
−0.16
−0.08
0.00
0.08
0.12
0.16
0.20
Z
C_l=0.10 c_l=1.00 c_u=1.00 C_u=0.10 b=40.00 ρ=0.15 σ=0.10
ϑ= 0.09 0.06 0.03 0.00
L
mean
l
u
U
0.00
0.02
0.04
0.06
0.08
ϑ
−0.12
−0.04
0.04 0.08 0.12 0.16 0.20
Z

amounts of uncertainty are sufficient to generate "large" inaction ranges (see Dixit 1989e) and important dynamic deviations from the frictionless optimum, in the long run positive and negative deviations tend to cancel out. The effect of adjustment costs on long-run average deviations from the frictionless optimum is, therefore, one order of magnitude smaller than that on inaction ranges. In the context of the symmetric cost-minimization problem considered above, this is quite intuitive; the optimizer attempts to track the frictionless optimum as closely as possible, and deviations from it are equally penalized in both directions.[5]

2.4 APPLICABILITY OF THE RESULTS

Given the analytical expressions in the Appendix, numerical solution of (2.2) and (2.3) yield action and return points as functions of adjustment costs, uncertainty, concavity (summarized in b), and drift. We now need to discuss applicability of the simplified model above in specific examples, noting that exact solutions are typically available for constant elasticity (loglinear) models (Grossman and Laroque 1990, Bentolila and Bertola 1990, Bertola 1989a, Dixit 1989a) and that numerical solutions may be obtained, adapting the methods outlined above, for more complex functional assumptions as well.

The problem solved above is simplified in many respects. In particular, the basic framework is such that in the absence of adjustment costs the dynamic optimization problem would collapse to a sequence of static choices. This simplification is harmless if no intertemporal linkages other than adjustment costs are present in the case under study. We discuss the further simplifying assumptions we made in this context, before turning to a discussion of other intertemporal links.

Consider, for example, a firm's labor demand policy. In the absence of turnover costs, employment should be chosen to set labor's marginal revenue product equal to unit wages. This defines $x^*(y)$, with y including wages, prices of intermediate materials, productivity, and output prices. If x_t^* represents the logarithm of desired employment, Brownian motion dynamics may be a good approximation if the increments in the rate of growth of prices, wages, and productivity are approximately independent over time. The parameters and variables of the simplified optimization problem are then readily interpreted in terms of real-world

5. To the extent that average effects are relevant, however, functional forms different from the symmetric one considered here—asymmetries in adjustment costs, discounting, and drifts—all have a role in determining them (see Bentolila and Bertola 1990, Bertola 1989b). The positive effects of inflation on welfare identified by Diamond (1988) and Benabou (1989) in a search context depend heavily on sharp asymmetries in flow objective functions and on interactions through general equilibrium conditions.

quantities; lump-sum and proportional adjustment costs apply to proportional employment changes, and z measures log-deviations of employment from the level that would maximize the flow of operating cash flow $\Pi(.,.)$. If a year is the time unit, and the yearly wage of a unit of labor is the numeraire, then ρ, ϑ, and σ^2 are in time units; the concavity index b measures lost cash flows in the same units in which the yearly wage bill is measured; and C_u and C_l refer to the lump-sum cost, in those same units, to be paid when changing the logarithm of employment—namely, a *fraction* of the current wage bill has to be paid whenever employment is changed by any nonzero amount.[6]

In partial equilibrium models of the firm, the required rate of return ρ may well subsume all intertemporal aspects other than adjustment costs; thus problems of menu pricing, investment, and inventory management can be similarly framed in terms of the model proposed above. In many cases, however, intertemporal linkages would not disappear if adjustment costs were removed. For example, even when labor turnover is costless a firm should adopt forward-looking employment policies in the presence of learning-by-doing, or of strategic interactions with potential and actual competitors. More to the point, a consumer's portfolio and consumption choices are subject to the intertemporal budget constraint in the absence of adjustment costs.

In such situations, x_t^* should be understood to represent the state- and time-contingent choice that would be optimal if adjustment costs were removed while maintaining the other intertemporal linkages.[7] Such a x_t^* process need not be readily expressible as a function of *exogenous* y_t variables. For example, optimal consumption rules have not been derived for constant relative risk aversion utility under incomplete markets.[8] Consumption-portfolio problems, however, imply that some variable follows a martingale; if x_t^* corresponds to this variable, the assumption of Brownian motion dynamics is justifiable—to some extent—even when all exogenous variables are stationary. For example, if x_t^* represents desired consumption, it should be a martingale when the

6. Such an assumption may be realistic, in fact, if the lump-sum component of firing costs represents production lost because of strikes or other disruption of labor relations, and that of hiring costs represents workdays lost in training the new hirees.
7. When the additional source of intertemporal linkages is a budget constraint, wealth should be adjusted to allow for adjustment costs.
8. A consumer's infinite horizon, frictionless optimization problem could formally be expressed as in (2.1) by defining $\Pi(x_t, y_t) = U(x_t) - \lambda_t x_t$, where λ_t denotes the shadow value of wealth in the absence of transaction costs. An appropriate choice of frictionless price between the selling and buying price would guarantee that the budget constraint is satisfied along the optimal path in the presence of transaction costs. Still, the λ_t process cannot, strictly speaking, be an element of y_t (it is not exogenous when considering the adjustment policy).

Hall (1978) assumptions are satisfied. In a slightly more general framework, x_t^* may be taken to represent marginal utility.

We postpone further discussion of specific applications to Section 4 below, and to future research. In concluding our overview of microeconomic optimization techniques, we note that the simple quadratic-deviation model illustrates the general features of similar, more complex models, and provides a convenient starting point for moving on to aggregation and empirical work in the next sections. The policy followed by an optimizer in the framework explored above is expressed in terms of log-deviations of actual from "desired" state variables, an intuitively appealing rule of thumb. When a solution can be found for more general and sophisticated models, it must be quite similar in character to the one we discussed—though trigger and return points may be defined in a space that is not independent of the structural parameters we subsume in $x^*(y_t)$.[9]

3. *Macroeconomics*

Though microeconomic agents faced by kinked adjustment cost functions often choose inaction and may react disproportionately to innovations when they do act, neither inaction nor instantaneous sharp reactions are typically observed at the aggregate level. This has led macroeconomists to devise assumptions that would make smooth, partial adjustment optimal at the microeconomic level as well—namely, to assume unit adjustment costs to be increasing in the speed of adjustment, assumptions hardly justifiable at the microeconomic level. As we show in this section, however, macroeconomic data may well be consistent with realistic microeconomic behavior. To reconcile microeconomic behavior and aggregate evidence, and to understand how and to what extent microeconomic rigidities work their way into the macroeconomy, it is crucial to assess the degree of coordination of individual actions at all points in time (Caballero and Engel 1989b,c).

Two polar cases highlight the importance of aggregation and coordination issues. At one extreme, if the individuals in a group are identical, in all respects the aggregate should behave like each of the individuals. Symmetric, perfectly bunched equilibria of this type have been studied in static macroeconomic models of sticky prices (e.g., Mankiw 1985, Akerlof and Yellen 1985, Blanchard and Kiyotaki 1987, Rotemberg 1987); in a dynamic setting, perfect bunching in state space would imply that

9. For example, in the loglinear models of Grossman and Laroque (1990) and Bertola (1989a), trigger and return points are defined in terms of *marginal* contributions of the endogenous state variable to flow benefits.

all units take similar actions at the same time. At the other extreme, however, if a large group of agents follow one-sided rules and are uniformly spread in the state space, their actions are perfectly uncoordinated and the aggregate is unaffected by microeconomic rigidities, fully flexible, and smooth (Caplin and Spulber 1987).

In light of the stark contrast between these extreme cases, it is essential to model the determinants of cross-sectional distributions over the relevant state space. Blinder (1981) stresses this insight in his treatment of inventories, and Caplin (1985) takes the initial steps for a formal and systematic study of joint movements by heterogenous units that follow intermittent adjustment policies. The Caplin and Spulber (1987) steady state model is the first analytical study of the role of cross-sectional distributions. Recent work by Caballero and Engel (1989a,b,c) has provided a suitable framework for a formal study of nonsteady state aggregate dynamics in terms of the behavior of cross-sectional distributions. We review and extend these technical developments below.

A complete treatment of the endogenous determination of cross-sectional distributions should take into account strategic and general equilibrium interactions between individual decisions, structural dissimilarities across units, and imperfect cross-sectional correlation of stochastic factors affecting individual units. Results on the role of structural heterogeneity are briefly reviewed in (3.3) below, and the concluding Section 5 discusses the role of interactions in the type of models we study. The present section, however, focuses on the correlation of factors affecting individual units. We model the processes taken as exogenous at the individual level in terms of two sources of uncertainty: one common to all units in the group under consideration ("aggregate shocks"), the other uncorrelated across units ("idiosyncratic shocks"). We further distinguish among two types of aggregate shocks: "large" or "structural" ones—like an oil shock or a permanent change in monetary policy rules or wage-setting practices—and "continuous" or "smooth" ones—the common component of ongoing fluctuations.

The basic lesson of the models below is that idiosyncratic shocks tend to smooth out microeconomic rigidities by spreading agents in state space, while aggregate shocks (especially large ones) tend to coordinate individual units' actions, thus allowing microeconomic inaction to affect the dynamic behavior of aggregate time series. This point is illustrated in three stages. We first consider a model in which all uncertainty is idiosyncratic. In this framework, we discuss conceptual links between *probability distributions* at the individual unit's level and *empirical* or *cross-sectional* distributions at the aggregate level, and we discuss the role of microeconomic frictions and idiosyncratic sources of uncertainty in determining the aggregate dynamic response to a once-and-for-all aggregate shock. This thought experiment serves as an introduction to treatment of ongoing aggregate uncertainty, which we discuss in the second stage in the well-explored case of one-sided (S,s) rules and, in the third stage, in the general band-policy case. A general stylized model of unsynchronized band-policy adjustment makes it clear that, even when no large aggregate shocks occur, the effect of microeconomic frictions on aggregate dynamics is an increasing function of the relative importance of common and idiosyncratic uncertainty.

3.1 IDIOSYNCRATIC UNCERTAINTY AND THE AGGREGATE

Consider a large number n of economic agents indexed by i, $i = 1, \ldots, n$, and suppose that the path of each agent's endogenous state variable (e.g., capital, cumulative orders, prices, workers, cash balances, etc.) would be described by

$$x^*_{it} = \vartheta t + \sigma W_{it}, \tag{3.1}$$

in the absence of adjustment costs. Here, ϑ is a constant aggregate drift and W_{it} is a stochastic process whose increments are independent across i (idiosyncratic) as well as across time.

For simplicity, we shall conduct our analysis in terms of a discrete time, discrete state-space Markov chain equivalent of the continuous processes assumed in Section 2. The discrete representation of (3.1) is

$$x^*_{it+dt} = \begin{cases} x^*_{it} + \eta, & \text{with probability } p \equiv \frac{1}{2}\left(1 + \frac{\vartheta dt}{\eta}\right); \\ x^*_{it} - \eta, & \text{with probability } (1 - p) \equiv \frac{1}{2}\left(1 - \frac{\vartheta dt}{\eta}\right). \end{cases} \tag{3.2}$$

If we let $\eta \equiv \sigma\sqrt{dt}$, as $dt \rightarrow 0$ this process converges to Brownian motion with drift, consistently with the specification of Section 2 above (see, e.g., Ross 1983).

Let all agents follow identical (L,l,u,U) control policies of the type discussed in Section 2. We again denote by x_{it} the actual value of agent i's state variable, and by z_{it} its deviation from the level that would be optimal in the absence of adjustment costs:

$$z_{it} \equiv x_{it} - x^*_{it}.$$

To economize on notation, we do not explicitly allow for depreciation of the actual stock here; with $\delta = 0$, $\vartheta = \vartheta^*$ denotes the drift of the desired stock as well as that of its deviation from the actual stock.

