

# **Errors in the Measurement of the Output Gap and the Design of Monetary Policy\***

**Athanasios Orphanides, Richard D. Porter, David Reifschneider,  
Robert Tetlow, and Frederico Finan**

Board of Governors of the Federal Reserve System  
Washington, D.C. 20551

August 1999

## **Abstract**

We exploit data on historical revisions to real-time estimates of the output gap to examine the implications of measurement error for the design of monetary policy, using the Federal Reserve's model of the U.S. economy, FRB/US. Measurement error brings about a substantial deterioration in economic performance, although the problem can be mitigated somewhat by reducing the coefficient on the output gap in policy rules. We also show that it is usually optimal to place some weight on the level of the output gap in the conduct of policy, but under extreme conditions it may be preferable to focus on output growth.

**Keywords:** interest rate rules, policy evaluation, output gap measurement

---

\*The authors would like to thank, without implication, Lewis Alexander, Dale Henderson, Donald Kohn, David Lindsey, Michael Prell, Brian Sack, David Small, Joel Wesley, John Williams, and an anonymous referee for comments and suggestions and to Ken Kuttner for generously providing us data on his estimates of trend output. The opinions expressed are those of the authors and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System.

## **I. Introduction**

Decision making by the Federal Reserve takes place in the context of considerable uncertainty about the current as well as prospective state of the economy. Data on economic activity and prices are incomplete and subject to substantial revision; moreover, the correct interpretation of these data is often unclear. Key elements in this interpretation are current readings on aggregate resource utilization: To the extent that they are accurate, readings combining actual data and inferences about the economy's productive potential are useful indicators of future inflationary tendencies. Unfortunately, at times their reliability is questionable. This would seem to have been the case during the late 1990s when inflation was coming in persistently lower than would have been expected given most estimates of resource utilization.

This paper examines the role of the output gap and other measures of aggregate resource utilization in the conduct of monetary policy, given that readings on the state of excess demand are inherently imprecise. In particular, we examine how the usefulness of the estimated output gap declines when its measurement error increases in magnitude, and how macroeconomic performance can suffer as a result, even if policymakers respond to increased uncertainty by reducing the weight placed on resource utilization in the formulation of policy.

The problem we investigate is by no means new; in discussions of monetary policy and macroeconomic stabilization, researchers have recognized the difficulties associated with the measurement of potential output and resource utilization for years, beginning at least with Friedman (1947, 1953). Many authors have highlighted the issue since then. For example, Kuttner (1994) and McCallum (1998) emphasize that monetary policy is complicated by mismeasurement of the output gap. Estrella and Mishkin (1999) and Staiger, Stock and Watson (1997) draw attention to the corresponding problems associated with errors in estimates of the unemployment gap. However, quantitative assessments of the policy implications of mismeasurement have appeared only recently. Smets (1998) integrates estimation of the output trend in a policy evaluation model, and illustrates the effect of errors in this estimation on the choice of efficient simple policy rules. Using estimates of measurement error derived from real-time estimates of the output gap, Orphanides (1998) shows that such errors lead to a significant deterioration of feasible policy outcomes and cause efficient policies to be less activist. Finally,

Rudebusch (1999) shows that these considerations are essential for reconciling estimated policy reaction functions and optimal policy.

The contribution of this paper is twofold. First, in order to assess the measurement error problem, we follow Orphanides (1998) in taking advantage of the real-time estimates of the output gap that were historically available to the Federal Reserve. Unlike preceding papers, however, we use Board of Governors' staff econometric model, FRB/US to analyze the implications of output gap mismeasurement in the policy process. The size and scope of the model ensure that a large range of disturbances and propagation mechanisms in the economy are incorporated into the analysis. Second, we specify that agent expectations are model-consistent. This assumption assures that private agents take into account the response of the Federal Open Market Committee (FOMC) to uncertainty and that any results we obtain are not a manifestation of arbitrary or unstated assumptions of policymakers' ability to systematically mislead private agents. This formulation makes our analysis less vulnerable to the Lucas critique.

In this analysis, we use simple interest-rate reaction functions to characterize the setting of the federal funds rate. In particular, our focus is on rules that respond to the measured *level* of resource utilization (among other factors), although we also consider alternative policies that respond to *growth* of actual and potential output. Such simple policy rules, the Taylor rule for example, have been shown to perform well at stabilizing the economy in experiments with a variety of economic models, at least if the output gap is measured well.

Based on the experience of the last 30 years, real-time assessments of the output gap have often fallen wide of the mark, as judged in hindsight with the benefit of more complete information and analysis. By contrast, revisions to measures of inflation, the other input to these rules, are much smaller. Hence, we assume that the monetary authority has only an approximate knowledge of the true current level of resource utilization in the economy, but retains a correct understanding of the true inflation outcomes. The degree to which it errs in judging the true state of excess demand is calibrated to match the historical record. Because there is some question about the actual degree of measurement uncertainty the FOMC is likely to face prospectively, we consider alternative magnitudes of measurement error, again based on historical experience.

To presage our results, we find that measurement error has important effects for the appropriate conduct of the monetary authority as well as for policy performance. In most circumstances, an authority that takes explicit account of the uncertainty in the environment in which it operates is more successful than an authority that turns a blind eye to the issue. Under some circumstances it may also be optimal for policymakers to ignore almost completely measures of the level of resource utilization in the conduct of policy.

The rest of this paper is organized as follows: Section II summarizes the historical record regarding revisions to the output gap. Section III discusses several features of the FRB/US model that are relevant to our analysis, and details how model simulations are used to evaluate the effect of measurement error on the macroeconomic performance yielded by alternative policy rules. Section IV presents our findings. We close with a summary of our main conclusions and some directions for future research.

## **II. Mismeasurement of the Output Gap and Capacity Utilization**

Preliminary measurements of economic conditions often differ from later ones as government statistical agencies incorporate more complete source information into the published data on GDP, inflation, and other macro series. From these evolving data sources economists attempt to extract estimates of underlying concepts such as the NAIRU (the level of unemployment consistent with stable inflation), trend factor productivity, and potential real GDP - key elements of the supply side of the economy. These latter estimates undergo their own set of revisions in reaction to updates to previously published data, as well as to the receipt of new readings and the use of new estimation techniques, which attempt to provide better indications of the long-run trends in the economy. From these various sets of data, a measure of the output gap — the percentage difference between the levels of actual and potential real GDP — can be defined. Like its underlying source material, it is always under revision, sometimes by a very large amount. Primarily because of difficulties in measuring potential output, historical revisions to the output gap have dwarfed those recorded for estimates of inflation and GDP.<sup>1</sup>

---

1. In part, these difficulties spring from a certain fundamental ambiguity in the concept of potential

In theory, the larger the output gap — that is, the farther actual output lies above potential output — the more resources are stretched, and the stronger are inflation pressures. Accurate and timely measures of the gap would therefore appear to be an important ingredient for effective policy. Unfortunately, the historical accuracy of contemporaneous measures of the output gap that were available to the Federal Reserve System is not encouraging. Orphanides (1999) compares current estimates of the gap for the period from the mid-1960s to the mid-1990s and finds that the output gap as measured today lies almost uniformly above the contemporaneous estimates: The real-time estimates of potential output over this period were systematically overly optimistic. The late 1960s furnish a particularly striking example. At that time, the data show that the output gap was believed to be at about zero. With the benefit of hindsight, almost any estimate now would place the gap then at nearly 5 percentage points. Taken at face value, this divergence would imply that policymakers did not recognize the considerable upward inflationary pressure that the economy was subject to at that time.<sup>2</sup> Moreover, the divergence between contemporaneous and after-the-fact views of resource utilization widened during the following decade; in 1975, for example, the real-time estimate of the output gap stood at -15 percent, but today most conventional methods for measuring potential would estimate the depth of the 1975 recession to be appreciably smaller.

Errors of this magnitude may seem unduly large, but they arose from assumptions regarding the NAIRU and trend growth that were regarded as mainstream at the time. Coming off of the experience of the 1950s and 1960s, most economists would have regarded the NAIRU

---

aggregate output. Although an individual manufacturing plant may have a well-defined normal operating rate from an engineer's standpoint, the same is not true for the economy as whole. Rather, the economy's productive potential is typically defined as the trend component of actual output, with the trend estimated in various ways. Alternatively, potential output is inferred from the behavior of inflation.

2. In fact, the output gap measures used in the 1960s were not constructed by the Federal Reserve itself but rather came from the Council of Economic Advisors. The staff at the Federal Reserve Board began to compute their own estimates beginning in the early 1980s. A discussion of the nature and causes of historical mismeasurement of the output gap is beyond the scope of this paper. For more on the subject, see Orphanides (1999).

to be in the neighborhood of 4 percent and growth in potential GDP to be over 4 percent.<sup>3</sup> Thus, large errors in estimates of the output gap in the 1970s reflected changes in these two key supply-side parameters that were not detected immediately, but instead were only gradually distinguished in incoming data. Gradually the revision errors became smaller as these more fundamental misconceptions about potential GDP were discarded.<sup>4</sup>

The upper portion of Table 1 presents some summary statistics for the difference between the current revised estimate and the initial estimate of the output gap for two periods – from the 1980s through the early 1990s, and over the full period beginning in 1966 and ending in 1994. The full-period results, which are importantly influenced by large measurement errors in the 1970s, exhibit higher mean errors and root mean square errors; in fact, the variation in the full period is over twice that of the more recent period. In addition, the errors decayed more slowly in the earlier period, as evidenced by the higher sample autocorrelations.<sup>5</sup>

Of course the output gap is not the only measure of resource utilization that could be used to guide policy. The most accepted substitute for the output gap would be the unemployment gap, the difference between the unemployment rate and the estimated NAIRU. Unfortunately,

---

3. For example, on page 16 the 1973 *Economic Report of the President* stated that: “Both the ‘full employment budget’ and the ‘GNP potential’ were [during most of the 1960s], and continue to be, defined in terms of a 4 percent unemployment rate.” This estimate is substantially below current estimates of the NAIRU for that period; for example, the estimate embedded in the FRB/US model would suggest a figure closer to 6 percent. As regards estimates of trend growth, according to Perry (1971), “... potential [output] grows at a rate of between 4.2 and 4.4 percent each year of the 1970s.” Taking actual output at the end of the 1970s as approximately equal to potential, Perry’s forecast error would have been in the neighborhood of 20 percent. Errors of this magnitude were by no means unusual during this period; more remarkable, perhaps, is the seemingly high precision of the estimates suggested by the narrow range of growth rates given by Perry.

