Chapter 3

Optimal Consumption with Estimated Earnings Processes

3.1 Introduction

In the previous chapter, we found that relaxing the assumption of normality dramatically changes predictive distributions for future earnings. For example, using an AR(1) specification with random effects and general form for the error, the error distribution was estimated to be much more heavy-tailed than the normal distribution. In some cases, the semiparametric estimates of the autoregressive coefficients were much higher than the estimates under normality. In this chapter we attempt to evaluate the consequences these differences could have for stylized models of optimal consumption behavior. The goal is to develop alternative, economically motivated measures for comparing statistical models for earnings. We will only develop solutions in very simple cases; numerical solutions could be extended to more realistic situations, although the computations can quickly become very expensive. In spite of their simplicity, the examples worked out below clearly indicate that the discrepancy between the parametric and semiparametric approaches to modeling earnings dynamics is sizable, when measured by their implications for consumption smoothing.

We begin in the next section by setting up a stylized consumption problem, in which a consumer faces uncertainty about future earnings and uses a riskless asset to carry wealth into future periods and smooth fluctuations in income. Recently there has been renewed interest in models of optimal consumption because of the possibility that incorporating liquidity constraints and a precautionary motive for saving leads to theoretical predictions which track observed behavior fairly well. (Deaton (1991), Carroll (1992), Carroll (1997)). In addition, consumers with a precautionary motive can exhibit "Keynesian" responses to an increase in future taxes, thus breaking Ricardian equivalence (Barsky, Mankiw, and Zeldes (1986), Kimball and Mankiw (1989)). See Deaton (1992) for an overview of the literature on consumption.

Before attempting numerical solutions we look at some simple cases, for which closed-form solutions can be obtained. In the first case, utility has a logarithmic form and income can be insured against, so that the consumer consumes out of a predetermined endowment. The second case allows for uninsurable risky income, but assumes that income in each period is i.i.d., and the consumer has exponential utility. This allows an explicit expression to be derived for the value function in every period.

Next we bring in autocorrelated earnings, liquidity constraints and general isoelastic utility. This lets us use risk calculations to make comparisons between different statistical models of earnings. Dynamic programming is used to obtain optimal consumption policies in a finite horizon model, where the value function in the terminal period could be interpreted as giving continuation payoffs. We compare optima using the parametric and semiparametric correlated random effects models for college-educated male heads of household as the specification for the earnings process. The loss from using the "wrong" earnings dynamics model is calculated in certainty equivalent terms.

In the second set of numerical exercises, based on the work of Smith (1991), we consider a limited family of consumption policies, and try to find the best consumption policy within this restricted class. Looking only at a limited set of "rules of thumb" allows us to deal with longer horizons quite easily, and also to use predictive distributions that incorporate parameter uncertainty. In essence, we simply use the predictive draws from the previous chapter to perform Monte Carlo integration for expected utilities under a given decision rule.

The two models for earnings yield noticeably different optimal consumption policies. This seems to result mainly from the very different autoregressive coefficients estimated under the two assumptions about the model errors. That the estimated coefficients differ so greatly is in itself an empirical puzzle, so we reconsider this issue and try to provide an interpretation based on heuristic large-sample considerations.

3.2 A Stylized Consumption Problem

Consider a consumer who seeks to maximize expected utility:

$$E\left[\sum_{t=1}^{T-1} \beta^{t-1} u(c_t) + \beta^{(T-1)} V_T(x_T)\right], \tag{3.1}$$

where V_T gives terminal payoffs, and x_T as cash on hand at period T. The expectation is with respect to the joint distribution of an earnings sequence $(y_1, \ldots, y_T)$.

The consumer starts out with cash on hand of x_0 and also receives income y_1 in period 1, giving cash on hand at the beginning of period 1 of $x_1 = x_0 + y_1$. For $t = 2, \ldots, T$, cash on hand is given by $x_t = (1 + \rho)(x_{t-1} - c_{t-1}) + y_t$.

Different models for the income process can be interpreted as giving different joint distributions $F(y_1, \ldots, y_T)$. Denote the history of earnings at time t by $h_t = (y_1, \ldots, y_t)$. The consumer's problem is to choose a collection of functions

$$c_t(x_t, h_t) \in \Re,$$

for $t = 1, \ldots, T - 1$, to maximize expected utility. A natural restriction is to require consumption to be nonnegative; if the consumer cannot borrow against future income, we would also impose the restrictions that $c_t \leq x_t$. Let the feasible set for consumption be given by $c_t \in \Gamma(x_t, h_t)$.

Under fairly weak assumptions this problem has a solution which can be calculated via backwards induction. See Hernández-Lerma and Lassere (1996) for results on existence of optima in discrete-time dynamic programs. Starting at period $T-1$, we can recursively define the value functions:

$$V_t(x_t, h_t) = \max_{c_t \in \Gamma(x_t, h_t)} \left\{ u(c_t) + \beta \int V_{t+1}((1+\rho)(x_t - c_t) + y_{t+1}, h_{t+1}) dF(y_{t+1}|h_t) \right\} \quad (3.2)$$

At each stage, the maximizing argument c_t of the right hand side, regarded as a function of x_t and h_t, will give the optimal consumption policy. This does not require that the solution be an interior solution, nor that the consumption policy be differentiable in cash on hand.

This dynamic programming characterization of the optimum also provides a way to approximate a many-period model by a model with a shorter horizon, by interpreting the terminal payoff function V_T as continuation payoffs. For example, Gourinchas and Parker (1996) use this idea to avoid having to specify preferences for consumption after retirement, essentially calibrating the continuation payoffs from observed consumption behavior after retirement. In the next section, we will try to solve some simple consumption problems using short horizons ($T = 10$ and $T = 3$) using numerical methods. Before developing the numerical solution, we will consider some simplified versions of the consumption model, for which closed-form solutions are easily obtained, in order to get some intuition for reasonable choices for the continuation payoffs.

First consider the case without liquidity constraints, and suppose that $u(\cdot)$ and $V_t(\cdot)$ are concave. We will also assume that there are complete markets in contingent claims, and let x_0 denote the price of the uncertain future income stream, regarded as a bundle of contingent claims. If utility in each period is logarithmic, the analysis of Samuelson (1969) (see also Ingersoll (1987)) shows that

$$V_{T-k} = (1 + \beta + \cdots + \beta^k) \log(x_{T-k}) + q,$$

which has the same form as V_T. (q is an arbitrary constant.) Taking the limit as $k \to \infty$, this becomes $\frac{1}{1-\beta} \log(x) + q$. Analogous results can be obtained for a general CRRA utility function; see Ingersoll (1987). This suggests that if we are going to interpret V_T as continuation payoffs, and T is small relative to the "true" horizon, we could use the same function as u, except we would need to scale it up by approximately $1/(1-\beta)$. For example, if $\beta = .91$, then $1/(1-\beta) \approx 11$. However, in this setting x_t should be interpreted as containing the present value of future earnings, in addition to cash on hand.

If labor income cannot be insured against, it does not seem possible to obtain useful closed-form expressions with CRRA utility. However, one can use exponential utility, which exhibits constant absolute risk aversion. The results below are similar to those in Cantor (1985); there the income process is allowed to be a Gaussian linear process, and optimal consumption is characterized, although an expression for the value function is not explicitly derived. See also Caballero (1990), for an analysis in discrete time with an infinite horizon, and Kimball and Mankiw (1989), for similar results in a continuous time model.

If we assume that income y_t is i.i.d., $\beta = 1/(1+\rho)$, and utility takes the exponential form

$$u(c) = V_T(c) = -e^{-ac},$$

then the following result is easy to verify by induction.

