

# NOTES

## ON ADJUSTING THE HODRICK-PRESCOTT FILTER FOR THE FREQUENCY OF OBSERVATIONS

Morten O. Ravn and Harald Uhlig\*

*Abstract*—This paper studies how the Hodrick-Prescott filter should be adjusted when changing the frequency of observations. It complements the results of Baxter and King (1999) with an analytical analysis, demonstrating that the filter parameter should be adjusted by multiplying it with the fourth power of the observation frequency ratios. This yields an HP parameter value of 6.25 for annual data given a value of 1600 for quarterly data. The relevance of the suggestion is illustrated empirically.

### I. Introduction

THE Hodrick and Prescott (1980, 1997) filter (hereafter, the HP filter) has become a standard method for removing trend movements in the business cycle literature. The filter has been applied both to actual data (Backus & Kehoe, 1992; Blackburn & Ravn, 1992; Brandner & Neusser, 1992; Danthine & Donaldson, 1993; Danthine & Girardin, 1989; Fiorito & Kollintzas, 1994; Kydland & Prescott, 1990) and in studies in which artificial data from a model are compared with the actual data (Backus, Kehoe, & Kydland, 1992; Cooley & Hansen, 1989; Hansen, 1985; Kydland & Prescott, 1982).

Although the use of the HP filter has been subject to heavy criticism (Canova, 1994, 1998; Cogley & Nason, 1995; Harvey & Jaeger, 1993; King & Rebelo, 1993; Söderlind, 1994), it has withstood the test of time and the fire of discussion remarkably well. Thus, although elegant new bandpass filters are being developed (Baxter & King, 1999; Baxter, 1994; Christiano & Fitzgerald, 1999), it is likely that the HP filter will remain one of the standard methods for detrending.

Most applications of this filter have been to quarterly data, but data is often available only at the annual frequency, whereas in other cases monthly data might be published. This raises the question of how one can adjust the HP filter to the frequency of the observations so that the main properties of the results are conserved across alternative sampling frequencies. Although most researchers have followed Hodrick and Prescott (1980, 1997) and used the value of 1600 for the smoothing parameter when using quarterly data, there is less agreement in the literature when moving to other frequencies. Backus and Kehoe (1992) use a value of 100 for annual data, whereas Correia, Neves, and Rebelo (1992) and Cooley and Ohanian (1991) suggest a value of 400.

Received for publication October 22, 1999. Revision accepted for publication May 10, 2001.

\* London Business School and Centre for Economic Policy Research, and Humboldt University and Centre for Economic Policy Research, respectively.

We are grateful to James Stock, three anonymous referees, Albert Marcet, and Dan Knudsen for useful comments. We also thank Dave Backus for provision of data.

Baxter and King (1999) have recently shown that a value of around 10 for annual data is much more reasonable. They arrive at this value by visually inspecting the transfer function of the HP filter for annual data and comparing it to a bandpass filter. Hassler et al. (1992) had already obtained a similar value by investigating the average cycle length obtained in a time series of output.

This paper complements these insights using two different analytical approaches. The first approach uses the time domain and focuses on the ratio of the variance of the cyclical component to the variance of the second difference of the trend component: this ratio is often used for calculating the smoothing parameter. For a particular benchmark stochastic process, it is shown that time aggregation changes this ratio by the fourth power of the observation frequency. The second approach uses the frequency domain and investigates the transfer function of the HP filter, thereby obtaining a general result. Again, a change-of-variable argument shows that one should adjust the HP parameter with approximately the fourth power of the frequency change. Both approaches therefore yield a value of approximately  $1600/4^4 = 6.25$  for annual data, which is close to the value of 10 given by Baxter and King (1999).

We then show that our recommendations work extremely well on U.S. GDP data: using a value of the smoothing parameter of 6.25 for annual data and 1600 for quarterly data produces almost exactly the same trend. This leads us to reconsider the business cycle “facts” reported in earlier studies. As an example, we cast doubt on a finding by Backus and Kehoe (1992) on the historical changes in output volatility and return instead to older conventional wisdom (Baily, 1978; Lucas, 1977): output volatility turns out to have decreased after World War II.

The remainder of the paper is organized as follows. Section II presents the HP filter and provides the first, time domain-based approach, whereas section III provides the second, frequency domain-based approach. In section IV, we recompute some facts about business cycles. Finally, section V concludes.

### II. A Time Domain Perspective

The HP filter removes a smooth trend  $\tau_t$  from some given data  $y_t$  by solving

$$\min_{\tau_t} \sum_{t=1}^T ((y_t - \tau_t)^2 + \lambda((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2).$$

The residual  $y_t - \tau_t$  (the deviation from the trend) is then commonly referred to as the *business cycle component*.

The filter involves the smoothing parameter  $\lambda$ , which penalizes the acceleration in the trend relative to the business cycle component. Researchers typically set  $\lambda = 1600$  when working with quarterly data. However, data does not always come at quarterly intervals. It may even be desirable to move to annual, monthly, or some other time interval of observation instead.

Thus, the question arises how the HP filter should be adjusted for the frequency of observations, and this question is the focus of this paper. We do not investigate whether the HP filter is desirable per se or aim at a comparison to some optimal bandpass filter as in Baxter and King (1999). Rather, we take it as granted that a researcher wishes to filter the data using the HP filter, and ask how the parameter  $\lambda$  should be adjusted when changing the sampling frequency.

A popular perspective on the smoothing parameter in the literature is to consider the decomposition of some given time series  $y_t$  into a trend  $\tau_t$  and a cycle  $c_t$ :

$$y_t = \tau_t + c_t \quad (1)$$

If  $c_t$  as well as the second difference of  $\tau_t$  are normally and independently distributed, then the HP filter is known to be optimal, and  $\lambda$  is given as the ratio of the two variances,  $\lambda = \sigma_c^2 / \sigma_{\Delta^2 \tau_t}^2$  (Hodrick & Prescott, 1980, 1997; King & Rebelo, 1993). However, even if the HP filter is optimal for equation (1), it is unlikely to be optimal when time aggregating the process (1) because time aggregation usually introduces moving average terms. As our focus is on adjusting  $\lambda$ , when changing the frequency of observation, we shall however ignore the issue of optimal filtering and instead simply focus on the question of how the ratio of the variances change.

