# Nonlinearities and the Macroeconomic Effects of Oil Prices*

James D. Hamilton

jhamilton@ucsd.edu

Department of Economics

University of California, San Diego

December 9, 2009

Revised: November 15, 2010

ABSTRACT

This paper reviews some of the literature on the macroeconomic effects of oil price shocks with a particular focus on possible nonlinearities in the relation and recent new results obtained by Kilian and Vigfusson (2009).

*I thank Rob Vigfusson for graciously sharing his data and helping me follow his code. An earlier version of this paper was circulated under the title, "Yes, the Response of the U.S. Economy to Energy Prices is Nonlinear."

# 1 Overview.

I noted in a paper published in the *Journal of Political Economy* in 1983 that at that time, 7 out of the 8 postwar U.S. recessions had been preceded by a sharp increase in the price of crude petroleum. Iraq's invasion of Kuwait in August 1990 led to a doubling in the price of oil in the fall of 1990 and was followed by the ninth postwar recession in 1990-91. The price of oil more than doubled again in 1999-2000, with the tenth postwar recession coming in 2001. Yet another doubling in the price of oil in 2007-2008 accompanied the beginning of recession number 11, the most recent and frightening of the postwar economic downturns. So the count today stands at 10 out of 11, the sole exception being the mild recession of 1960-61 for which there was no preceding rise in oil prices.

Oil shocks could affect the economy through their consequences for both supply and demand. On the supply side, consider a firm whose output $Y$ depends on inputs of capital $K$, labor $N$, and energy $E$:

$$Y = F(K, N, E).$$

Suppose that the capital stock is fixed in the short run and that wages adjust instantly to ensure that labor demand equals a fixed supply $\overline{N}$. Then if $X$ denotes the price of energy relative to the price of output,

$$\frac{\partial Y}{\partial X} = \frac{\partial F}{\partial E}\frac{\partial E}{\partial X}. \tag{1}$$

Multiplying (1) by $X/Y$ results in

$$\frac{\partial \ln Y}{\partial \ln X} = \frac{\partial F}{\partial E}\frac{E}{Y}\frac{\partial \ln E}{\partial \ln X}. \tag{2}$$

If the marginal product of energy equals its relative price ($\partial F/\partial E = X$), then the first terms on the right side of (2) will be recognized as the energy expenditure share

$$\frac{\partial F}{\partial E}\frac{E}{Y} = \frac{EX}{Y} = \gamma$$

where $\gamma$ denotes the firm's spending on energy relative to the value of its total output. Letting lower-case letters denote natural logarithms, (2) can be written

$$\frac{\partial y}{\partial x} = \gamma \frac{\partial e}{\partial x}. \tag{3}$$

In other words, the elasticity of output with respect to the relative price of energy would be the energy expenditure share $\gamma$ times the price-elasticity of energy demand.

The energy expenditure share is a small number. In 2009, the U.S. consumed about 7.1 billion barrels of petroleum products, which at the current $80/barrel price of crude corresponds to a value around $570 billion. This would represent only 4% of U.S. GDP. Moreover, the short-run price-elasticity of petroleum demand is extremely small (Dahl, 1993), so that expression (3) implies an output response substantially below 4%. For this reason, models built around this kind of mechanism, such as Kim and Loungani (1992), imply that oil shocks could only have made a small contribution to historical downturns. Note also that (3) implies a linear relation between $y$ and $x$; an oil price decrease should increase output by exactly the same amount that an oil price increase of the same magnitude would decrease output.

To account for larger effects, it would have to be the case that either $K$ or $N$ also adjust in response to the oil price shock. Finn (2000) analyzed the multiplier effects that result if firms adjust capital utilization rates in order to minimize depreciation expenses. Leduc and Sill (2004) incorporated this utilization effect along with labor adjustments resulting from sticky wages. Again these models imply a linear relation between $y$ and $x$, though Atkeson and Kehoe's (1999) treatment of putty-clay investment technology produces some nonlinear effects.

Davis (1987a, 1987b) stressed the role of specialized labor and capital in the transmission mechanism. If the marginal product of labor falls in a particular sector, it can take time before workers relocate to something more productive, during which transition the economy will have some unemployed resources. Moreover, these effects are clearly nonlinear. For example when energy prices fell in 1985, some workers in the oil-producing sector were forced to find other jobs. As a result, it is possible in principle for aggregate output to fall temporarily in response to an oil price decrease just as it does for an oil price increase

Although many discussions (e.g., Kilian and Vigfusson, 2010) treat this relocation of workers as the sole source of asymmetry introduced by allocative disturbances, my 1988 paper demonstrated that unemployment could result not just from workers who are in transition between sectors but also from workers who are simply waiting until conditions in their sector once again improve. In such models, idle labor and capital rather than decreased energy use as in (1) account for the lost output.

An alternative mechanism operates through the demand side. An increase in energy prices leaves consumers with less money to spend on non-energy items and leaves an oil-importing country with less income overall. If a consumer tries to purchase the same quantity of energy $E$ in response to an increase in the relative price given by $\Delta X$, then saving or expenditures on other items must fall by $E \cdot \Delta X$, with a proportionate effect on demand given by

$$\frac{\Delta Y^d}{Y} = \frac{E}{Y} \Delta X.$$

If as in a fixed-price Keynesian model demand $Y^d$ is the limiting determinant of total output, we would have

$$\frac{\partial y}{\partial x} = \gamma,$$

so that by this mechanism the effect once again is linear and bounded[1] by the expenditure share $\gamma$.