4. One might argue that monetary policymakers did not pay much attention to the output gap at this time, and that if the gap had played a more central role in decision making, a more serious effort would have been made to measure it, resulting in smaller historical revisions. However, large historical measurement errors may simply reflect the statistical difficulty of extracting trend estimates of potential GDP in real time. For example, Orphanides and van Norden (1999) show that *ex post* revisions of the output gap are the same order of magnitude as the output gap itself for a wide variety of univariate detrending methods.

5. In part, these differences may reflect the source of the estimates in the two periods, as noted above in footnote 3.

**Table 1**  
**Statistics on Historical Errors**  
**in the Measurement of the Output Gap and Capacity Utilization**

Resource Utilization Measure	Mean Error	Root Mean Squared Error	Autocorrelation
Output gap <sup>1</sup>			
Full period (1966:Q1 – 1994:Q4)	3.2	4.2	0.919
Late period (1980:Q1 – 1994:Q4)	1.1	1.9	0.785
Capacity utilization (rescaled) <sup>2</sup>			
Full period (1968:Q1 – 1994:Q4)	0.9	1.9	0.921
Late period (1980:Q1 – 1994:Q4)	-.1	0.8	0.796

1. Output gap errors are defined as the difference between current Federal Reserve Board staff estimates of the historical output gap, and the real-time estimates made by the Council of Economic Advisors prior to the early 1980s, and the Federal Reserve Board staff thereafter, as reported by Orphanides (1999).
2. Capacity utilization errors are defined as the difference between current estimates of the historical capacity utilization rate in the manufacturing sector, and the first quarterly estimate published in the *Federal Reserve Bulletin*. Because industrial production and capacity utilization are more cyclically variable than GDP and the output gap, the errors are rescaled to make them comparable to output gap errors. The multiplicative rescaling factor equals 0.7, the ratio of the standard deviation of the output gap to the standard deviation of the capacity utilization rate over the period 1968 to 1994.

we lack the complete set of real-time data needed to evaluate the precision of this measure, as concurrent estimates of the NAIRU were not provided regularly by either the Council of Economic Advisors, the Congressional Budget Office, or the Federal Reserve.

However, if we confine our attention to the manufacturing sector, we can use the capacity utilization index as an alternative measure of aggregate resource utilization. Koenig (1996) discusses the usefulness of capacity utilization in this context. To a large degree, current and real-time estimates of capacity utilization mirror corresponding estimates of the output gap. Like the gap, in the latter part of the 1960s concurrent estimates of capacity utilization were lower and

closer to the NAICU than they are now thought to be.<sup>6</sup> More generally, when measured on both a real-time and a current basis over the 1968 to 1994 period, the correlation between capacity utilization and the output gap is 0.80 and 0.89, respectively; revisions to the two series are also significantly correlated (0.67).<sup>7</sup> However, as indicated by the mean errors reported in the bottom portion of Table 1, errors in the initial estimates of capacity utilization have not been nearly as one sided as those recorded for the output gap. Furthermore, the root mean squared errors for capacity utilization are relatively low, suggesting that inherent measurement error in this measure of aggregate resource utilization might be less of a problem.<sup>8</sup>

The appropriate lesson to draw from this varied experience is not clear. Should we assume that revisions to the output gap will be on the order of that experienced during the 1970s, or take heart from the more benign experience encountered subsequently? To the extent that the profession has learned a valuable lesson from the experience of the 1970s, one might expect that a more appropriate benchmark would be the smaller-sized revisions that are more in line with the post 1970s experience. However, the experience of the earlier period suggests that it would be foolish to overlook the possibility that gap estimates could go very much awry once again — particularly if the economy were to experience disruptive shifts in technology and labor markets that required a considerable elapse of time before their implications for the NAIRU, trend factor

---

6. The NAICU is the acronym for the index of capacity utilization that is comparable to the NAIRU.

7. It might be argued that the general similarity between the two sets of revisions reflects a common problem of circularity. Specifically, it is possible that estimates of resource utilization are revised over time to make them more consistent with subsequent inflation outcomes, when in fact apparent inconsistencies between initial utilization estimates and later inflation developments actually reflect other causal factors that are not well understood. This argument may have some validity in the case of the output gap, where some techniques for estimating potential output exploit information on inflation. However, the argument is not relevant for capacity utilization, because the procedures used to estimate industrial capacity are independent of inflation developments, as they depend only on plant-level estimates of sustained practical capacity, Corrado and Matthey (1997).

8. In computing the mean errors and root mean squared errors for capacity utilization, the historical revisions have been re-scaled to make them comparable to the output gap revisions. See the notes to Table 1 for further details.



productivity, and the output gap could be ascertained. In any event, the results using the industrial capacity index are more encouraging than those for the output gap.

In light of these divergent results, in the quantitative analysis of the following sections we use several different estimates of the degree of measurement error faced by policymakers. Two of the estimates are based on the historical revisions to the output gap, as recorded for the 1980-1994 period and the full sample. The third estimate of measurement error draws upon historical revisions to capacity utilization. Taken together, these estimates give us a broad range of benchmarks on the magnitude of uncertainty regarding the output gap that might be encountered going forward.

### **III. Evaluating the Effect of Measurement Error: Methodology**

To assess the implications of measurement error for monetary policy, we compare the simulated performance of the economy under alternative policy strategies as we vary the degree of measurement error. In this section we discuss the specifics underlying our approach. We begin with a brief overview of the macro model used in the analysis, FRB/US, and highlight several of its important features. Next, we discuss the type of policy rules that we consider in our analysis, and the nature and degree of mismeasurement to which they are susceptible. Finally, we describe our use stochastic simulations in evaluating the macroeconomic performance produced by alternative policy strategies, and we discuss the concept of the *policy frontier*: the macroeconomic performance delivered by the set of rules giving the best possible combinations of output and inflation variability.

#### **The FRB/US Model**

FRB/US is a large, open-economy model of the United States used by Board staff as a tool for forecasting and policy analysis. The model's main equations are derived from standard economic theory, and are estimated using data from the 1965-1997 period. These equations provide a detailed empirical accounting of the monetary policy transmission mechanism. Because documentation for the model's design and basic properties can be found in two recent *Federal Reserve Bulletin* articles, as well as in a number of recent studies, we restrict ourselves to a brief

discussion of two features that are important for our analysis, inflation dynamics and expectations.<sup>9</sup>

*Resource utilization and the inflation process.* In the model, policymakers are able to influence resource utilization through open-market operations that indirectly affect bond rates and other asset prices, which, in turn, influence aggregate demand and inflation. In particular, an above-average degree of resource utilization causes inflation to rise; slack in labor and product markets has the opposite effect. This characterization of the inflation process implies that policymakers face a long-run constraint: If they wish to stabilize the rate of inflation, eventually they must bring about a convergence of actual output to potential through adjustments to the real funds rate. As will be discussed below, this requirement is not necessarily an easy task: Because trend productivity and other elements of potential output are subject to unpredictable shifts, there is the possibility that the monetary authority may significantly mismeasure the output gap.

*Expectations.* Expectations play a key role in FRB/US, especially in the valuation of bond yields and other asset prices. In the version of the model used for this study, we assume that households, firms and investors have model-consistent expectations — that is, they understand fully how the economy works, at least as specified in FRB/US, and use that knowledge to predict the future.<sup>10</sup> This assumption ensures that, in simulating the model, the public's expectations are fully consistent with the specific monetary policy rule in effect -- a desirable feature in a study that attempts to measure the relative long-run performances produced by different policies.<sup>11</sup> Dropping this assumption and replacing it with some form of adaptive expectations would imply a

---

9. In the first Bulletin article, Brayton, Mauskopf, Reifschneider, Tinsley and Williams (1997) discuss the role of expectations in the FRB/US model. In the second article, Reifschneider, Tetlow and Williams (1999) review the basic properties of the model and its predictions for the macroeconomic effects of various shocks. More complete discussion of the design, properties, and uses of the model can be found in Brayton and Tinsley (1996) and in Brayton, Levin, Tryon, and Williams (1997).

10. However, they do not know in advance what shocks to aggregate demand and supply will occur, and therefore base their expectations on the assumption that all future shocks equal zero.

11. A corollary of this assumption is that policy is perfectly credible; for example, in the simulations discussed in this report the public has no doubts about the monetary authority's inflation objectives and its commitment to achieving them.

public that would often be mistaken in its expectations about the size and timing of the response of monetary policy to changes in macroeconomic conditions.

### Level policy rules and measurement error

To generate quantitative estimates of the macroeconomic performance associated with a particular strategy for conducting monetary policy, and to evaluate how performance is affected by measurement error, we need a policy rule indicating how the federal funds rate is adjusted in response to changes in perceived macroeconomic conditions.<sup>12</sup> An example is the well-known Taylor rule,

$$(1) \quad r_t = rr^* + \pi_t + .5(\pi_t - \pi^*) + .5y_t,$$

where  $r$  denotes the nominal federal funds rate;  $rr^*$  is the equilibrium real rate;  $\pi$  is the four-quarter moving average of inflation;  $\pi^*$  is the target rate of inflation sought by policymakers; and  $y$  is the output gap.

For the most part we concentrate on rules that, like the Taylor rule, set the funds rate in response to, among other factors, current estimates of the *level* of the output gap. Research carried out at the Federal Reserve and elsewhere has shown that policies of the type

$$(2) \quad r_t = (1 - \alpha_r)(rr^* + \pi_t) + \alpha_r r_{t-1} + \alpha_\pi(\pi_t - \pi^*) + \alpha_y y_t,$$

do a good job of stabilizing the economy. Specifically, values for the coefficients  $\alpha_r$ ,  $\alpha_\pi$ , and  $\alpha_y$  can be found that, based on stochastic simulations of a variety of standard econometric models, yield relatively low variability in output and inflation about their norms (Levin, Wieland and Williams, 1999). As Williams (1999) shows, for the FRB/US model under model-consistent expectations, a large portion of the superior performance of this type of rule relative to the standard Taylor rule is attributable to the use of the lagged funds rate as an additional

---

12. In all the model simulations, we focus on errors in the measurement of the output gap (and related variables) alone, and ignore errors in the measurement of inflation, on the grounds that the latter have been shown to be relatively minor by Rudebusch (1999) and Orphanides (1997,1999).

determinant.<sup>13</sup> However, including additional variables (beyond the funds rate) in level rules of this type, such as the exchange rate or oil prices, does little to improve macroeconomic performance.