It is convenient to consider a benchmark continuous-time version of equation (1) that satisfies the conditions previously stated, that is, where the cycle as well as the second difference of the trend are independently and normally distributed, taking the form of Brownian motion increments.<sup>1</sup> We then analyze the change in the variances when observing the process at discrete time intervals. Let  $y_t$  be the "flow"  $dz_t$  of some stochastic process  $z_t$  with

$$dz_t = \tau_t dt + \sigma_c dW_t^1 \quad (2)$$

where

$$d\tau_t = \mu_t dt, \quad d\mu_t = \sigma_\tau dW_t^2 \quad (3)$$

and  $dW_t^1$  and  $dW_t^2$  are two independent Brownian motions. There are two possibilities for observing the process at some discrete time interval  $\alpha$ : these observations may be time aggregated (or time averaged) or they may be sampled at these discrete time intervals. (See Christiano and Eichenbaum (1986).)

<sup>1</sup> See the appendix of Ravn and Uhlig (2001) for a discrete time analysis and for an extended discussion of the links with optimal filtering.

Consider time aggregation first; that is, for some length  $\alpha > 0$ , consider observing

$$y_{t;\alpha} = \int_{s=0}^{\alpha} dz_{t-s} = \tau_{t;\alpha} + c_{t;\alpha}$$

where

$$\tau_{t;\alpha} = \int_{s=0}^{\alpha} \mu_{t-s} ds,$$

$$c_{t;\alpha} = \int_{s=0}^{\alpha} \sigma_c dW_t^1.$$

For any stochastic process  $x_t$ , define the  $\alpha$ -differencing operator

$$\Delta_\alpha x_t = x_t - x_{t-\alpha}.$$

We are interested in how

$$\lambda_\alpha = \frac{\sigma^2(c_{t;\alpha})}{\sigma^2(\Delta_\alpha^2 \tau_{t;\alpha})}$$

changes with  $\alpha$ .<sup>2</sup>

Clearly,

$$\sigma^2(c_{t;\alpha}) = \alpha \sigma_c^2 = \alpha \sigma^2(c_{t;1}).$$

For  $\Delta_\alpha^2 \tau_{t;\alpha}$ , introduce first  $x_t = \Delta_\alpha \tau_{t;\alpha}$  and write it as

$$\begin{aligned} x_t &= \int_{s_1=0}^{\alpha} (\mu_{t-s_1} - \mu_{t-\alpha-s_1}) ds_1 \\ &= \int_{s_1=0}^{\alpha} \int_{s_2=0}^{\alpha} d\mu_{t-s_1-s_2} ds_1. \end{aligned}$$

Substitute  $d\mu_{t-s_1-s_2} = x_{t-s_1-s_2} ds_2$  and repeat this calculation to obtain an expression of the second  $\alpha$  difference,

$$\begin{aligned} \Delta_\alpha^2 \tau_{t;\alpha} &= \sigma_\tau \int_{s_1=0}^{\alpha} \int_{s_2=0}^{\alpha} \int_{s_3=0}^{\alpha} dW_{t-s_1-s_2-s_3}^2 ds_2 ds_1 \\ &= \sigma_\tau \int_{s=0}^{3\alpha} A(s; \alpha) dW_{t-s}^2, \end{aligned}$$

where

<sup>2</sup> One can equally well divide the processes by  $\alpha$  to obtain time averaging rather than time aggregation: this makes no difference for  $\lambda_\alpha$  and the calculation is very similar.

$$A(s; \alpha) = \int_{s_1=0}^{\alpha} \int_{s_2=0}^{\alpha} 1_{[0,\alpha]}(s - s_1 - s_2) ds_2 ds_1$$

and where the last equality was obtained by a change of variables,  $s = s_1 + s_2 + s_3$ . The variance is therefore given by

$$\sigma^2(\Delta_{\alpha}^2 \tau_{t;\alpha}) = \sigma_{\tau} \int_{s=0}^{3\alpha} A(s; \alpha)^2 ds. \tag{4}$$

Although one could calculate  $A(s; \alpha)$ , one does not have to. Simply observe that

$$A(s; \alpha) = \alpha^2 A(s/\alpha; 1).$$

With one more change of variable to  $\bar{s} = s/\alpha$  in equation (4), we finally find

$$\sigma^2(\Delta_{\alpha}^2 \tau_{t;\alpha}) = \alpha^5 \sigma_{\tau} \int_{\bar{s}=0}^3 A(\bar{s}; 1)^2 d\bar{s} = \alpha^5 \sigma^2(\Delta_1^2 \tau_{t;1}),$$

and hence

$$\lambda_{\alpha} = \frac{1}{\alpha^4} \lambda_1.$$

That is, the HP parameter  $\lambda$  should be adjusted with the fourth power of the frequency change. This finding will be reconfirmed in section III, using another approach.

For sampling at discrete time intervals  $\alpha$ , the calculations become simpler yet. Suppose we observe the flow  $y_t = dz_t$  at intervals  $\alpha$ .<sup>3</sup> The diffusion part still has variance  $\sigma_c^2 dt$ . What needs to be calculated is the variance of  $\Delta_{\alpha}^2 \tau_t$ . The same calculation as before leads to

$$\begin{aligned} \Delta_{\alpha}^2 \tau_t &= \int_{s_1=0}^{\alpha} \int_{s_2}^{\alpha} \sigma_{\tau} dW_{t-s_1-s_2}^2 \\ &= \int_{s=0}^{2\alpha} B(s; \alpha) dW_{t-s}, \end{aligned}$$

where

$$B(s; \alpha) = \int_{s_1=0}^{\alpha} 1_{[0,\alpha]}(s - s_1) ds_1 = \alpha B(s/\alpha; 1)$$

Similar to the calculation above,

<sup>3</sup> Observing should be understood here in the sense that the continuous-time limit approximates some discrete time process at very small time intervals.

TABLE 1.—OPTIMAL POWER ADJUSTMENT AT FREQUENCY  $\omega$  FOR AN ADJUSTMENT LOCALLY AROUND A QUARTERLY SAMPLING RATE

$\omega$	0	$\pi/20$	$\pi/10$	$\pi/5$
$m(1, \omega)$	4	3.992	3.967	3.868

As one can see, the optimal adjustment is generally between 3.8 and 4.0 at the relevant frequencies.

$$\lambda_{\alpha}^{(s)} = \frac{\sigma_c^2 dt}{\sigma^2(\Delta_{\alpha}^2 \tau_t)} = \frac{1}{\alpha^3} \lambda_1^{(s)}.$$

That is, the smoothing parameter for the HP filter should be adjusted using the third power of  $\alpha$ . This result differs from the fourth-power result for the previous time-averaged data, but it also differs from the literature suggestion of adjusting with the second or the first power of  $\alpha$ .