Though both x^*_{it} and x_{it} are nonstationary, z_{it} only takes values on a bounded state space $[l,U]$ if each unit follows the general band-policy rules of Section 2. Letting $k \equiv (U - L)/\eta + 1$ be an integer, the discrete representation of the relevant state space is a $k \times 1$ vector

$$s \equiv [L, L + \eta, \ldots, l, \ldots, u, \ldots, U - \eta, U]'.$$

Now denote with a $1 \times k$ vector f_{i0} the probability density for each unit's position in state space at time zero, z_{i0}:

$$f_{i0} \equiv [f_{i0}(L), f_{i0}(L + \eta), \ldots, f_{i0}(l), \ldots, f_{i0}(u), \ldots, f_{i0}(U - \eta), f_{i0}(U)].$$

For example, if the position of unit i is known exactly, only one element of f_{i0} is positive and equal to one. Given the time-zero information, the relationship in (3.2) and the band-policy adjustment rule imply that probability densities at successive instants are linked by the recursion

$$f_{it} = f_{it-dt}P, \tag{3.3}$$

where P (equal for all units) denotes the transition matrix over s implied by the x^*_{it} transition probabilities in (3.2) and by the (L,l,u,U) adjustment rule that maps x^*_{it} into z_{it}:

$$P = \begin{array}{c} \\ L \\ L+\eta \\ L+2\eta \\ \vdots \\ U-2\eta \\ U-\eta \\ U \end{array} \begin{array}{c} \begin{array}{ccccccccccccc} L & L+\eta & L+2\eta & \ldots & l & \ldots & u & \ldots & U-2\eta & U-\eta & U \end{array} \\ \begin{bmatrix} 0 & 1-p & 0 & \ldots & p & \ldots & 0 & \ldots & 0 & 0 & 0 \\ p & 0 & 1-p & \ldots & 0 & \ldots & 0 & \ldots & 0 & 0 & 0 \\ 0 & p & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & . & \vdots & . & \vdots & . & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & 1-p & 0 \\ 0 & 0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & p & 0 & 1-p \\ 0 & 0 & 0 & \ldots & 0 & \ldots & 1-p & \ldots & 0 & p & 0 \end{bmatrix} \end{array}$$

We can iterate (3.3) forward from time zero to obtain

$$f_{it} = f_{i0}P^t,$$

and it is easy to show that the Markov chain under consideration is ergodic: starting from any f_0, f_t eventually converges to a unique, invariant steady-state probability density f, the discrete counterpart of the ergodic density discussed in Section 2, which satisfies $f = fP$, $\Sigma_1^k f_j = 1$, and is the same for all units.

From the macroeconomic point of view, we are not concerned with the position or probability density of each individual unit. Rather, we would like to characterize the *empirical, cross-sectional* distribution, i.e., the realization of all z_{it} positions at each point in time, which we denote with $\tilde{f}_t$. The elements of this $1 \times k$ vector measure the fraction of units located at every point in state space at time t. Defining the aggregate time series, X_t, as the mean of the actual positions of the x_{it} (e.g., actual capital, price, etc.):

$$X_t = \frac{1}{n} \sum_{i=1}^{n} x_{it}, \tag{3.4}$$

the mean deviation of agents from their frictionless position is simply given by

$$\bar{z}_t = \tilde{f}_t s,$$

Recalling that $x_{it} = x^*_{it} + z_{it}$, we obtain

$$X_t = \vartheta t + \bar{z}_t. \tag{3.5}$$

Our interest in individual probability distributions arises from the fact that when the number n of units is large, conceptual links can be established between the Markov chain relevant to an individual's probability density on the one hand, and a vector difference equation describing the path of the whole cross section on the other. Specifically, if we assume that the initial empirical distribution $\tilde{f}_0$ is given and that each unit's initial probability density f_{i0} is the same (f_0), and we consider a larger and larger n, than $\tilde{f}_0$ can be made arbitrarily close to f_0, and $\tilde{f}_t$ can be made arbitrarily close to $f_t = f_0 P^t$. This is a simple application of the Glivenko-Cantelli theorem (see, e.g., Billingsley 1986); heuristically, when the total number of units n tends to infinity the number of units in each state-space location becomes large enough that, by a strong law of large numbers, the probabilities associated to each position in state space coincide with the actual fractions of units located in the same states. As this happens at all point in times, the fraction of units moving between positions in the state space must coincide with the probabilities in the units' transition matrix.

This insight makes it possible to characterize aggregate dynamics when adjustment costs are present ($L < U$) and a large group of units are distributed over $[L, U]$ in some arbitrary fashion. By (3.5) and an application of (3.3) to the empirical distribution, the aggregate follows

$$X_{t+dt} = \vartheta t + \tilde{f}_{t+dt}s = \vartheta t + \tilde{f}_t Ps.$$

Thus, as long as $\tilde{f}_t \neq f$, the dynamic behavior of X_t differs from that of a frictionless economy.

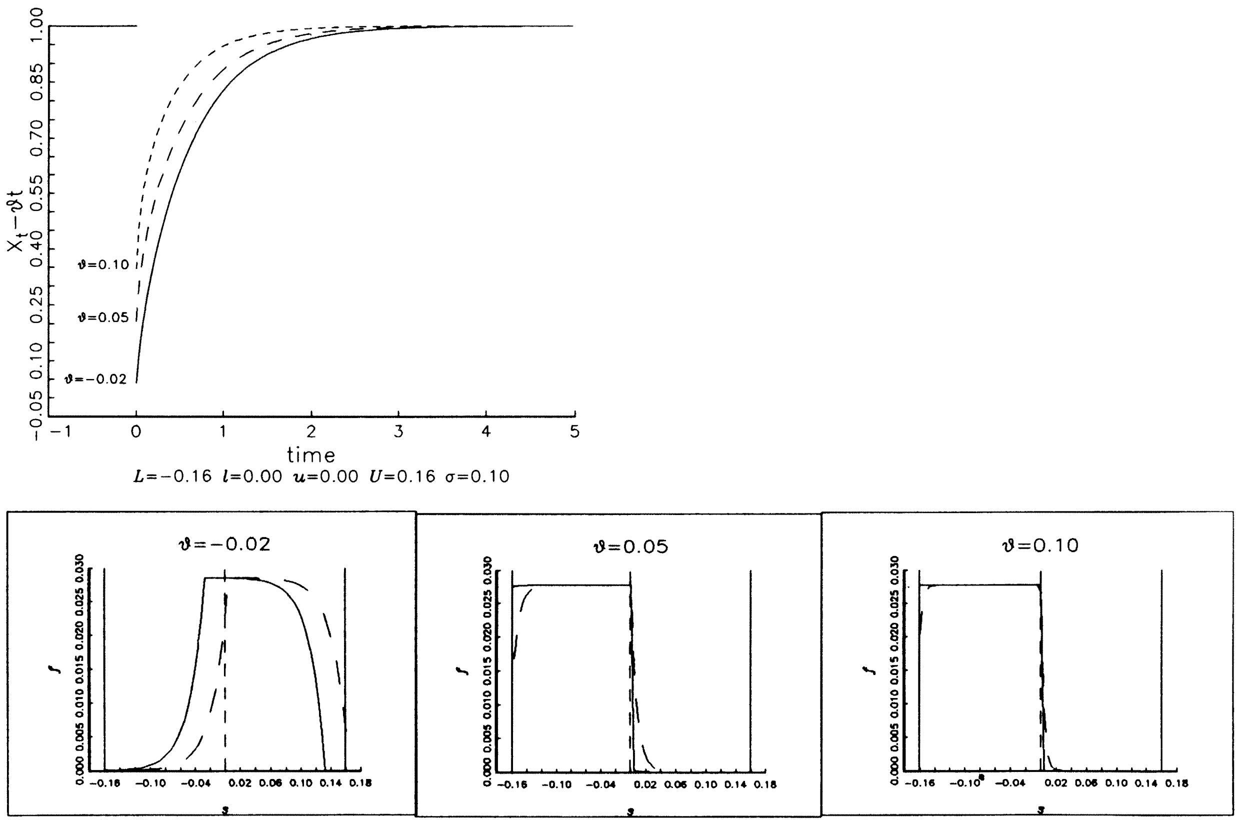
The Markov chain describing the probability density of individual units is ergodic and, if n is large, the same is true of the *empirical* distribution. "Ergodicity" of the empirical distribution means that the *actual realization* of the cross-sectional distribution becomes stationary; thus, if only idiosyncratic sources of uncertainty are present, then $\bar{z}_t$ eventually converges to the constant $z_\infty \equiv f\,s$ starting from *any* initial distribution $\tilde{f}_0$. We normalize this constant to zero in what follows.[10] It is important to make it clear that individual z_{it} deviations from the frictionless optimum are in general not zero in the long-run steady state, and convergence of $\bar{z}_t$ does *not* mean that once the steady state is reached microeconomic activity should cease. Rather, individual units continuously move and change their relative positions; but in steady state the *fraction* of agents that leave each position is equal to the fraction that arrives to it. Outside the steady state the empirical density $\tilde{f}_t$ changes over time, and so does its first moment, $\bar{z}_t$.

It is interesting to study in some detail the role of ϑ, the aggregate drift, in determining the size and character of the aggregate impulse response after a once-and-for-all structural change that moves the empirical distribution away from its steady state. We showed in Section 2 that when the ratio of drift to variance is large, the long-run probability distributions for the position in state space of an individual unit's z_{it} are skewed. When dealing with a large number of similar individuals, their empirical distribution can be similarly characterized by the results above. In a menu-cost pricing framework, for example, if trend money growth has been strongly positive then we would expect relatively many units to be near the point that triggers price increases; few should be close to the point that triggers price reductions. Consider now a sudden, temporary acceleration of money growth or an unanticipated increase in the money level; this would trigger price adjustment by many units, and elicit a small output response. Conversely, a negative monetary surprise would trigger few downward price adjustments and have a large, negative impact on output.

This insight is illustrate in Figure 6 (Tsiddon 1988 makes a similar point). Starting from the stable distribution, we plot the aggregate response to a once-and-for-all aggregate shock of size 0.03; such a shock

10. z_∞ should in fact be quite close to zero on the basis of the microeconomic results of Section 2.3.

Figure 6



$X_t - \vartheta t$
1.00
0.85
0.70
0.55
0.40
0.25
0.10
-0.05
ϑ=0.10
ϑ=0.05
ϑ=-0.02
-1
0
1
2
3
4
5
time
L=-0.16 l=0.00 u=0.00 U=0.16 σ=0.10
ϑ=-0.02
ϑ=0.05
ϑ=0.10
f
s
0.000 0.005 0.010 0.015 0.020 0.025 0.030
-0.16 -0.10 -0.04 0.02 0.06 0.10 0.14 0.18

could be modeled either as a discrete increase in the level of all x^*_{it} by the same amount, which would bunch a discrete mass of units at the l return point, or as an accumulation of small shocks in infinitesimal time, which by inducing units to act sequentially would preserve the relative positions of the adjusting units. The distinction is not important for the issue at hand. We choose the latter alternative—the smaller panels in the figure display the resulting cross-sectional distribution of units just after the shock as a solid line, and the stable distribution to which they will eventually return as a dashed line.

The solid, dashed, and dotted lines in the main panel of the Figure 6 represent detrended X_t for ϑ's equal to -2, 5, and 10 percent, respectively. For all three cases, we assume $L = -U$, $l = u = 0$ (adjustment takes z to zero from symmetric trigger points). The assumption of similar behavior in the presence of different drifts is obviously unwarranted; if adjustment costs are the same in the three environments, the drift affects the location of trigger and return points (see Figure 3 above). However, microeconomic optimality of adjustment policies has second-order importance for the aggregate response to a one-time common shock, starting from the steady state.[11]

On the one hand as argued above, a positive shock has a large impact on $\bar{z}_t$ (i.e., a smaller impact on X_t) when the drift is negative; in steady state, only 1.8% of the units are close enough to the lower trigger point for a 3% shcok to induce them to adjust x_{it} to x^*_{it}. On the other hand, when $\vartheta = 10$ almost all the units are between L and l in steady state, a 3% shock triggers action by 7% of the units, and as a result one-third of the aggregate shock passes through at time zero. In the aftermath of the one-time shock, all three paths converge exponentially back to the steady state. The speed at which this happens is an increasing function of σ, i.e., of the size of the idiosyncratic shocks reshuffling the cross-sectional distribution (see Caballero and Engel 1989a, and our discussion below).

3.2 ONGOING AGGREGATE UNCERTAINTY

Aggregate impulse responses to one-time shocks highlight important insights, but fall short of providing operational tools for an analysis of aggregate data. For this purpose, it is necessary to model explicitly the probability structure governing aggregate uncertainty and to replace

11. In the experiment considered, the initial aggregate response is equal to the product of the size of jump and the fraction of agents exercising control: given a drift and variance, if the optimal jump were *larger* than the one arbitrarily assumed in Figure 6 then the ergodic distribution associated to the optimal control would concentrate *fewer* units in the neighborhood of the trigger point.

equation (3.1) with a system incorporating a stochastic aggregate component, A_t:

$$x_{it}^* = A_t + \sigma_I W_{it}$$
$$A_t = \vartheta t + \sigma_A W_{at} \tag{3.6}$$

Here, W_{it} and W_{at} are independent random variables whose increments have zero mean and unitary variance per unit time.

If ongoing aggregate uncertainty is present, it does not wash out when averaging across units, no matter how many. Thus, in the absence of adjustment costs the aggregate defined in (3.4) would now be stochastic:

$$X_t^* = A_t = \vartheta t + \sigma_A W_{at}. \tag{3.7}$$

Taking the fluctuations of W_{at} as exogenously given, we would like to study how the actual path X_t differs from the path in (3.7) in the presence of microeconomic adjustment costs. The same steps that led to equation (3.5) establish that $X_t = X_t^* + \tilde{f}_t s$, and we proceed to study the evolution of $\tilde{f}_t$.

From the point of view of each individual unit, the source of uncertainty is irrelevant and the equations in (3.6) can be combined to yield

$$x_{it}^* = \vartheta t + \sigma W_{iat}$$

where $\sigma = \sqrt{\sigma_A^2 + \sigma_I^2}$ and W_{iat} is a random variable with the same univariate probability structure as W_{it} and W_{at}. This equation is analogous to (3.1) above. Thus, for any i the evolution of f_{it} can be characterized along the lines of Section 3.1. The fact that now the innovations in σW_{iat} are correlated across units (with correlation coefficient equal to σ_A/σ) is irrelevant when considering an individual unit.