Proposition 1 *Given the assumptions above, and assuming that the first order conditions characterize the optima in every period, the value function in period $T-k$ can be written as*

$$V_{T-k}(x_{T-k}) = -(1+R+R^2+\cdots+R^k)e^{-\frac{ax_{T-k}}{1+R+\cdots+R^k}} \cdot B + q,$$

where q is an arbitrary constant and

$$B = \prod_{s=1}^{k} \left[E(e^{-\frac{ay}{1+R+\cdots+R^{s-1}}})\right]^{\frac{R^k+R^{k-1}+\cdots+R^{k-s+1}}{1+R+\cdots+R^k}},$$

where $(1+R+\cdots+R^{s-1}) = 1$ if $s=1$ and $(R^k+R^{k-1}+\cdots+R^{k-s+1}) = R^k$ if $s=1$. y is a random variable with the same distribution as y_t, $t=1,\ldots,T$.

This characterization does not rule out negative consumption (and with normally distributed income, no amount of saving can guarantee nonnegative consumption). With exponential utility, this still leads to a well-defined solution, though clearly this model should not be taken too literally.

One can then make some further assumptions about the distribution of y_t and the other model parameters in order to see how V_t varies in t. For example, suppose income has a normal distribution with mean μ and variance σ^2. In the college graduates sample used in the previous chapter, the mean of labor earnings was about \$40,000, and the standard deviation was approximately \$18,000. The factor B is straightforward to evaluate using standard results on moment-generating functions. We assume that $T = 40$, and use the previous results to generate value functions and their derivatives at various choices for k.

Figure 3.1 shows the value function for $k = 0, 3, 10, 20, 37$ (setting $q = 0$ in each case). We set the coefficient of absolute risk aversion $a = 3/50,000$, and $R = \beta = 1/(1.05)$. In the terminal period utility is quite sensitive to changes in wealth when wealth is close to 0. However, in earlier periods this sensitivity declines markedly, because borrowing against future (expected) income can be used to correct temporary dips in fortunes.

3.3 Optimal Consumption with Liquidity Constraints

3.3.1 Ten Period Model

The analysis in the previous section does not apply in the case where the consumer is not allowed to borrow. Then the first-order conditions may not apply, though the value function characterization of the optimum in equation 3.2 is still appropriate. We can use equation 3.2 to numerically solve for the optimal consumption rule; moreover we can employ more realistic specifications for earnings, and compare in a limited way some of the different approaches to inference for earnings dynamics developed in the previous chapter. However, computational constraints severely limit the range of models that can be solved using this method. We will focus on very simple cases, where relatively straightforward methods can provide answers quickly. More complex models can be handled by direct extension of the methods explained below.

We will assume that the felicity function $u(\cdot)$ and terminal payoffs V_T take the CRRA form, with coefficient of relative risk aversion γ:

$$u(c) = \frac{u^{1-\gamma}}{1-\gamma}.$$

We also assume that the consumer cannot borrow. Thus $\Gamma_t(x_t, h_t) = [0, x_t]$. This is essentially the same as the setup in Deaton (1991), except for the finite horizon.

We first consider a ten-period model, in which the earnings process is given by:

$$\log y_t = 10.5 + .51(\log y_{t-1} - 10.5) + \epsilon_t,$$

where

$$\epsilon_t | y_1, \ldots, y_{t-1} \sim \mathcal{N}(0, .24^2).$$

The parameter values of the earnings process were chosen to mimic the distribution of earnings in the college subsample. In particular, that subsample had a mean log earnings of 10.5, the estimate (posterior mean) of the autoregressive coefficient in the parametric correlated random effects (CRE) AR model was 0.51, and the estimated standard deviation of the error term was 0.24. We are implicitly assuming that the individual-specific intercept is known and equal to 0; later on in this section we will consider how the optimum would change if we altered the choice of the intercept.

Because the earnings model is Markov and we are assuming for now that the agent knows the process generating the data, the only part of the earnings history h_t that is relevant for expected utility at time t is y_t. Thus we can write $V_t(x_t, h_t)$ as $V_t(x_t, y_t)$; the state space in each period will be only two-dimensional.

A more realistic model for earnings would incorporate an age profile, rather than assume earnings to be stationary over the horizon, and the numerical procedures discussed below could be easily modified to incorporate this feature. However, over the relatively short horizons being considered here, the assumption of stationarity might be reasonable. Our results turn out to be similar to other work which explicitly models the life-cycle pattern of earnings.

We set: $\beta = (1/1.1) \approx .91$, $\rho = .05$, and $\gamma = 3$. These choices are similar to those in Deaton (1991) and other recent work on precautionary saving. The terminal period payoff V_T is specified to be $u(x_T)$, where u is the isoelastic felicity function as before.

Numerical methods have been widely used to solve dynamic consumption problems. Recent work along these lines includes Deaton (1991), Carroll (1997), Hubbard, Skinner, and Zeldes (1994), Hubbard, Skinner, and Zeldes (1995), Engen and Gruber (1995), Gourichas and Parker (1996), and Cocco, Gomes, and Maenhout (1997).

The computational algorithm proceeds as follows: for $t = T-1, T-2, \ldots, 1$

1. Generate a grid of values for x_t, y_t.[1]

2. At each point in this grid, set up a grid of feasible values for c_t; this should reflect the liquidity constraints.

[1] The size of these grids will depend on t. We started by choosing minimum and maximum values for $\log(x_1)$ and $\log(y_1)$ (9.5 to 11.5 in both cases), then working out the corresponding ranges for later periods to ensure that no extrapolation of the value function would be needed. The step size of the grid was .1 (in log points) along both dimensions and for all t, although it would be straightforward to allow the step size to vary. To implement the calculations on a computer, it is useful to use data structures that can aggregate and index grids of different sizes efficiently, such as Matlab's cell arrays or suitably general container classes in a language such as C++. There seems to be additional scope to develop new data types to efficiently perform these recursive calcuations in future applications.

3. For each possible combination of x_t, y_t, and c_t, evaluate the expected utility, as in the right hand side of equation 3.2:

$$u(c_t) + \beta \int V_{t+1}((1+\rho)(x_t - c_t) + y_{t+1}, y_{t+1}) dF(y_{t+1}|y_t).$$

4. The expectation in the preceding expression is calculated using Gauss-Legendre quadrature with ten abscissae.[2] Except in period $T-1$, where interpolation of V_T is not needed, the value function V_{t+1} is interpolated linearly.

5. The maximizing choice of c_t within the grid set up previously is saved, along with the maximized value function $V_t(x_t, y_t)$, and the algorithm proceeds back to period $t-1$.

Figure 3.2 shows the optimal consumption policy in period 1, for three values of y_1: \$13,360 (in logs, 9.5), \$36,320 (10.5), and \$89,320 (11.4). These correspond approximately to the .025, .5, and .975 quantiles of earnings in the college subsample (aggregated over the 10 years of data) from which the AR parameter estimates were taken. The optimal policy is to consume all of current wealth up to some "break point," which depends on the current draw for y. At the median level of earnings, the break point occurs around $x_1 = \$33,000$. If earnings is at the .975 quantile, then the break point is higher because

[2]Since the conditional distribution is normal in this case, it would be more natural to use Gauss-Hermite quadrature. However, the Gauss-Legendre rule seems to give adequate results, and can be used without modification for the nonnormal conditional densities considered later. In future work we plan to investigate the use of adaptive quadrature methods, and make use of error bound analysis to ensure numerical accuracy.

future earnings are expected to be high, and hence there is less need to store value in case of low earnings draws.

For comparison, Figure 3.3 shows optimal consumption in period 9. The qualitative shape is similar to the initial period case; however, the slope is much higher after the break, so that a higher fraction of wealth is consumed for any values of x_t, y_t. This is quite sensible given that all of x_T will simply be consumed, so that any savings carried forward into the last period only serves to buffer against a low draw for y_T.

Figure 3.4 shows V_T, the isoelastic terminal period payoff function, and V_1, the value function in the initial period, for median levels of current income. The dramatic flattening of the value function which was observed in the exponential utility case with i.i.d. income and no liquidity constraints (Figure 3.1) is not evident here. The initial period value function looks much like the terminal period payoffs, although its slope is slightly lower at small values of x and slightly higher at large values of x. The introduction of liquidity constraints increases marginal utility at low levels of wealth, because a low income draw cannot easily be smoothed by borrowing. In addition, the impatience of the agent ($\beta < 1/(1+\rho)$), effectively shortens the relevant horizon of the problem. So the loss in generality from considering a ten-period problem appears to be relatively small.