In practice, one may therefore wonder whether adjustment with the fourth or the third power is more appropriate. Our recommendation here is to always use the fourth power rather than the third. First, most macroeconomic time series are time averaged, so that the preceding calculation would suggest adjusting with the fourth power anyhow. But, even for the sampling case, simulations of this process shows that adjusting with the fourth power rather than the third produces essentially the same trend. The next section can be read as an explanation why this is the case.

### III. A Frequency Domain Perspective

An alternative way to look at the issue is from a frequency domain perspective, which allows us to provide a general result, as we no longer need to assume the special structure (2) and (3). The transfer function of the HP filter is given by (King & Rebelo, 1993)

$$h(\omega; \lambda) = \frac{4\lambda(1 - \cos(\omega))^2}{1 + 4\lambda(1 - \cos(\omega))^2} \tag{5}$$

This filter is similar to a high-pass filter. (See, for example, Ravn and Uhlig (1997) or Baxter and King (1999) for a plot of the transfer function.) Choosing different values for  $\lambda$  is comparable to choosing different values for the cutoff point of the high-pass filter.

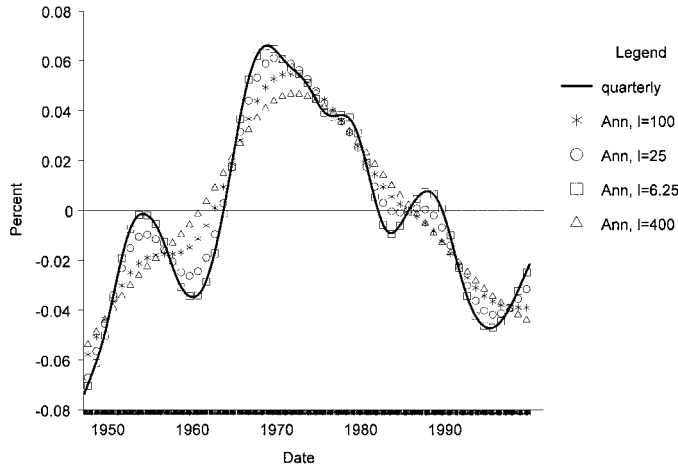
Let  $h(\omega; \lambda_1)$  be the filter representation for quarterly data and let  $h(\omega/s; \lambda_s)$  be the filter representation for an alternative sampling frequency,  $s$ , where we let  $s$  be the ratio of the frequency of observation compared to quarterly data ( $s = 1/4$  for annual data or  $s = 3$  for monthly data). Then, ideally, we would like to have

$$h(\omega; \lambda_1) \approx h(\omega/s; \lambda_s). \tag{6}$$

Although this cannot hold exactly for all  $\omega$ , it should hold at least approximately.<sup>4</sup> To derive the appropriate adjustment

<sup>4</sup> By this equation we do not mean to say that the HP filter is “optimal” in any sense; rather, it says that, as the frequency of the observations is altered, the filter—being optimal or not—should have approximately the same properties.

FIGURE 1.—TREND COMPONENTS OF US REAL GDP



The figure illustrates the HP filter trend components of U.S. real GDP sampled either at the quarterly frequency and using  $\lambda_{quarterly} = 1600$  (the solid line) or at the annual frequency using alternative values for  $\lambda_{annual}$ . For  $\lambda_{annual} = 6.25$ , the trend components are practically identical. To make the figure clearer, we have taken a linear trend out of the HP filter trend components.

rule  $\lambda_s$ , one could, in principle, find  $\lambda_s$  as to minimize some distance metric between  $h(\omega; \lambda_1)$  and  $h(\omega/s; \lambda_s)$ . However, we take a shortcut to this and specify a simple functional rule for this adjustment process: we apply the simple criterion to multiply  $\lambda$  with some power of the frequency adjustment, that is, to choose

$$\lambda_s = s^m \lambda_1. \tag{7}$$

Thus, the problem is to choose  $m$  so as to fit equation (6).

Consider a marginal change in the observation frequency ratio  $s$  around  $s = 1$ , and look at its differential impact on the HP filter. For the correct adjustment, it should be the case that

$$\frac{d}{ds} h(\omega/s; \lambda_s) \approx 0 \tag{8}$$

where  $\frac{d}{ds}$  denotes the total derivative with respect to  $s$ . For each  $\omega$  and  $s$ , this equation can be solved for the parameter  $m = m(s, \omega)$ : one finds that

$$m(s, \omega) = 2 \frac{\omega/s \sin(\omega/s)}{1 - \cos(\omega/s)}. \tag{9}$$

If the power specification is appropriate, then this expression should be approximately constant over the range of “relevant” frequencies,  $\omega$ . Inspection of the transfer function shows that it suffices to restrict attention to values  $0 \leq \omega \leq \pi/5$  (Ravn & Uhlig, 1997). Table 1 lists values of  $m = m(1, \omega) = m(s, \omega s)$  for  $\omega$  in this range. The values in this table suggest that  $m = 4$  (or something close to it) is an excellent choice if one wishes to make the transfer function invariant to the frequency of observation, thereby reconfirming the results of section II for time-aggregated data. The analysis furthermore shows that  $m = 4$  is the exact outcome only at  $\omega = 0$ : otherwise, a slightly lower number between, say,  $m = 3.8$  and  $m = 4$  might be more appropriate.

Thus, for  $\lambda_{quarterly} = 1600$ , this implies that  $\lambda_{annual} = 1600/4^4 = 6.25$  (or 8.25 for  $m = 3.8$ ) and  $\lambda_{monthly} = 1600 \cdot 3^4 = 129600$  (104035 for  $m = 3.8$ ).

Given these results, we now check how well this adjustment rule works in practice. We examine U.S. real GDP from the Bureau of Economic Analysis for the period 1947–2000 sampled at the quarterly and the annual frequency. We compare the trend component of the quarterly data using  $\lambda_{quarterly} = 1600$  with the trend components of the annual data using  $\lambda_{annual} = 400, 100, 25$ , or 6.25. The results are shown in figure 1.<sup>5</sup> This figure clinches our case once more: the trend component of the quarterly data using  $\lambda_{quarterly} = 1600$  and the trend component of the annual data using  $\lambda_{annual} = 6.25$  are practically identical, whereas large differences are visible for  $\lambda_{annual} = 400, 100$ , or 25.

#### IV. Recomputing the Facts

Based on the preceding analysis, it seems natural to ask whether the modification of the rule for adjusting the smoothing parameter matters for reported business cycle “facts.” For an application, we recompute some of the

<sup>5</sup> To make the results visually clearer, we have removed a linear trend from the HP filter trend components.