However, the source of uncertainty has a crucial role in determining cross-sectional distributions. To see this, consider the extreme cases: If there were only idiosyncratic uncertainty ($\sigma_I > 0$, $\sigma_A = 0$), the cross section would be closely related to the probability density of individual units, as shown in Section 3.1. But if only aggregate uncertainty existed ($\sigma_I = 0$, $\sigma_A > 0$), then the probability density of a single unit and the cross section would bear no relationship to each other. For example, if all units start together and there is no idiosyncratic uncertainty, the empirical distribution remains concentrated in a spike wandering through the state space forever, driven by aggregate shocks—although probabilistic statements about an individual unit's position should ultimately be based on the ergodic distribution, as before. In the context of menu-cost pricing, Caplin and Leahy (1990) construct a model in which a nondegenerate and self-replicating *family* of empirical distributions exists in the absence of idiosyncratic uncertainty, and yields a convenient statistical representation for the *aggregate* price process.

In general, no cross-sectional distribution is invariant to aggregate shocks: thus, the empirical distribution cannot converge to a limit. In specific applications of this general principle, it is important to take into account not only the relative importance of aggregate shocks, but the form of adjustment policies and the character of exogenous processes. We proceed to highlight the latter insights by a review of available results on one-side (S, s) adjustment policies, perhaps the best known among kinked adjustment cost models. This case provides an exception to the general rule: If no large shocks occur, and ongoing uncertainty is continuous and monotonic, a steady-state empirical distribution exists and a large group of individuals will converge under fairly weak conditions.

3.3 ONE-SIDED (S,S) RULES

If the path of x^*_{it} is monotonic, i.e., all its changes are in one direction, adjustment under lump-sum and proportional adjustment costs can be described by two points in state space, customarily denoted s and S, such that when $z_{it} = s$ action brings it instantaneously to S. In the framework of Section 2, the optimal solution converges to such one-sided rules when $\xi \equiv 2\vartheta/\sigma^2$ is large in absolute value. The results reviewed below should be understood to apply in situations where the drift dominates the variance of x^*_{it}. While optimizing agents would also perform unidirectional corrective actions when adjustment in the other direction is prohibitively costly (e.g., the irreversible investment case of Pindyck 1988 and Bertola 1989), the results below would *not* be applicable as long as x^*_{it} movements occur in both directions.

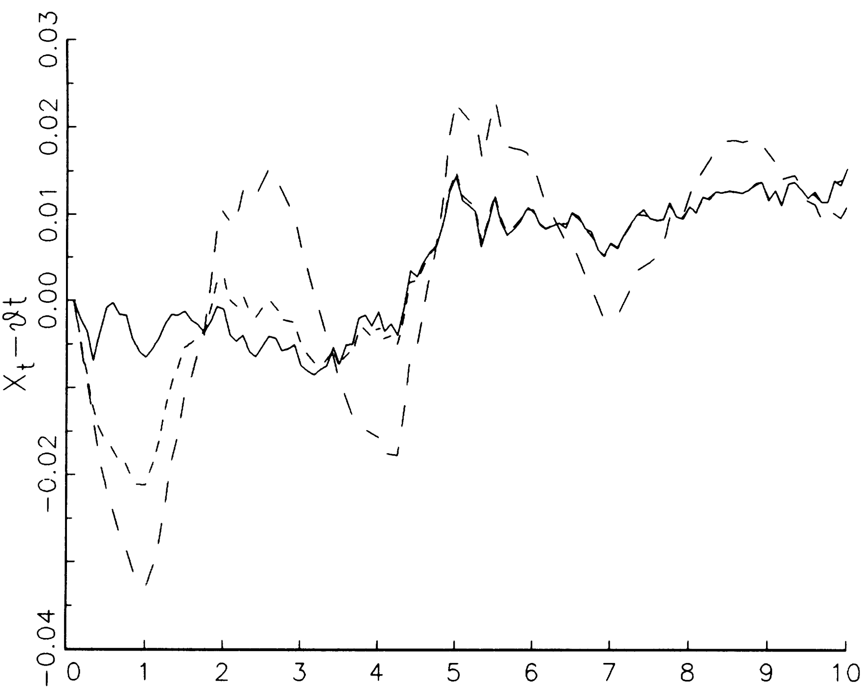
Caplin and Spulber (1987) discuss a striking feature of the one-sided model. They present an example in which no idiosyncratic uncertainty exists, s and S are the same across units, and the initial cross-sectional distribution of prices is uniform on the $[S, s)$ interval; and they show that the uniform cross-sectional distribution is unaffected by monotone, continuous increases in the quantity of money. Thus, $\bar{z}_t$ is identically constant at all t (and equal to zero under the normalization $s = -S$), and microeconomic frictions have no effect on the aggregate path. To see this, suppose the z_i deviations are uniformly distributed on $(-S,S]$, and let the aggregate move continuously by ΔA_t over an interval of time Δt, where Δ denotes change in the variable. This shifts the whole distribution down by $\Delta A_t/2S$, leaving an empty space of equal length on the top—but units that were within distance $\Delta A_t/2S$ of s before the shock have moved into the space at the top, thus preserving the uniform empirical distribution: $\bar{z}_t$ remains unchanged, and $\Delta X_t = \Delta X_t^* = \Delta A_t$.

The ergodic probability density of each unit's z_i, f, is also uniform on $(s,S]$ when exogenous shocks are monotonic and adjustment is of the (S,s) type. Thus, *in steady state* the probability density of an individual unit and the empirical cross section coincide, as they did in Section 3.1, even in the presence of aggregate shocks. Outside the steady state, however, the resemblance fades. Caballero and Engel (1989a,b) show that if nonstationary idiosyncratic shocks are added to Caplin and Spulber's model, the empirical distribution of the z_is converges to a uniform distribution starting from any initial distribution, but more slowly than each unit's probability density converges to the ergodic one; aggregate shocks affect all units x_{it}^* equally, and do not *mix* their cross-sectional distribution. Only idiosyncratic shocks aid convergence of the empirical distribution to the stationary, uniform one, while convergence of the probability density of a single unit depends on the total uncertainty it faces, including the aggregate component.

Even though aggregate shocks do not aid convergence, they do affect the mean of the cross section, $\bar{z}_t$, outside the steady state. In Figure 7 we illustrate this insight by plotting the detrended path of X_t in the aftermath of a "large" aggregate shock, namely a structural change that doubles the absolute value of both S and s (from $\tau_0 = 0.08$ to $\tau_1 = 0.16$). The solid line refers to the frictionless case $S = s = 0$, while the short- and long-dashed lines refer to cases in which idiosyncratic uncertainty is, respectively, large and small. After the structural shock, the initial cross-sectional distribution is uniform on a subinterval of the new $(s, S]$ interval. The cross section eventually converges to a uniform on the whole new $(s, S]$ interval; in the shorter run, however, microeconomic rigidities have a substantial effect. We can see that while the economy with large idiosyncratic uncertainty (short dashes) converges relatively quickly, when idiosyncratic uncertainty is small (long dashes) convergence is slow and departures from the frictionless path can be long-lasting.

Several other extensions have been considered by Caballero and Engel (1989a,b,c) in the context of one-sided adjustment. On the one hand, heterogeneous behavior across units (needed for Caplin and Spulber's steady state to be a relevant benchmark) can result from structural heterogeneity rather than from differences in the exogenous processes' realizations; if units' (S,s) bands are different, the aggregate dynamics replicate the frictionless path whenever the cross-sectional distribution of $\frac{S_i - s_i}{2} - \frac{z_i}{S_i - s_i}$ is uniform on the unit interval, and these normalized deviations converge to uniformity even when idiosyncratic uncertainty is negligible. On the

Figure 7



ϑ=0.05 τ_1=0.16 τ_0=0.08 σ_A=0.02 σ= 0.01 0.04

other hand, the uniform distribution is not invariant to "large" aggregate shocks, i.e., discrete changes that discontinuously alter all units' position. Since "small" shocks have a limited role in the one-sided case, this is a natural framework for exploring the consequences of *recurring*, probabilistic regime changes. Suppose, for example, that A_t—following a more general process than the one in (3.6)—may at times move instantaneously by the finite amount ΔA_t; even if the cross section were uniform before the shock, the discrete shift would concentrate the finite fraction $\Delta A_t/(2S)$ in a spike at the single point S. In the aftermath of such a large shock, idiosyncratic shifts would spread the spike and the cross section would tend toward the uniform, steady-state distribution—but further large shocks would undo the gains in that direction and rebunch some agents anew. In this situation, there would be a continuous tension between the endogenous tendency toward uniformity, due to heterogeneity, and relatively infrequent structural changes that prevent the cross section from ever reaching a steady state in which the path of X_t coincides with that of X_t^*.

3.4 GENERAL POLICY RULE

When exogenous events can make adjustment in either direction desirable, no cross-sectional distribution is invariant to aggregate shocks even when only "small" ones can occur. To some extents the insights of the previous section are still useful in this case: a large group of individuals subject to idiosyncratic uncertainty (or heterogeneous in other respects) will display a tendency to converge toward the cross-sectional distribution that would be stable in the absence of coordinating aggregate shocks. This tendency, however, is hampered by *ongoing* common shocks; in the model we develop below, the relative strength of the forces at work in the two directions is summarized by the ratio of the variance per unit time of the idiosyncratic and common components of uncertainty. The tension between aggregate and idiosyncratic shocks noted in the one-sided (S,s) case arises here even when no "large" shocks occur.

Consider again a large group of individuals following discontinuous adjustment rules, with the same trigger and return points (L,l,u,U), and assume both idiosyncratic and aggregate uncertainty to be present as per equations (3.6) and (3.7). We need to extend the discrete time representation in equation (3.2) to take both stochastic components into account. Let A_t be a simple binomial random walk,

$$A_{t+dt} = \begin{cases} A_t + v, & \text{with probability } q; \\ A_t - v, & \text{with probability } 1 - q. \end{cases} \tag{3.8}$$

We shall refer to positive aggregate shocks as "booms," and to negative ones as "recessions." Assuming

$$v = \sigma_A \sqrt{dt}, \qquad q \equiv \frac{1}{2}\left(1 + \frac{\vartheta dt}{v}\right) \tag{3.9}$$

the aggregate process converges to Brownian motion with drift ϑ and standard deviation σ_A as $dt \to 0$.

Now let us write the innovation of each x^*_{it} process *conditional* on whether the economy is in a boom or a recession:

$$\begin{aligned} \Delta A_{t+dt} = v &\Rightarrow x^*_{it+dt} = \begin{cases} x^*_{it} + \eta & \text{with probability } p_b \\ x^*_{it} - \eta & \text{with probability } 1 - p_b \end{cases} \\ \Delta A_{t+dt} = -v &\Rightarrow x^*_{it+dt} = \begin{cases} x^*_{it} + \eta & \text{with probability } p_r \\ x^*_{it} - \eta & \text{with probability } 1 - p_r \end{cases} \end{aligned} \tag{3.10}$$

where we define:[12]

$$p_b = \frac{1}{2}\left(1 + \frac{\sigma_A}{\sigma}\right), \quad p_r = \frac{1}{2}\left(1 - \frac{\sigma_A}{\sigma}\right), \quad \eta = \sigma\sqrt{dt}, \quad \sigma = \sqrt{\sigma_A^2 + \sigma_I^2}. \tag{3.11}$$

Aggregating over units and over time, it is straightforward to verify that the aggregate stochastic process and each of the individual processes all converge to Brownian motion as $dt \to 0$, and that the increments of A_t and of $(x^*_{it} - A_t)$ are independent for all i.

As long as we *condition* on the realization of the aggregate, we can again translate probability statements at the individual unit's level into statements about the cross-sectional behavior of a large number of units. From this point of view, p_b represents the fraction of the many units in each state that receive a positive shock during a boom, and p_r represents the analogous fraction during a recession. The expressions above for p_b and p_r simply reflect the fact that more units are affected by positive shocks during booms than during recessions, and that the difference between booms and recessions becomes more pronounced as the variance of aggregate shocks rises relative to that of idiosyncratic shocks.

Given the transition probabilities in (3.10), the form of the adjustment policy, and the realization of aggregate shocks, it is possible to characterize the evolution of the cross-sectional distribution of the z_{it}s, and of $\bar{z}_t$. In contrast to Section 3.1 above, where only idiosyncratic shocks were present, the change of empirical distribution at successive instants depends on whether a "boom" or "recession" is occurring. Given $\tilde{f}_t$, we have

$$\tilde{f}_{t+dt} = \begin{cases} P_b\tilde{f}_t & \text{during a boom} \\ P_r\tilde{f}_t & \text{during a recession} \end{cases} \tag{3.12}$$

where P_b and P_r are the transition matrices during booms and recessions for the individual unit's z_{it}; these matrices can be written out, using (3.10) and the band-policy parameters, in the form of P above—with (respectively) p_b and p_r in place of p. The difference between the elements of P_b and P_r is increasing in the ratio of σ_A to $\sigma = \sqrt{\sigma_I^2 + \sigma_A^2}$; defining

$$\gamma = \frac{\sigma_A}{\sigma},$$

12. Note that, given dt, it would be necessary to use the alternative definition $v \equiv \sqrt{\sigma_A dt + \theta^2(dt)^2}$ for consistency across equations (3.8–3.11). The $(dt)^2$ term has no role in the continuous limit if $\sigma_A > 0$; still, it is necessary to take it into account when using a finite dt—as we do in Section 4 below.

it is straightforward to verify that $P_b - P_r = \gamma D$, for D a matrix whose elements are 0, 1, or -1.