We next compute the optimal consumption policy calibrated with the posterior mean of the semiparametric CRE model. This has an AR coefficient of .75, and a heavy-tailed error distribution. (We use the same mean of 10.5 as in the parametric model.) The error distribution has the countable normal mixture form, with density

$$\sum_{j=1}^{\infty} p_j \phi(\cdot|\mu_j, \sigma_j^2);$$

where $\sum_{j=1}^{\infty} p_j = 1$ and $\phi(\cdot|\mu, \sigma^2)$ denotes a normal density function with mean μ and variance σ^2. The parameters are set to their posterior means from the analysis in Chapter 2, and we are assuming that the consumer knows this distribution. Figure 3.5 compares the error distributions used in the parametric and semiparametric specifications for earnings. Figure 3.6 shows the optimal consumption policy in the first period. If current earnings is around the median, then the optimal consumption policy looks similar to the optimum under the parametric earnings model. However, if current earnings are high, then the higher persistence implies that future earnings are higher in expectation. Thus the consumer can consume more of current cash on hand. If current earnings are low, future earnings are lower in expectation than under the parametric model, so it is prudent to save more under the semiparametric model for earnings. The consumer saves some of cash on hand even when it is between \$15,000 and \$20,000, whereas under the parametric earnings model the consumer would consume all of cash on hand in that range.

The preceding discussion suggests that the higher AR coefficient in the semiparametrically calibrated model could be driving the difference in consumption policies. We also solved the consumption problem under the parametric earnings model, but changing the AR coefficient to .75. Figure 3.7 shows optimal consumption in this model. This is very similar to the optimum in the semiparametric model; however, the break point at the high earnings draw is slightly lower. So much of the difference between the two earnings models comes from the different AR coefficients, a point we will discuss further below.

3.3.2 Truncation at Three Periods

If a ten-period model is appropriate, is there much lost by truncating the problem at even shorter horizons, say three periods? We return to using the earnings model with an autoregressive coefficient of 0.51 and normally distributed errors. Figure 3.8 shows the results from the three period model using the original earnings process, where we set $V_T = 2 \cdot u$. The consumption functions are qualitatively similar to those in Figure 3.2, although the slopes are somewhat higher than in the ten-period case. The analogous plot using the semiparametric earnings model is given in figure 3.9. Again, the optimal consumption rule in period 1 has break points similar to the optimum with $T = 10$ (using the semiparametric earnings model), but the slope of the consumption function above the break point appears to be higher in the three-period case.

We also consider how the consumption policies change with various modifications of the parameters of the decision problem. First we set $V_T = 10 \cdot u$. As Figure 3.10 shows, this results in break points being lower. The consumer wants to have more wealth available for consumption in the final period (or to carry forward into later periods, if V_T is interpreted as continuation payoffs), so consumption in earlier periods must adjust accordingly.

Next we return to using $V_T = 2 \cdot u$ but set γ (the coefficient of relative risk aversion) to 15 instead of 3. This is shown in Figure 3.11. The break points for the consumption function have shifted down. The consumer, being more prudent, saves more of current cash on hand to guard against low income realizations in future periods. For example, with a low earnings draw of \$13,360, the consumer with $\gamma = 3$ will consume the first \$20,000 of cash on hand, while the consumer with $\gamma = 15$ will begin to save some of any cash on hand greater than \$15,000.

We also try setting the mean for log earnings to 9.5 and 11.5; this is shown in Figures 3.12 and 3.13, respectively. As would be expected from the basic permanent-income model, an individual with very high permanent earnings will consume more of current wealth, for any level of current income; an individual with low permanent earnings will tend to consume less.

3.3.3 Calculating Certainty Equivalents

The results up to now suggest that the difference between the parametric and semiparametric models is possibly quite large, in the sense that optimal behavior looks quite different under the two specifications for earnings. On the other hand, it may be that the utility loss from using the "wrong" earnings model may be rather small. A more theoretically sound comparison is to ask how much the agent would be willing to pay to use the "right" model instead of the wrong one.

The same backward induction argument developed for the calculation of optimal consumption can be used to calculate the expected utility associated with an arbitrary consumption rule. Given a consumption rule, a $T-$tuple $\{\tilde{c}_t(x_t, y_t), t = 1, \ldots, T\}$, define

$$\tilde{V}_T(x_T, y_T) = V_T(\tilde{c}_T(x_T, y_T))$$

and recursively define

$$\tilde{V}_t(x_t, y_t) = u(\tilde{c}_t(x_t, y_t)) + \beta \int \tilde{V}_{t+1}((1+\rho)(x_t - \tilde{c}_t(x_t, y_t)) + y_{t+1}, y_{t+1}) dF(y_{t+1}|y_t).$$

Then $\tilde{V}_1$ will give the overall expected utility associated with the decision rule, given y_1

and $x_1 = x_0 + y_1$. Alternatively, taking the expectation of $\tilde{V}_1$ over the distribution for y_1 gives the expected utility given only initial endowment x_0.

Given two consumption rules $\{\tilde{c}_t(x_t, y_t), t = 1, \ldots, T\}$ and $\{\hat{c}_t(x_t, y_t), t = 1, \ldots, T\}$, where $\tilde{V}_1 > \hat{V}_1$, the consumer would be willing to pay CE in period 1 to use $\tilde{c}$ instead of $\hat{c}$, where CE solves

$$\tilde{V}_1(x_1 - CE, y_1) = \hat{V}_1(x_1, y_1).$$

Figure 3.14 shows the V_1 (at three different values for y_1) under the semiparametric earnings model, along with $\hat{V}_1$ which incorrectly assumes the normal AR model for earnings. In general, the loss from using the incorrect policy is fairly small, often too small to measure accurately given the grid size of the numerical solution method. However, at the highest earnings draw the consumer would be willing to pay a sizable premium to use the correct consumption policy, because V_1 is very flat at medium to high values of current cash on hand. For example, suppose that $y_1 = \exp(11.4) \approx \$89,000$ and $x_1 = \exp(10.9) \approx \$54,000$. Using the incorrect policy gives $\tilde{V}_1 = -1.202 \times 10^{-9}$, whereas as using the correct policy gives $V_1 = -1.152 \times 10^{-9}$. For comparison, using the optimal policy when $x_1 = \exp(10.8) \approx \$49,000$ gives an expected utility of $V_1 = -1.190 \times 10^{-9}$. So in this case the consumer should be willing to pay over $5,000 to use the correct policy. Intuitively, when current income is high, the incorrect consumption policy underestimates future earnings, and conserves too much of current wealth for consumption in later periods.

Figure 3.15 shows the reverse situation, in which the true model for earnings is the normal AR model, and we compare the optimal consumption policy to the policy computed under the semiparametric earning model. The figure shows V_1 and $\tilde{V}_1$ at the median level of current income; other choices for y_1 give similar results. The extremely large difference between the two curves suggests that using the wrong consumption rule can be very costly. For example, if $y_1 = \log(10.5)$, then the expected utility from incorrectly using the policy computed under the semiparametric earnings model gives an expected utility of -4.9×10^{-9}, if cash on hand is $x_1 = \$89,000$. Using the optimal consumption rule, expected utility is -4.6×10^{-9} when $x_1 = \$15,000$. So in this situation the consumer would be willing to pay over $74,000 to use the correct decision rule over the ten periods. Here, using the wrong model leads the consumer to smooth consumption too little; combined with liquidity constraints, this means that the consumer can have very low consumption in some periods. Since utility is unboundedly negative as consumption approaches 0, this seems to result in a huge loss from using the incorrect earnings model.

We have obtained similar results if the consumer uses the policy calculated under a normal model with incorrect AR coefficient .75, when the correct AR coefficient is .51. This suggests that overestimating the AR coefficient is much worse than underestimating it, if the parameter estimates are going to be used to formulate policies for consumption smoothing.