TABLE 2.—OUTPUT VOLATILITY

	Standard Deviations (%)			$n = 4$		$n = 2^*$	
	I. Prewar	II. Interwar	III. Postwar	I/III	II/III	I/III	II/III
Australia	3.77 (0.37)	2.47 (0.35)	1.40 (0.14)	2.69	1.77	3.3	2.5
Canada	3.13 (0.27)	5.06 (0.77)	1.50 (0.21)	2.09	3.38	2.0	4.4
Denmark	2.20 (0.17)	2.45 (0.37)	1.35 (0.15)	1.63	1.82	1.6	1.8
Germany	2.32 (0.21)	5.26 (0.88)	1.80 (0.24)	1.29	2.92	1.5	4.4
Italy	2.13 (0.20)	2.60 (0.30)	1.51 (0.14)	1.41	1.72	1.2	1.8
Japan	2.10 (0.27)	2.47 (0.38)	1.45 (0.18)	1.45	1.70	0.8	1.0
Norway	1.07 (0.09)	2.89 (0.56)	1.06 (0.12)	1.01	2.72	1.1	2.0
Sweden	1.73 (0.22)	2.41 (0.47)	1.03 (0.09)	1.68	2.34	1.7	2.6
United Kingdom	1.54 (0.16)	2.50 (0.30)	1.27 (0.17)	1.21	1.97	1.3	2.1
United States	3.30 (0.35)	4.91 (0.70)	1.58 (0.17)	2.09	3.11	1.9	4.1

Numbers from Backus and Kehoe (1992). Numbers in parentheses are standard errors computed from GMM estimations of the unconditional 1 moments.

TABLE 3.—THE CORRELATION OF PRICES AND OUTPUT

	$n = 4$			$n = 2^*$		
	I. Prewar	II. Interwar	III. Postwar	I. Prewar	II. Interwar	III. Postwar
Australia	0.29 (0.14)	0.30 (0.18)	-0.26 (0.18)	0.60 (0.10)	0.59 (0.12)	-0.47 (0.11)
Canada	0.11 (0.15)	0.69 (0.12)	-0.01 (0.15)	0.41 (0.13)	0.77 (0.08)	0.12 (0.16)
Denmark	0.18 (0.12)	0.02 (0.26)	-0.60 (0.09)	0.18 (0.12)	-0.26 (0.25)	-0.48 (0.11)
Germany	0.04 (0.13)	0.86 (0.06)	-0.17 (0.14)	-0.01 (0.15)	0.71 (0.09)	0.01 (0.16)
Italy	0.01 (0.10)	0.14 (0.15)	-0.33 (0.14)	-0.02 (0.11)	0.58 (0.09)	-0.24 (0.14)
Japan	-0.49 (0.11)	-0.18 (0.25)	-0.37 (0.18)	-0.45 (0.11)	0.03 (0.22)	-0.60 (0.10)
Norway	0.47 (0.11)	0.16 (0.16)	0.57 (0.10)	0.65 (0.08)	0.16 (0.19)	-0.63 (0.08)
Sweden	-0.08 (0.17)	0.23 (0.09)	-0.38 (0.09)	0.15 (0.13)	0.30 (0.10)	-0.53 (0.07)
U.K.	0.16 (0.14)	0.14 (0.24)	-0.72 (0.08)	0.26 (0.12)	0.20 (0.21)	-0.50 (0.14)
U.S.	0.05 (0.11)	0.75 (0.09)	-0.25 (0.21)	0.22 (0.11)	0.72 (0.13)	-0.30 (0.16)

Numbers taken from Backus and Kehoe (1992). Numbers in parentheses are standard errors.

results reported by Backus and Kehoe (1992) for a cross section of OECD countries using historical annual data. These authors used  $\lambda_{annual} = 100$ , whereas we shall use  $\lambda_{annual} = 6.25$ .

One of Backus and Kehoe's (1992) most interesting findings was that output volatility was higher in the interwar period than during the postwar period, but that there is no general rule as far as a comparison of the postwar period with the prewar (prior to World War I) period is concerned. This result is in contrast to the conventional wisdom of, for example, Burns (1960), Lucas (1977), and Tobin (1980) that output volatility declined after World War II relative to both earlier periods. Another interesting result was that prices changed from generally being procyclical before World War II to being countercyclical thereafter.

Table 2 lists the results for output volatility when using our recommended value for the smoothing parameter. We find that the difference in volatility between the prewar and the postwar period generally narrows and that, for most countries, there has been a decline in volatility in the postwar period relative to either the interwar period or the prewar period.<sup>6</sup> In contrast to Backus and Kehoe (1992), these results are in line with the traditional wisdom previously quoted. This is an important result that Baily (1978) and Tobin (1980) have interpreted in terms of stabilization policy.

Table 3 reports the results for the cyclical behavior of the price level. There, and except for Norway, our results reconfirm the finding of Backus and Kehoe (1992), that prices have become countercyclical in the postwar period and that the interwar period historically was the period in which procyclicality was most pronounced. That is, this result seems to be fairly robust to the choice of the smoothing parameter. These results are also in line with other studies, such as Cooley and Ohanian (1991) and Ravn and Sola (1995).

<sup>6</sup> By this we do not mean to challenge Romer's, 1989 argument that the high prewar volatility is due to measurement error. However, one should notice that, for example, UK data do not suffer from these measurement problems.

## V. Conclusions

This paper provides an analytic investigation into how the smoothing parameter,  $\lambda$ , of the HP filter should be adjusted when changing the frequency of observation. The major conclusion is that the  $\lambda$  parameter should be adjusted according to the fourth power of a change in the frequency of observations. For annual observations, this suggests setting  $\lambda = 6.25$ , which is close to the value found in Baxter and King (1999), but different from the value  $\lambda = 100$  or  $\lambda = 400$  typically found in the literature. Some well-known comparisons of business cycles moments across countries and time periods have been recomputed using the recommended fourth-power adjustment. In particular, we cast doubt on a finding by Backus and Kehoe (1992) and return instead to older conventional wisdom (Baily, 1978; Lucas, 1977; Tobin, 1980): based on the new HP filter adjustment rule, output volatility turns out to be lower in the postwar period compared to the prewar period.