Past evaluations of the performance delivered by this class of level rule have tended to assume that policymakers know the current value of the output gap precisely. However, there are two sources of error in any estimate of the gap. First, it is well understood that policymakers do not observe the current-quarter value of real GDP but must instead estimate it. Because of lags in the receipt of monthly data as well as subsequent revisions, their estimates are subject to significant error. Nevertheless, such errors appear to have little effect on macroeconomic performance: Levin *et al* show that the performance of level rules of the type described by equation 2 deteriorates only a little if the lagged output gap is used in place of its current-quarter value. By contrast, the second source of error -- misestimates of trend productivity and the value of the NAIRU -- is much more important. Historically, these supply-side mistakes have led to errors in the measurement of the output gap that dwarf the errors associated with revisions in GDP data.

To gauge the effects of error in the measurement of resource utilization, we assume as a base case that policymakers observe the gap with an error calibrated to have the same properties as that of the historical revisions to the output gap for the period 1980 to 1994. Let  $\xi$  be the time-series process describing the revisions in the estimate of the output gap. As Table 1 demonstrates, such revision errors are highly persistent. To model them we assume that  $\xi$  follows a first-order autoregressive process,  $\xi_t = \rho \xi_{t-1} + v_t$ , where the disturbance term  $v_t$  is assumed to be a Gaussian white-noise process.<sup>14</sup> As shown in the upper portion of Table 2, even

---

13. For a discussion of why inclusion of the lagged funds rate in policy rules helps to stabilize the economy – a result that holds for a number of other models beside FRB/US – see Sack and Wieland (1999), Tetlow and von zur Meuhlen (1999), and Woodford (1999).

14. Such a first-order process represents a simplification of the true revision process in the data. However, it offers a parsimonious way of capturing the size and persistence of the revisions--the key features for our analysis.

**Table 2**  
**Alternative Measurement Error Processes,**  
**Estimated from Historical Revisions to the Output Gap and Capacity Utilization**

**Assumed Model:**  $\xi_t = \rho \xi_{t-1} + v_t$

Measurement Error for:	$\hat{\rho}$	Standard Deviation of $v$	Unconditional Standard Deviation Of $\xi$
<b>Level of resource utilization <sup>1</sup></b>			
Base case – output gap revisions, 1980:Q1 to 1994:Q4	0.84	0.97	1.77
Worst case – output gap revisions, 1966:Q2 to 1994:Q4	0.96	1.09	3.77
Best case – capacity utilization revisions, 1980:Q1 to 1994:Q4	0.80	0.51	0.86
<b>Gap between actual and trend growth <sup>2</sup></b>			
Base case – output gap revisions, 1980:Q1 to 1994:Q4	0.52	0.71	0.83
Worst case – output gap revisions, 1966:Q2 to 1994:Q4	0.60	0.83	1.04

1. To estimate the error model, historical errors in the level of output are approximated by historical revisions to the level of either the output gap or rescaled capacity utilization.
2. To estimate the error model, historical errors in the growth gap – the difference between actual and trend real GDP growth – are approximated by revisions to the four-quarter change in the output gap.

though the late-period revisions are considerably smaller than those for the full sample, they are still quite substantial: In our base-case simulations, the unconditional standard deviation of measurement error for the output gap is 1.8 percentage points – a sizable number in comparison to the standard deviation of the actual gap itself over the past 30 years, 3.2 percentage points.<sup>15</sup>

As noted in the previous section, while on an ex post basis we can measure the historical degree of error in measures of aggregate resource utilization, it is an open question how much

---

15. Throughout our analysis, we assume that policymakers know the parameters of this process, even if they do not actually observe the errors themselves in real time. Of course, it is unlikely that policymakers would know with certainty the statistical process generating measurement error. However, it is beyond the scope of this paper to provide a more detailed treatment of this form of policymaker uncertainty, given the difficult learning issues entailed in such an analysis.

uncertainty policymakers face going forward. After all, knowledge of the economy does accumulate and the Federal Reserve might do better in the future than it has in the past. However, there also is a chance that unforeseen shifts in productivity growth and the structure of labor and product markets could lead to errors larger than those experienced over the 1980-1994 period.

Consequently, we evaluate the sensitivity of our results to varying degrees of measurement error by considering two additional cases. For our worst-case scenario, we calibrate the gap errors in the simulations to the revisions to the output gap made over the entire period from 1966 to 1994. As shown in Table 2, the implied magnitude of this degree of error is much larger than that of the base case, with a standard deviation equal to 3.8 percentage points, twice that of the base case. By contrast, our best-case scenario posits a degree of measurement error one-half as large as the base case. Unlike the other two cases, here the error process is not calibrated from historical revisions to estimates of the output gap, but instead to revisions to capacity utilization over the 1980 to 1994 period.<sup>16</sup>

## Growth rules

Although our primary focus is on the effect of measurement error on the performance of policies that respond to the estimated level of resource utilization, we are also interested in alternative strategies that ignore the output gap altogether, given the difficulty of measuring it accurately. In the absence of measurement problems, such strategies appear to be inferior to strategies that rely on the output gap, but they may be superior once measurement problems are taken into account. One such alternative is for policymakers to concentrate on the difference

---

16. The three calibrations of measurement error differ not only in the unconditional standard deviation of  $\xi$ , but also in persistence ( $\rho$ ). Our focus in this study is primarily on the consequences of the overall magnitude of measurement error for monetary policy, but it is worth noting that increases in persistence, holding the unconditional standard deviation of measurement error unchanged, also could effect the ability of level rules to stabilize the economy. Our investigation of this issue, however, indicates that variation in persistence of the degree shown in Table 2 has little effect on the effectiveness of such rules. However, we do find that the performance of level rules improves noticeably as  $\rho$  approaches zero, when the standard deviation of  $v$  is increased to keep the unconditional variance of  $\xi$  constant.

between the actual and potential growth rates of the economy. In this case, policymakers would react to the rate of inflation and the rate of *growth* of output instead of its *level*. Since the sum of inflation and real output growth is simply the growth of nominal income, this approach is closely related to nominal income growth targeting, which has often been recommended as an alternative to money growth targeting. The specific rule we examine is

$$(3) \quad r_t = (1 - \alpha_r)(rr^* + \pi_t) + \alpha_r r_{t-1} + \alpha_\pi (\pi_t - \pi^*) + \alpha_g (g_t - g_t^*),$$

where  $g$  denotes the four-quarter growth rate of real GDP, and  $g^*$  is the estimated trend growth rate of potential GDP. This alternative specification follows the natural growth rate targeting rule proposed by Orphanides (1999), an approach that is considerably less sensitive to measurement problems than level rules that respond to resource utilization gaps.

The performance of this growth rule provides a standard against which the performance of level rules can be compared, with and without error in the measurement of the output gap. Of course, growth rules are subject to their own set of measurement errors as well, because initial readings on actual GDP growth are subject to revision, and the long-run growth rate of potential output is not known with certainty. To ensure we compare level and growth-rate rules on a similar standard, we allow for errors in the measurement of  $g - g^*$  in simulations of growth rate rules. Specifically, our calibration of these errors is based on historical revisions to the four-quarter change in the output gap, as discussed in Orphanides (1999); estimates of first-order autoregressive processes fitted to these data are reported in the lower portion of Table 2.

### **Stochastic simulations and policy frontiers**

To carry out our analysis, we rely on “stochastic simulations” that trace out the developments of a FRB/US economy continually buffeted by disturbances to goods, labor, financial, and international markets. The disturbances themselves are drawn from a large sample of random shocks, whose pattern of cross-correlations within a quarter and correlation over time corresponds to that estimated for the residuals to the model’s equations observed over the period

from late 1979 to mid-1998.<sup>17</sup> This period was chosen to correspond roughly with the period from which the base-case degree of measurement error is derived.

In the model simulations, we make the simplifying assumption that the measurement errors made by policymakers are uncorrelated with the true disturbances to productivity, prices, aggregate spending, and so forth that hit the economy. This assumption is unrealistic; for example, the large errors for the output gap recorded in the mid-1970s are no doubt partly attributable to the slow recognition of the slowdown in productivity growth that began a few years earlier. Unfortunately, the assumption is unavoidable. Given that we do not have enough information to fully understand the origin of the errors observed historically, we cannot plausibly link the measurement error to aggregate demand and supply shocks.

Holding the monetary policy strategy fixed, one can use the artificial economy generated by stochastic simulations of FRB/US to compute the average economic performance delivered by that strategy. Because the policies we examine succeed on average in achieving their objective, we do not measure performance by *first moments*, that is, the mean outcomes for output, unemployment, and so forth. Both the level and growth rules ensure that resource utilization averages to its “natural” level over extended periods of time, and both produce an average long-run rate of inflation equal to the assumed (constant) target rate sought by policymakers. Instead, our definition of performance concerns *second moments*: the variability of output, inflation and interest rates.<sup>18</sup>

---

17. Although the standard version of FRB/US measures trend productivity using deterministic time trends – which by definition are not subject to random shocks – for this paper the model was modified slightly to allow the level of trend productivity to be buffeted by highly persistent stochastic disturbances. Kuttner (1992) calibrated the magnitude of these productivity shocks to match the stochastic process for potential GDP.

18. Because stochastic simulations are computationally intensive when expectations are model consistent, we linearize FRB/US to take advantage of the AIM algorithm, an efficient solution procedure discussed by Anderson and Moore (1985) and Anderson (1999); to explore this algorithm and examine up-to-date versions of the AIM code see <http://www.bog.frb.fed.us/pubs/oss/oss4/aimindex.html>. A side benefit of this approach is that it enables us to determine analytically the macroeconomic performance of specific policy rules through intermediate computations of the unconditional variance-covariance matrix of the system. Thus, in computing variances for output, inflation, and other variables, we do not literally generate long time series of simulated data. However, if we were to do so, the results would be identical to those we



Policies that generate lower macroeconomic variability are assumed to be preferable to ones that cause the economy to be more volatile. In particular, in choosing between any two policies yielding the same standard deviation for changes in the funds rate, the strategy associated with lower variability of both output and inflation will be deemed to be superior. Of course, it is not true that for any pair of policies, one will invariably produce smaller unconditional standard deviations for both output and inflation; for some pairs, one policy will do better in one dimension and worse in the other. In such cases, neither policy dominates the other.

Performance comparisons can be carried out for all possible rules of a given type, to find the subset that are not dominated by other rules of the same type. For example, one can take the level rule described by equation 2 and compute the unconditional standard deviations of output and inflation associated with every possible combination of values for its three coefficients,  $\alpha_r$ ,  $\alpha_\pi$ , and  $\alpha_y$ . Discarding those coefficient sets that are associated with greater variability in both dimensions yields a collection of *efficient rules*, in the sense that no other specific rule of that type can deliver lower variability of inflation without an accompanying increase in output variability. The tradeoff between output and inflation variability described by this set of efficient rules is called the *policy frontier*.