3.4 Evaluating rules of thumb

We now wish to consider the case with parameter uncertainty. Even under the simple normal AR model $y_t | h_{t-1} \sim \mathcal{N}(\alpha + \rho y_{t-1}, \sigma^2)$, if the model parameters $\theta = (\alpha, \rho, \sigma)$ are unknown, then the correct predictive distribution, taking into account parameter uncertainty, has density

$$f(y_1, \ldots, y_T) = \int \prod_{t=1}^{T} \phi(y_t|\alpha + \rho y_{t-1}, \sigma^2) dP(\theta).$$

Here $\phi(\cdot|\mu, s)$ refers to the density of a normal random variable with mean μ and variance s, and $P(\theta)$ is the probability measure of θ, which could be a posterior distribution if we are implicitly conditioning on past data.

This joint distribution for y will not be first-order Markov, or even stationary. Hence the dimensionality of the state space will explode as the horizon increases, and for even moderate values of T the approach taken in the previous section will not be computationally feasible. Therefore a different strategy is needed.

One approach is to look for the best consumption policy $\{\hat{c}_t(x_t, h_t) : t = 1, \ldots, T\}$ within a restricted class. Given the relatively simple form of the optimal consumption functions we found in the previous analysis, such "rules of thumb" may be fairly close to optimal. This approach is developed in Smith (1991), who gives conditions under which such an approximation to the optimal policy can be made arbitrarily precise as the class of policy functions is made progressively larger. In addition, he shows in a numerical example that even relatively simple classes of functions can do quite well in practice.

The results of the previous section, as well as other work on optimal consumption under liquidity constraints, appear to suggest that a relatively simple function might be adequate. (See for example the plots in Deaton (1991).) We start with the following function:

$$\hat{c}(x; a_1, a_2) = x + 1(x > a_1)(a_2 - 1)(x - a_1).$$

This has consumption equal to cash on hand up to a_1. If cash on hand is greater than a_1, then consumption increases only by $a_2 \in [0,1]$ for each additional unit of cash on hand. There is no dependence on time, or current income realization, at this point.

For a given choice of a_1, a_2, and a set of Monte Carlo draws for $(y_1, \ldots, y_T)$, the expected utility associated with the consumption policy function can be evaluated quite easily. Let $y_1^{(j)}, \ldots, y_T^{(j)}$ denote the j-th Monte Carlo draw for the future earnings sequence. Form $x_1^{(j)} = x_0 + y_1^{(j)}$, and for $t = 2, \ldots, T$, recursively calculate

$$x_t^{(j)} = (1+\rho)(x_{t-1}^{(j)} - \hat{c}(x_{t-1}^{(j)}; a_1, a_2)) + y_t^{(j)}.$$

Then the utility associate with policy $\hat{c}(\cdot; a_1, a_2)$ and earnings sequence $y_1^{(j)}, \ldots, y_T^{(j)}$ is given by

$$\sum_{t=1}^{T-1} \beta^{t-1} u(\hat{c}(x_t^{(j)}; a_1, a_2)) + \beta^{(T-1)} V_T(x_T^{(j)}).$$

The preceding expression can be averaged over the entire set of Monte Carlo draws $j = 1, \ldots, J$ to obtain an approximate expected utility associated with $\hat{c}(\cdot; a_1, a_2)$. Hence one can look for "optimal" choices for a_1 and a_2 under different distributions for earnings.

We evaluated the expected utility on a grid of values for a_1 and a_2. After some experimentation with wider ranges to narrow the size of the grids, we used a grid which ran from 10,000 to 40,000 in steps of 500 for a_1 and a grid which ran from 0 to .5 in steps of .025 for a_2. We set $\beta = 1/(1.1)$, $\gamma = 3$, and $\rho = .05$ as in the previous section. The initial cash-on-hand was set to \$5,000, and we set $y_1 = \exp(10.5)$.

Under the parametric normal model with no parameter uncertainty, the best function over this grid has $a_1 = 33,000$, and $a_2 = 0.3500$. The expected utility from this policy is $EU = -2.55 \times 10^{-9}$. The shape of the consumption function is similar to the period

1 optimum in Figure 3.2. The certainty equivalent loss relative to the exact optimum is difficult to measure precisely, because of the discreteness of the grid used in the calculations of the exact optima, but using log-linear interpolation suggests that the loss is under \$1,300. This is fairly small given how much we have restricted the class of possible decision rules.

We then considered generating draws for earnings from the semiparametric AR model, again with no parameter uncertainty. The initial condition, however, was drawn from the stationary distribution under the parametric model for simplicity. This resulted in a "best" policy of $a_1 = 22,500$, $a_2 = 0.4250$, with an associated expected utility of $EU = -3.43 \times 10^{-9}$. The break point occurs earlier than in the parametric case. The expected utility is lower, as one would expect from a more persistent income process in which buffering is less effective.

We then allow for parameter uncertainty in the parametric AR model. As in the previous two cases, we assume $y_1 = \exp(10.5)$. Draws for $y_2, \ldots, y_{10}$ were generated using the draws for ρ and σ from the posterior distribution in the parametric CRE model estimated on the college graduates subsample. (The appendix to Chapter 2 describes the construction of the predictive draws in both the parametric and semiparametric CRE models.) As in the previous two cases, the intercept term of the AR model was chosen so that the process (in logs) would have stationary mean of 10.5 given the draw for ρ.

The best choices for the consumption policy parameters were $a_1 = 33,500$ and $a_2 = 0.3250$, very similar to the parametric model with no parameter uncertainty. The expected utility under this policy function was -2.55×10^{-9}. The certainty equivalent loss from using the policy calculated without parameter uncertainty is essentially zero. So here the incorporation of parameter uncertainty into the predictive distributions makes little difference, given this class of functions. In the case of the semiparametric model, adding parameter uncertainty changes the best policy rule, but the differences are still fairly small. The best choices for a_1, a_2 were $(24,500, 0.4500)$, which gave an expected utility of -2.64×10^{-9}. The certainty equivalent loss from using the rule calculated without parameter uncertainty (with $a_1 = 22,500$ and $a_2 = 0.425$) is less than $1,100.

Next we allow the consumption rule to depend on the most recent draw for earnings. We set

$$a_1 = b_1 + b_2 \log(y),$$

since the break points in the exact optima calculated in the previous section seem to move nearly linearly in the log of first period income. Using the parametric AR model, with no parameter uncertainty, the optimal policy has $b_1 = 33,000$, $b_2 = 0$, and $a_2 = .35$. This decision rule is plotted, for three values of current income ($\log(y) = 9.5$, $\log(y) = 10.5$, and $\log(y) = 11.4$), in the upper left plot in Figure 3.16. This is exactly the same as the optimum in the previous set of experiments, where b_2 is restricted to be zero from the outset.

Under the semiparametric model, with no parameter uncertainty, the selected decision rule is also insensitive to current income; it has $b_1 = 17,000$, $b_2 = 550$, and $a_2 = 0.425$. So the break point is slightly lower than in the parametric case, and there is some sensitivity to current income, although it is quite small. This is shown in the upper right portion of Figure 3.16. When the predictive distribution (taking into account parameter uncertainty) from the parametric earnings model is used to evalute the decision rules, the selected rule is similar to the rule selected in the first case–the parametric model with no parameter uncertainty. It has $b_1 = 32,500$, $b_2 = 100$, and $a_2 = 0.325$. This is shown in the lower left plot in Figure 3.16. Likewise, introducing parameter uncertainty to the semiparametric specification for earnings gives a decision rule with $b_1 = 23,500$, $b_2 = 400$, and $a_2 = 0.425$, again similar to the case without parameter uncertainty, although the break point does occur at a noticeably higher level of cash on hand. This is shown in the lower right plot in Figure 3.16.

The insensitivity of the decision rule to current income is puzzling. This could be an artifact of the relatively inflexible parameterization, which restricts the break point to be linear in the log of current income.

There seems to be an excessive amount of Monte Carlo error in these results as they are currently calculated. The expected utility surface turns out to have ridges along which rules which appear to be quite different can be nearly equivalent in terms of payoffs, so that repeating the exercise with a new set of draws can give different results. The distinctions between rules of thumb made above should therefore be taken as preliminary.

