ELSEVIER Journal of Monetary Economics 34 (1994) 89–99

JOURNAL OF Monetary ECONOMICS

Money may matter, but how could you know?

Kevin D. Hoover*, Stephen J. Perez

Department of Economics, University of California, Davis, CA 95616-8578, USA

(Received February 1993, final version received March 1994)

Abstract

Christina and David Romers' reply to our article '*Post Hoc Ergo Propter Hoc* Once More' misses the point. Our argument was never that monetary policy did not matter, but that their methods could not provide useful evidence that it did. Yet, they offer additional evidence of the same type with respect to the efficacy of monetary shocks without effectively replying to the criticisms of their methods. We show point by point that such responses as they give leave our original conclusion intact: their narrative/statistical approach is a complicated version of the fallacy *post hoc ergo propter hoc*; and, as such, will not sustain inferences with respect to the direction and strength of the causes of output fluctuations.

Key words: Monetary policy; Economic methodology; Business fluctuations

JEL classification: B41; C19; E32; E52

1. Introduction

In their reply to our article '*Post Hoc Ergo Propter Hoc* Once More' Christina and David Romer (1993) mischaracterize our argument in a number of ways. Their mischaracterization begins with their title, 'Monetary Policy Matters', for we never said that it did not matter. Instead, we argued that the Romers'

*Corresponding author.

We appreciate the detailed comments of Thomas Mayer, James Hartley, Jane Millar, Kevin Salyer, Steven Sheffrin, and the participants in the U.C. Davis Macroeconomics Brownbag Workshop.

SSDI 0304393294011 51 U

methods cannot provide logically compelling evidence that it does. The Romers' reply does nothing to diminish the force or character of our original arguments.

The bulk of the Romers' reply is an extension of the results in their original paper to post-1987 data and an attempt to demonstrate the robustness of those results to the possibility that oil shocks are also important in causing recessions. They present us as offering oil shocks as an alternative to monetary shocks as the cause of recessions. But again they mischaracterize our argument. We repeatedly stated that we did not conclude that oil shocks cause recessions and that monetary policy does not; but rather that the Romers' methods were incapable of discriminating between different possible causes. Comparison of Tables 1 and 2 in their reply with Table 2 in our article clearly shows that the Romers' additional results (including their several variations on measuring oil and monetary shocks) are no different in character from those that we already reported. (Indeed, in one respect they are less informative, because, unlike us, they use oil shocks only as control variables, and do not present impulse responses for them.)

The Romers do not defend their methods. Rather they use those methods to generate more of the same sort of evidence. The addition of more unsound evidence to an unsound argument cannot make the argument sound. Our arguments, however, were, and are, directed at their methods. Our claim is that the Romers commit the fallacy *post hoc ergo propter hoc*. The Romers take us to task for referring to their methodology as new, reiterating their view that it is venerable and commonplace. And they are right; it is not new. The fallacy *post hoc ergo propter hoc* is extremely old and frequently encountered. What is new is to see it explicitly defended.

The Romers give much less attention to those of our arguments unrelated to oil shocks. Since the issues they raise with respect to these other arguments are more relevant and more profound, let us consider them before returning to the matter of oil shocks.

2. Is all the world AR(24)?

The Romers dismiss the results of our simulation study of their methods as '... having little bearing on [their] conclusions'. There are two issues that need to be separated: first, the question whether a univariate autoregressive model with twenty-four lags can control for the natural dynamics of an economic system; second, the importance of nonlinearities in achieving our results.

The Romers say nothing about the second issue, which is the methodologically more important point of our simulation study: the demonstration that their univariate methods cannot control for other relevant sources of fluctuations. Nonlinearities have nothing to do with this point at all.