## REFERENCES

- Backus, David K., and Patrick J. Kehoe, "International Evidence on the Historical Properties of Business Cycles," *American Economic Review* 82:4 (1992), 864-888.
- Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, "International Real Business Cycles," *Journal of Political Economy* 100:4 (1992), 745-775.
- Baily, Martin N. "Stabilization Policy and Private Economic Behavior," *Brookings Papers on Economic Activity* 1 (1978), 11-50.
- Baxter, Marianne, "Real Exchange Rates and Real Interest Rate Differentials: Have We Missed the Business Cycle Relationship?" *Journal of Monetary Economics* 33:1 (1994), 5-37.
- Baxter, Marianne, and Robert King, "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *this REVIEW* 81:4 (1999), 573-593.
- Blackburn, Keith, and Morten O. Ravn, "Business Cycles in the U.K.: Facts and Fictions," *Economica* 59 (1992), 383-401.
- Brandner, Peter, and Klaus Neusser, "Business Cycles in Open Economies: Stylized Facts for Austria and Germany," *Weltwirtschaftliches Archiv* 128 (1992), 67-87.
- Burns, Arthur, "Progress Towards Economic Stability," *American Economic Review* 50:1 (1960), 1-19.
- Canova, Fabio, "Detrending and Turning-Points," *European Economic Review* 38:3/4 (1994), 614-623.
- , "Detrending and Business Cycle Facts," *Journal of Monetary Economics* 41:3 (1998), 475-512.

- Christiano, Lawrence J., and Martin S. Eichenbaum, "Temporal Aggregation and Structural Inference in Macroeconomics," NBER working paper no. T0060 (1986).
- Christiano, Lawrence J., and Terry J. Fitzgerald, "The Band Pass Filter," Northwestern University manuscript (1999).
- Cogley, Timothy, and James M. Nason, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control* 19 (1995), 253–278.
- Cooley, Thomas F., and Gary D. Hansen, "The Inflation Tax in a Real Business Cycle Model," *American Economic Review* 79:4 (1989), 733–748.
- Cooley, Thomas J., and Lee E. Ohanian, "The Cyclical Behavior of Prices," *Journal of Monetary Economics* 28:1 (1991), 25–60.
- Correia, Isabel H., Joao L. Neves, and Sergio T. Rebelo, "Business Cycles from 1850–1950: New Facts about Old Data," *European Economic Review* 36:2/3 (1992), 459–467.
- Danthine, Jean-Pierre, and John B. Donaldson, "Methodological and Empirical Issues in Real Business Cycle Theory," *European Economic Review* 37:1 (1993), 1–35.
- Danthine, Jean-Pierre, and Michel Girardin, "Business Cycles in Switzerland: A Comparative Study," *European Economic Review* 33:1 (1989), 31–50.
- Fiorito, Ricardo, and Tryphon Kollintzas, "Stylized Facts of Business Cycles in the G7 from a Real Business Cycles Perspective," *European Economic Review* 38:2 (1994), 235–269.
- Hansen, Gary D., "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics* 16:3 (1985), 309–327.
- Harvey, Andrew C., and A. Jaeger, "Detrending, Stylized Facts and the Business Cycle," *Journal of Applied Econometrics* 8 (1993), 231–247.
- Hassler, John, Petter Lundvik, Torsten Persson, and Paul Söderlind, "The Swedish Business Cycle: Stylized Facts over 130 Years," monograph no. 22, IIES, Stockholm University (1992).
- Hodrick, Robert J., and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," Carnegie Mellon University discussion paper no. 451 (1980).
- , "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking* 29:1 (1997), 1–16.
- King, Robert G., and Sergio T. Rebelo, "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control* 17:1/2 (1993), 207–231.
- Kydland, Finn E., and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica* 50:6 (1982), 1345–1370.
- , "Business Cycles: Real Facts and a Monetary Myth," *Federal Reserve Bank of Minneapolis Quarterly Review* 14 (Spring 1990), 3–18.
- Lucas, Robert E., Jr., "Understanding Business Cycles" (pp. 7–29), in K. Brunner and Allan H. Meltzer (Eds.), *Stabilization of the Domestic and International Economy*, Carnegie-Rochester Conference Series 5 (Amsterdam: North-Holland, 1977).
- Ravn, Morten O., and Martin Sola, "Stylized Facts and Regime Changes: Are Prices Procyclical?" *Journal of Monetary Economics* 36:3 (1995), 497–526.
- Ravn, Morten O., and Harald Uhlig (1997), "On Adjusting the HP-Filter for the Frequency of Observations," CentER discussion paper no. 50.
- , "On Adjusting the HP-Filter for the Frequency of Observations," CEPR discussion paper no. 2858 (2001).
- Romer, Christina, "The Prewar Business Cycle Reconsidered: New Estimates of GNP 1869–1908," *Journal of Political Economy* 97:1 (1989), 1–37.
- Söderlind, Paul, "Cyclical Properties of a Real Business Cycle Model," *Journal of Applied Econometrics* 9 (1994), S113–S122.
- Tobin, James, *Asset Accumulation and Economic Activity* (Chicago: University of Chicago Press, 1980).

## IDIOSYNCRATIC RISK AND VOLATILITY BOUNDS, OR CAN MODELS WITH IDIOSYNCRATIC RISK SOLVE THE EQUITY PREMIUM PUZZLE?

Martin Lettau\*

### I. Introduction

RECENTLY, there has been a lot of interest in computing asset prices in incomplete market models; see, for example, Constantinides and Duffie (1996), Heaton and Lucas (1996), den Haan (1996), Krusell and Smith (1997) and Storesletten, Telmer, and Yaron (1997). These papers have shown that market incompleteness can affect prices of financial assets qualitatively. In this paper, I propose a simple method to check whether these effects are quantitatively important enough to solve the equity premium puzzle.

Received for publication April 20, 1999. Revision accepted for publication May 10, 2001.

\* Federal Reserve Bank of New York and Centre for Economic Policy Research.

This paper was written during a visit at the Department of Economics at New York University; I am grateful for its hospitality. John Campbell, Mark Gertler, Blake LeBaron, Sydney Ludvigson, Anthony Lynch, Harald Uhlig, two anonymous referees, and seminar participants at Humboldt University, New York University, and the Federal Reserve Bank of New York provided helpful comments. The views are those of the author and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

The main argument is as follows. Most incomplete market models specify endogenous endowment (labor income) shocks that are not fully insurable. Agents are allowed to trade in a small number of securities and solve for their optimal portfolio and consumption policies. It is difficult to test these types of models directly because the quality of household-level consumption data is very poor.<sup>1</sup> Instead of this direct approach using consumption data, I use data on individual income, which is measured more precisely than is individual consumption. In other words, I assume that agents cannot smooth idiosyncratic income shocks at all and are forced to consume their endowment. If agents were allowed to trade using some restricted set of securities, they would be able to smooth, at least partially, their individual shocks. Hence, the income process provides an upper bound on the volatility of individual consumption. If models with idiosyncratic risk are not able to generate large risk premia, they will most likely not be able to perform better with consumption data. I find even very volatile income shocks

<sup>1</sup> One exception is Cogley (1998).

are not able to quantitatively to generate high-enough risk premia.