3.5 Reexamining the Correlated Random Effects Models

Many of the results of the previous sections seemed to be largely driven by the different estimates for the autoregressive coefficient in the parametric and semiparametric models. In this section we try to interpret this result from the point of view of model selection.

Suppose that the true data-generating process (DGP) is equal to the correlated random effects model with normal errors, for some choice of parameter values. Then we would expect the estimate of the AR coefficient under the parametric model to be reasonably accurate; in particular it will converge in probability to the true value as n (the number of individuals in the data set) becomes large, holding T fixed. We would also expect the semiparametric model to do well, since it essentially nests the parametric model as a special case. To verify this latter point, we generated a single artificial data set using the parametric CRE model as the DGP. The parameter values were set to be equal to the posterior means from the estimates using the college graduates subsample, and the sample size and initial conditions (y_{i1} for $i = 1, \ldots, n$) were set equal to their values in the college graduates subsample. So the autoregressive coefficient in the DGP was set to 0.51. Using the semiparametric CRE model, with the same prior distribution as in the previous chapter, gave a posterior mean (standard deviation) for the autoregressive coefficient of 0.54 (0.04), reasonably close to the true value used in the DGP. Also, inference using the parametric model on the same data set gave exactly the same results for the autoregressive coefficient.

What if the DGP was a member of the semiparametric CRE family, but not a member of the parametric CRE family? We would expect inference using the semiparametric CRE model to work reasonably well.[3] Using the parametric model would give consistent estimates of ρ, even though it is misspecified, since the parametric model, being based

[3]However, it is difficult to make this statement more precise. In particular, large-sample properties of the semiparametric models under consideration appear to be limited to arguments which hold for a set of parameter values of prior measure 1. The particular parameter value associated with the "true" DGP is not generally guaranteed to be in this set.

on normality, would only use the first two moments of the data to estimate ρ, and both the parametric and semiparametric models have the same implications for the first two moments of the data for any given value of the autoregressive coefficient.

To verify this, we generated an artificial college graduates data set, using the posterior means from the semiparametric CRE model. Thus the "true" autoregressive coefficient was 0.75. The posterior mean for the AR coefficient, based on the parametric likelihood, was 0.78, with a standard deviation of 0.04.

This suggests that what is happening is that *both* models are misspecified. We begin by using the "error components" model of the previous chapter, in which income is the sum of a persistent autoregressive process and white noise:

$$y_{it} = v_{it} + \epsilon_{it},$$

$$v_{it} = \rho v_{i,t-1} + w_{it},$$

$$v_{i1} \sim \mathcal{N}(0, \sigma_v^2), \;\; w_{it} \sim \mathcal{N}(0, \sigma_w^2), \;\; \gamma_i \sim \mathcal{N}(0, \sigma_\gamma^2).$$

First, we consider the parametric version of this model, in which all innovation terms are assumed to be normally distributed:

$$\epsilon_{it} \sim \mathcal{N}(0, \sigma^2).$$

We generate an artificial data set based on the parametric estimates for college graduates, setting $\rho = .92$, $\sigma_w = .17$, $\sigma_v = .42$, and $\sigma = .16$. Then we ran both the parametric and semiparametric CRE routines on this data. The results are essentially the same: using the parametric CRE model the posterior mean (standard deviation) for the AR coefficient is

0.56 (0.04); using the semiparametric CRE model the posterior mean (standard deviation) is 0.56 (0.04).

Next, suppose that instead of normally distributed error terms, all the innovations (v_{i1}, w_{it}, and ϵ_{it}) were generated by a heavy-tailed distribution. We used t distributions with 3 degrees of freedom, scaled to have the same standard deviations as in the normal case. Using the same data set, the estimate from the parametric model was 0.56 (0.04), while the estimate from the semiparametric model was 0.59 (0.04). So the large difference between the estimates using the actual PSID data on college graduates remains puzzling.

We could have used other, more complicated specifications for earnings in the consumption applications considered in this chapter. Calculation of the optima using dynamic programming methods would become harder, because the state space would be larger, but the same basic ideas discussed in the previous sections would still apply. Still, practically any feasible statistical model for earnings dynamics would likely be imperfect. So the possibility of model misspecification would still need to be taken seriously, and introducing explicit, if stylized, utility calculations can bring some clarity to these issues. For more on the connections between model choice and decision theory, see Chamberlain (1998) and Bernardo and Smith (1994), Ch. 6.

3.6 Conclusion

Recent work on optimal consumption has calibrated earnings dynamics in a variety of ways, but approach has been essentially atheoretical: point estimates are obtained using conventional loss functions, and when there is a question as to which model for earnings is “best,” the choice has often been made informally, or by appealing to generic statistical criteria rather than by focusing on how the models might be used.

This chapter has used simplified optimal consumption programs, combined with estimates based on the college graduates subsample of the PSID, to provide economically interpretable measures of the differences between the conventional parametric approach to inference in dynamic panel data models, and the semiparametric Bayesian approach developed in the previous chapter. The ultimate goal, only partly realized in the analysis here, is to view econometric model choice as a formal decision problem, with loss functions based on economic rather than ad hoc statistical considerations. For example, we have found that in formulating consumption plans, overestimating the persistence of earnings is much worse than underestimating it. This suggests that working with symmetric loss functions, as much applied econometric analysis implicitly does, may not be appropriate, at least when there is enough parameter uncertainty. More generally, the semiparametric model for earnings leads to quite different consumption-smoothing policies than the parametric model, whether measured by the shape of the implied consumption functions, or in certainty equivalent terms. However, much of this difference comes from very different estimates of the autoregressive coefficient, which in turn raises the issue of why the estimates should be so different.

We also found that the “rule of thumb” approach was difficult to implement in practice, because unsuitable parameterizations of the decision rule could give results quite far from the unconstrained optimum. Further work is needed to see if generalizing the class of decision rules under consideration could give more useful results.

So the question of how best to deal with parameter uncertainty in complex dynamic programs remains open. An alternative approach, which we plan to consider in future research, is to look for approximate sufficient statistics for the earnings process, which could be used to reduce the dimensionality of the state space. This idea was discussed and used in Barberis (1996) in a model of portfolio choice. One could base these approximations on related models that have usable sufficient statistics (e.g. many linear Gaussian models).

While outside the scope of this study, there are many issues in the modeling of observed consumption behavior that remain open, and for which good modeling of individual earnings risk could be relevant. For example, relatively little of the recent work on optimizing models of consumption directly addresses the issue of when and how liquidity constraints might arise, and which individuals can be expected to be liquidity constrained in this manner. Clearly the model with liquidity constraints we considered above cannot apply to all consumers in a general equilibrium setting. Empirical work, such as Zeldes (1989), has tried to determine which consumers behave as if they were liquidity constrained, but more could be done to extend consumption models like those considered in this chapter, by explicitly incorporating contracting in order to endogenize liquidity constraints.