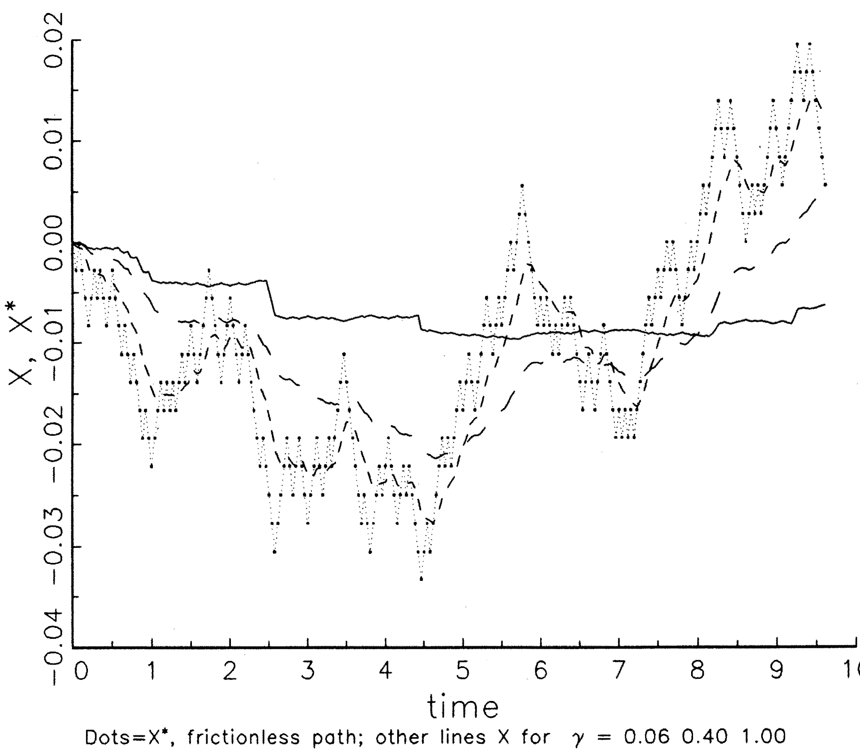
Now, any instant is a boom with probability q, a recession with probability $(1 - q)$. Iterating the transitions forward, we get

$$\tilde{f}_{t+dt} = \tilde{f}_0 \prod_{h=dt}^{t+dt} P_h, \; P_h \quad \begin{cases} P_b & \text{with probability } q \\ P_r & \text{with probability } (1-q) \end{cases} \forall h \tag{3.13}$$

where P_h denotes the realization of the transition matrix at time h. Since P_h alternates randomly between the two values P_b and P_r, as of time zero $\tilde{f}_t$ is a random vector for all t if $\sigma_A > 0$; the empirical distribution does not converge to a steady state. Consequently, $\bar{z}_t$ fluctuates forever, reflecting the impact of microeconomic frictions on the dynamics of aggregate variables.

Figure 8 illustrates how aggregate dynamics depend on the value of γ. The dotted line plots a frictionless aggregate sample path (the accumulated aggregate shocks, A_t or X_t^*); this would be the path of the endogenous variable X_t if no adjustment costs were present. The other lines plot the aggregate path in the presence of adjustment costs and of idiosyncratic uncertainty. The variance of aggregate shocks, the realization of A_t, and adjustment costs are the same for all paths. For each γ value, total uncertainty faced by individual units is σ_A/γ, and we compute the optimal adjustment policy for this value of σ, keeping all other parameters constant; the larger is γ, the narrower is the inaction range. We start each path assuming that the initial empirical distribution is the individual unit's ergodic one, and we use (3.13), (3.5), and the definition of $\bar{z}_t$ to plot the aggregate path. Although the inaction range is wider when γ is small, more uncertainty at the individual level unambiguously implies that units change their prices more often; if this were not the case, i.e., if the barriers were so much widened by higher uncertainty as to imply unchanged or even lower average adjustment costs per unit time, concavity of the flow benefit function $\pi(.)$ would imply large flow losses. Larger idiosyncratic uncertainty implies that reshuffling of the cross-sectional position of individual units is faster, and that the cross-sectional distribution is less sensitive to aggregate shocks; thus, the smaller is γ, the closer the aggregate tracks the frictionless path. In Figure 8, when aggregate uncertainty accounts for only 6% of the uncertainty faced by each individual unit ($\gamma = 0.06$), the aggregate path responds promptly to almost every aggregate innovation; when $\gamma = 0.40$, the aggregate path is quite sluggish and smooth instead; and when $\gamma =$

Figure 8



1, the aggregate path is approximately constant for long periods of time, reflecting the inaction at the individual level, and displays sharp but infrequent movements.

It is apparent from the figure that, depending on the value of γ, the model may be able to explain the dynamic patterns of macroeconomic data without resorting to the microeconomically unrealistic assumptions of conventional representative agent/convex adjustment cost models. The parameters of a model taking into explicit account microeconomic inaction and idiosyncratic uncertainty are "deeper" than those of dynamic optimization models based on ad hoc functional forms. In particular, a crucial role is assigned to the relative importance of common and idiosyncratic shocks, indexed by γ. Information about this parameter in different circumstances, countries, and periods should be extremely important in macroeconomic applications. Preliminary steps in the direction of empirical work are taken in the new section.

4. *Empirical implementation and durable goods consumption*

Empirical work on dynamic problems at all levels of aggregation typically adopts strictly convex adjustment cost functions, most often quadratic ones. This assumption yields easily estimable partial-adjustment dynamic relationships, and linear-quadratic models are considerably simpler than the more realistic microeconomic models reviewed in Section 2. The structural interpretation of partial-adjustment coefficients, however, is often unclear. Autoregressive representations—or more generally, the covariogram—are a convenient way to describe the data but do not provide an *economic* interpretation of their dynamics. Blinder (1981), Bar-Ilan and Blinder (1987), Hamermesh (1989), and others have noted the tension between empirical tractability and microeconomic realism. Still, it has proven very difficult to interpret available data (aggregated over individuals, over time, and over heterogeneous endogenous variables) in terms of optimal microeconomic behavior.

This section uses the stochastic aggregation model of Section 3 to study expenditure on durable goods in the United States. We aim to illustrate the explanatory power of the framework we propose, rather than provide a detailed study of the many issues involved. Consumer durables are a natural candidate for a first application of the techniques we propose. Mankiw (1982) finds these data in gross violation of the restrictions imposed by frictionless optimization in a permanent income framework, and the stock-adjustment model estimated by Bernanke (1985) does not succeed in rationalizing the dynamics of the data in an optimization framework. Caballero (1990) shows that the data could be interpreted in terms of different reaction lags to innovations across consumers, without violating the basic permanent income hypothesis in the long run. Grossman and Laroque (1990), Bar-Ilan and Blinder (1987), and Lam (1989) note the realism of discontinuous adjustment models in the context of individual durable goods purchases and discuss their implications for aggregate expenditure, without, however, addressing the problem of stochastic aggregation.

4.1 METHODOLOGY

Our framework lends itself naturally to an integrated treatment of data at different levels of aggregation. Here, however, we use the tight stochastic specification in 2.2 above to interpret aggregate time series only, seeking a structural interpretation of dynamic relationships between endogenous and exogenous variables at the aggregate level. For expositional clarity, we discuss empirical problems and solutions in three separate steps.

Step A: inference about the frictionless model. The difference between an endogenous aggregate state variable (X_t) and its hypothetic frictionless counterpart (X_t^*) is the basic determinant of the model's dynamics. Of course, X_t^* is unobservable, and its behavior needs to be inferred from the economic structure of the problem. A functional relationship between X_t^* and other variables can be specified on theoretical grounds, and, as we show next, it may be possible to use low-frequency information about observable endogenous variables to estimate its parameters.

We assume that, in the frictionless case, the durables stock to wealth ratio would be a function of the relative price of durables and nondurables. Specifically, we let

$$x_{it}^* = h_{it} + \alpha P_t + c_t, \tag{4.1}$$

where x_{it}^* is the logarithm of the frictionless durable stock of individual i at time t, h_{it} is the logarithm of her wealth, P_t is the logarithm of the relative price of durables and nondurables, and c_t is a deterministic function of time meant to capture secular changes in tastes and technology.

If the parameters in (4.1) are common across individuals, and the geometric mean of individuals' relative shares in wealth and (desired) durables is approximately constant over time (or its variation can be absorbed in the other regressors), then it is straightforward to obtain from (4.1) an expression relating averages at the aggregate level;[13] and recalling that $X_t = X_t^* + \bar{z}_t$, we obtain a relationship between observable variables,

$$X_t = H_t + \alpha P_t + C_t + \bar{z}_t, \tag{4.2}$$

where X_t and H_t are the logarithm of the average durables stocks and wealth, C_t absorbs c_t as well as secular terms possibly arising from the difference between geometric and arithmetic means, and $\bar{z}_t$ is the dynamic error term introduced by adjustment costs.

The X_t series we use is constructed from National Income real expenditure data. Assuming a 2% quarterly depreciation rate, we obtain an initial stock for the first quarter of 1954 by averaging expenditure on

13. Note that we abstract from the traditional, static aggregation issues that would arise even in the absence of adjustment costs. These problems have been extensively studied. See, for example, Stoker (1984). A model of aggregate data should also, in principle, address issues of aggregation across different types of durable goods.

durable goods from the third quarter of 1952 to the second quarter of 1954 and dividing by the depreciation rate, and we produce a quarterly stock series up to the fourth quarter of 1988. We approximate H_t by the accumulated innovations of an estimated ARI(1,1) representation of (log) personal income levels (see Campbell and Deaton 1989), and P_t is the log-difference of the implicit deflators. All data are seasonally adjusted and, since (4.2) applies to per capita quantities, we adjust all series and the depreciation rate for population growth.[14] As to C_t, we fit a piecewise linear trend with a break in the first quarter of 1975. In the theoretical model, the structural break represents a one-time shift in the secular components and in ex-ante real interest rates.[15]

We estimate α and the parameters of C_t by running OLS on (4.2). The unobservable $\bar{z}_t$ has quite complex univariate dynamics in the presence of adjustment costs, and it is obviously *not* independent of current and lagged values of exogenous variables. $\bar{z}_t$, however, necessarily has finite unconditional variance in the framework considered here; the important dynamic effects of transactions and adjustment cost must wash out over long-time averages in a rational maximization framework (see Section 2.3 above). Thus, if X_t^* is an integrated variable, then X_t and X_t^* are cointegrated, and the relationship between X_t^* and exogenous processes cointegrated with it can be recovered from a regression of X_t on the same processes. In the case at hand, the right-hand variables in (4.2) can be shown to be integrated, and not cointegrated among themselves. Thus, we can obtain superconsistent estimates of α and the parameters in C_t from a regression of $X_t - H_t$ on P_t and a broken trend. Cointegration tests and regression results are reported in Table 1.[16]

To proceed, we treat the predicted values from the cointegrating regression as a X_t^* series, and its residuals as a $\bar{z}_t$ series. To simplify the notation, we make no distinction between these estimates and underlying "true" values.

Step B: booms and recessions. From the X_t^* sequence and the assumed quarterly depreciation rate of the per capita stock of durables ($\delta = 0.02$

14. We also remove the deterministic component of per capita personal income growth. This is intended to capture the role of finite horizons in OLG models with productivity growth.
15. When entered as a separate regressor, ex post real interest rates in terms of durable goods are insignificant and have no important effect on our estimates.
16. Unfortunately, critical values for multivariate cointegration models including time trends depend crucially on the specific characteristics of the model considered. The 3.13 value reported in Table 1 should only be used as a reference threshold. Furthermore, in light of the large serial correlation in $\bar{z}_t$ that naturally arises from the theoretical model, the tests may have low power to reject the non-cointegration null.

Table 1

	H	t	$\tilde{t}$	P	ADF
X	1	−0.001	0.004	−0.448	−3.377
X	1	0.000	0.005	—	−2.400
X	1	—	—	—	0.907
H	—	—	—	—	−1.640
P	—	−0.003	−0.001	—	−2.633
P	—	—	—	—	0.081

All equations include a constant.
ADF is augmented (three lags) Dickey-Fuller test.
The 5% critical value is −3.13.
t and $\tilde{t}$ represent the trend and additional trend from 75:01 on, respectively.

plus the rate of population growth) we can construct a series for accumulated aggregate shocks, A_t:

$$A_t \equiv X_t^* + (\delta t + \log \text{(population at } t)).$$

We fit a random walk with drift to the A_t sequence in the discrete time binomial framework of Section 3, assuming that four Bernoulli innovations occur between observations. This yields estimates for the (annualized) drift and the standard deviation of the binomial aggregate process; in the durables application, these turn out to be $\theta = 0.10$ and $\sigma_A = 0.040$. We also obtain a period-by-period estimate of the number of boom subperiods (positive aggregate shocks) within each quarter. For example, if the realized $A_t - A_{t+1}$ is abnormally high we may infer that three or all of the four shocks were positive, while if $A_t - A_{t+1}$ is close to zero, we may deduct that the quarter contains two booms and two recessions.