Formally, computation of a policy frontier involves minimizing a loss function of the form

$$(4) \quad L = \varphi V(y) + (1 - \varphi)V(\pi),$$

subject to the model and the variance-covariance matrix of stochastic shocks, where  $V()$  denotes the unconditional variance of the variable within braces. Points along the frontier are found by minimizing  $L$  for values of the parameter  $\varphi$  between zero and one. To ensure that all frontier policies produce similarly sized fluctuations in the funds rate (and are therefore comparable when viewed in the space of outcomes for the unconditional standard deviations of output and inflation), we impose an additional constraint on the minimization problem. Specifically, we only consider policies that yield a standard deviation of the quarterly change in the federal funds rate that is no greater than the standard deviation in the data over the past 30 years. This constraint

---

report.

has the effect of restricting our attention to frontier rules that display realistic movements in the stance of monetary policy — that is, ones within the range of historical experience.

Because different types of rules yield different frontiers, they provide a convenient way of comparing each type's overall performance: As will be seen, under some conditions one class of rule will clearly dominate another, in the sense that all points along its frontier will lie to the southwest of the other. Frontiers also can be used to summarize the effect of measurement error on the performance of rules by showing how the position of the frontier is affected by increases in the degree of mismeasurement of the output gap and potential output growth.

In the next section, we move in stages to full-scale frontier analysis under various assumptions about both the size of output gap mismeasurement faced by the monetary authority and its ability to cope with this uncertainty. We begin by considering the performance of the Taylor rule and consider how it is altered by measurement error. Next, we consider a class of Taylor-type rules in which the responsiveness of the funds rate to the output gap is gradually increased. This analysis serves to illustrate the marginal benefit of basing policy actions in part on measured resource utilization and shows how this benefit falls as measurement error increases. From there we move to a comparison of actual policy frontiers and show how the performance delivered by efficient level rules declines as the output gap is measured with less accuracy. As part of this discussion, we also consider how the nature of these efficient rules changes in the face of uncertainty about the true state of resource utilization. Finally, we consider the comparative performance of alternative efficient rules that ignore the output gap altogether and instead concentrate on the growth rate of the economy.

#### **IV. Simulation Results**

This section presents our quantitative findings on the stochastic performance of the economy under alternative monetary policy rules. These rules will vary according to the relative importance placed by policymakers on output and inflation stabilization, on the specific factors incorporated into the rule, and on the degree of error underlying measures of resource utilization. This focus is meant to be broad enough to encompass the range of choices and situations faced by policymakers.

## Taylor rules

Table 3 shows the implications for economic performance of errors in the measurement of the output gap under two policy rules: the original version of the Taylor rule, and a revised version recently put forth by Taylor in which the coefficient on the output gap is set to unity instead of 0.5.<sup>19</sup> For purposes of comparison, the table also contains information on the actual performance of the U.S. economy over the 1980-1998 period.<sup>20</sup> We choose these two rules as illustrative examples because they are well known; in addition, the original version of the rule “fits” the data, at least for the period examined by Taylor.<sup>21</sup> Results are shown for the three different assumptions about the magnitude of output gap error: a base case, derived from empirical evidence on output gap mismeasurement for the period from 1980 to 1994; a small-errors case, calibrated to the revision history of capacity utilization; and a large errors case, derived from data on output gap mismeasurement over the entire period from 1966 to 1994.

The first row of the table shows the standard deviations of the output gap, inflation, and the change in the federal funds rate as generated by stochastic simulations of the FRB/US model under the Taylor rule, equation (1). In these simulations, the FRB/US economy is subjected to stochastic shocks representative of those that have hit the economy over approximately the past 20 years. Although these disturbances include shocks to potential output, we assume initially

---

19. Citing Brayton *et al.* (1997), Taylor (1998) advocates this alteration as having the advantage of producing a more procyclical funds rate.

20. Care should be taken in comparing historical standard deviations with the stochastic simulation statistics, since the two are not strictly comparable. The Volcker disinflation gave rise to a pronounced decline in the trend rate of inflation over the first half of the 1980s, and this decline causes the historical variability of inflation to be overstated relative to the stochastic simulation results, which do not incorporate similar disinflationary episodes. In the model simulations, the long-run trend rate of inflation is stationary, because the target rate of inflation is always held constant.

21. Taylor (1993) noted that his rule fit the data for the 1987-1992 period – that is, the rule's errors were roughly zero over the period, based on the output gap estimates then available.

**Table 3**  
**Performance of Original and Revised Taylor Rules**

Rule	Magnitude Of Error	Unconditional Standard Deviations, in percent		
		Actual Output Gap	PCE Inflation	Quarterly change in Funds Rate
Original Specification	Without errors	2.4	1.8	1.0
	small errors	2.5	1.9	1.0
	base-case errors	2.6	1.9	1.1
	large errors	3.6	3.1	1.2
Revised Specification	Without errors	1.8	2.0	1.6
	small errors	1.9	2.0	1.7
	base-case errors	2.2	2.1	1.9
	large errors	3.6	4.4	2.1
Memo: the data 1980:Q1-1998:Q4		2.1	2.3	1.4

that the FOMC perceives these shocks as they occur and thus measures the output gap correctly. (To keep matters straight in this and all other cases, the unconditional standard deviation of the output gap should be viewed as that of the correct *ex post* measure of the gap, notwithstanding any *ex ante* misperceptions.)

The second, third and fourth rows of the table, in comparison with the first, show the growing deterioration of performance associated with increasing mismeasurement of the output gap. For the small-errors case, the degradation in performance is minor, and even for the base case it is relatively modest. However, the variability of output and inflation rises dramatically if measurement error is on the order characterized by the large errors, our worst-case example.

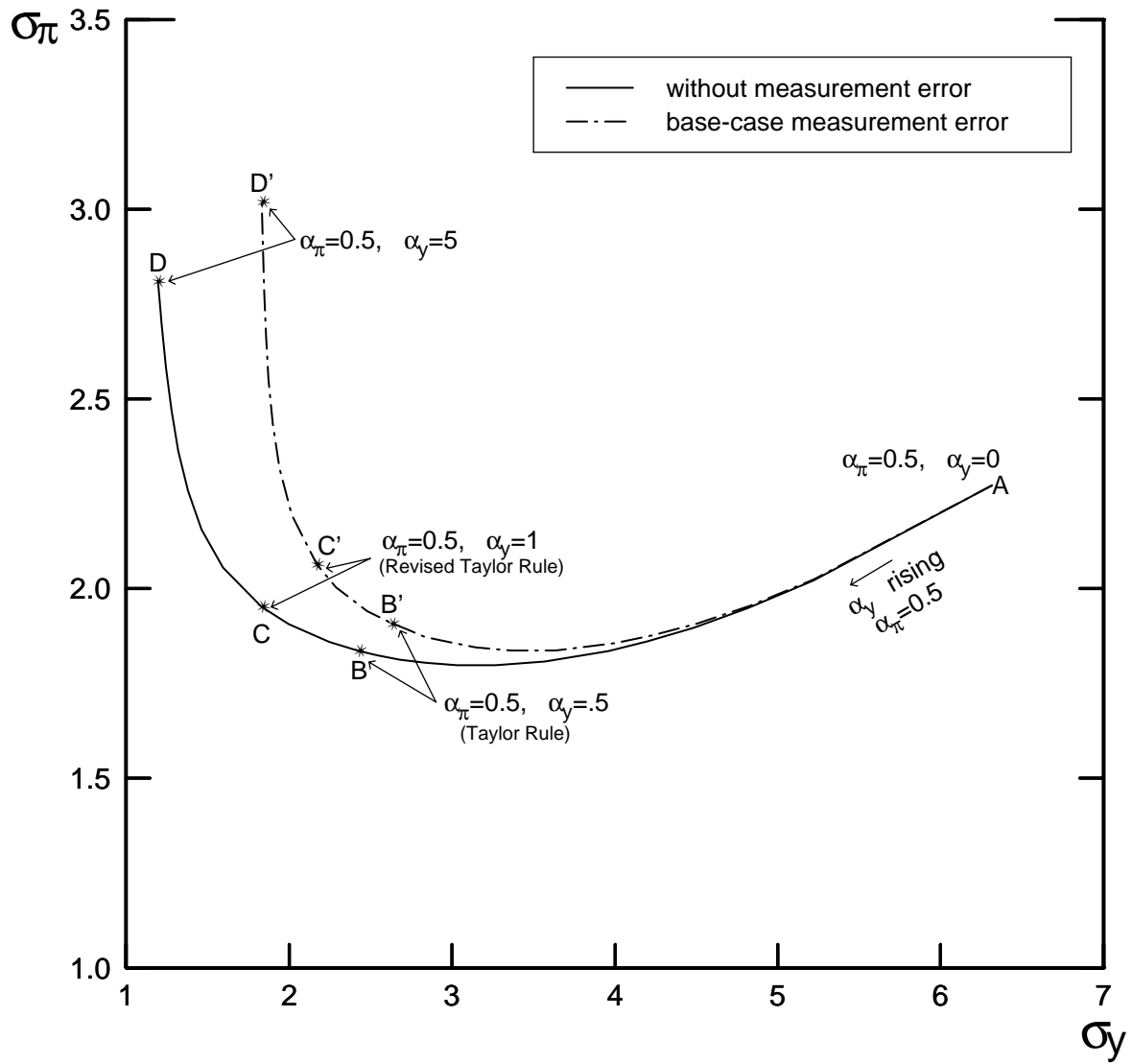
The bottom half of the table focuses on the revised Taylor rule, which is more responsive to the output gap than Taylor's original specification. If there is no measurement error, the revised rule yields substantially better performance in terms of reducing output variability at the cost of a small deterioration in inflation control. This improvement over the original Taylor rule is achieved through an increase in the variability of changes in the funds rate. With measurement error, the improvement in reduced output variability is less and comes at the cost of a significantly greater deterioration in inflation variability: Comparing outcomes under the base case to performance with no measurement error, a two-tenths increase in the standard deviation of the output gap under the original rule is doubled under the revised specification. As one might expect, the deterioration in performance varies directly with the weight placed on the output gap in the setting of the funds rate.

If assigning more weight to the output gap in the rule yields deterioration in economic performance, should we give any positive weight to the output gap (or other measures of resource utilization)? That is, what is the proper response when mismeasurement is a serious problem? We begin a systematic investigation of this question with the aid of Figure 1. Suppose a monetary authority wishes to use a rule like the Taylor rule to guide monetary policy but is concerned that mismeasurement of potential output may lead to policy mistakes. One extreme option is to set the coefficient on the output gap to zero – that is, to not respond at all to movements in the perceived output gap. The performance of such a rule is shown as point A on the right-hand side of the figure, where performance is measured in terms of the unconditional standard deviations of the output gap ( $\sigma_y$ ) and inflation ( $\sigma_\pi$ ), both expressed in percentage points.