Figure 3.1: Value Functions, Exponential Utility

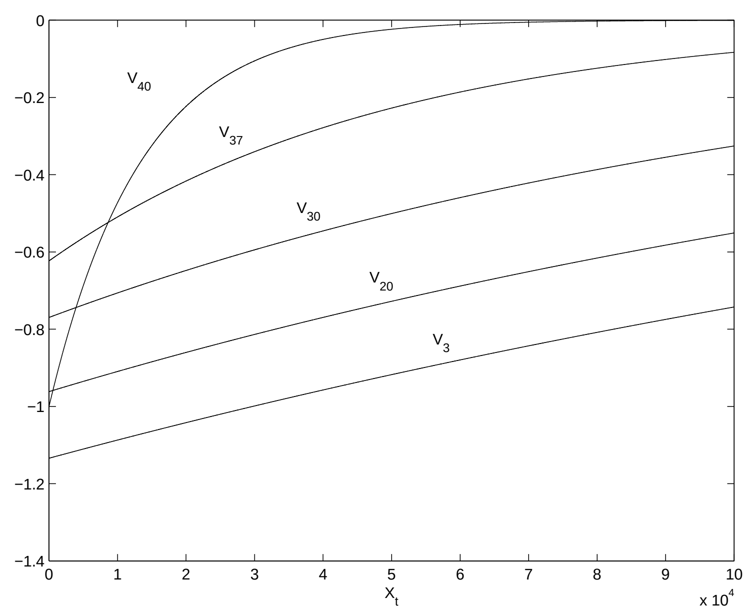

Figure 3.2: Normal AR Model, 10 Periods

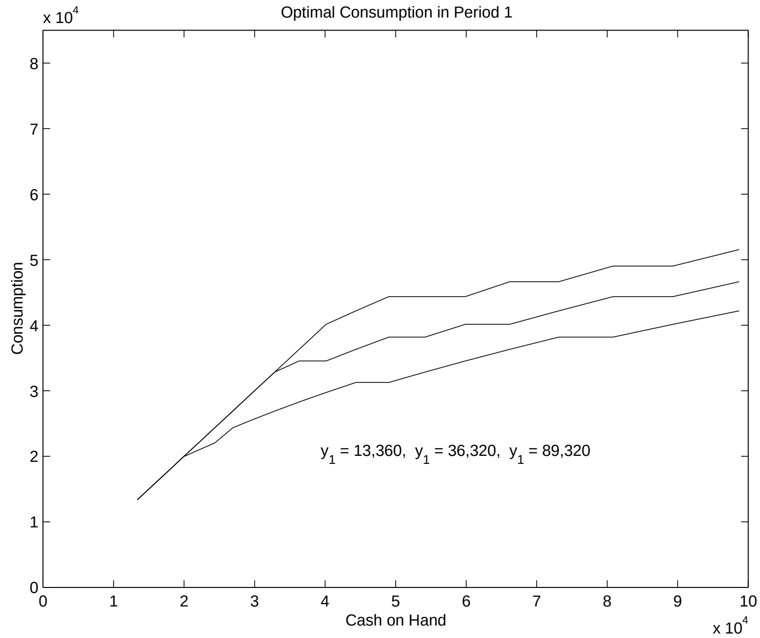

Figure 3.3: Normal AR Model, 10 Periods

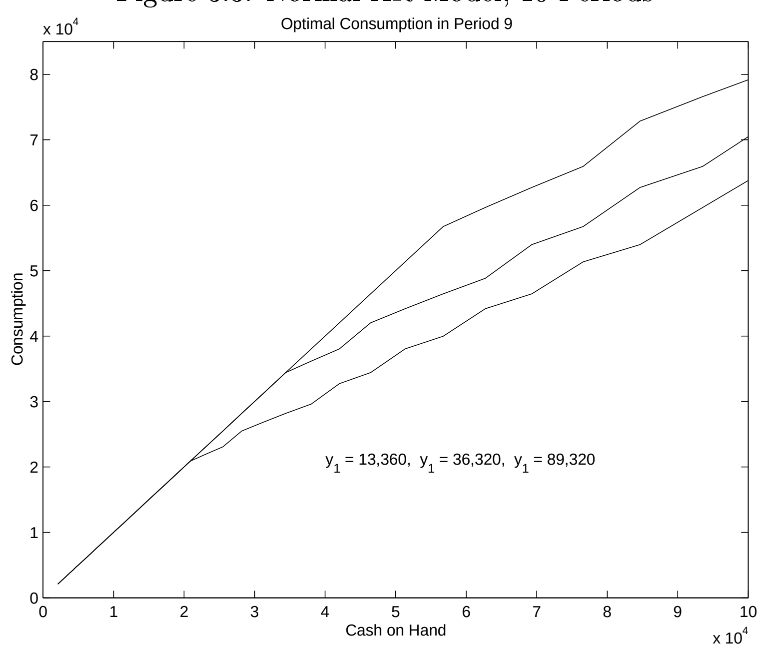

Figure 3.4: Value Functions, Normal AR Model, 10 Periods

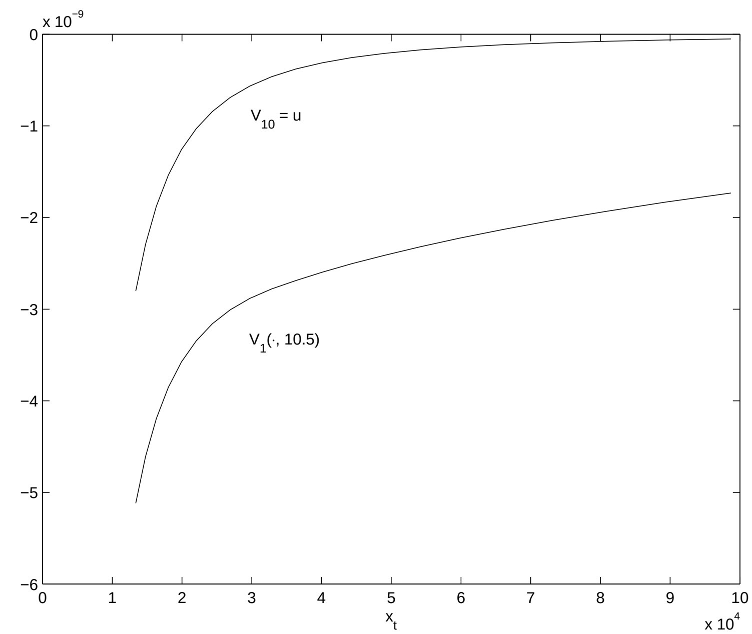

Figure 3.5: Error Distributions, Normal and Semiparametric Earnings Models

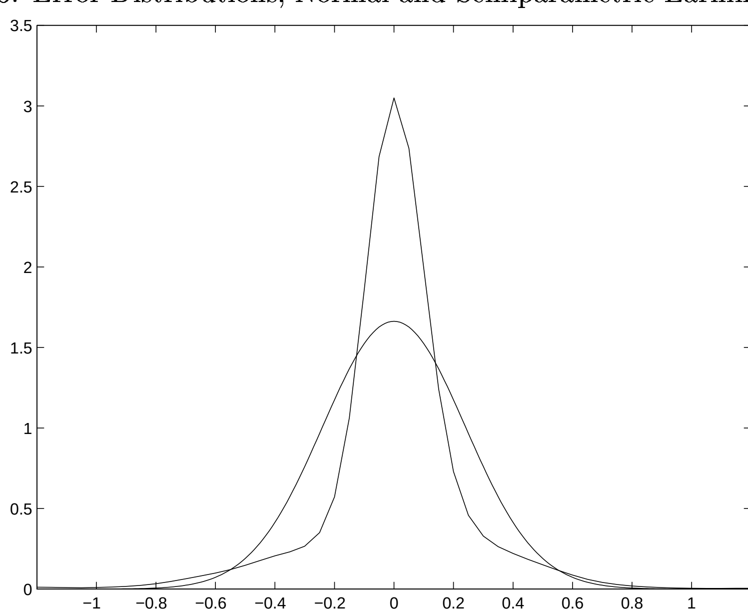