Step C: inaction range and uncertainty decomposition. In the framework of Section 3, the estimated process from Step B indicates how many of the P_h, $h = 1,2,3,4$ should equal P_b for each quarter, though not the order in which they occur; the other transition matrices are equal to P_r, and the empirical distributions at successive observations is linked by the recursion

$$\tilde{f}_{t+1} = \tilde{f}_t \prod_{h=1}^{4} P_h. \tag{4.3}$$

To proceed, we use a nonlinear optimization routine to maximize the fit of the model as measured by the mean square of the prediction errors

$$e_t \equiv \bar{z}_t - \tilde{f}_t \hat{s},$$

where $\tilde{f}_t$ is, at every point in the sample, an estimate of the distribution of individual z_{it} deviations over their state space, $\hat{s}$. The estimated $\tilde{f}_t$ sequence must respect the recursion contraint in (4.3).

The free parameters at this stage are those entering the two transition matrices P_r and P_b; given ϑ and σ_A from Step B, these matrices depend on the relative importance of aggregate uncertainty in total unit-level uncertainty, $\gamma = \sigma_A/\sqrt{\sigma_A^2 + \sigma_I^2}$, and on the form of the individual unit's adjustment rule, the (L,l,u,U) quadruple. Our procedures allows us to estimate γ. As to the adjustment rule, the arguments in footnote 11 suggest that aggregate data are unlikely to convey information on the four points separately. We simplify the estimation procedure by assuming that

$$(L,l,u,U) = \tau(-0.50,0.00,0.45,0.50),$$

and we estimate τ, the overall width of the inaction band, instead of the four separate parameters (L,l,u,U). The assumption that $l = 0$ is simply a normalization, and has no substantive implications in the loglinear model we use; and the distance of the return points from the trigger points is assumed a priori to be strongly asymmetric.

To interpret these parameters and the assumptions we make about them, it is easiest to think of purchases of durable goods in the framework of Grossman and Laroque (1990). An individual can upgrade her durable good but, because of transaction costs, she does so by discrete jumps of size $\tau/2$ (the absolute value of $L - l$). Given the strong drift due to depreciation, *on average* consumers are unlikely to contemplate downgrading their stock of durables, and in fact a reduction (beyond depreciation) of the aggregate stock of durables should realistically be ruled out. Society as a whole cannot disinvest (or can do so in return for only dismal scrap values). If durables accumulation is literally irreversible, u and U both approach infinity; we can set them to a reasonably finite number without affecting the results, however, because the strong drift implies that they should seldom be approached in the sample path.[17]

The interval $(l,U]$ is important because in its absence adjustment would follow a one-sided (S,s) rule, and the type of aggregate uncertainty we allow for would not generate any interesting dynamics. The

17. Note that *individuals*, hit by idiosyncratic as well as aggregate shocks, *can* and will downgrade their durables. Such transactions occur on the secondhand market, and are irrelevant from the point of view of National Accounts data; still, transaction costs on used goods are the determinant of infrequent adjustment at the microeconomic level. Careful modeling of used goods transactions is left to future research.

key insight is that the evolution of wealth and relative prices would sometimes make disinvestment desirable (though impossible) in an average sense, and this is captured by allowing individual units to go beyond l when the random shocks they receive are so negative (positive for z_{it}) as to offset the strong negative drift in z_t due to wealth growth and durable goods depreciation.

The number of partitions of the state space, k, is determined in the estimation procedure through the relation

$$k = 2\left[\frac{\tau}{2\sqrt{((\sigma_A/\gamma)^2)\,dt + \theta^2\,(dt)^2}}\right] + 1,$$

where $[x]$ denotes the integer part of x.

In practice, we choose starting values for γ and τ and assume that the initial distribution is the one that would be stable if A_t grew linearly at rate ϑ with no uncertainty; we disregard the first 10% of the residuals to obtain an essentially random initial condition. For given γ and τ, our estimation program chooses the order in which booms and recessions occur within each quarter so as to minimize the absolute value of each e_t residual, and generates a sequence of empirical distributions based on this best-fit criterion. (We have experimented with programs that do not allow any freedom of choice as to the unobservable sequence of within-quarter innovations; the basic results are not sensitive to this.) We then feed the sum of the squared residuals (setting the first 10% equal to zero) to a standard minimization routine, which iterates to convergence over γ and τ.

4.2 RESULTS

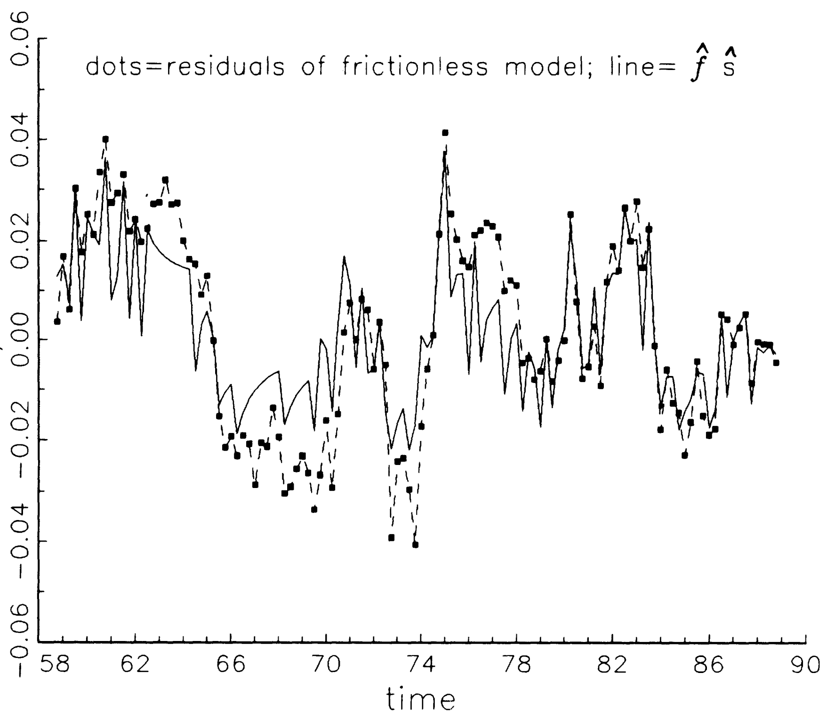
Table 2 presents the results. The bandwidth (τ) is about 52%, to imply that consumers typically wait for their durable stock-to-desired stock ratio to fall by about 26% before upgrading. This estimate should be confronted with the predictions of theoretical models such as that of Grossman and Laroque (1990), and with evidence on real-life transaction costs on typical durable purchases. On both counts, a 26% jump in the value of durable goods when adjustment is undertaken does not seem unreasonable at the individual level. The estimate of γ suggests that common shocks account for about 30% of the total uncertainty faced by individual consumers of durables.[18] The third row in Table 2 reports the R^2 measur-

18. In terms of the binomial model of Section 3, row 7 indicates that any given instant is a boom with probability 0.765, in which case 67% of the agents experiment with a positive shock.

Table 2

	Estimate	Std. Deviation
τ	0.514	0.027
γ	0.299	0.052
R^2	0.707	—
v	0.012	—
θ	0.100	—
σ_A	0.040	—
q	0.765	—
p_b	0.673	—
p	0.592	—
m	15	—

Figure 9



ing our model's fit of $\{\bar{z}_t\}$. The fit is quite encouraging; the infrequent adjustment model, even when "large" shocks are ruled out and the type of rule is arbitrarily specified, explains about 88% of the standard deviation of the component left unexplained by a frictionless model. Figure 9

Table 3

	MA(1)	Q(30)
ΔE_t	−0.071	46.2
	(0.092)	(0.030)
ΔE_t^*	−0.937	42.7
	(0.040)	(0.066)
$\Delta \hat{E}_t^*$	−0.933	61.2
	(0.034)	(0.007)

In parenthesis: standard deviations for the MA coefficients, significance levels for the Q portmanteau static.

illustrates this by plotting $\tilde{z}_t$ as estimated in Step B above along with $\hat{z}_t = f_t\hat{s}$, the predicted value form the estimation procedure.

As an alternative way to highlight the dynamic explanatory power of the model, consider its implications for the time series of expenditures. Mankiw (1982) argues that if the stock of durables—like nondurables consumption—follows an approximate random walk, as they should under the PIH, and if durable goods depreciate geometrically, then the first difference of *expenditures* should follow an MA(1) with a negative MA coefficient equal to the depreciation rate (plus population growth, when considering per-capita series) minus one. Thus, we would expect a negative MA coefficient with an absolute value in the order of 0.95 (or larger) in quarterly data. Mankiw found that expenditures display no such negative MA component, and his basic result is reproduced in the first row of Table 3; the MA coefficient for the first difference of the data we use (the period 59:1 to 88:4) is −0.07 and insignificant. In the presence of transaction costs, however, Mankiw's observations should apply to the expenditure series implied by the X_t^* series constructed above,[19] not the actual expenditure series.

We can first check whether Step A does deliver an X_t^* with the appropriate stochastic properties, estimating an IMA(1,1) process for the frictionless expenditure series implied by X_t^*, i.e., on

$$E_t^* \equiv e^{X_t^*} - (1 - \delta)e^{X_{t-1}^*}.$$

The results, reported in the second row of Table 3, are comforting. The MA coefficient emerges clearly and its magnitude is about right.

19. Note that Mankiw's derivation does not include price effects. Given our specification of X_t^* and the stochastic properties of the price series, however, this does not change the basic interpretation of the results.

It is more important and interesting to check whether the frictionless expenditure series recovered from Step C,

$$\hat{E}_t^* \equiv e^{X_t - \hat{z}_t} - (1-\delta)e^{X_{t-1} - \hat{z}_{t-1}},$$

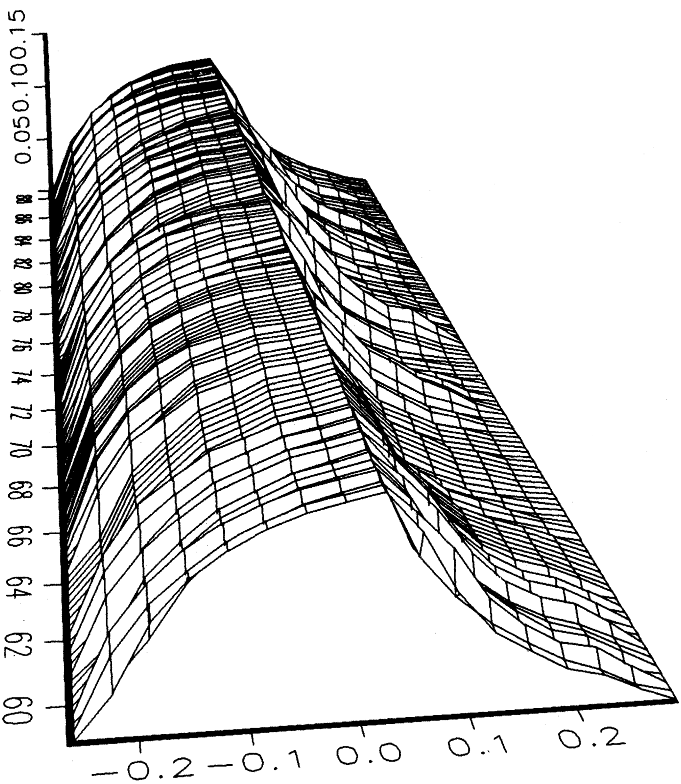
has the appropriate MA(1) structure. The third row in Table 3 shows that the estimation procedure's outcome is overall consistent with the basic implications of intertemporal optimization models, although the Portmanteau statistic suggests that more complex dynamics are present as well. Thus, once the dynamics captured by a model of infrequent, unsynchronized adjustment are removed, the residual satisfies the implications of a frictionless model.[20]

These results highlight the importance of cross-sectional developments across heterogeneous individuals for an understanding of aggregate dynamics. Adjustment costs have an important role in the short and medium run—although, as we repeatedly noted above, frictionless models should not mispredict actual data by very much in the long run. In the durables case, in fact, the frictionless PIH model fails to predict short-run dynamics but is not rejected in the long run (Caballero 1990). In the model we propose, the dynamics are generated by continuous shifting and reshaping of the cross-sectional distribution by the aggregate process and by idiosyncratic shocks. Aggregate expenditures are triggered when units reach L and move back to l.

To interpret the results, it is helpful to inspect visually the estimated cross-sectional distributions that our procedure allows us to infer from aggregate data. Figure 10 shows a three-dimensional view of the $\tilde{f}_t$ sequence, and Figure 11 plots the dynamic path of its percentiles. The dynamics of the distribution are not very pronounced, which is not surprising since "large" aggregate shocks and structural changes have been excluded from our empirical model. The responsiveness of durable stocks (and expenditures) to innovations does vary through time, always depending nonlinearly on the recent history of wealth and price innovations. When the aggregate stock of durables is low relative to the desired level, our model interprets the evidence as a shift of the distribution toward L; when the stock of durables is high, the model concentrates more units near U. Absent further aggregate developments, idiosyncratic shocks would tend to reshape these off-steady state distributions and to produce exponential impulse responses similar to those of Figure 6.