The paper is organized as follows. First, I briefly demonstrate how idiosyncratic shocks can affect the Sharpe ratio.<sup>2</sup> Intuitively, idiosyncratic shocks can affect asset prices if their distribution depends on aggregate state variables (Mankiw, 1986; Constantinides & Duffie, 1996). Second, I use the estimated processes for idiosyncratic income estimated by Heaton and Lucas (1996), Krusell and Smith (1997), and Storesletten et al. (1997) and compute the Sharpe ratios as if agents had to consume their endowment. I find that none of the idiosyncratic income processes is able to generate a Sharpe ratio that is found in the data.<sup>3</sup>

## II. The Sharpe and Idiosyncratic Shocks

Let  $C_t^j$  denote consumption of consumer  $j$  in period  $t$ . Lowercase letters denote log variables, so  $c_t^j$  is log consumption of agent  $j$ .  $M_{t+1}^j$  is the stochastic discount factor (SDF) of agent  $j$ ,  $R_{t+1}$  is the return of an asset, and  $RP_{t+1}$  is the risk premium. The first-order condition is

$$E_t[M_{t+1}^j R_{t+1}] = 1. \quad (1)$$

The Sharpe ratio is defined as the price of risk of the market portfolio and depends on the volatility of the SDF and its correlation with the market portfolio:

$$\begin{aligned} SR(M_{t+1}^j) &= \max_{\text{all assets}} \frac{E_t[RP_{t+1}]}{\sigma_t[RP_{t+1}]} \\ &= -\rho_t(M_{t+1}^j, RP_{t+1}^m) \frac{\sigma_t[M_{t+1}^j]}{E_t[M_{t+1}^j]} \leq \frac{\sigma_t[M_{t+1}^j]}{E_t[M_{t+1}^j]}. \end{aligned} \quad (2)$$

Note that, in an economy with complete markets, the SDF of the representative agent is perfectly negatively correlated with the market portfolio, so that the Sharpe ratio is determined only by the volatility of the SDF. This is, however, no longer true in a model with idiosyncratic shocks because returns on aggregate assets cannot be correlated with idiosyncratic shocks. Hence, despite the fact that the SDF of an individual might be very volatile, the net effect on the Sharpe ratio is ambiguous.

I assume that consumption growth is of the form

$$\Delta c_{t+1}^j = E_t \Delta c_{t+1}^j + \epsilon_{t+1} + \eta_{t+1}^j,$$

with  $\epsilon_{t+1} \sim N(-\sigma_\epsilon^2/2, \sigma_\epsilon^2)$  and  $\eta_{t+1}^j \sim N(\sigma_{\eta,t+1}^2/2, \sigma_{\eta,t+1}^2)$ . To illustrate that it is crucial that the variance of idiosyncratic shocks depends on the aggregate state, I consider the simplest example:

<sup>2</sup> I focus on the Sharpe ratio instead of the equity premium because it is a measure of the risk-return tradeoff that is independent of any specific asset. Computing prices of individual asset would require the specification of a dividend process.

<sup>3</sup> The working-paper version of this paper, Lettau (2001), contains a more detailed analysis.

$$\sigma_{\eta,t+1}^2 = a_0 + a_1 \epsilon_{t+1}. \quad (3)$$

Lettau (2001) shows that

$$SR(M_{t+1}^j) \approx \gamma \sigma_\epsilon \left( 1 - a_1 \frac{1 + \gamma}{2} \right).$$

The Sharpe ratio increases if the variance of idiosyncratic shocks depends negatively on the aggregate state. This is the same mechanism highlighted by Mankiw (1986) and Constantinides and Duffie (1996).

## III. Volatility Bounds with Income Data

Next, I evaluate the empirical importance of the mechanism presented in the previous section. The problem with a direct empirical analysis of this channel is that individual consumption data (for example, in the PSID) is known to be of poor quality. In addition, consumption data in the PSID span only a short time series. Hence, estimating equations like equation (3) directly from consumption data is flawed with measurement error. However, individual income data in the PSID is of much higher quality, and there is a large empirical literature that estimates income processes for individual households (Abowd & Card, 1989; Carroll, 1992; MaCurdy, 1982). As an alternative to a direct test using consumption data, I use these income processes instead of evaluating the volatility bounds on an individual level. This strategy is appropriate if individual consumption is smoother than individual income. Of course, because there is no data set with a long time series of high-quality household consumption data, it is impossible to check directly whether individual consumption is indeed smoother than income. However, we can find theoretical conditions under which we expect this to be the case. I will argue next that these conditions are likely to be fulfilled in the data.<sup>4</sup>

Start with an individual who consumes her permanent income. This is a useful point of departure because much of the consumption literature is cast in this framework. (See Deaton (1992) for a survey.) Intuitively, if income growth is negatively autocorrelated, consumption is smoother than income because a shock to income has less than a one-to-one effect on permanent income. On the other hand, if income growth is positively autocorrelated, then permanent income is more volatile than current income. Because, theoretically, permanent income can be more or less volatile than income, it is an empirical question which case is more realistic. The evidence in the empirical literature points overwhelmingly in the direction of negatively autocorrelated income growth on the household level (Abowd & Card, 1989; Carroll, 1992; MaCurdy, 1982; Pischke, 1995). In this case, permanent income is smoother than current

<sup>4</sup> Lettau (2001) contains a more detailed derivation of the following statements.

income, and therefore consumption of a permanent income agent will be smoother than her income.<sup>5</sup>

If consumption is indeed smoother than income, then risk premia and volatility bounds computed using income data provide a useful lower bound. If risk premia are low for income data, then there is surely little hope in generating higher-risk premia using smoother consumption data. Hence, this approach provides a straightforward benchmark to evaluate the potential of models with idiosyncratic risk. The punchline is that risk premia are too small even if individual income data are used to compute them. The reason is that the covariance of the volatility of idiosyncratic income risk and the aggregate state is too small to increase risk premia, even when income is used instead of consumption.