Zeroing out the output coefficient, however, may be an overreaction to the problem. We can investigate the consequences of incrementally adjusting upward the output coefficient, while holding the inflation coefficient fixed at the value in the Taylor rule ( $\alpha_\pi = 0.5$ ) and computing the implications for economic performance. This experiment traces out two curves, beginning from point A in Figure 1, one for the case where there are no measurement errors (the solid line) and the other for base-case mismeasurement of the output gap (the dot-dashed line). At points

Figure 1

Improvement in Macroeconomic Performance  
 from Increasing the Sensitivity of Policy to the Output Gap  
 With and Without Measurement Error  
 (axes in percent)



close to A, it makes little difference whether or not there is measurement error for the decision to place some weight on the perceived output gap in setting policy. The two curves show that, up to a point, placing some weight on resource utilization improves macroeconomic performance in terms of a reduction in the variability of both output and inflation simultaneously.<sup>22</sup> Eventually, however, any further decreases in output variability can only come at the expense of an increase in inflation variability. Comparing the dot-dashed and solid lines shows that the presence of mismeasurement makes the trade-off between the standard deviation of output and the standard deviation of inflation less favorable than it otherwise would be.

Although it is clear from Figure 1 that measurement error decreases the incremental benefit from responding to the output gap, it is not clear at this point how sensitive this result is to changes in the magnitude of error. Figure 2 illustrates this sensitivity, by plotting the same two curves shown in Figure 1 along with two others. These additional curves show the same incremental benefit of output gap responsiveness, but with large and small degrees of measurement error, as defined earlier. As can be seen, the deterioration in the benefit from increased responsiveness to the output gap is minor for the small-errors case, but very large for the large-errors example.

### **Frontiers with level rules**

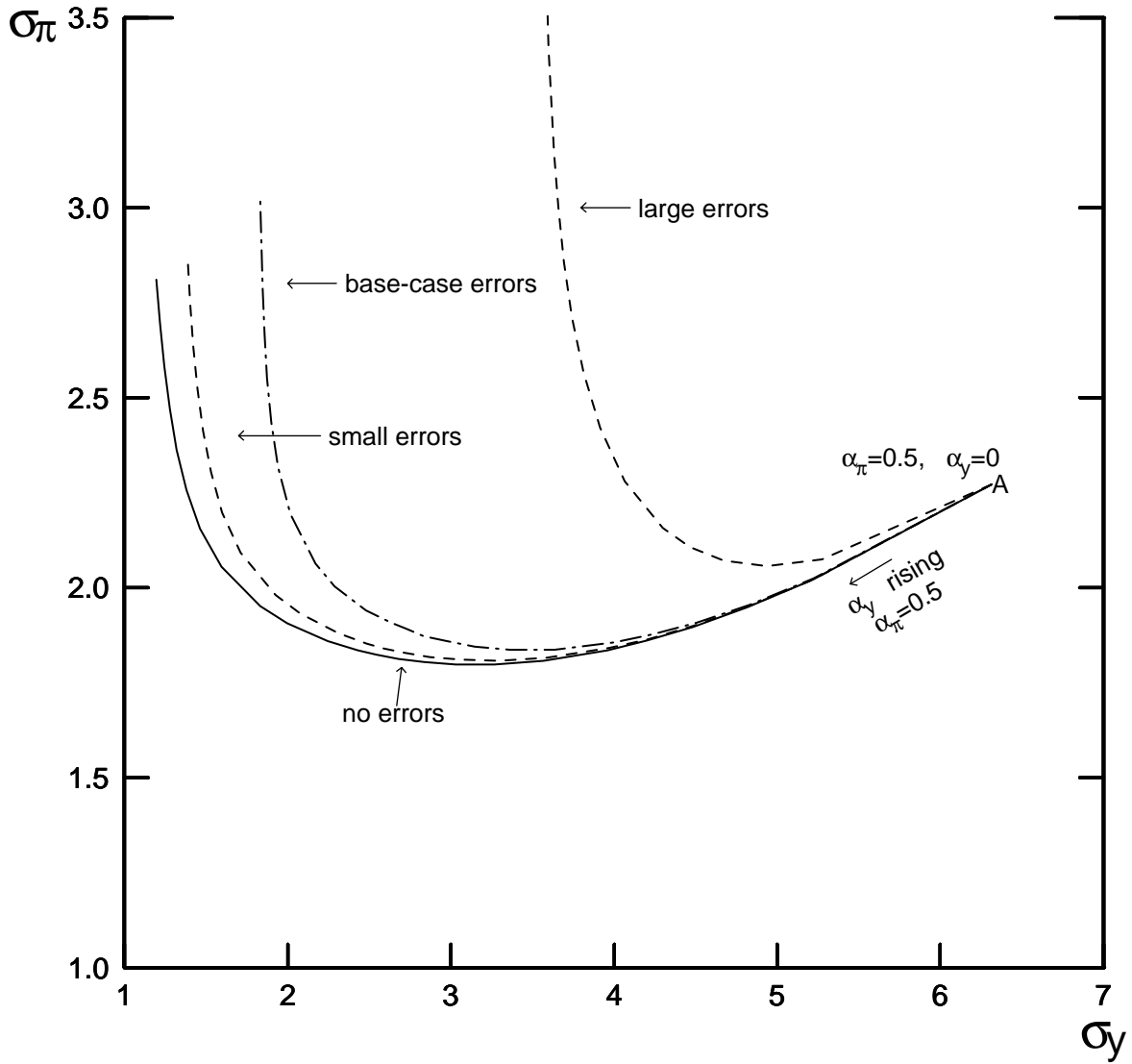
The preceding discussion is useful for judging the marginal contribution of the output gap to policy design, holding the responsiveness of the funds rate to inflation unchanged. The figures tell us that policymakers who place any significant weight on stabilizing output as an objective of policy would not want to abandon the output gap entirely -- at least, not in the absence of some alternative measure of resource utilization. The exercise embodied in these figures was, however, a restricted one. We now relax some of the restrictions by making comparisons across policy frontiers, with and without mismeasurement of the output gap. To make these

---

22. This is not a general result. In particular, if at the initial point the funds rate is already very responsive to inflation – for example, if  $\alpha_\pi = 4$  – then adding some consideration of resource utilization will reduce output variability only at a cost of increased fluctuations in inflation.

Figure 2

Effect of Increasing Uncertainty on the  
Marginal Contribution of Increased Sensitivity  
to the Output Gap  
(axes in percent)





comparisons, we generalize the policy rule by allowing a coefficient on the lagged federal funds rate, thereby allowing choices over three coefficients,  $\{\alpha_r, \alpha_\pi, \alpha_y\}$ , as shown in equation 2.

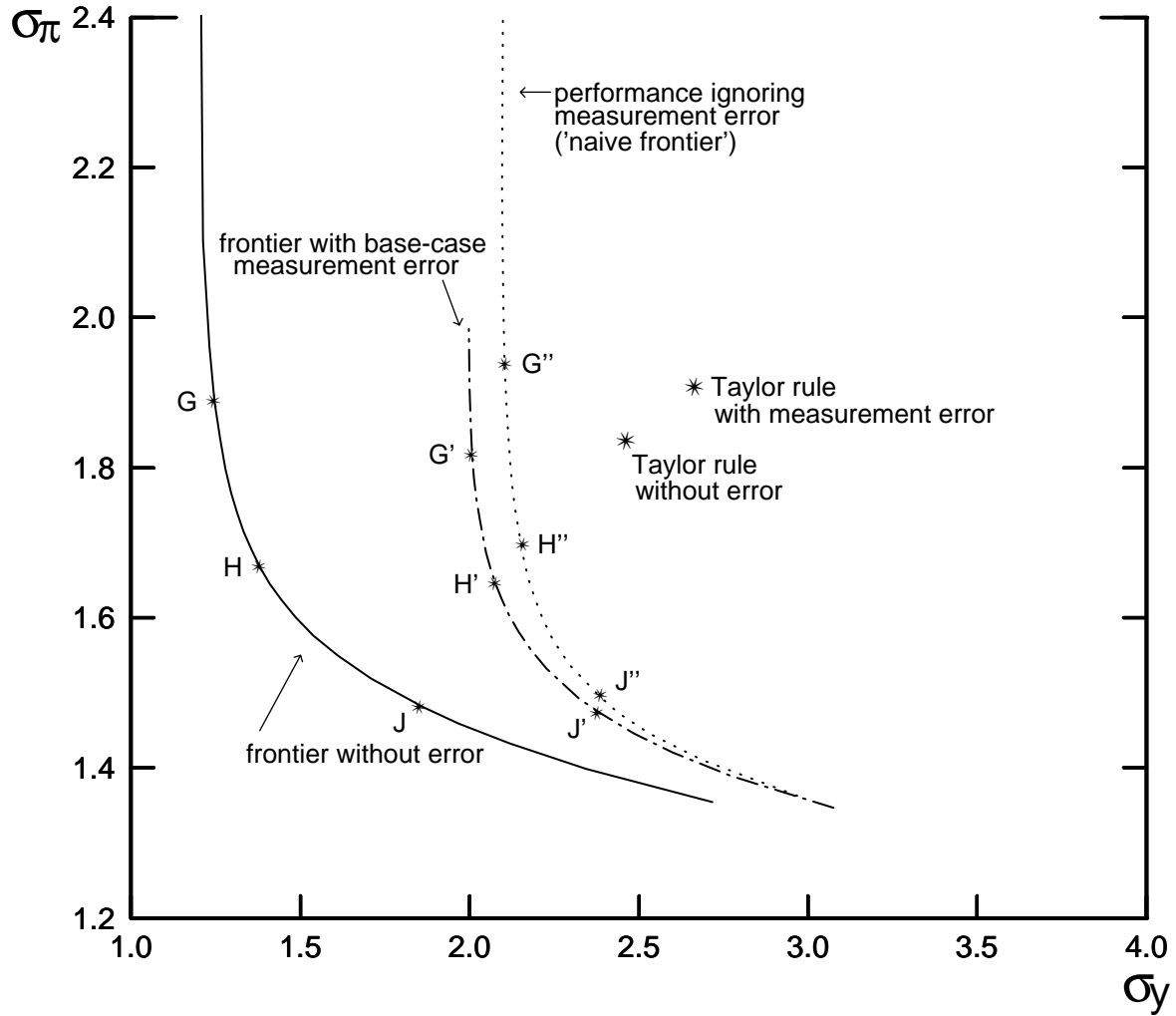
Recall that a policy frontier is the locus of points showing the combinations of the unconditional standard deviations of output and inflation produced by rules that are efficient, in the sense that no alternative set of coefficients for the given type of rule can produce a reduction in the variability of output (inflation) without an increase in the variability of inflation (output), or an increase in the standard deviation of the change in the federal funds rate. The solid line to the southwest of Figure 3 shows the frontier for the FRB/US model for the level rule, using the same set of shocks described above but with no mismeasurement of the output gap. This curve describes the best set of trade-offs available to the monetary authority, assuming that it makes no errors in measuring the output gap whatsoever.