20. Note that the results in Table 3 are just an alternative measure of the fit highlighted in Table 2: As the R^2 goes to one in Table 2, row 3 in Table 3 converges to row 2. Alternatively, as the R^2 in Table 2 goes to zero, the third row in Table 3 converges to the first row in the same table.

Figure 10

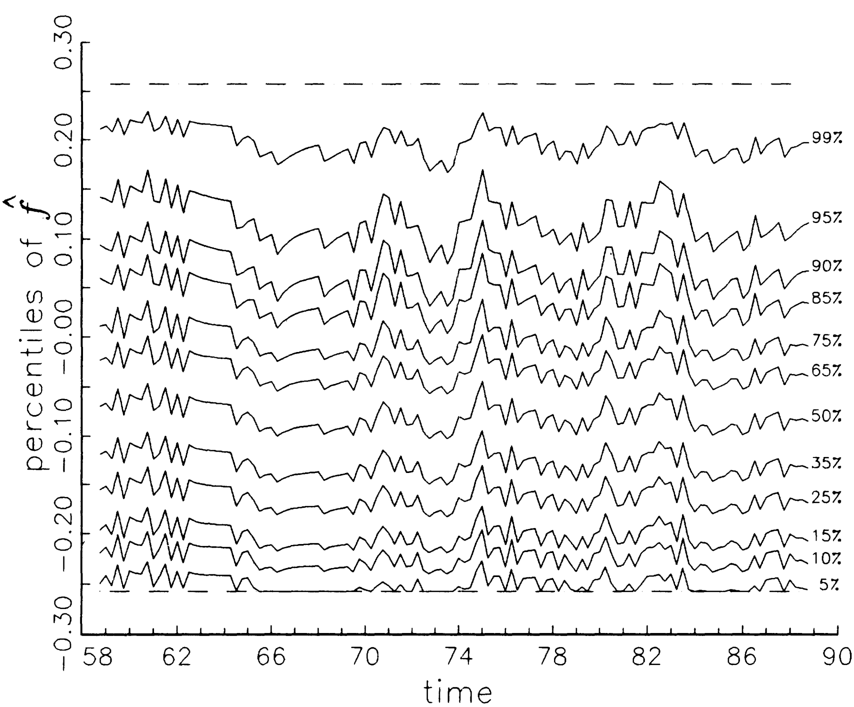


5. *Concluding remarks*

This paper studies microeconomic optimization, dynamic aggregation, and empirical estimation using relatively simple models. The concluding section addresses some of the more complex issues we have disregarded, discusses how they could be dealt with in future research, and notes that the most important insights appear robust to these and other extensions.

Consider again the durables consumption goods application of Section 4. As argued above, by taking into account the discontinuous nature of adjustment at the individual level the techniques we propose should provide an interpretation of aggregate dynamics that is "deeper," or more structural, than that obtained by representative agent models of dynamic optimization under convex adjustment costs. A truly structural model of durable goods consumption should, however, take into consid-

Figure 11



eration issues of general equilibrium interactions, endogenous coordination via strategic complementarities, and structural heterogeneity. We illustrate these issues by discussing the role of γ in the model above; this parameter measures the correlation between desired durables purchases across different units. One might think of confronting the statistical results obtained from estimation on aggregate data with microeconomic evidence, such as that obtainable from panel studies of income dynamics. It is important to realize, however, that a finding of $\gamma = 0.3$ does *not* imply that wealth innovations have a 0.3 correlation across individuals.

On the one hand, the results of Caballero and Engel (1989c) suggest that cross-sectional heterogeneity in behavioral parameters would bias γ toward lower values, attributing to idiosyncratic uncertainty the low degree of coordination due to different adjustment policies. More generally, redistribution effects due to heterogeneous parameters in equation (4.1) would not be properly recognized by our estimation procedure. Heterogeneity of this type presents a problem for any structural macroeconomic model, and we do not have much to say on this score.

On the other hand, there are at least two mechanisms by which the desired stock of durables may be found to covary much more strongly than individual incomes or other wealth innovations. First, movements of the price of durables (relative to nondurables and other points in time) are common across units. The price of durables should, of course, be endogenous in a completely specified model, and in future work it will be necessary to take into account intertemporal substitution and transaction costs on the supply side as well as the demand side of the market for new and used durable goods. Second, individual decisions may be endogenously coordinated if one unit's optimal actions depend "strategically" on other's actions—for example, bandwagon effects may be present in durables consumption; more interestingly, strategic interactions would need to be taken into account in models of price setting. Endogenous coordination—whether through supply constraints or strategic complementarities—would generally emphasize truly exogenous common shocks, via a multiplier effect. It would also make it much more difficult to derive optimal microeconomic adjustment rules, because the parameters of the processes taken as given by individual units would need to be determined endogenously in terms of the optimal adjustment rules themselves. The importance of these issues needs to be explored on a case-by-case basis; rules of the band-policy type with fixed parameters may be close to optimal, for example, if strategic complementarities are weak or supply is elastic relative to demand.

The simplifying assumptions made in the formal work above allow a tight characterization both of the microeconomic optimization problem and the aggregation process. Many insights are much more general than the specific models we have used to illustrate them, however, and more realistic, less tractable models would share many of the general features noted above. At the individual unit's level, kinked adjustment cost functions are realistic; this implies that optimal adjustment should be infrequent, interspersed with long periods of inaction, possibly lumpy, and these features are consistent both with casual empiricism and available disaggregated data. As to the dynamics of aggregate data, microeconomic inaction implies that close attention should be paid to the degree of coordination across units, and to the extent to which their actions are synchronized. In general, these issues can be modeled in terms of a distinction between common and idiosyncratic forces driving dynamic adjustment. As to empirical applications, information about the position and shape of cross sectional distributions is crucially important for a better understanding of macroeconomic fluctuations in the presence of adjustment costs. Such information can be obtained from the dynamics of aggregate variables themselves, as shown above. Although the implied dynamic reaction to shocks may or may not be similar to that generated by more standard (e.g., autoregressive) models, depending for example on the relative importance of "large" events, our approach still recommends itself for its microeconomic foundations, and may make it possible to exploit in macroeconomic applications the information provided by disaggregated data.

APPENDIX

The control problem

Define two processes $\{M_\tau\}$ and $\{N_\tau\}$ denoting the cumulative amount of (respectively) upward and downward adjustment performed on z up to time τ. By this definition,

$$dM_\tau \geq 0, \quad dN_\tau \geq 0 \quad \forall \tau$$

where dM_τ represents the differential of a continuous sample path or the discrete increment $M_t - M_{t-}$ when this is finite. Also define the sets of times $\{i\}$ and $\{j\}$ where the time path of $\{z_\tau\}$ is discontinuous:

$$\{i | N_{t+\tau_i} > N_{t+\tau_i-dt}, \tau_i > t\}, \quad \{j | M_{t+\tau_j} > M_{t+\tau_j-dt}, \tau_j > t\}.$$

It is then possible to represent "adjustment" formally:

$$dz_\tau = -\vartheta dt + \sigma dW_\tau + dM_\tau - dN_\tau \tag{A.1}$$

$$v(z_t) \equiv \max_{\{M_\tau\},\{N_\tau\}} E_t \left\{ \int_t^\infty e^{-\rho(\tau-t)} \left(-\frac{b}{2}(z_\tau)^2 \right) d\tau \right.$$

$$\left. - \int_t^\infty e^{-\rho(\tau-t)}(c_u dN_\tau + c_l dM_\tau) - \sum_{i=1}^\infty e^{-\rho(\tau_i-t)} C_u - \sum_{i=1}^\infty e^{-\rho(\tau_j-t)} C_l \right\}. \tag{A.2}$$

As long as the regularity condition of footnote 3 is satisfied, $v(z)$ is bounded above by zero. It is bounded below as well if $\rho > 0$ and ϑ and σ are finite.

Smooth pasting

Differentiability of $v(.)$ at the trigger points is endogenous when adjustment costs are not differentiable; see Dixit (1989d) for a proof that this is the case when fixed trigger and return points are optimal. Given differentiability, the conditions in (2.2) imply those in (2.3): consider $x \approx l$, $y \approx L$. For adjustment taking z from y to x not to dominate the candidate band-policy, by a Taylor approximation it must be the case that

$$v(l) + v'(l)(x - l) - (v\Sigma(L) + v'(L)(y - L)) \leq C_l + c_l(x - y).$$

But $v(l) - v(L) = C_l + c_l(U - u)$. Thus, we require

$$(v'(l) - c_l)\,(x - l) - (v'(L) - c_l)\,(y - L) \leq 0,$$

which is satisfied $\forall x > y$ only if the first line of (2.3) holds true. Similar considerations apply to downward adjustment.

The functional form of v(z)

When $dM_\tau = dN_\tau = 0$, z_t follows a Brownian motion process with drift $-\vartheta$ and standard deviation σ. An application of Ito's lemma yields an expression for the expected change of $v(z)$:

$$E_t\{dv(z_t)\} = v'(z_t)(-\vartheta)\,dt + \tfrac{1}{2}v''(z_t)\sigma^2 dt$$

In an ongoing optimization program, these expected "capital gains" plus current flow costs are equal to the required return on the current value of the program, $\rho\, v(z)dt$. Thus,

$$\tfrac{1}{2}v''(z)\sigma^2 - v'(z)\vartheta - \rho v(z) = \frac{bz^2}{2}\,. \tag{A.3}$$

It is easy to verify that

$$v^p(z_t) \equiv -\frac{b}{2}\left(\frac{z_t^2}{\rho} + \frac{\sigma^2 - 2z_t\vartheta}{\rho^2} + \frac{\vartheta^2}{\rho^3}\right)$$

is a particular solution of this differential equation. This is the present discounted value of flow losses if no adjustment is ever undertaken; its maximum value is $-\vartheta^2/\rho - \sigma^2 < 0$ at $z = \vartheta\rho$.

All solutions of (A.3) can be obtained by adding $v^p(z)$ to a function that solves the homogeneous part of (A.3). Such a function can be written

$$A_1 e^{\alpha_1 z_t} + A_2 e^{\alpha_2 z_t}$$

where α_1, α_2 solve the characteristic equation $\alpha\vartheta + \frac{1}{2}\alpha^2\sigma^2 - \rho = 0$ and A_1, A_2 are constants of integration. As long as inaction is indeed optimal for finite periods of time, the value function resulting from the problem in (A.1, A.2) must be a solution of (A.3); therefore, it can be written in the form

$$v(z_t) = -\frac{b}{2}\left(\frac{z_t^2}{\rho} + \frac{\sigma^2 - 2z_t\vartheta}{\rho^2} + \frac{\vartheta^2}{\rho^3}\right) + A_1 e^{\alpha_1 z_t} + A_2 e^{\alpha_2 z_t}. \tag{A.4}$$

The A_1 and A_2 constants are to be chosen so as to satisfy the conditions imposed on $v(z)$ at the boundaries of the domain over which (A.3) is valid.

The ergodic distribution

We approximate Brownian motion by a discrete random walk to make use of standard results from the theory of Markov chains. In the interior of the inaction region (i.e., for $L < z_t < U$) let

$$z_{t+dt} = \begin{cases} z_t + dz, & \text{with probability } \frac{1}{2}(1 - \vartheta\, dt/dz); \\ z_t - dz, & \text{with probability } \frac{1}{2}(1 + \vartheta\, dt/dz). \end{cases} \tag{A.5}$$

The $\{z_t\}$ process never leaves the bounded state space $\{L, L + dz, \ldots, U - dz, U\}$, hence it is ergodic if it can be shown to possess a unique invariant distribution, and it does possess a unique invariant probability distribution since all its states are positive recurrent (Ross 1983, p. 109). If downward adjustment of x_t is impossible, U approaches infinity; in this case, positive recurrency and ergodicity require $\vartheta < 0$.

The binomial random walk converges to Brownian motion as dz and dt approach zero, provided that $(dz)^2 = \sigma^2 dt$. Its invariant distribution over the discrete states $\{L, L+dz, \ldots, U - dz, U\}$ similarly converges to the invariant distribution of the continuous-time process in (A.1) over its continuous state space (L, U).

The ergodic distribution can be derived exploiting its invariance property. At every point z in the interior of the inaction range, except l and u, the discrete steady-state probability distribution function should satisfy the balance equation

$$f(z) = f(z - dz)\tfrac{1}{2}(1 - \vartheta\, dt/dz) + f(z + dz)\tfrac{1}{2}(1 + \vartheta\, dt/dz).$$

Rearranging,

$0 = (f(z + dz) - f(z)) - (f(z) - f(z - dz)) + \vartheta \frac{dt}{dz} [(f(z + dz) - f(z)) + (f(z) - f(z - dz))].$

Dividing by dz and taking the limit, we find that $f(z)$ is continuously differentiable. Dividing by $(dz)^2$, and using $dt/(dz)^2 = \sigma^{-2}$, we have in the limit

$$\frac{\sigma^2}{2} f''(z) = \vartheta f'(z). \tag{A.6}$$

The general solution of this functional equation has the form

$$\begin{aligned} f(z) &= Az + B \quad \text{if} \quad \vartheta = 0 \\ &= Ae^{\xi z} + B \quad \text{if } \vartheta \neq 0, \text{ for } \xi \equiv \frac{2\vartheta}{\sigma^2}. \end{aligned}$$

To determine which values of A and B are appropriate at every point, we make use of the balance equations at the trigger and return points.