Eqs. (1) and (2) of '*Post Hoc Ergo Propter Hoc*' imply that Y can be expressed as an ARMA(3, 2). The Romers' methodology could work only if a univariate AR with long enough, but finite lags, were always adequate to represent such a process. But this is wrong. What drives our results is that Z is cyclical and Granger-causes Y, but is not Granger-caused by Y. Passing familiarity with the analysis of Granger-causality (Granger, 1969; Sims, 1972; Sargent, 1979, pp. 279–290; Engle et al., 1983) shows that reducing a bivariate representation to a univariate ARMA will not remove the incremental predictive power of Z for Y, which is the mechanism that triggers the dummies.[1] Even without formal analysis, this is easily seen by running Granger-causality tests on the residuals of our simulated ΔY from an ARMA(3, 2) filter: In practice, Z continues to Granger-cause Y at high levels of statistical significance. Furthermore, it is easy to see that, if any bivariate system could be adequately represented by an AR process such as the AR(24) that the Romers use, then nothing would Granger-cause anything else; for, if an AR model with a large finite number of lags could remove the contribution of all other variables to incremental predictability, Granger-causality tests would always accept the null of no Granger-causality except in bad draws. This we regard as an argument *reductio ad absurdum* against the Romers' conclusion.

To check against the possibility that our results are a bad draw, we re-simulated the model 500 times. On average this produced 3.3 identified monetary shocks. The average impulse response to a unit monetary shock was a 9.6 percent decline in Y at the maximum and a 5.9 percent decline at thirty-six months. This from a model in which monetary policy has no causal efficacy whatsoever.[2]

The Romers admit that their results could fail in the presence of nonlinearities; but, they claim, their selection method for monetary shocks avoids the

[1] It is of course true that the dummies are a deterministic function of Y. Were they a linear function of Y that would imply that there was no well-defined Granger-causal relation between them and Y at all; for Granger-causality is a fundamentally indeterministic notion (Granger, 1969, p. 378). But of course they are nonlinear functions of Y. The fact then that Z Granger-causes Y means that past history of Y itself cannot fully predict the current behavior of Y even up to a white-noise error. Precisely, the times when Y is unusually high, and therefore Π is unusually high, are the times when Z is cyclically high. The dummies thus are triggered when Z is high and ready to turn down; and, since Y does not Granger-cause Z, the past history of Y cannot control for behavior of Z nor for the behavior in the dummies that it triggers. The general point that the Romers' methods do not support the collapsing of a multivariate system to a univariate one is made somewhat differently in Leeper (1993).

[2] It turns out on these results, that the simulation reported in our original article is an unusually dramatic instance. Clearly, however, an average simulation would have been just as devastating to the Romer's conclusions. And, in any case, their asseveration to the contrary notwithstanding, when the question is one of the logic of their method, a single counterexample illustrates its ineffectiveness as well as 500.

problem. We deal with the selection of shocks in Section 4 below. It is important to observe that we need nonlinearities for one purpose only: to mimic the Romers' pattern of Federal Reserve reactions, there must sometimes be episodes of higher than normal inflation associated with cyclically high output. The only relevant nonlinearity in our model is the curvature of the aggregate-supply curve near full employment; see Eq. (3) in our original article.

The Romers' footnote 13 is misleading on this point. That Y is set to 22,999 whenever Eqs. (1) and (2) imply that it exceeds 23,000 is a *minor technical point.* It merely prevents prices from becoming negative when aggregate demand crosses the asymptote implied by Eq. (3). In the simulation in our original paper, this situation did not arise; it, therefore, cannot explain our results. It does sometime arise in repeated simulations; but it is *not* the source of our conclusions; it is merely a simple way to capture an upper limit to Keynesian supply responses. Nevertheless, it is only one possible way of generating data to make our general point. We have also simulated a model with the same aggregate supply curve and a downward-sloping aggregate demand curve in which the solution for Y is bounded above by 23,000, and in which, therefore, this issue does not arise. The results of that model are precisely of the same character as those reported in our paper. In any case, this strikes us an odd point for the Romers to raise. We would have thought that the existence of an aggregate-supply curve that becomes more steep as full employment is approached was folk wisdom, at least among those who believe that the Federal Reserve uses recessions as an instrument to control inflation.

The Romers hold Sichel (1992) up as evidence of the nonexistence of relevant nonlinearities. But this is misleading. Sichel (1992) discusses only the time shape of cyclical fluctuations in GNP, and says nothing about prices, which is the relevant point. Were we to replace the processes that generate Y [Eqs. (1) and (2) in our original article] with ones that better mimic Sichel's findings, while retaining the nonlinear price relation, we would still be able to mimic the Romers' description of Federal Reserve behavior.