By contrast, the dotted curve of the figure — the one farthest to the northeast — shows the actual performance delivered by the exact same set of rules (i.e., rules with the same coefficient values) underlying the no-errors frontier, but with policymakers actually facing the base-case degree of measurement error. In model simulation exercises, such rules might be chosen by a naive monetary authority that fails to take into account mismeasurement of the output gap in gauging its policy stance. We refer to the dotted curve as the “naive frontier,” although in truth it is not a true frontier — that is, none of the rules underlying the curve are efficient when the output gap is known to be estimated imprecisely.

By most measures, the difference between the solid and dotted lines would have to be described as large. For policies towards the top of the curves — policies selected because of a strong aversion to output variability — the gap between the frontier without errors and the naive frontier is about a percentage point, measured in terms of the standard deviation of output. Towards the bottom of the two curves, which represent policies that are more tolerant of output variability, the differences are considerably smaller. This finding is consistent with the results of Table 3, which show that the Taylor rule is less affected by measurement error than its revised version, which has a larger coefficient on the output gap.

Figure 3

Effect of Measurement Error on Policy Frontiers  
for the Output Level Rule  
(axes in percent)



The dot-dashed line in Figure 3, which lies between the frontier without errors and the naive frontier, is the *re-optimized frontier*. In contrast to the naive frontier, the specific rules underlying this curve have been recomputed to find those that are efficient in the face of measurement error. There are two important lessons to be drawn from this intermediate frontier. First, the horizontal distance between the re-optimized frontier and the naive frontier is not large. That is, re-optimization results in an improvement in output variability, but not a large one, and the improvement is substantially smaller than the deterioration caused by measurement error itself.<sup>23</sup> Second, the vertical distance measuring inflation variability between the re-optimized frontier and the naive frontier is considerably larger, at least for ranges of preferences that heavily favor output stabilization.

To expand on these two points, consider the effect of measurement error on the performance of particular rules, that is, level rules with specific numerical values for the coefficients on output, inflation, and the lagged funds rate. Point G on the solid line represents the choice that might be made by an authority that estimates the output gap without error, and places a fairly strong weight on output stabilization relative to that placed on inflation stabilization. (In the context of the loss function, such a preference corresponds to a value for  $\varphi$  equal to 0.85.) But if measurement error is a problem, the performance generated by rule G is G'; relative to what is attainable in the absence of error, there is a large deterioration in output variability, with little change in inflation stability. Given that measurement error applies only to the output gap, this result is not surprising. Now consider the performance of a rule chosen by policymakers with the same preferences but who take into account the existence of measurement error in their search for efficient policies.<sup>24</sup> The resulting frontier point is G' on the dot-dashed curve. As can be seen, policymakers can substantially reduce the variability of inflation if they

---

23. However, other simulation exercises reveal that if measurement error is extremely large, as was the case during the 1970s, then the improvement from re-optimization can be quite substantial.

24. Formally, for  $\varphi = 0.85$  rule G' minimizes the loss function described by equation 4 in the presence of base-case measurement error, while rule G minimizes the same loss function if the absence of error.

choose rules with an eye to measurement error; however, they gain little in terms of reduced output variability. This result holds because there is little a monetary authority can do to counter the effect of mismeasurement on the variability of the output gap itself; that effect is too direct to be mitigated substantially. But inflation variability can be brought down significantly simply by managing output less actively, as occurs with the move from G'' to G'.<sup>25</sup>

### **Policy Attenuation**

To move from point G'' to point G' in Figure 3, the policy rule was altered in recognition of the presence of measurement error in the output gap. We now take a closer look at the nature of this adaptation, and in particular consider the typical quantitative effect of measurement error on the coefficients of efficient policy rules. Our focus in this discussion is whether measurement error leads in general to attenuation – that is, to a decrease in the responsiveness of policy to estimates of current economic conditions.

To answer this question, consider the upper panel of Table 4, which lists the coefficients of three different rules, G, H and J. The performance of these rules in the absence of error is shown in Figure 3; as can be seen, all three lie on a frontier if the output gap is measured accurately. In choosing one rule on the frontier over another, a policymaker would demonstrate an implied tolerance for output variability relative to inflation variability, as indexed by the parameter  $\varphi$  in the loss function (equation 4). For example, rule G would be chosen by policymakers who are relatively intolerant of output variability ( $\varphi = 0.85$ ), while rule J would be chosen by policymakers who have the opposite preferences ( $\varphi = 0.15$ ). Rule H would be chosen by a monetary authority that dislikes fluctuations in output and inflation equally ( $\varphi = 0.5$ ). Not

---

25. Points H and J together with corresponding points on the re-optimized and naive frontier make the same point, except the extent of the improvement – particularly the improvement in terms of reduction in inflation variability – is a function of preferences ( $\varphi$ ). Preferences toward the northern end of the solid line – that is, preferences that are strongly adverse to output variability – tend to produce the largest deterioration in output variability from ignoring measurement error; they also tend to produce the largest improvements in terms of reduced inflation variability from re-optimization.

**Table 4**  
**The Effect of Measurement Error**  
**on Policy Attenuation for the Level Rule,**

$$r_t = (1 - \alpha_r)(rr^* + \pi_t) + \alpha_r r_{t-1} + \alpha_\pi (\pi_t - \pi^*) + \alpha_y y_t,$$

(selected relative weights for output variability)

Relative weight on output Variability ( $\varphi$ )	Location on Figure 3	Optimal Coefficients		
		$\alpha_r$	$\alpha_\pi$	$\alpha_y$
Without measurement error				
0.15	J, J''	0.96	1.22	0.88
0.50	H, H''	0.99	0.63	1.12
0.85	G, G''	0.98	0.33	1.17
base-case measurement error				
0.15	J'	0.81	1.10	0.61
0.50	H'	0.77	0.62	0.73
0.85	G'	0.75	0.38	0.77

surprisingly, the specific rules associated with these preferences have rising coefficients on inflation, and declining coefficients on the output gap, as one moves along the frontier from point G to point J. This pattern shows up in the first three rows of Table 4. In all three cases, the coefficient applied to the lagged federal funds rate is close to unity.<sup>26</sup>

Now consider how the specific rules produced by these different tolerances change when (base-case) measurement error is present and taken into account. The performance associated with this re-optimization, using the same three assumptions for  $\varphi$ , are points G', H' and J' on the

---

26. Values for  $\varphi$  do not map directly into the optimal coefficients on output and inflation in frontier rules, because optimal coefficients also depend on the magnitude of the stochastic disturbances as well as the dynamics of the macroeconomy embedded in the model.

dot-dashed curve in Figure 3; the coefficients of these three rules are shown in the bottom panel of Table 3. Several interesting observations can be gleaned from a comparison of these coefficients with those derived in the absence of measurement error. First, as the relative weight on output stabilization increases, the coefficient on output rises. This rise occurs with or without measurement error, although the increases are less pronounced under base-case measurement error. Similarly, with measurement error the coefficient on inflation falls less rapidly as the relative tolerance of output variability declines.

Second, the general effect of measurement error is to reduce the size of the coefficients on output and inflation in optimized policy rules. Without dwelling on specific pair-wise comparisons, the table shows that the presence of measurement error always results in an attenuation of policy responses to perceived movements in the output gap, with attenuation greatest for rules associated with a low tolerance for output variability. The results also indicate some attenuation in the responsiveness of the funds rate to inflation, although  $\alpha_\pi$  actually increases slightly if a high enough priority is placed on output stabilization.<sup>27</sup>

Finally, measurement error noticeably reduces the coefficient on the lagged funds rate, which implicitly represents the effect of past values of the output gap and inflation on the current setting of the funds rate. Because past values of the gap are not known with certainty even well after the fact, their value in determining the proper stance of monetary policy today is reduced by increased measurement error.

From the viewpoint of the optimal control literature, it may appear surprising that policy attenuation is the efficient response to uncertainty in the output gap. In particular, if the problem were one of additive uncertainty, then application of the standard separation theorem would suggest that the policy rule should remain invariant to this uncertainty. There are, however, two problems with this argument in the present circumstance. First, the certainty equivalence result is a property of optimal control rules, not simple rules such as the ones examined here. In the

---

27. An alternative interpretation takes account of the fact that  $(1-\alpha_\pi)$ , the coefficient on the equilibrium funds rate plus inflation, goes up. This shift, in combination with the change  $\alpha_\pi$ , suggests a net substitution in response to measurement error away from reliance on the output gap to greater reliance on inflation.

context of the FRB/US model, rules that are optimal in this sense would require conditioning on literally hundreds of variables; only then would additive uncertainty not influence the optimal rule. The second problem is that measurement error is not additive uncertainty as such, because our measure of the output gap does not represent the optimal within-quarter forecast of the output gap. To qualify as such a measure, the output gap would have to be estimated using an optimal filter defined in the context of the FRB/US model, and as such would look quite different from the output gap measures currently or historically available to policymakers.

Before leaving this subsection, we note that none of the rules in Table 4 look much like the original Taylor rule: Nearly all of them have larger impact coefficients, and all, through the mechanism of the lagged funds rate, display more persistence. Figure 3 shows that, in terms of standard deviations of inflation and the output gap, the Taylor rule is not a particularly attractive rule in the FRB/US model: Its performance is dominated not only by optimal rules but also by suboptimal rules on the naive frontier.