Figure 3.6: Semiparametric AR Model, 10 Periods

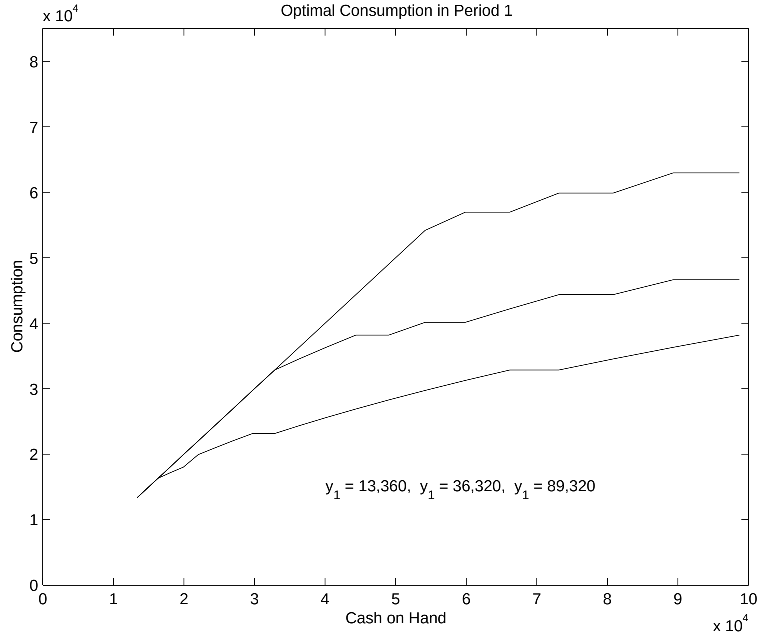

Figure 3.7: Normal AR Model with AR Coefficient = .75

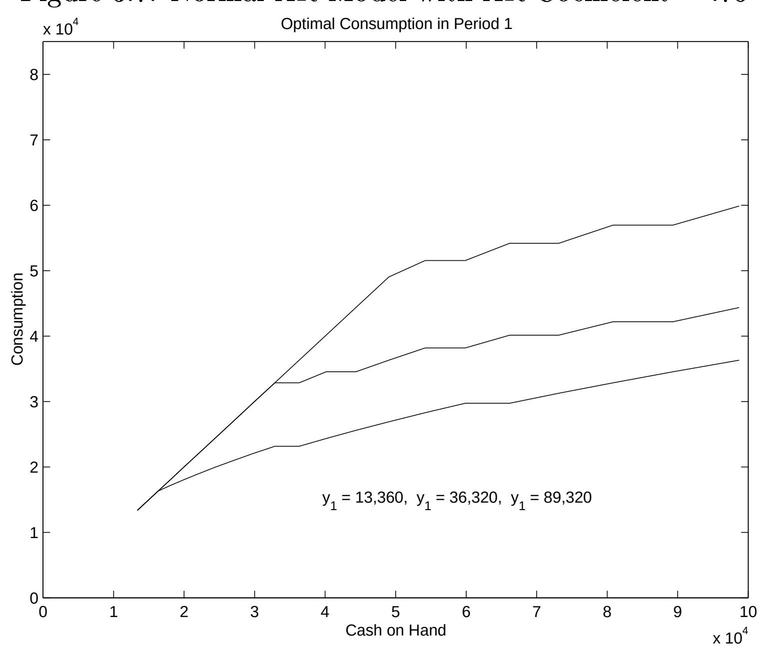

Figure 3.8: Normal AR Model, 3 Periods

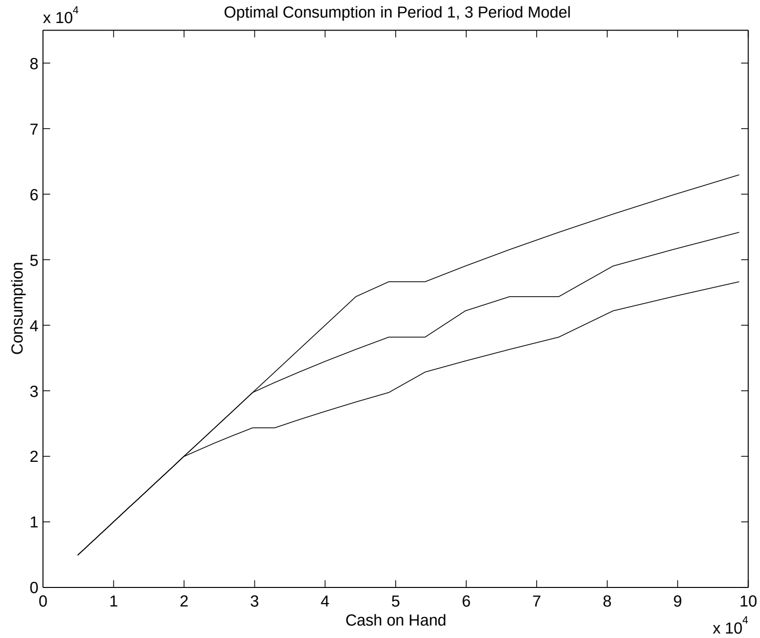

Figure 3.9: Semiparametric AR Model, 3 Periods

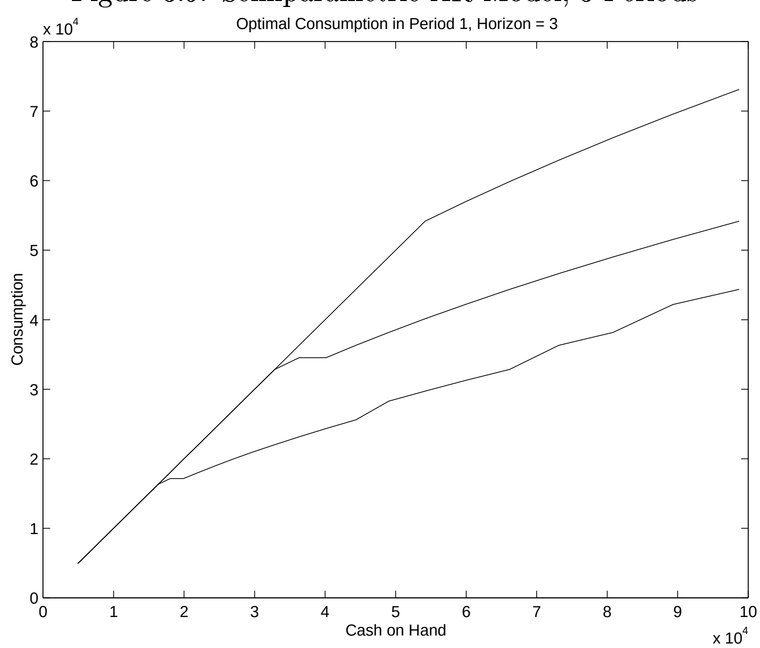

Figure 3.10: Normal AR Model, 3 Periods, Different V_T

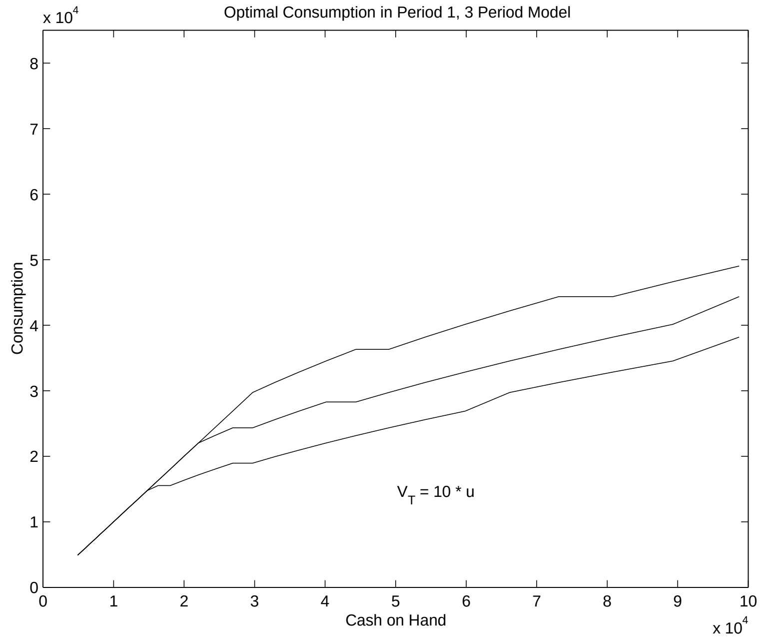

Figure 3.11: Normal AR Model, 3 Periods, $\gamma = 15$