Consider first the case $L < l$, with strict inequality. The discrete process (A.5) never reaches point L, and jumps to l instead. If $z_t = L + dz$,

$$z_{t+dt} = \begin{cases} L + 2\,dz, & \text{with probability } \frac{1}{2}(1 - \vartheta\, dt/dz); \\ l & \text{with probability } \frac{1}{2}(1 + \vartheta\, dt/dz). \end{cases} \tag{A.7}$$

Hence, $f(L) = 0$, and z can reach l not only from $l - dz$ and $l + dz$, but from $L + dz$ as well. For invariance of $f(z)$, it must be that

$$\begin{aligned} &f(L + dz) = f(L + 2\,dz)\tfrac{1}{2}(1 + \vartheta dt/dz) \\ &f(l) = f(l - dz)\tfrac{1}{2}(1 - \vartheta\, dt/dz) + f(L + dz)\tfrac{1}{2}(1 + \vartheta\, dt/dz) + f(L + dz) \\ &\qquad \tfrac{1}{2}(1 + \vartheta\, dt/dz). \end{aligned}$$

In the limit, $f(z)$ is continuous at l and L but need not be continuously differentiable. Making use of $f(L) = 0$, the balance equations can be rearranged to read

$$\begin{aligned} f(l) - f(l - dz) = {}& (f(l + dz) - f(l)) + (f(L + dz) - f(L)) \\ & -\vartheta\, dt/dz\, [(f(l - dz) - f(l + dz)) - (f(L + dz) - f(L))]. \end{aligned}$$

Dividing by dz, using $(dz)^2 = \sigma^2 dt$ and denoting the left- and right-hand side derivatives with $f'_{(-)}(.)$ and $f'_{(+)}(.)$, in the limit we find

$$f'_{(-)}(l) = f'_{(+)}(l) + f'_{(+)}(L).$$

Consider next the case $L = l$ (reflecting barrier), which is the case when control incurs linear adjustment costs with no fixed component. Then, when $z_t = L$,

$$z_{t+dt} = \begin{cases} L + dz, & \text{with probability } \frac{1}{2}(1 - \vartheta\, dt/dz); \\ L & \text{with probability } \frac{1}{2}(1 + \vartheta\, dt/dz). \end{cases}$$

Although $f(L_{(-)}) = 0$, $f(L) > 0$ in this case. As it is possible to reach point L both from $L + dz$ and from L itself, the invariant distribution must satisfy the balance equation

$$f(L) = f(L + dz)\,\tfrac{1}{2}(1 + \vartheta\, dt/dz) + f(L)\,\tfrac{1}{2}(1 + \vartheta\, dt/dz).$$

Right-continuity of $f()$ at L follows in the limit. Dividing by dz and taking the limit,

$$f'_{(+)}(L) = \frac{2\vartheta}{\sigma^2} f(L).$$

Similar computations provide boundary conditions for the stable density at the upper trigger and return point. These boundary conditions and the adding-up constraint

$$\int_L^U f(z)dz = 1,$$

form a rank-deficient system of linear equations in (at most six) A and B constants, and a solution can always be derived in closed form.

REFERENCES

Abel, A. B., and O. J. Blanchard. 1986. The present value of profits and cyclical movements in investment. *Econometrica* 53: (2) 249–73

Akerlof, G., and J. Yellen. 1985. Can small deviations from rationality make significant differences to economic equilibria? *American Economic Review* 75 (4): 708–21

Arrow, K. J. 1968. Optimal capital policy with irreversible investment. In *Value, capital and growth. Papers in honor of Sir John Hicks.* edited by J. N. Wolfe. Edinburgh: Edinburgh University Press

Baldwin, R., and P. Krugman. 1989. Persistent trade effects of large exchange rate shocks. *Quarterly Journal of Economics* 104 (4): 635–54

Bar-Ilan, A., and A. Blinder. 1987. The life cycle permanent income model and consumer durables. NBER Working Paper No. 2149.

Barro, R. J. 1972. A theory of monopolistic price adjustment. *Review of Economic Studies* 39 (1): 17–26.

Bather, J. A. 1966. A continuous time inventory model. *Journal of Applied Probability* 3: 538–49.

Baumol, W. 1952. The transactions demand for cash. *Quarterly Journal of Economics* 67 (4):545–56.

Benabou, R. J. 1989. Some welfare effects of inflation. MIT Working Paper No. 542, November.

Bentolila, S. 1988. Firing costs and labor demand in Europe. Ph.D. Thesis, MIT.

Bentolila, S., and G. Bertola 1990. Firing costs and labor demand: How bad is Eurosclerosis? *Review of Economic Studies*. Forthcoming.

Bertola, G. 1989a. Irreversible investment. Mimeo.

Bertola, G. 1989b. Labor turnover costs and average labor demand. Mimeo.

Bernanke, B. 1985. Adjustment costs, durables, and aggregate consumption. *Journal of Monetary Economics* 15 (1): 41–68.

Billingsley, P. 1986. *Probability and measure*, 2d ed. New York: John Wiley & Sons.

Blanchard, O. J., and N. Kiyotaki. 1987. Monopolistic competition and the effects of aggregate demand. *American Economic Review* 77 (4): 647–66.

Blinder, A. S. 1981. Retail inventory behavior and business fluctuations. *Brookings Papers on Economic Activity* 2: 443–505.

Caballero, R. J. 1990. Expenditure on durable goods: A case for slow adjustment. *Quarterly Journal of Economics*. Forthcoming.

Caballero, R. J., and E. M. R. A. Engel. 1989a. The *S-s* economy: Aggregation, speed of convergence and monetary policy effectiveness. Columbia University Working Paper No. 420.

Caballero, R. J., and E. M. R. A. Engel. 1989b. Heterogeneity and output fluctuations in a dynamic menu cost economy. Mimeo, October.

Caballero, R. J., and E. M. R. A. Engel. 1989c. Aggregation, convergence and coordination. Mimeo, December.

Campbell, J., and A. Deaton. 1989. Why is consumption so smooth? *Review of Economics Studies* 56 (3): 357–74.

Caplin, A. 1985. Variability of aggregate demand with (S,s) inventory policies. *Econometrica* 53:1395–1409.

Caplin, A., and Krishna, K. 1986. A simple dynamic model of employment. Mimeo, Harvard University.

Caplin, A., and J. Leahy. 1990. State-dependent pricing and the dynamics of money and output. *Quarterly Journal of Economics*. Forthcoming.

Caplin, A., and D. Spulber. 1987. Menu costs and the neutrality of money. *Quarterly Journal of Economics* 102 (4): 703–26.

Constantinides, G. M. 1986. Capital market equilibrium with transaction costs. *Journal of Political Economy* 94 (4): 842–62.

Constantinides, G. M., and S. F. Richard. 1978. Existence of optimal simple policies for discounted-cost inventory and cash management in continuous time. *Operations Research* 26 (4): 620–36.

DeLong, B., and L. H. Summers. 1986. Are business cycles symmetrical? In *The*

American business cycle, edited by R. Gordon. Chicago: University of Chicago Press.

Diamond, P. 1988. Search, sticky prices, and inflation. MIT Working Paper No. 509, December.

Dixit, A. K. 1989a. Entry and exit decisions under uncertainty. *Journal of Political Economy* 97 (3): 620–37.

Dixit, A. K. 1989b. Hysteresis, import penetration, and exchange rate pass-through. *Quarterly Journal of Economics* 104 (2): 206–28.

Dixit, A. K. 1989c. Intersectoral capital reallocation under price uncertainty. *Journal of International Economics* 26:309–25.

Dixit, A. K. 1989d. A simplified exposition of the theory of optimal control of Brownian motion. Mimeo.

Dixit, A. K. 1989e. Analytical approximations in models of hysteresis. Mimeo.

Dumas, B. 1988. Perishable investment and hysteresis in capital formation. Mimeo.

Frenkel, J. A., and B. Jovanovic. 1980. On transactions and precautionary demands for money. *Quarterly Journal of Economics* 95 (1): 25–43.

Gavin, M. K. 1986. Labor market rigidities and unemployment. The case of severance costs. FRBG International Finance Discussion Paper No. 184.

Grossman, S. J., and G. Laroque. 1990. Asset pricing and optimal portfolio choice in the presence of illiquid durable consumption goods. *Econometrica*. Forthcoming.

Hall, R. E. 1978. Stochastic implications of the life cycle–permanent income hypothesis: Theory and evidence. *Journal of Political Economy* 86 (5): 971–87.

Hamermesh, D. S. 1989. Labor demand and the structure of adjustment costs. *American Economic Review* 79 (4): 674–89.

Hansen, L., and K. Singleton. 1983. Stochastic consumption, risk aversion, and the temporal behavior of assets returns. *Journal of Political Economy* 91 (2): 249–65.

Harrison, J. M. 1985. *Brownian motions and stochastic flow systems*. New York: John Wiley & Sons.

Harrison, J. M., T. M. Sellke, and A. J. Taylor. 1983. Instantaneous control of Brownian motion. *Mathematics of Operations Research* 8:439–53.

Holt, Ch. C., et al. 1960. *Planning production, inventories, and work force*. New Jersey: Prentice-Hall.

Kemp, M. C., and H. Y. Wan, Jr. 1974. Hysteresis of long-run equilibrium from realistic adjustment costs. In *Trade, stability and macroeconomics. Essays in honor of Lloyd A. Metzler*, edited by G. Horwich and P. A. Samuelson. New York: Academic Press.

Lam P. 1989. Irreversibility and consumer durables expenditures. *Journal of Monetary Economics* 23 (1): 135–50.

Mankiw, N. G. 1982. Hall's consumption hypothesis and durable goods. *Journal of Monetary Economics* 10 (2): 417–25.

Mankiw, N. G. 1985. Small menu costs and large business cycles: A macroeconomic model of monopoly. *Quarterly Journal of Economics* 100 (2): 529–39.

Miller, M. H. and D. Orr. 1966. A model of the demand for money by firms. *Quarterly Journal of Economics* 80 (3): 413–35.

Neftci, S. N. 1984. Are economic time series asymmetric over the business cycle? *Journal of Political Economy* 92 (3): 307–28.

Nickell, S. J. 1974. On the role of expectations in the pure theory of investment. *Review of Economic Studies* 36 (1): 1–19.
Nickell, S. J. 1978. Fixed costs, employment, and labour demand over the cycle. *Economica* 45 (3): 329–45.
Pindyck, R. S. 1988. Irreversible investment, capacity choice, and the value of the firm. *American Economic Review* 78 (5): 969–85.
Ross, S. M. 1983. *Stochastic Processes*. New York: John Wiley & Sons.
Rotemberg, J. 1987. The new Keynesian microeconomic foundations. *NBER Macroeconomics Annual* 1987: 69–104.
Rothschild, M. 1971. On the cost of adjustment. *Quarterly Journal of Economics* 85 (4): 605–22.
Scarf, H. 1960. The optimality of (S,s) policies in the dynamic inventory problem. In *Mathematical methods in the social sciences*, edited by K. J. Arrow, S. Karlin, P. Suppes. Stanford: Stanford University Press.
Sheshinski, E., and Y. Weiss. 1977. Inflation and costs of price adjustment. *Review of Economic Studies* 44 (2): 287–303.
Sheshinski, E., and Y. Weiss. 1983. Optimum pricing policy under stochastic inflation. *Review of Economic Studies* 50 (3): 513–29.
Smith, G. W. 1989. Transactions demand for money with a stochastic, time-varying interest rate. *Review of Economic Studies* 56 (4): 623–33.
Stoker, T. M. 1984. Completeness, distribution restrictions, and the form of aggregate functions. *Econometrica* 52: 887–908.
Tobin, J. 1956. The interest elasticity of the transactions demand for cash. *Review of Economics and Statistics* 38: 241–47.
Tsiddon, D. 1988. On the stubborness of sticky prices. Working Paper No. 1974, Hebrew University.

Comment

ANDREW CAPLIN

I would like to congratulate the organizers of the conference and the authors, for a first-rate piece of work on the general topic of aggregation with indivisibilities. The topic is close to my heart, and reading this work has increased my confidence in the overall research program. In this comment I will point out how their work expands our knowledge, and make a few suggestions on future directions of research.