### **Sensitivity analysis**

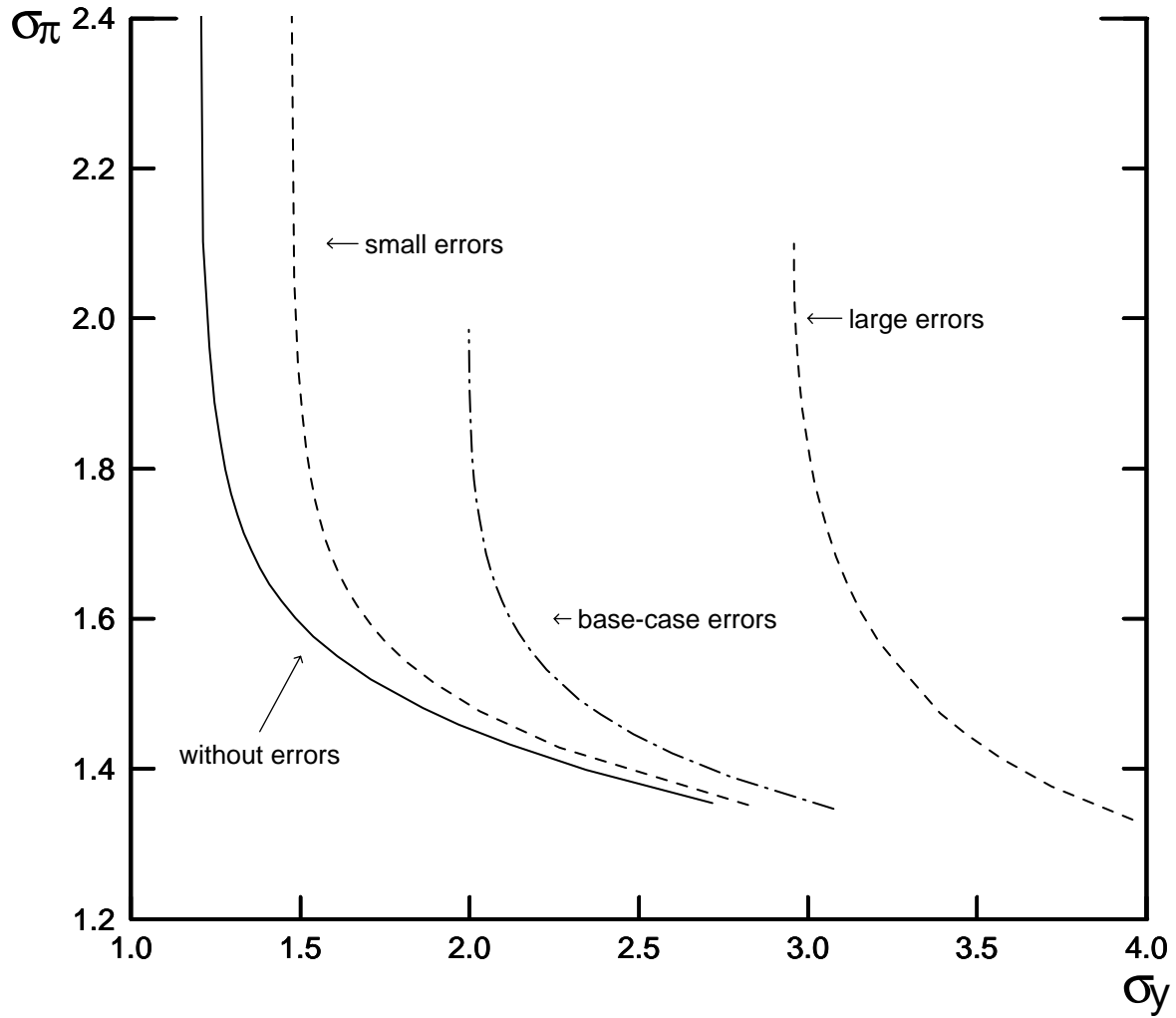
As stressed in section 2, it is not clear how much uncertainty underlies current estimates of resource utilization: Although it may be on a par with that observed over the more recent 1980-1994 period for the output gap (our base-case assumption), arguments can be made for both a larger and smaller degree of measurement error. Thus, it is of interest to examine the sensitivity of our results to alternative assumptions about the likely magnitude of mismeasurement going forward.

Figure 4 outlines the implications of alternative degrees of measurement error on economic performance by considering two additional frontiers computed using the large and small errors discussed earlier, with full re-optimization to the new degree of mismeasurement.

The sensitivity analysis added by these two curves (the dashed lines) helps put the base-case performance in perspective. Relative to the no-errors case, a small degree of measurement error – namely, errors that are one-half the magnitude of those of the base case – produces a shift to the northeast that is significantly less than one-half the size of the shift found in the base case.

Figure 4

Implications for Efficient Output Level Policies  
of Alternative Magnitudes of Measurement Error  
(axes in percent)





At the same time, measurement error on the large scale recorded for the entire 1966-1994 period suggests the possibility of very poor performance indeed for level rules.

Table 5 shows the degree of attenuation required by re-optimization under these different degrees of error. As one would expect, results obtained with small errors lead to less attenuation on all fronts. But, were the economy to experience errors like those of the large-errors case, a re-calculation of the optimal policy rule would yield a large reduction in output responsiveness: Relative to the no-errors case, coefficient values would be more than cut in half. This result is important, not so much because the possibility of errors this large is that great, but because any failure to recognize the mismeasurement of the output gap, and attenuate policy accordingly, would result in even poorer performance than that shown by the large-errors frontier in Figure 4.

### **Frontiers with growth rules**

The range of measurement error just discussed, and the accompanying variation in economic performance, suggest that it would be useful to consider the comparative performance of rules that do not rely on accurate measures of the level of resource utilization. If a rule that disregards the output gap altogether could be shown to perform reasonably well, then mismeasurement of the level of potential output – and the uncertainty regarding the magnitude of that mismeasurement – could be ignored. One rule that does not respond to levels of resource utilization is the growth rule described by equation 3. Of course, rules that rely on growth rates are also subject to measurement error. But the nature of growth gap error is less problematic than output gap error because, for example, an error in the estimated rate of trend growth does not have the same cumulative effect on the former as it does on the latter.

To examine the comparative performance of growth rules, we compute three frontiers – one without mismeasurement of the growth rate gap ( $g - g^*$ ), one with measurement error of the magnitude experienced over the 1980 to 1994 period, and one with error on the scale encountered over the last 30 years. As noted earlier, for our base case we assume that errors in the estimated gap between actual and trend output growth have an unconditional standard deviation of about 0.8 percent. Abstracting from revisions to actual GDP, this degree of error is

consistent with a sizable degree of uncertainty about the rate of potential GDP growth: For example, if the point estimate of trend growth were 3 percent, the 95 percent confidence interval would extend from 1-1/2 to 4-1/2 percent.

**Table 5**  
**The Sensitivity of Policy Attenuation**  
**To Varying Degrees of Measurement Error for the Level Rule,**

$$r_t = (1 - \alpha_r)(rr^* + \pi_t) + \alpha_r r_{t-1} + \alpha_\pi(\pi_t - \pi^*) + \alpha_y y_t,$$

(selected relative weights for output variability)

Measurement Error	Location on Figure 3 if available	Optimal coefficients		
		$\alpha_R$	$\alpha_\pi$	$\alpha_y$
Relative weight on output variability equals 0.15				
No error	J, J''	0.96	1.22	0.88
Small errors	N/A	0.90	1.16	0.78
Base-case error	J'	0.81	1.10	0.61
Large errors	N/A	0.62	1.64	0.34
Relative weight on output variability equals 0.50				
No error	H, H''	0.99	0.63	1.12
Small errors	N/A	0.90	0.60	0.96
Base-case error	H'	0.77	0.62	0.73
Large errors	N/A	0.61	1.28	0.49
Relative weight on output variability equals 0.85				
No error	G, G''	0.98	0.33	1.17
Small errors	N/A	0.89	0.32	1.01
Base-case error	G'	0.75	0.38	0.77
Large errors	N/A	0.62	1.08	0.56

Figure 5 shows several of the same level rule frontiers that were plotted in Figure 4 (the dashed lines), and compares them to frontiers for the growth rules (the solid lines). Two important points stand out from this graph. First, whether one might do better – given error in the measurement of the gap – ignoring levels altogether depends on the extent of the measurement problem. For example, if our base-case calibration of output gap error is representative of the uncertainty faced by the FOMC, then level rules dominate growth rules in terms of economic performance. However, if uncertainty about the true level of the output gap is very large, then it is possible to obtain the opposite result. In particular, if mismeasurement of the level and growth of potential output is on the order experienced over the past 30 years, then optimized growth rules would tend to outperform optimized level rules. This change in dominance occurs because, unlike level rules, the performance of growth rules is not particularly sensitive to the magnitude of errors to estimates of potential output growth – a characteristic demonstrated by the small distance separating the three growth rule frontiers.<sup>28</sup>

#### **IV. Conclusions**

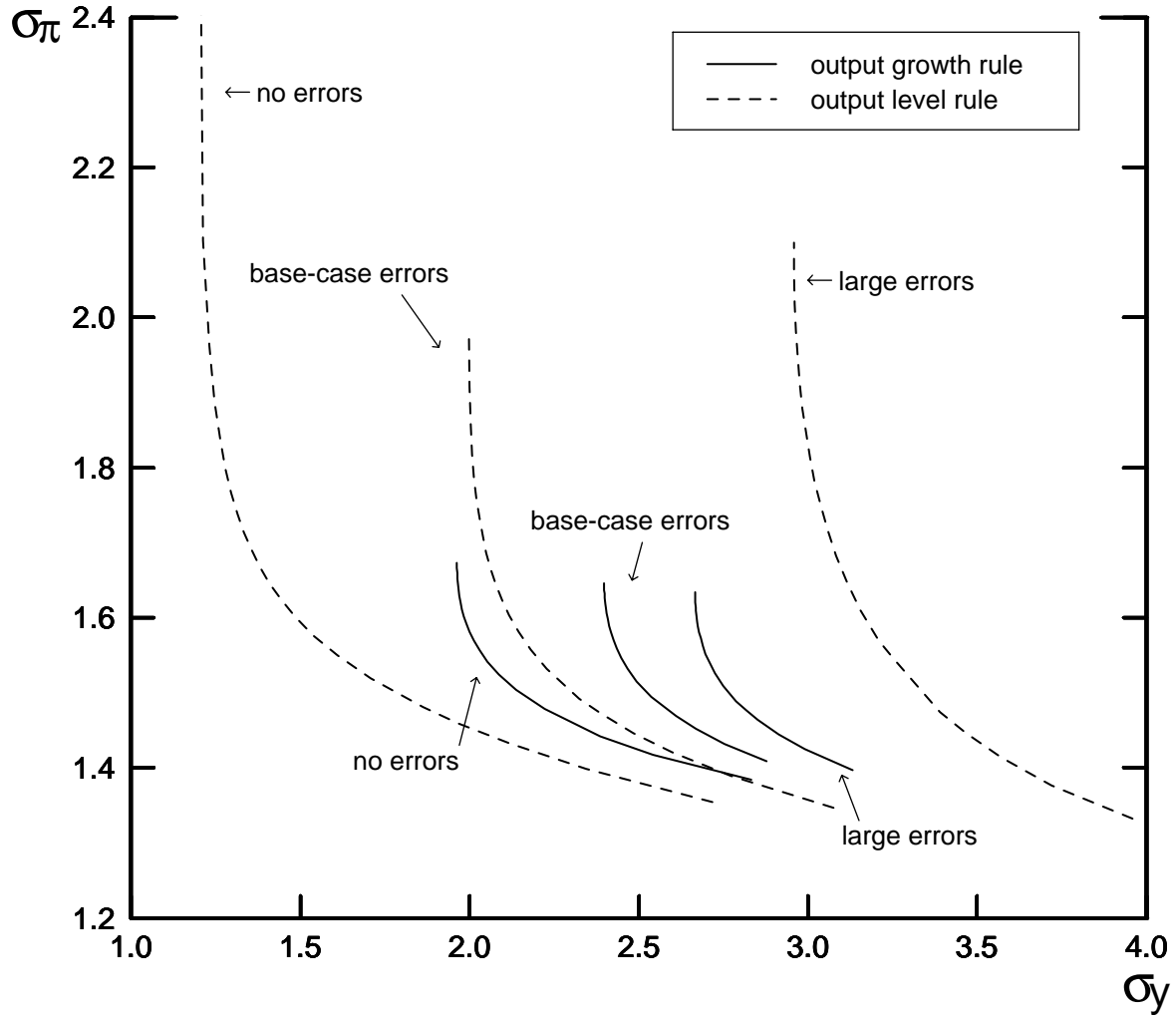
This report has examined the implications of imprecision in output gap estimates for the design of monetary policy. Broadly speaking, we find that measurement error of the magnitude experienced historically has important consequences for policy, as evaluated in the context of policy rules that respond to the estimated level of aggregate resource utilization (among other factors). In particular, such error makes the output gap a less useful guide to setting the funds rate, and leads to an increase in the variability of output and inflation. Based on the experience of the 1980 to 1994 period, the consequences of measurement error are significant but not that