The Romers argue that changes in expected inflation and supply shocks could imply accelerating prices without the nonlinear aggregate supply curve. They appear to miss the logic of the argument in a way that undercuts their own position. We have formulated a model in which accelerating inflation signals a real downturn and have shown that their methods fail in such a model. They suppose their methods, however, to be generally successful. It is sufficient to show that they do not work in this case to show that they are not general. It *reinforces* our point – it does not weaken it – to show that supply shocks or expectational shifts would also cause accelerating prices and downturns. We have no doubt that a similar demonstration of the inefficacy of their methods could be given in such cases as well. But it would just be icing on the cake; the general point is already established. It is especially odd for the Romers to mention real shocks in this context; for, as discussed in Section 5 below, it is precisely these shocks that their narrative method attempts to exclude. Thus, in attempting to impugn our illustrative model, they raise questions about their own method.

3. Dynamic forecast errors

The Romers concede the analytical conclusions of our discussion of dynamic forecasting: that dynamic forecasts are simply a combination of the information of the ordinary forecast errors and that the estimators impose structure on the errors unrelated to economic behavior. The reader may judge whether this impairs their usefulness or not. We are content to accept the Romers' stipulation that their quantification of the effects of monetary policy shifts does not rely on the use of dynamic forecasts. For we believe that the modesty of this conclusion would naturally escape anyone who had read the eight pages and twelve figures in their original article (Romer and Romer, 1989) devoted to the quantification of dynamic forecasts.

4. Confusions about the logic of causality

The Romers state that our claim that causal inference requires regime changes '.... reflects a basic misunderstanding'. First, randomness (e.g., a randomized experiment) is alone enough, they say, to permit causal inference. Second, '... if policy is responding to a variable that it is reasonable to think a priori does not directly affect real output, this portion of changes in policy can also be used to determine the effects of policy'. In either case, they argue regime changes are dispensable. The misunderstanding is the Romers' own, and rests on an equivocation between randomized experiments and residual random error.

Let us define a description of the world as consisting of three elements: parameters, which may be selected mutually independently by economic agents or nature; random shocks, which are white noise innovations; and variables, which are the product of the interaction of parameters and random shocks, and which are at best indirectly controllable through control over parameters. Any complete description of the world can be put into this canonical form through a suitable transformation.

Now consider two data-generating processes:

Process I

$$\Pi = a \rightarrow \varepsilon, \tag{1}$$

$$M = b\Pi + v, \tag{2}$$

$$Y = cM + \eta; \tag{3}$$

Process II

$$Y = d + \omega, \tag{4}$$

$$\Pi = eY + \zeta, \tag{5}$$

$$M = f\Pi + \delta. \tag{6}$$

One may think of Π as inflation, M as monetary policy, and Y as real output; lower case Roman letters are parameters; lower case Greek letters are random shocks (white noise innovations). Using the recursiveness criterion of Simon (1953), in process I inflation causes monetary policy which causes real output, while in process II real output causes inflation which causes monetary policy. Without knowledge of the parameters one cannot distinguish whether either of these causal orders or some other order is correct: The two processes have precisely the same reduced forms. This is the famous problem of observational equivalence (cf. Basmann, 1965, 1988).[3] If the direction of causal order is not known in advance, randomness in the forms of the white noise innovations on each of the equations is of no help in establishing causal order. Regime changes, however, might be. To take a simple example, let the innovations all be normally distributed. Then the marginal probability densities of output for each process (the density estimated by the regression corresponding to the common reduced form of the two processes) may be given for process I as

$$D(Y) = \mathcal{N}(cba, c^2b^2\sigma_\varepsilon^2 + c^2\sigma_v^2 + \sigma_\eta^2), \tag{7}$$

and for process II as

$$D(Y) = \mathcal{N}(d, \sigma_\omega^2). \tag{8}$$

Both densities can be estimated by the same reduced form:

$$Y = A + \Omega, \tag{9}$$

where A is a constant coefficient and Ω is a residual.[4] Now suppose that we know from historical narrative or other independent information that there is a regime change in the monetary policy process. If process I were the true ordering, then such a regime change would be a change in either b or the moments of v. In that case, Eq. (7) shows that Eq. (9) would show a structural break. If process II were the true ordering, such a regime change would be a change in either f or the moments of δ. Eq. (8) shows that Eq. (9) would not

[3] The example here is simple, involving only contemporaneously related variables, but the problem generalizes, for example, to vector autoregressions (cf. Cooley and LeRoy, 1985).

