Chapter Title: Optimal Misspecified Beliefs

Book Title: The Conquest of American Inflation

Book Author(s): Thomas J. Sargent

Published by: Princeton University Press

Stable URL: http://www.jstor.com/stable/j.ctv39x7tk.9

6
Optimal Misspecified Beliefs

Equilibrium with mistakes

This chapter describes three key conceptual issues:

1. How to formulate equilibria where agents have a common misspecified least squares forecasting model.

2. How expectations can contribute independent dynamics within equilibria.

3. How within such equilibria the classic adaptive expectations scheme can use second moments to approximate a first moment.

These issues recur throughout this essay. To expose them, I temporarily set aside the Phillips curve example, and formulate the issues in terms of Bray's (1982) simple model of the price of a single good. This model is a work horse for studying bounded rationality. I alter Bray's model to illustrate an equilibrium concept that merges aspects of rational and adaptive expectations in a new way. Later I apply this equilibrium concept to the Phillips curve model.

I focus on market equilibrium with optimal but misspecified forecasts. Optimal means that the free parameters of the forecasting scheme are chosen by maximum likelihood or generalized least squares. Misspecified means that the forecasting model is wrong either in functional form or in using too small an information set. The true model depends on how market participants' model is misspecified.

An experiment in Bray's lab

I follow Bray (1982) and assume that

$$p_t = a + bp_{t+1}^e + u_t \tag{40}$$

where u_t is i.i.d., with mean zero and variance σ_u^2; $a > 0, b \in (0,1)$; p_t is the market price of a good; and p_{t+1}^e is the market's expectation of the price. The rational expectations equilibrium has $p_{t+1}^e = \frac{a}{1-b}$, and $p_t = \frac{a}{1-b} + u_t$.

In place of rational expectations, Bray posited that p_{t+1}^e is the empirical average of past prices, which can be represented recursively as $p_{t+1}^e = p_t^e + t^{-1}(p_{t-1} - p_t^e)$. With these expectations, Bray showed that when $0 < b < 1$, p_{t+1}^e converges almost surely to the rational expectation $\frac{a}{1-b}$. During a transition to rational expectations, while learning continues, the state variable p_t^e contributes dynamics and makes the price serially correlated. But Bray showed that these dynamics are transitory.[1] At the rational expectations equilibrium, p_t is a constant plus a serially uncorrelated shock.

To let expectations impart persistent serial correlation to the price, I depart from Bray and assume that the market has adaptive expectations of the form

$$p_{t+1}^e = Cp_t + (1 - C)p_t^e, \tag{41}$$

where $|C| < 1$. Notice how, in effect, Bray assumed a version of (41) in which $\frac{1}{t}$ replaces C, which makes p_{t+1}^e become a sample average of past p's. In contrast, the algorithm with a fixed C as in (41) is known as a fixed gain algorithm. Fixing C has the effect of discounting past observations, relative to Bray's scheme. If $C \in (0,1)$, (41) is a version of adaptive expectations, past observations discounted at a rate $C - 1$. Discounting past observations arrests convergence to rational expectations

[1] See Sargent (1993, Chapter 5) for a graph of these transient dynamics.

and prevents the state variable p^e_{t+1} from converging to a constant. Instead, it can converge to a serially correlated stationary stochastic process.

Equation (41) would be the linear least squares forecast if the price were to follow the process

$$p_t = p_{t-1} + \epsilon_t - (1 - C)\epsilon_{t-1}, \tag{42}$$

where ϵ_t is a serially uncorrelated process that equals the one-step ahead error in forecasting p_t linearly from its own past. With this specification the market sees the price as composed of purely permanent and transitory components.[2]