Because agents receive different income in each period and they cannot trade with each other, they will in general disagree on asset prices. One agent might require an expected return of 5% for some asset given her particular income process, and another agent might require only a 3% expected return for the same asset. If there were open asset markets, they would trade the asset and smooth their consumption paths so that in equilibrium they agree on expected returns of all assets. Because I abstract from any asset trading in this paper, I compute the expected returns implied by the consumption process (which is equal to the income process by assumption) of each individual agent. To obtain an upper bound for the ability of idiosyncratic risk to increase risk premia, I consider the highest expected return of a given asset across all agents. In terms of the Sharpe ratio, using the individual income processes as consumption will result in different Sharpe ratios for different agents. An upper bound for the Sharpe ratio that would result from trading is given by the highest Sharpe ratio computed from the individual income processes. To give the models the best possible shot, I report this maximal Sharpe ratio across agents. The Sharpe ratio, estimated from postwar data using S&P 500 excess returns, is 0.27 per quarter or 0.54 per annum.

#### A. Income Process from Heaton and Lucas

Heaton and Lucas (1996) (hereafter, HL) present a model with idiosyncratic labor income shocks, transaction costs, borrowing, and short-sale constraints with eight discrete states. Table 1 summarizes the results for idiosyncratic income shocks estimated by HL. They estimate  $\sigma_{\eta,t+1} = \tilde{\alpha}_0 + \tilde{\alpha}_1 \log(\gamma_{t+1}^a)$ , where  $\gamma^a$  is the growth rate of aggregate

<sup>5</sup> One possible objection to these arguments is that consumption behavior in models with incomplete markets might deviate from permanent income behavior. But, even in more-complex models, it is unlikely that consumption is more volatile than current income if income growth is negatively autocorrelated.

TABLE 1.—THE SHARPE RATIO WITH HEATON-LUCAS INCOME SHOCKS

Income Process	Heaton and Lucas	
	With Dividends	No Dividends
	$\gamma = 1$	
HL1: $\tilde{\alpha}_1 = 0$	0.0110	0.0111
HL2: $\tilde{\alpha}_1 = -1.064$	0.0299	0.0301
HL3: $\tilde{\alpha}_1 = -4.450$	0.0767	0.0769
	$\gamma = 5$	
HL1: $\tilde{\alpha}_1 = 0$	0.0552	0.0553
HL2: $\tilde{\alpha}_1 = -1.064$	0.1813	0.1820
HL3: $\tilde{\alpha}_1 = -4.450$	0.5773	0.5785

Table reports the average Sharpe ratio across the eight states of the economy weighted with the probabilities of the stationary distribution. The maximal Sharpe ratio across agents is used in each state. The model is calibrated for annual data.

income.<sup>6</sup> HL consider three sets of estimates:  $\tilde{\alpha}_0 = 0.251$ ,  $\tilde{\alpha}_1 = -0$ ;  $\tilde{\alpha}_0 = 0.360$ ,  $\tilde{\alpha}_1 = -1.064$ ; and  $\tilde{\alpha}_0 = 0.290$ ,  $\tilde{\alpha}_1 = -4.450$ . The second set of parameters corresponds to the PSID regressions results. The aggregate shocks are chosen to match the moments of aggregate consumption. In each of the eight states, I compute the Sharpe ratio for both types of agents. Then I compute the average Sharpe ratio (weighted with the probabilities of the stationary distribution) across states using the higher of the two Sharpe ratios. The transitions probabilities are taken from HL, who consider two cases: one with aggregate dividends and one without.

For low risk aversion of unity, the picture looks very bleak. None of the three cases reach a Sharpe ratio of 0.1. Note that the effect of dividend uncertainty is very small. It is clear that higher risk aversion is needed to match the Sharpe ratio in the data. The bottom panel uses  $\gamma = 5$ . Here, Sharpe ratios increase substantially when  $\tilde{\alpha}_1$  is negative. For the case corresponding to the PSID data, the Sharpe ratio is approximately 0.18. This value is still only about one-third of the Sharpe ratio in the data. Only if  $\tilde{\alpha}_1 = -4.45$  (that is, more than four times the PSID value) is the Sharpe ratio high enough to be consistent with the data.

Is this a success? As mentioned before, this analysis computes only a very loose upper limit for the Sharpe ratio: consumers cannot trade against their idiosyncratic labor shocks at all; the Sharpe ratio of the poor agents is used despite the fact that those agents are likely to be borrowing constraint and hence do not determine asset prices; and, lastly, the required value for  $\tilde{\alpha}_1$  is approximately four times the value found in the data. Only in this case does the model generate a reasonable Sharpe ratio. Thus, it is not surprising that HL found low Sharpe ratios and risk premia once they allow agents to trade. The model is not able to produce higher risk premia even if agents are forced to consume their labor income.

<sup>6</sup> HL specify the standard deviation of the idiosyncratic shock as a linear function of the aggregate growth rate. This implies  $\sigma_{\eta,t+1}^2 = \tilde{\alpha}_0^2 + 2\tilde{\alpha}_0\tilde{\alpha}_1 \log(\gamma_{t+1}^a) + \tilde{\alpha}_1^2(\log(\gamma_{t+1}^a))^2$ .



TABLE 2.—THE SHARPE-RATIO WITH KRUSELL-SMITH INCOME SHOCKS

Agent	Aggregate State	
	Bad	Good
Employed	0.0641	0.1088
Unemployed	0.1759	0.2840
No employment shocks	0.0037	0.0037

Table reports the Sharpe ratio with KS unemployment shocks. The model is calibrated for quarterly data.

### B. Income Process from Krusell and Smith

Krusell and Smith (1997) (hereafter, KS) propose a model in which agents have a higher probability to become unemployed when aggregate times are bad. This yields again a negative covariance of idiosyncratic uncertainty and state aggregate that has the potential to increase risk premia. Let aggregate output per capita be given by  $Y^a$ . If an agent becomes unemployed, she receives unemployment insurance of  $gY^a$ . Hence, an employed agent's income is  $(1 - g)Y^a$ . Aggregate income can take on two values, so that each agent has four possible states.

KS calibrate the model as follows. The unemployment rate in the good aggregate state is 4% and 10% in the low state. Unemployed agents receive  $g = 9\%$  unemployment insurance. The Markov transition probabilities are chosen so that unemployment shocks are fairly persistent: that is, an unemployed agent today has a smaller probability to find a job tomorrow than an agent who is employed today. Again, I refer to KS for the exact transition probabilities. All other parameters are taken from KS, who calibrate the model for quarterly data.

Table 2 reports the Sharpe ratios in this economy. (Recall that the Sharpe ratio in quarterly data is 0.27.) Unemployed agents have a much higher Sharpe ratio than do employed agents because their current and expected future income is much lower. Note that the Sharpe ratio of both types is higher in good aggregate times; that is, it is procyclical. This comes from the specific choice of the transition probabilities. However, the lowest Sharpe ratio (of the unemployed agents in bad times) just matches the Sharpe ratio in the data.