[4] The coefficient A must not be confused with the parameters in processes I and II; similarly Ω must not be confused with the random shocks. A and Ω are generally functions of the underlying parameters and random shocks.

show a structural break. Thus, the stability or instability of the estimated probability distribution in the face of regime changes of known types provides evidence bearing on causal direction. The example given here is one of many relevant stability implications that bear upon causal direction; see Hoover (1990a, 1991, 1993) for a more detailed treatment. Simple as this example is, however, it serves to demonstrate the sense in which a regime change may be necessary to infer causal direction.

The Romers' argument gains its plausibility from the intended analogy between their own work and randomized experiments in laboratory conditions. Randomization in that context is a selection rule for a parameter rather than for a shock. For example, if monetary policy in process I above is randomized, the parameter b in Eq. (2) might be chosen according to a random rule. Contrary to the Romers' implication, one is not then dispensing with regime changes; rather one is choosing them in a particular way. There is no analogy between this process and anything in the Romers' method.

Although they incorrectly deny the need for regime changes, one might think that the Romers' method of selecting monetary shocks could be interpreted as isolating instances of parameter changes in the Federal Reserve's reaction function. Unfortunately the evidence that the Romers deploy is irrelevant to that question. In their reply, they say that they look for '... a variable that it is reasonable to think a priori does not directly affect real output'. They argue that inflation is such a variable, that '.... it is not plausible that inflation has large effects on real output except through its impact on policy' (cf. Romer and Romer, 1989, p. 134). To see that this is irrelevant, consider Fig. 1, a slightly generalized version of the Romers' understanding of the economy, showing the causal ordering of the variables described above. The Romers provide an *a priori* argument that causal link (ii), 'Π causes Y', is weak to nonexistent. So what? That is not a point of dispute; for all it would imply were it a powerful channel of causal influence is that real output had multiple causes (in this case, a direct and an indirect channel for inflation). Despite their claim that their method can isolate the direction of causation (Romer and Romer, 1989, pp. 121, 122) the Romers provide no evidence on links (iii), 'M causes Y', and (iv), 'Y causes Π', which bear on the direction of causation. With respect to link (iii), 'M causes Y', they beg the question. They simply assume that causality runs from policy to output and use their statistics to attempt to measure the strength of the effect.[5] Even more important, they provide no evidence with respect to link (iv), 'Y causes Π', On the one hand, they must suppose that link (iv) is operative, otherwise there would be no point in the Federal Reserve's attempting to control inflation through induced recessions. On the other hand, an operative

[5] See Hoover (1990, 1994) for discussions of the issues of causal direction vs. causal strength.

link (iv) implies that inflation is endogenous and that a Federal Reserve response to inflation is also a response to output. That this renders the Romers' methods useless with respect to controllability and causal direction is the message of our simulation model. All of the assertions (or for that matter empirical evidence) with respect to link (ii), 'Π causes Y', have no bearing on link (iv), 'Y causes Π'. And on link (iv), the Romers offer nothing.

5. Monetary dummies vs. oil dummies

The Romers object to our representing oil shocks by fixed dummies. They notice that the price changes that we use to localize those shocks (see Fig. 1 in our original article) are of very different sizes. And no theory that they know would suggest that such disparate price changes would have similar effects. We must recapitulate our reasons for using oil dummies. First, we sought to mimic the Romers' procedures exactly to demonstrate their lack of discrimination. Oil was not central to our goal; but Hamilton's (1983) analysis of oil was convenient. It provided a seriously intended alternative story to monetary shocks and a set of dates independent of our subjective choice. Hamilton (1983, p. 230) argues that it is quantity shocks in oil, not price shocks, that are the proximate causes of recessions. The sharp, transient spikes in oil prices (see Fig. 1 in our original article) are *indicators* of these quantity effects induced by state regulatory structures (e.g., by the behavior of the Texas Railroad Commision and similar state agencies). The distinction between shocks and propagation mechanisms is relevant here. Hamilton's argument is only that oil shocks trigger recessions, and not that there are no other propagation mechanisms that magnify and prolong the initial cause.