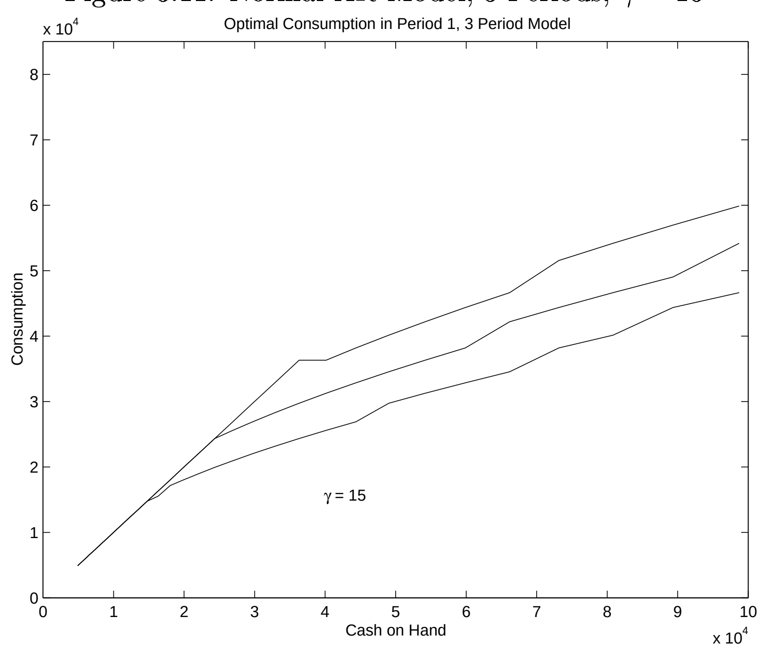

Figure 3.12: Normal AR Model, 3 Periods, $E(\log y) = 9.5$

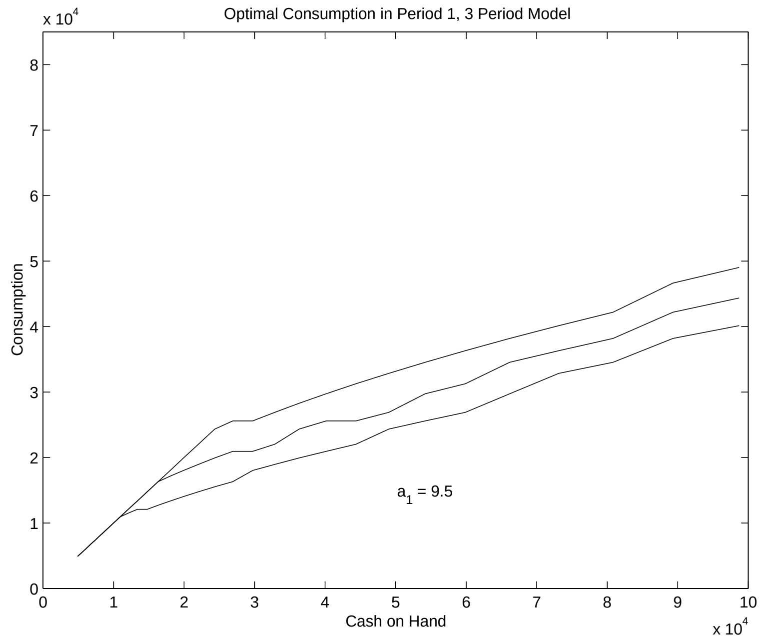

Figure 3.13: Normal AR Model, 3 Periods, $E(\log y) = 11.5$

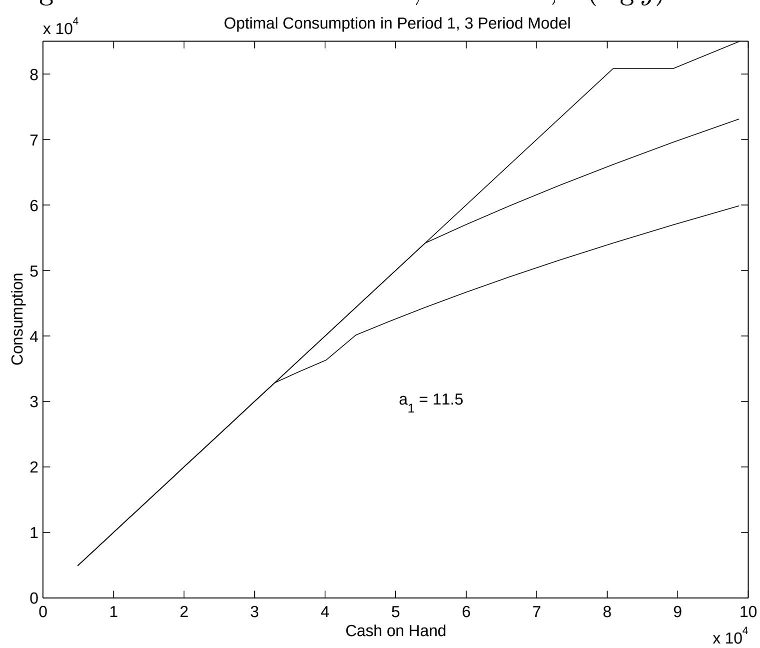

Figure 3.14: V_1 and $\tilde{V}_1$, Semiparametric AR Earnings Model

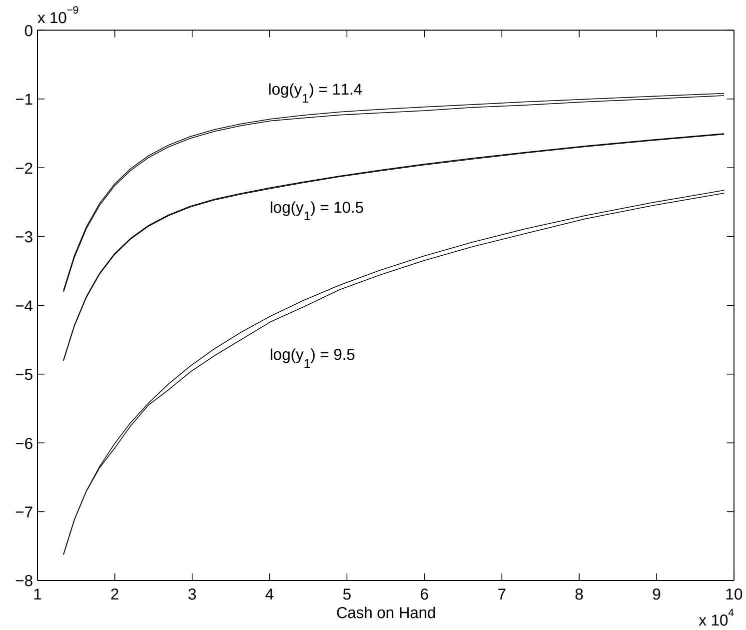

Figure 3.15: V_1 and $\hat{V}_1$, Normal AR Earnings Model

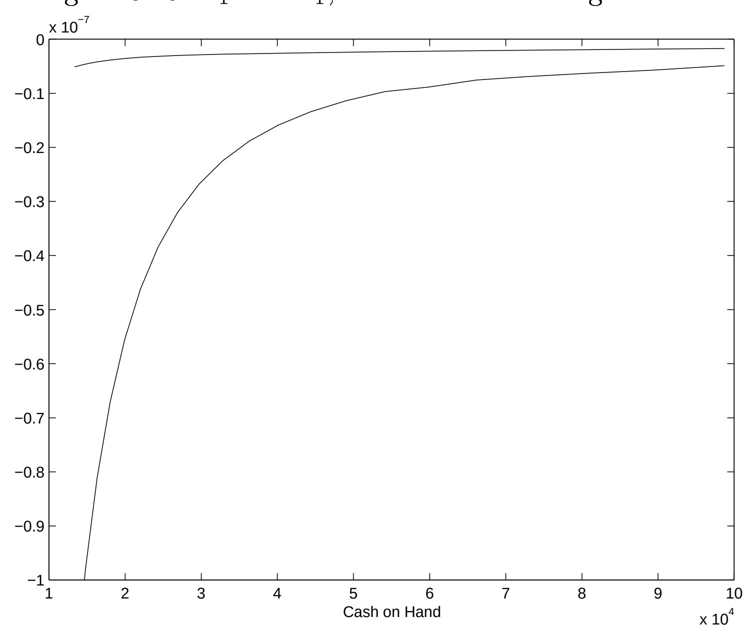

Figure 3.16: Rules of Thumb