---

28. As with level rules, the performance of growth rules does not appear to be particularly sensitive to variations in the persistence of measurement error, at least over a plausible range of values. Specifically, we find that raising the autocorrelation coefficient in the error process from 0.5 to 0.9 -- holding the unconditional variance of measurement error constant -- leads to only a modest increase in the variability of output and inflation associated with optimized growth rules. Thus, even if our calibration of the error process understates the persistence that might arise from periodic unanticipated shifts in the trend rate of growth -- given that such shifts would likely generate long strings of one-sided errors -- our conclusions concerning the relative benefits of pursuing a growth rate strategy would be little affected. Note even for the large errors case, the estimated autocorrelation coefficient was not that large, only 0.6.

Figure 5

Comparative Performance of Efficient Policies:  
Output Growth Rule versus Output Level Rule  
(axes in percent)



pronounced. However, the experience of the 1970s suggests that there are circumstances in which the output gap is an extremely poor guide to policy – for example, when the economy undergoes significant structural change, and real-time estimates of potential output are seriously flawed. Under such conditions, the monetary authority could impart considerable unnecessary variation in output and inflation by responding strongly to (mis)estimates of excess demand and supply.

One appropriate response to rising uncertainty about resource utilization is *attenuation*, that is, a reduction in the coefficient on the output gap in the interest rate policy rule. However, while attenuation mitigates the increase in output and inflation variability that arises from measurement error, the extent of mitigation is limited. Another possible response to uncertainty is to ignore the estimated level of resource utilization altogether in setting policy, and to instead concentrate on other economic indicators, such as output growth. Our analysis suggests that such a strategy — specifically, the use of a funds rate rule that responds to the difference between actual and trend real GDP growth in place of the output gap — yields inferior macroeconomic outcomes if the magnitude of measurement error is moderate (the base-case or small errors cases discussed in the text). However, if uncertainty about the level and growth rate of potential GDP is especially pronounced, then growth rules are superior to rules that rely on the output gap.

In this paper, we calibrated the properties of measurement error to that of historical revisions to real-time estimates of the output gap and capacity utilization — an approach that strengthens the empirical relevancy of our results. However, in doing so we side-stepped the difficult issue of the causal links between historical revisions and the fundamental shocks to productivity and other factors that have hit the economy over the past 30 years: For tractability, we assumed that measurement error and fundamental shocks are uncorrelated in the simulations. An interesting (albeit ambitious) avenue for future research would be to model these links empirically, and to incorporate the complicated filtering exercise involved in such a problem into the stochastic simulations.

Perhaps the most important lesson to be drawn from our analysis, however, is simply that successful monetary policy design needs to be founded on realistic informational assumptions about what policymakers can and do know when policy decisions are made. Unless such realism

is incorporated into the analysis, the benefits from following activist stabilization policy may appear greater than what is likely to be achievable in practice.

## References

- Anderson, Gary (1999), "A Reliable and Computationally Efficient Algorithm for Imposing the Saddle-Point Property in Dynamic Models," working paper.
- Anderson, Gary and George Moore (1985), "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models," *Economics Letters*, vol. 17, 247-252.
- Brayton, Flint; and Peter Tinsley, eds., (1996), "A Guide to FRB/US: A Macroeconomic Model of the United States," Finance and Economic Discussion Series working paper no. 1996-42, Board of Governors of the Federal Reserve System.
- Brayton, Flint; Eileen Mauskopf, David Reifschneider, Peter Tinsley, and John Williams (1997), "The Role of Expectations in the FRB/US Macroeconomic Model," *Federal Reserve Bulletin*, vol. 83, April, pp. 227-245.
- Brayton, Flint, Andrew Levin, Ralph Tryon and John Williams (1997) "The Evolution of Macro Models at the Federal Reserve Board," *Carnegie-Rochester Conference Series on Public Policy*, vol. 47, pp. 43-81.
- Corrado, Carol and Joe Matthey, "Capacity Utilization," *Journal of Economic Perspectives*, vol 11, Number 1, Winter 1997, pp. 151-167.
- Estrella, Arturo and Frederic Mishkin (1999), "The Role of NAIRU in Monetary Policy: Implications of Uncertainty and Model Selection," in John B. Taylor, ed., *Monetary Policy Rules* (Chicago: University of Chicago Press.).
- Friedman, Milton (1947) "Lerner on the Economics of Control," *The Journal of Political Economy*, vol. 55, No. 5. (Oct., 1947), pp. 405-416.
- \_\_\_\_\_, (1953) "The Effects of Full-Employment Policy on Economic Stabilization: A Formal Analysis, in .Milton Friedman, *Essays in Positive Economics* (Chicago, Chicago University Press), pp. 117-132.
- Koenig, Evan F. (1996), "Capacity Utilization as a Real-Time Predictor of Manufacturing Output," Federal Reserve Bank of Dallas *Economic Review*, 3rd Quarter, pp. 16-23

- Kuttner, Kenneth N. (1994) "Estimate Potential Output as a Latent Variable," 12, *Journal of Business and Economic Statistics*, 3 (July), pp. 361-368.
- Kuttner, Kenneth N. (1992) "Monetary Policy with Uncertain Estimates of Potential Output," vol. 16, Federal Reserve Bank of Chicago *Economic Perspectives*, vol. 1 (January/February), pp. 2-15.
- Levin, Andrew; Volker Wieland, and John C. Williams (1999), "Robustness of Simple Monetary Policy Rules Under Model Uncertainty," in John B. Taylor, ed., *Monetary Policy Rules*, (Chicago: University of Chicago Press).
- McCallum, Bennett (1998) "Issues in the Design of Monetary Policy Rules," Working Paper.
- Orphanides, Athanasios (1997) "Monetary Policy Rules Based on Real-Time Data," Finance and Economics Discussion Series working paper no. 1998-03, Board of Governors of the Federal Reserve System (December).
- Orphanides, Athanasios (1998) "Monetary Policy Evaluation with Noisy Information," Finance and Economics Discussion Series working paper no. 1998-50, Board of Governors of the Federal Reserve System (October).
- Orphanides, Athanasios (1999) "The Quest for Prosperity without Inflation," unpublished manuscript, Division of Monetary Affairs, Board of Governors of the Federal Reserve System (May).
- Orphanides, Athanasios and Simon van Norden (1999) "The Reliability of Output Gap Estimates in Real Time," working paper, Board of Governors of the Federal Reserve System.
- Perry, George L. (1971) "Labor Force Structure, Potential Output and Productivity," *Brookings Papers on Economic Activity*, vol. 3, pp. 533-65.
- Reifschneider, David; Robert Tetlow and John Williams (1999), "Aggregate Disturbances, Monetary Policy, and the Macroeconomy: The FRB/US Perspective," *Federal Reserve Bulletin*, vol. 85, January, pp. 1-19.
- Rudebusch, Glenn D. (1999), "Is the Fed Too Timid? Monetary Policy in an Uncertain World," Federal Reserve Bank of San Francisco working paper no. 99-05 (March).
- Sack, Brian and Volker Wieland (1999) "Interest-Rate Smoothing and Optimal Monetary Policy: A Report on Recent Empirical Evidence," unpublished manuscript, Division of Monetary Affairs, Board of Governors of the Federal Reserve System (April).

Smets, Frank (1998) "Output Gap Uncertainty: does it matter for the Taylor rule?" Bank of International Settlements Working Paper, No. 60 (November).

Staiger, Douglas, James H. Stock and Mark W. Watson (1997) "The NAIRU, Unemployment and Monetary Policy," vol. 11, *Journal of Economic Perspectives*, 1 (Winter), pp. 33-50.

Taylor, John B. (1993) "Discretion versus Policy Rules in Practice," vol. 39, *Carnegie-Rochester Conference Series on Public Policy*, 2 pp. 195-214.

Taylor, John B. (1998) "An Historical Analysis of Monetary Policy Rules," in John B. Taylor (ed.), *Monetary Policy Rules*, Chicago: University of Chicago Press (forthcoming).

Tetlow, Robert; and Peter von zur Muehlen (1999), "Simplicity Versus Optimality: The Choice of Monetary Policy Rules When Agents Must Learn," Finance and Economics Discussion Series working paper no. 1999-10, Board of Governors of the Federal Reserve System (January).

Williams, John (1999) "Simple Rules for Monetary Policy," Finance and Economics Discussion Series working paper no. 99-12, Board of Governors of the Federal Reserve System (February).

Woodford, Michael (1999) "Optimal Monetary Policy Inertia" unpublished manuscript, Princeton University, January 28, 1999.