There is, of course, a more general question: Are the Romers correct to insist that causes (i.e., shocks) must be proportional to their ultimate effects? An exploding gasoline tanker might destroy a house; Mrs. O'Leary's cow kicking over a lantern destroyed 17,000 houses. A tug of a finger on a trigger may do nothing, or may kill a man, or may, if it kills an Austrian archduke, redirect the history of a century.

The Romers protest that they never maintained that monetary shocks were the only cause or even the largest cause of recessions. While it is true (and we refer to their robustness tests) that the Romers do not insist on a monocausal explanation of recessions in general, they do claim to look for particular episodes in which the economy would not have turned down anyway without the monetary shock. Their narrative/statistical method seems to require that they isolate some instances of monetary monocausality. Otherwise, it is unclear how they can be sure that – by mistake or design – the Federal Reserve is not simply reacting to a nonmonetary cause. Hamilton's oil shocks and the Romers' monetary shocks generally coincide; so despite the Romers' assertions to the contrary, how do we know that the Federal Reserve is not clamping down in response to the oil price shock which leads the ultimate quantity effect on output?[6] Wanting to have it both ways again, the Romers criticize our simulation model (see fn. 12 in their reply; cf. Section 2 above) because it does not account for the possibility of shocks to supply or to expectations of inflation. Yet such shocks are recession-inducing, so that episodes which involve them (ones in which the economy is set to turn down independently of monetary policy) are precisely the ones the Romers' own narrative methods are supposed to avoid. The attack on our argument is thus made at the expense of the coherence of their own argument.

In a shift of ground, the Romers argue that the issue is not monetary policy shocks versus oil shocks, but the incremental contribution of monetary shock data to the predictability of output conditional on other variables including oil shocks. This is a shift of ground both because it relinquishes the assertion of monocausality in the particular episodes that they identify and because it changes the operative definition of causality. An appropriate measure of incremental predictability if the Granger-causality test. The Romers claim that we offer no evidence on the question of incremental predictability. This is false. The F-tests in Table 4 in our original article are almost equivalent to Granger-causality tests. But, in any case, the question that is relevant is whether Federal Reserve actions induce recessions: a question of control. Granger-causality is irrelevant to a notion of causality as controllability (Granger, 1980, p. 1; Hoover, 1988, pp. 173–174).

With respect to oil shocks, what the Romers have done is to check the robustness of their regressions to alternative specifications. What we did was to check the robustness of their methods to alternative specifications. Again, let us make it plain, we do not (and never did) claim that oil shocks are the cause of recessions. The Romers' evidence is thus deployed against a point that we do not maintain. We demonstrated that had the Romers set out to prove the efficacy of oil shocks instead of monetary shocks, they would have suceeded in their own eyes. Their methods simply do not discriminate.

6. Post hoc ergo propter hoc still

The Romers' reply mostly adds evidence irrelevant to the methodological issue at hand. The Romers hardly answer the objections that we raised in our original paper. Such answers as they give do not succeed in undermining the force of our objections.

The Romers conclude their reply by characterizing us as maintaining that we believe that constructing any counterexample in which their methods fail or

[6] James Hamilton made a similar observation in private correspondence with us.

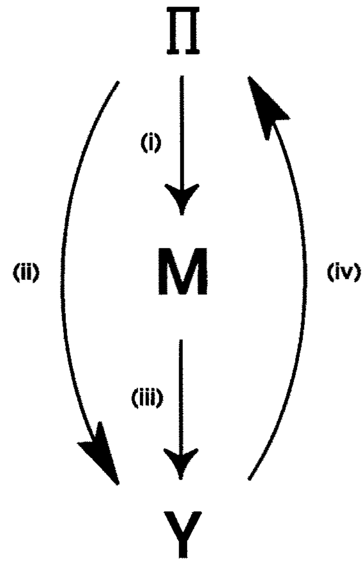

Fig. 1. A causal ordering of inflation (Π), monetary policy (M), and real output (Y).

finding any variable the effects of which cannot be distinguished from a monetary shock imply that their methods are entirely uninformative. To this we plead guilty. If the Romers' methods fail in the transparent, ideal test case offered by our simulation, why should we trust them in the real world? If they cannot discriminate between monetary shocks and a seriously offered alternative explanation, why should we take them as useful evidence for monetary shocks?