I want to formulate a mapping whose fixed point delivers a reasonable concept of equilibrium under misspecification. I begin by describing how the price actually moves when the market's beliefs are (41). Equation (41) can be rearranged to be

$$p^e_{t+1} = \frac{C}{1-(1-C)L} p_t \tag{43}$$

where $C \in (0,1)$. Substituting (43) into (40) shows that when people in the market believe that the price should be forecast according to (43), their actions make the actual law of motion for price become

$$p_t = \frac{a}{(1-b)} + \frac{1}{(1-bC)} \left[\frac{1-(1-C)L}{1-\frac{1-C}{1-bC}L} \right] u_t, \tag{44a}$$

or

$$p_t = \nu + f(L)u_t \tag{44b}$$

where $\nu, f(L)$ are defined to match the last two equations. The first and second moments of p_t are described by the stationary mean ν and the spectrum

$$F(\omega) = f(\exp(i\omega))f(\exp(-i\omega))\sigma_u^2, \;\; \omega \in [-\pi, \pi].$$

[2] See Muth (1960).

Notice that $F(\omega)$ depends on C, through f.

When C is a small positive number, the perceived law of motion (42) differs from the actual one (44a) in interesting ways. Equation (44b) has a constant, plus a mixed moving average, autoregressive piece. The perceived law has to emulate a constant through its unit root.

Misspecification

To motivate a restriction on C, note two important facts about (44): (a) given that the price obeys (44), the linear least squares one-step forecasting rule, conditioned on the infinite history of p_t, is not a geometric distributed lag like (41). And (b), even if we were to restrict the expectations rule to take the form (41), the best forecasting rule of this class would make C solve a forecast error minimization problem and thereby make C an outcome, not a parameter. A rational expectations equilibrium repairs both of these dimensions. I soften the equilibrium concept by leaving feature (a) untouched, while fixing dimension (b).

Think of putting a single individual into this market and of constraining him to use a rule of the form (41). Suppose that everyone else, called the market or the representative agent, uses C, making the equilibrium price obey (44).[3] We assume that the single individual chooses a c to yield the best fitting model of the form:[4]

$$p_t = \frac{1-(1-c)L}{1-L}\epsilon_t \tag{45a}$$

[3] This 'big C, little c' formalism parallels constructions of Stokey (1989) and many others, but here the decision rule is a forecast function. Evans and Honkapohja (1993) formulate and compute a closely related approximate equilibrium under a constant gain learning rule. Their δ plays the role of our c. Their economic model is nonlinear and has multiple rational expectations equilibria. They use computer simulations to approximate an equilibrium setting of a gain parameter. Marcet and Nicolini (1997) also use computer simulations to approximate equilibria under a constant gain algorithm.

[4] I should defend dividing through by (1-L). Two methods in the literature are to divide through by $(1-\rho L)$, where ρ is made to approach 1 from below; or to set initial conditions that initialize the $\{\epsilon_t\}$ process to be zero before some date.

or

$$p_t = g(L)\epsilon_t. \tag{45b}$$

The individual uses a forecasting rule parameterized by c, and chooses c to minimize the one-step ahead forecasting error.[5]

The market sets C, and the individual sets c, given C. Proceeding in the spirit (but not the letter) of Lucas and Prescott (1971) and Brock (1972), I propose the following:

DEFINITION: Given C and the consequent stochastic process for the price (44), an individual's best forecast parameter $c = B(C)$ is the nonlinear least squares estimator of c in (45a), where the data are generated by (44).

Given (a, b, σ_u^2), the nonlinear least squares problem induces a mapping $c = B(C)$. To complete our equilibrium concept, we shall use this best-estimate map like a best-response map. Given C, c solves the following minimum variance problem:

$$\hat{c} = B(C) = \operatorname{argmin}_c \left\{ E\left[g(L;c)^{-1}\left(\nu + f(L;C)u_t\right)\right]^2 \right\}, \tag{46}$$

where the expectation is taken with respect to the distribution of the u_t's. I deduced this expression for $E\epsilon_t^2$ by inverting (45b) to write $\epsilon_t = g(L)^{-1}p_t$, then using (44b) for p_t. Let the minimized value of the criterion on the right side of (46) be denoted $\bar{\sigma}_\epsilon^2$. This is the *actual* variance of one step ahead forecast errors associated with using the misspecified model (45).[6]

[5] To facilitate the minimization that defines $B(C)$ using the frequency domain calculations in the footnote below, I approximated $g(L)$ in (45) by $\frac{1-(1-c)L}{1-\rho L}$ where $\rho < 1$ and set ρ close to 1. This approximation keeps the spectral density of the approximating model well defined.