### C. Income Process from Storesletten, Telmer, and Yaron

Storesletten, Telmer, and Yaron (1997) (hereafter, STY) study an OLG model with idiosyncratic income shocks and borrowing constraints. They calibrate various Markov processes for idiosyncratic shocks to match the PSID data. Labor supply of agent  $j$ ,  $N_t^j$ , is determined by (potentially multiple) idiosyncratic shocks:  $N_t^j = \exp(\eta_t^j)$ . The labor input is supplied inelastically and combined with a aggregate Cobb-Douglas production function to produce the consumption good.

STY propose different processes for  $\eta_t^j$ . Their first three processes differ mainly in the persistence of the idiosyncratic shocks, which turns out to be important for the risk

sharing in their OLG model. In their fourth version, they allow the aggregate state to affect the variance of the idiosyncratic shock. They estimate equation (3) using PSID data. Using two different estimation techniques, they estimate the coefficients as  $\hat{a}_0 = 0.054$ ,  $\hat{a}_1 = -0.011$ ; and  $\hat{a}_0 = 0.054$ ,  $\hat{a}_1 = -0.046$ . Both slope coefficients are not negative enough to increase the Sharpe ratio. Finally, STY propose a different technique to estimate cross-sectional dispersion: they estimate a model with a high and low idiosyncratic variance regime depending on whether the aggregate economy is above or below trend. STY find that the variance of idiosyncratic shocks is 0.032 when the aggregate economy is above trend and 0.184 when it is below trend. Table 3 reports the Sharpe ratio for all five cases considered by STY. Again, all other parameters are taken from STY.

The case without any idiosyncratic shocks generates a Sharpe ratio of 0.0594. The following three cases produce the same Sharpe ratio as without idiosyncratic risk. Here, the distribution of the idiosyncratic shocks are assumed to be independent of the aggregate state. The next two versions incorporate the dependence of the variance of the idiosyncratic shock on the aggregate state as estimated from the PSID data set. Because the estimated  $a_1$ 's are negative, the Sharpe ratio increases, but the increase is only minute because the parameters are not negative enough. The Sharpe ratio in the case with the changing variance regime is approximately 0.1, which is still only one-fifth of the required value. None of the income processes considered by STY can produce high Sharpe ratios even if agents cannot smooth them.

## IV. Conclusion

To evaluate the empirical relevance of idiosyncratic uncertainty, I compute Sharpe ratios using individual income processes from the PSID data set. Using income data provides an upper bound for individual consumption data in the absence of reliable consumption data. I take the estimated income processes from recent studies of Heaton and Lucas (1996), Krusell and Smith (1997), and Storesletten, Telmer,

TABLE 3.—THE SHARPE-RATIO WITH STORESLETTEN, TELMER, AND YARAN INCOME SHOCKS

Income Process	Sharpe Ratio
No idiosyncratic shocks	0.0594
Unit root	0.0594
High persistence	0.0594
Moderate persistence	0.0594
Low persistence	0.0594
PSID $a_1 = -0.011$	0.0604
PSID $a_1 = -0.046$	0.0635
High/low variance regimes	0.1013

Table reports the Sharpe ratio for income processes considered by Storesletten, Telmer, and Yaron (1997). The parameters for the different AR1 income shocks are as follows. STY approximate the processes with discrete Markov chains. Let  $\rho$  be the AR1 coefficient and  $\sigma_\eta$  the innovation standard deviation. Unit root:  $\rho = 1$ ,  $\sigma_\eta = 0.201$ ; high persistence:  $\rho = 0.929$ ,  $\sigma_\eta = 0.230$ ; moderate persistence:  $\rho = 0.529$ ,  $\sigma_\eta = 0.251$ . The two PSID cases use the unit root parameters and the estimated coefficients in equation (3). The last case assumes a variance of 0.032 when the economy is in the high-aggregate state and 0.184 when the aggregate state is low. The model is calibrated for annual data.

and Yaron (1997) and compute Sharpe ratios as if agents had to consume their endowment. Because there is no market that equilibrates the SDF of different agents, they will disagree on asset prices. To obtain an upper bound on Sharpe ratios, I compare the Sharpe ratio in the data to that of the maximal Sharpe ratio across agents. I find that Sharpe ratios are generally smaller than in the data; only some extreme cases just reach the data equivalent. Hence, it seems unlikely that idiosyncratic consumption risk can generate an SDF that pass the Hansen-Jagannathan bounds test.

## REFERENCES

- Abowd, John, and David Card, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica* 57:2 (1989), 411–445.
- Aiyagari, Rao, and Mark Gertler, "Asset Returns and Uninsured Individual Risk," *Journal of Monetary Economics* 27:3 (1991), 311–331.
- Carroll, Christopher D., "The Buffer Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers of Economic Activity* 2 (1992), 61–135.
- Cogley, Timothy, "Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey," Federal Reserve Bank of San Francisco mimeograph (1998).
- Constantinides, George M., and Darrel Duffie, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy* 104 (1996), 219–240.
- Deaton, Angus, *Understanding Consumption* (Oxford: Clarendon Press, 1992).
- Den Haan, Wouter, "Understanding Equilibrium Models with a Small and Large Number of Agents," UCSD working paper (1996).
- Heaton, John, and Deborah Lucas, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy* 104 (1996), 443–487.
- Krusell, Per, and Anthony A. Smith, "Income and Wealth Heterogeneity, Portfolio Selection, and Equilibrium Asset Returns," *Macroeconomic Dynamics* 1 (1997), 387–422.
- Lettau, Martin, "Idiosyncratic Risk and Volatility Bounds, or Can Models with Idiosyncratic Risk Solve the Equity Premium Puzzle?" Federal Reserve Bank of New York working paper (2001).
- MaCurdy, Thomas E., "The Use of Time-Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis," *Journal of Econometrics* 18 (1982), 83–114.
- Mankiw, N. Gregory, "The Equity Premium and the Concentration of Aggregate Shocks," *Journal of Financial Economics* 17 (1986), 211–219.
- Pischke, Jörn-Steffen, "Individual Income, Incomplete Information and Aggregate Consumption," *Econometrica* 63:4 (1995), 805–840.
- Storesletten, Kjetil, Chris Telmer, and Amir Yaron, "Persistent Idiosyncratic Shocks and Incomplete Markets," Carnegie Mellon University mimeograph (1997).