In their reply the Romers have put their methods to further use, but they have not offered any evidence that they in fact work; while we have offered substantial evidence that they do not. We must then conclude this response in much the same manner as we concluded our original article. The substance of the Romers' original article is in their Fig. 1, which plots industrial production and the unemployment rate against the Romers' shock dates, showing that economic downturns follow those dates. The rest of their paper is a sophisticated repackaging of that figure; it adds no substance to a transparently fallacious argument. Money and monetary policy may matter to the real economy, and we may have many reasons for thinking so, but the reasons offered by the Romers are not good ones.

References

Basmann, R.L., 1965, A note on the testability of 'explicit causal chains' against the class of 'interdependent' models, Journal of the American Statistical Association 60, 1080–1093.

Basmann, R.L., 1988, Causality tests and observationally equivalent representations of econometric models, Journal of Econometrics 39, 69–101.

Cooley, T.F. and S.F. LeRoy, 1985, Atheoretical macroeconometrics: A critique, Journal of Monetary Economics 16, 283–308.

Engle, R.E., D.F. Hendry, and J.-F. Richard, 1983, Exogeneity, Econometrica 51, 113–127.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, reprinted in: R.E. Lucas, Jr. and T.J. Sargent, eds., Rational expectations and econometric practice (George Allen and Unwin, London, 1981) 371–386.

Granger, C.W.J., 1980, Testing for causality: A personal vewpoint, Journal of Economic Dynamics and Control 2, 329–352.

Hamilton, J.D., 1983, Oil and the macroeconomy since World War II, Journal of Political Economy 91, 228–248.

Hoover, K.D., 1988, The new classical macroeconomics: A sceptical inquiry (Blackwell, Oxford).

Hoover, K.D., 1990a, The logic of causal inference: Econometrics and the conditional analysis of causation, Economics and Philosophy 6, 207–234.

Hoover, K.D., 1990b, Review of Nancy Cartwright's *Nature's capacities and their measurement*, Economics and Philosophy 6, 309–315.

Hoover, K.D., 1991, The causal direction between money and prices: An alternative approach, Journal of Monetary Economics 27, 381–423.

Hoover, K.D., 1993, Causality and temporal order in macroeconomics or why even economists don't know how to get causes from probabilities, British Journal for the Philosophy of Science 44, 693–710.

Hoover, K.D., 1994, Causation, estimation and statistics: Comments on Cartwright and Woodward, in: Daniel Little, ed., On the reliability of economic models: Essays in the philosophy of economics, forthcoming.

Leeper, E.M., 1993, Has the Romers' narrative approach identified monetary policy shocks?, Unpublished typescript, Feb. (Federal Reserve Band of Atlanta, GA).

Romer, C.D. and D.H. Romer, 1989, Does monetary policy matter? A new test in the spirit of Friedman and Schwartz, in: Olivier Blanchard and Stanley Fischer, eds., NBER macroeconomics annual, Vol. 4 (MIT Press, Cambridge, MA).

Romer, C.D. and D.H. Romer, 1994, Monetary policy matters, Journal of Monetary Economics, this issue.

Sargent, T.J., 1979, Macroeconomic theory (Academic Press, Boston, MA).

Sichel, D.E., 1992, Inventories and the three phases of the business cycle, Working paper no. 128, revised Aug. (Economic Activity Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System, Washington, DC).

Simon, H.A., 1953, Causal ordering and identifiability, in: Herbert A. Simon, Models of man, Ch. 1 (Wiley, New York, 1957).

Sims, C.A., 1972, Money, income and causality, reprinted in: R.E. Lucas, Jr. and T.J. Sargent, eds., Rational expectations and econometric practice (George Allen and Unwin, London, 1981) 387–403.