[6] The parameter σ_ϵ has been concentrated out in this nonlinear least squares problem, and can be determined as the optimized value of (46). In calculating, we follow Sims (1993) and Hansen and Sargent (1993) by working in the frequency domain. The free parameters in (45) are c, σ_ϵ^2. The process has mean 0 and spectrum $G(\omega) = g(\exp(i\omega))g(\exp(-i\omega))\sigma_\epsilon^2$. Notice that $G(\omega)$ depends on c, through $g(\omega)$. Following Hansen and Sargent (1993), the best approximating

For a given C, the right side of (46) is an approximation problem like the ones of Sims (1971) and White (1982, 1994), who in various contexts studied the behavior of maximum likelihood estimators of misspecified models. Their formulations apply here because the agents inside our model behave like econometricians with misspecified models.

We must go beyond Sims's and White's formulations for the following reason. In their formulations, the true model is fixed from outside and does not depend on estimates of the approximating model; but in our setting, it will. This can be seen directly from (44), where the expectations parameter C plays an important role. This equilibrium concept imposes two features: (1) given the stochastic process for the price, an individual market participant's expectation parameter c must satisfy (46); and (2) the representative agent must be representative, requiring $C =$ B(C). These features are captured in:

DEFINITION: An equilibrium under forecast misspecification is a fixed point of B.

Kalai and Lehrer (1993) applied a theorem of Blackwell and Dubins (1962) to establish general conditions under which forecasts eventually must merge with rational expectations. The present model shuts off the Blackwell-Dubins mechanism at the outset because the agents' model is wrong.[7] In an equilibrium

c, σ_ϵ^2 can be computed by minimizing with respect to c, σ_ϵ^2 the expression

$$A(c, \sigma_\epsilon^2) = \frac{1}{N} \sum_{j=0}^{N-1} \left\{ \log G(\omega_j, c) + [G(\omega_j, c)^{-1} F(\omega_j)] + \nu^2 G(0)^{-1} \right\}, \qquad (47)$$

where $\omega_j = \frac{2\pi j}{N}$, for $j = 0, \ldots, N-1$. The sum approximates $(2\pi)^{-1}$ times the integral from $-\pi$ to π across frequencies. After we calculate the equilibrium C and the associated σ_ϵ^2, we form $\bar{\sigma}_\epsilon^2 = \exp K$, where K is $A(C, \sigma_\epsilon^2)$ and $C = B(C)$. The quantity $\bar{\sigma}_\epsilon^2$ is the true prediction error variance associated with using the wrong model.

[7] The model puts zero probability on events that have positive probability under the truth, violating Blackwell-Dubins' absolute continuity condition. For

with forecast misspecification, sufficient data and appropriate statistical tests would eventually tell how forecasts could be improved. However, it could require a large data set. Information about how much data would be needed is contained in the relationship between the spectral densities of the true and forecasting models, which we now briefly study.

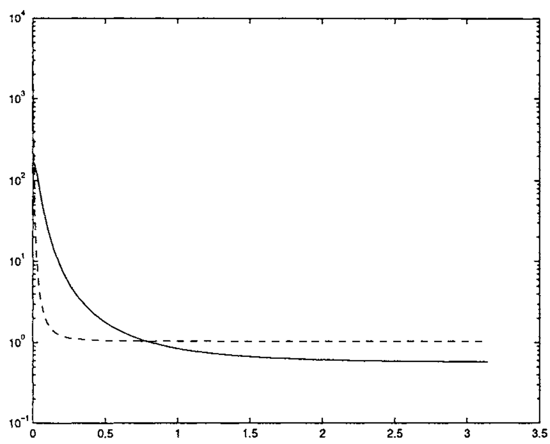

Figure 6.1. Log spectral densities of the true (solid) and forecasting (dotted line) models ($a = 1, b = .5, C = .081$). Angular frequency is the ordinate.