The paper has three parts. The first is a valuable survey of the microeconomics of adjustment costs. Here the authors produce a synthetic model that illustrates the impact of both nondifferentiability and discontinuity in adjustment costs. The second is an overview and extension of the literature on techniques for studying the aggregate distributional dynamics in models with microeconomic inertia. The final part is a pioneering attempt to bring this class of aggregative models to the data. I will focus most of my attention on this final part, since this is where the paper has its surprise value. I have spent some time trying to understand whether the apparent success of the authors' estimation procedure can be explained by certain idiosyncratic features. At this stage, I have failed to find any obvious candidates. It may just be that even in their current preliminary form, these aggregation models provide some worthwhile insights into macroeconomic fluctuations.

In the portion of the paper that deals with the microeconomics, Bertola and Caballero study an individual agent optimizing against an exogenous stochastic process when there are both fixed and linear costs of adjustment. They develop a canonical version of this model that exposes the central qualitative features of a wide variety of different models.

One element is a fixed cost of adjusting a microeconomic variable—for example, a cost of ordering new inventory, a cost of moving house, a cost of price adjustment, or a cost of entry and exit. The central qualitative feature of the optimal adjustment policy is then to make infrequent large changes in the state variable: small changes are simply not worthwhile. The second element is a linear adjustment cost with a kink at the origin, as when there is an imperfect capital market for the resale of a durable commodity, a transactions cost in asset purchases, or costs of hiring and firing workers. The basic conclusion is that there will be long periods of inaction in which it is not worth moving due to the difference between the value of upward and downward adjustment. When adjustment does take place, however, it may be on a small scale. When these two forms of adjustment cost are both present there is a hybrid policy that in its simplest form can be defined by four parameters. There is a range of inaction defined by two outer adjustment points: the agent allows state variable to diverge from its optimal value between lower and upper bounds L and U, respectively, $L < U$. There is adjustment from U to u and from L to l with $U > u > l > L$. The reason the variable is not adjusted to the same point from L and U is that the kink in the linear adjustment cost term discourages further adjustment at the margin.

The reason for the recent burst of activity in this class of microeconomic models is the use of the modern theory of optimal-control of continuous-time stochastic processes. This theory allows one to go beyond these simple qualitative conclusions to get precise characterizations of optimal strategies for each of the many problems that share this broad nature.

The macroeconomic development of these topics deals with the question of what happens at the aggregate level when agents face the kind of microeconomic circumstances that make them adjust in this frictional manner. In these settings, it becomes very difficult to use standard representative agent reasoning, especially when the shocks that agents face include both idiosyncratic and common elements. Instead, it becomes vital to shift the focus of macroeconomics toward distributional dynamics; the key issue is to assess the extent to which individual inertia is inherited at the aggregate level. This pushes us into a large number of technical boundaries. The current early stage of modeling these issues involved looking for simple stories that may capture qualitative features that will survive in more complete models.

The first way to get aggregate insights is to fix a strategy of the kind outlined beforehand. We also need to fix the extent of common and idiosyncratic shocks. We then consider the response of the economy to different paths of the common shock by studying the distribution of agents' positions within their range of inaction. Any well-trained economist can spot that there are a large number of missing elements. In certain cases we take the strategy as a primitive rather than solving for the optimal strategy given a complex shock process. We may also end up ignoring many general equilibrium effects whereby the strategies themselves influence the path of the supposedly exogenous shocks. Hence, these early models should be regarded as a way into the research rather than as the ultimate summary of how to aggregate when there are indivisibilities.

Bertola and Caballero boldly attempt to fit a model based on the fixed microeconomic strategies and a certain ratio of idiosyncratic to common shocks to data on aggregate durable purchases. My prior belief was that there would be a poor fit, but that this would be readily understandable in light of the preliminary nature of the models. It is Figure 9 that shows this expectation was not met. It appears that the model does a good job of fitting the data (although there are no formal procedures that allow us to assess the (S,s) model in a wider class of alternatives). It then becomes critical to detail the estimation procedure to see whether this good fit results from methodological idiosyncracies or fundamental economic forces.

The (S,s) theory explains the dynamics of the divergence between the actual and the "desired" stock of durables. The first task in the estimation procedure is to derive a time path for the desired stock, x^*, and the residual, $z = x - x^*$. At the same time this procedure characterizes the common shock that is driving the model; this is measured by the change in the desired stock. From then on the estimation involves essentially two free parameters: the fixed aggregate shock is run through an (S,s) aggregation model in which the width of the (S,s) bands and the extent of the idiosyncratic shocks are allowed to vary to match the data as well as possible.

One qualitative feature of these (S,s) models is a tendency to smooth out aggregate disturbances. In a two-sided (S,s) model, an increase in the desired stock of durables will raise the ratio of desired to actual durables, yielding a form of partial adjustment in the aggregate. The width of the (S,s) bands will influence the range of possible values of the ratio of actual to desired durables, as well as the periodicity of cycles in this ratio. The extent of the idiosyncratic shock has its main influence on the extent of the smoothing of aggregate shocks. Overall, it seems there are enough common features of (S,s) models to conclude that the very good fit uncovered by Bertola and Caballero is strong evidence in favor of these models. Before we can be confident of this interpretation, we must examine some special features of the estimation procedure to see if they can help account for the fit.

One potential explanation lies in the need to arbitrarily specify a starting point for the distribution of the residuals. It may be that a good fit could be artificially generated by choosing this unobservable in an advantageous manner. This turns out not to be relevant in the current procedure, in which Bertola and Caballero tie their hands by always using the individual firm steady state as the initial distribution. A second potential explanation is more subtle and hinges on the sample path dependence in (S,s) models. While observations are gathered only once a quarter, it is important to allow individuals to make decisions and experience shocks more frequently than this. Ideally one would allow continuous-time decisions; in their procedure, Bertola and Caballero subdivide a quarter into four subintervals. But there is significant path-dependence in these models. It is, therefore, possible that by appropriately ordering the unobservable within-quarter order of the shocks, one can greatly improve the model's fit. This explanation may be particularly potent if we find extremely wide and implausible (S,s) bands.

While this appears possible a priori, it does not appear to account for the fit in this particular case. I say this on the basis of two pieces of evidence—one private, one public. The private information is that I have seen a version of the model in which the freedom to alter the pattern of within-period shocks is removed. The end result is that the while the fit is not quite as good in the early part of the sample period, it is almost identical for the most recent ten-year period. The public information is that there is nothing obviously absurd about the predicted width of the (S,s) bands; when you buy a durable, it accounts for 26% of the total stock of durables that you hold.

This leaves open the hypothesis that the good fit has real economic causes. Figure 8 helps us understand why we might expect the model to yield a reasonable fit. The figure shows that as the extent of the idiosyncratic shocks is raised, the model's tendency to smooth out aggregate shocks also increases. But this is precisely the observation that explains the general success of partial adjustment models: large impulses tend to produce effects that get spread over time rather than being absorbed all at once.

This broad observation alone does not seem to be enough to explain the model's fit, especially for the last ten-year period. One other important qualitative feature of the (S,s) model is that it allows for a fairly rapid turnaround in the face of a change in the direction of the common shocks. Note that the (S,s) bands provide a bound on the maximum distance between desired and actual durables stocks. Therefore, a relatively short sequence of positive aggregate shocks may be enough to return the actual level of stocks close to its desired level even after a long string of negative aggregate shocks. This means there is the ability for a relatively rapid turnaround in these models, and this feature may also help improve the fit with actual data. Overall, Bertola and Caballero have left us with the unusual problem of trying to rationalize a surprisingly successful empirical exercise.

Finally, I would like to suggest a number of directions for future research in this area. First, it is surely desirable to redo this exercise using a more formal statistical approach in which alternatives are outlined and formal tests carried out. A potentially more important issue is to work with data tapes that include a greater level of microeconomic detail, so that one can actually observe the changing nature of the stochastic processes at different levels of aggregation. As for the theoretical models, it is important to develop models in which we can do more than pay lip service to general equilibrium considerations. Beyond this, as we deepen our exploration of macroeconomics without the representative agent whole new classes of questions will be opened up. One example is the interaction between information transmission and transactions costs: What effect does microeconomic inaction have on the ability of prices and other market data to transmit information? The work of Bertola and Caballero suggests that these topics will begin to enter the macroeconomic mainstream sooner rather than later.

Comment

ROBERT E. HALL

Bertola and Caballero make a substantial advance in this paper. They tackle a problem that many of us thought completely intractable—aggregation of nonconvex adjustment—and derive empirically useful results.

I see their paper as primarily a contribution to noise analysis. They show that we can view a time series as the sum of a value predicted by a neoclassical model plus a noise factor associated with nonconvex adjustment. Previous work on noise has often concluded that it is an important part of the overall pattern of movement of macrovariables. Bertola and Caballero cite Mankiw's investigation of noise in consumer durables as an example of the kind of problem their method can handle. Mankiw showed the importance of noise through the contrast between the stochastic process implied by consumer theory and the actual stochastic process of durables acquisitions. Theory predicts that the stock of durables should be a random walk, but in fact the flow of acquisitions of durables is close to a random walk. There is a big noise factor that accounts for the difference. With nonconvex adjustment at the level of the individual family, aggregate adjustment is smeared over time.

It may be helpful to summarize the three steps in the program recommended by Bertola and Caballero. First, estimate the neoclassical model without adjustment costs and calculate the noise series as the difference between the actual value of the series and the values predicted by the neoclassical model. Second, estimate the aggregate shock process and calculate the time series of the aggregate shocks. Third, estimate the parameters of the adjustment model from the relation between the noise series and the aggregate shock series.

Let me comment further on the empirical application to durables within this framework. From Mankiw, we know that the actual stock of durables is almost second-order integrated, not first-order integrated as predicted by theory absent adjustment costs. Changes in the stock of durables in response to wealth changes that theory predicts would occur instantaneously are actually delayed over time. The basic idea of Bertola and Caballero is that, with the right asymmetries, adjustment costs tend to delay adjustment and thus explain the empirical finding. The asymmetries are important and are discussed fully in this paper. Earlier work by Caplin and Spulber, with exact symmetry, gave an example where adjustment costs do not delay adjustment, after aggregation. The distributed lag pattern of delay identifies the parameters of adjustment.

Although aggregation of nonconvex adjustment rationalizes the lags found in data on consumer durables, there is no decisive evidence in favor of nonconvex adjustment as against other explanations of the persistence of durables investment. For example, the simple model in which families spend a fixed fraction of their incomes on durables also explains the basic facts. If we choose the aggregation-of-nonconvex-adjustment explanation, it is because we find its foundations in optimization more to our tastes, not because it beats the other model in any statistical way.

Figure 9 in the paper shows how the method explains the observed persistence and variance of the measured noise in durables investment. The rather considerable success shown in the figure demonstrates that the two parameters of the model can have values that make the model's persistence and variance match the data. We have to judge the results not so much by the good fit of Figure 9 but by how reasonable is the story told by the parameters. The story is the following: Families wait until their stocks of durables are 25% too low, if their fortunes are rising. At that threshold, they make a single purchase large enough to bring their stock up to its normal relation to wealth. If their fortunes are declining, they wait until their stock is 25% too high. At that threshold, they cut down to a stock that is 22% too high. At any given time, idiosyncratic shocks distribute families in the range from 25% too high to 25% too low. When a favorable aggregate shock comes along, the families that were on the low side respond by buying more immediately. In subsequent periods, idiosyncratic shocks are more likely than before the aggregate shock to push other families through the bottom threshold, at which point they will respond to the earlier aggregate shock. Because families are not uniformly distributed within the band between the two thresholds, the effect of the aggregate shock is spread over time. By contrast, under Caplin and Spulber's assumptions, there is no spreading because the exaggerated response of the families pushed over the edge exactly offsets the zero response from those who are not.

Although I find the theoretical work extraordinarily impressive and find the paper convincing that nonconvex adjustment can explain persistent noise, I am not yet persuaded that nonconvex adjustment will emerge as a major explanation of the noise in important macroaggregates. Surely consumer durables is the strongest application of the theory, because adjustment costs are largest in relative terms for the smallest decision makers. Even for durables, I am not sure I believe that the no-adjustment band is as wide as found in this paper. The narrower the band, the smaller and less persistent is the noise associated with nonconvex adjustment. For business investment and other nonhousehold variables, the band should be much tighter and thus nonconvex-adjustment noise much smaller. Yet noise seems particularly large for investment, especially inventory investment, where adjustment costs are probably small.

Even if other types of noise—possibly from sources of nonconvexities different from the one considered here—ultimately turn out to be more important than nonconvex-adjustment noise, I expect that the contribution of this paper will be a lasting one. I repeat my admiration for a remarkable step forward.

Discussion

Greg Mankiw suggested that Caballero and Bertola's theory has different implications about the number of units of durables bought and the average price paid per unit, which the authors should explore. Ben Bernanke suggested that the authors could examine panel data on automobile ownership and wealth to test these predictions. He also suggested forming out-of-sample forecasts for Z to test the accuracy of the model.

Olivier Blanchard asked if the residuals are completely accounted for each period by the estimation method. He also asked whether there was a fixed or variable number of shocks each period. Bertola responded that they allowed four innovations each period, which were chosen optimally.

Matthew Shapiro wondered whether the increased demand for smaller cars after the OPEC price increases led to a boom in automobile production, as the model predicts. Bertola answered that in general equilibrium, when the price of cars is endogenous, the predictions are not so clear. He also noted that the empirical work allowed for a change in intercept in 1975.