For parameters $a = 1, b = .5, \sigma_u^2 = 1$, I computed an equilibrium with misspecification to be $C = .0805, \sigma_\epsilon = 1.01579, \bar{\sigma}_\epsilon = 1.0579$. Remember that $\bar{\sigma}_\epsilon$ is the actual and σ_ϵ the believed one-step ahead standard error of forecast for the approximating model. For the true model, the standard error of forecast is $\frac{1}{1-b(1-C)} = 1.0208$. For these parameter values, Figure 6.1 plots the equilibrium spectral densities of the true and misspecified models. In interpreting this graph, it is important to remember that in minimizing (47), the approximating model is using the

discussions of merging results and their relevance for models of learning, see Kalai and Lehrer (1993) and Marimon (1997).

unit root to fit the mean.[8] Thus, the large gap between the spectral densities at low frequencies reflects how the approximating models fit first moment with features of second moments.[9] The least squares problem sets the one parameter $C = B(C)$ to achieve a compromise across frequencies.[10]

The true spectral density of the price process in Figure 6.1 has Granger's (1968) typical spectral shape, decreasing sharply with increases in frequency. This shape reveals substantial positive serial correlation in the price process. Agents' beliefs that the price is subject to permanent shocks, as reflected in (40), cause shocks to have persistent effects on price.

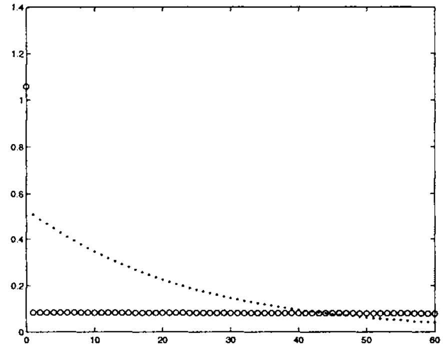

Figure 6.2. Impulse response function of true (dots) and approximating (circles) model.

[8] This is captured in the term $\nu^2 G(0)^{-1}$ in (46).

[9] It can be shown that using unconditional second moment properties to approximate conditional first moment properties underlies the example in section 4 of Sims (1993). Sims uses a linearly indeterministic model with a constant mean to approximate a model with a periodic mean (i.e., a seasonal dummy). Conceptually, the only difference between Sims's example and ours is that we have put the spike in the true spectral density at frequency zero rather than the seasonal frequencies that preoccupied Sims.

[10] We calculated $\bar{\sigma}_\epsilon$, the actual one-step ahead standard error of prediction from the misspecified model, to be 1.75, whereas the true model has a corresponding standard of prediction error of 1.02.

Figure 6.2 plots impulse response functions for the true and approximating models. The impulse response for the true model affirms the serial correlation in the price. The approximating model has a tendency to under predict short term consequences of a shock while over predicting long term ones, because the true function lies above the approximating one at first, then falls below it.

Lessons

The agents in the misspecified model are boundedly rational, where rational describes their use of least squares and bounded describes their model misspecification. Christopher Sims (1980) referred to bounded rationality as a wilderness for reasons that readers of this chapter may appreciate. Under the rational expectations assumption, there is only one model in play, though there can be different information sets. Under bounded rationality there must be at least two models, the one used by the boundedly rational agents, and the true one. These mutually influence each other, because the boundedly rational agents use their model to approximate the true one; and because the true one reflects the decisions of the agents. Both differ from the rational expectations model.

Beyond introducing a suitable equilibrium concept where agents use a misspecified model, the example illustrates how equilibrium serial correlation is affected as the misspecified model adjusts to match salient features of the data. The peculiar way that the adaptive expectations model uses a unit root to mimic a constant foreshadows aspects of a version of the Phillips curve model that will help vindicate econometric policy evaluation. I now return to the business of constructing that model.