Specialization of labor and capital could also be important for the transmission of demand effects as well. Demand for less fuel-efficient cars would be influenced not just by the consequences of an oil price increase for current disposable income but also by consideration of future gasoline prices over the lifetime of the car. Bernanke (1983) noted that uncertainty per se could lead to a postponement of purchases for capital and durable goods. A shift in demand away from larger cars seems to have been a key feature of the macroeconomic response to historical oil shocks (Bresnahan and Ramey, 1993; Edelstein and Kilian, 2009; Hamilton, 2009; Ramey and Vine, 2010), and Bresnahan and Ramey (1993) and Ramey and Vine (2010) map out in detail exactly how specialization of labor and capital in the U.S. automobile industry amplified the effects of historical oil price shocks and introduced nonlinearities of the sort anticipated by the sectoral-shifts hypothesis. In the model of Hamilton

[1] Price adjustment would make this effect smaller whereas the traditional Keynesian multiplier could make it bigger.

(1988), shifts in the advantages between sectors resulting from supply effects (greater production costs for sector 1 as a result of higher energy prices) or demand effects (less demand for the output of sector 1 as a result of higher energy prices) have identical macroeconomic consequences, operating in either case through idled labor in the disadvantaged sector.

In terms of empirical evidence on nonlinearity, Loungani (1986) demonstrated that oil-induced sectoral imbalances contributed to fluctuations in U.S. unemployment rates. Mork (1989) found that oil price increases have different predictive implications for subsequent U.S. GDP growth than oil price decreases. Other studies also reporting evidence that nonlinear forecasting equations do better include Lee, Ni and Ratti (1995), Balke, Brown, and Yücel (2002), and Hamilton (1996, 2003). Both Carlton (2010) and Ravazzolo and Rothman (2010) confirmed these predictive improvements using real-time data. Ferderer (1996) and Elder and Serletis (2010) demonstrated that oil-price volatility predicts slower GDP growth, implying that oil price decreases include some contractionary implications. Davis and Haltiwanger (2001) found nonlinearities in the effects of oil prices on employment at the individual plant level for U.S. data. Herrera, Lagalo, and Wada (2010) found a strong nonlinear response of U.S. industrial production to oil prices, with the biggest effects in industries the use of whose products by consumers is energy intensive. A nonlinear relation between oil prices and subsequent real GDP growth has also been reported for a number of OECD countries by Cuñado and Pérez de Gracia (2003), Jiménez-Rodrígueza and Sánchez (2005), Kim (2009), and Engemann, Kliesen and Owyang (2010).

By contrast, a prominent recent study by Kilian and Vigfusson (2009) found little evidence of nonlinearity in the relation between oil prices and U.S. GDP growth. In the next section I explore why they seem to have reached a different conclusion from many of the previous researchers mentioned above. Sections 3 and 4 note some of the further implications of their results for inference about nonlinear dynamic relations.

## 2 Testing for nonlinearity.

Let $y_t$ denote the rate of growth of real GDP, $x_t$ the change in the price of oil, and $\tilde{x}_t$ a proposed known nonlinear function of oil prices. The null hypothesis that the optimal one-period-ahead forecast of $y_t$ is linear in past values of $x_{t-i}$ is quite straightforward to state and test: we just use OLS to estimate the forecasting regression

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{p} \beta_i x_{t-i} + \sum_{i=1}^{p} \gamma_i \tilde{x}_{t-i} + \varepsilon_t \tag{4}$$

and test whether $\gamma_1 = \cdots = \gamma_p = 0$. As noted above, a large number of papers have tested such a hypothesis and rejected it. Kilian and Vigfusson's paper might leave the impression that these earlier tests were somehow misspecified or insufficiently powerful, and that the reason Kilian and Vigfusson reach a different conclusion from previous researchers is that they are proposing superior tests. Such a result would be surprising if true. For Gaussian $\varepsilon_t$ in (4), OLS produces maximum likelihood estimates which are asymptotically efficient, and the OLS $F$ test is the likelihood ratio test with well-known desirable properties. That some new test could be more powerful than the standard OLS test seems unlikely, and certainly if the OLS test rejects and the new test does not, the reconciliation cannot be based on the assertion that the new test is more powerful. Kilian and Vigfusson also include in their analysis some standard OLS tests, which offer further support for their conclusion that the relation appears to be linear. But insofar as these are the same OLS tests that have already produced rejections of the null hypothesis in previous studies, the difference in conclusions must come from a different data set or differences in the specification of the basic forecasting regression (4), and not from any superior properties of the new tests proposed in their paper.

Most of their paper explores the case in which $\tilde{x}_t$ is given by $x_t^+ = \max\{0, x_t\}$, the alternative hypothesis of interest taken to be that oil price increases have different economic effects from oil price decreases. This particular specification is one that previous researchers have found to be unstable over earlier data sets (e.g., Hooker, 1996; Hamilton, 2003), so it is unsurprising that Kilian and Vigfusson find that such a relation does not perform well on their sample either. My earlier investigation (Hamilton, 2003) concluded that the nonlinearities can be captured with a specification in which what matters is whether oil prices make a new 3-year high:

$$x_t^{\#} = \max\{0, X_t - \max\{X_{t-1}, ..., X_{t-12}\}\}$$

for $X_t$ the log level of the oil price. Below I reproduce the coefficients as reported in equation (3.8) of Hamilton (2003):

$$\begin{aligned} y_t \;=\;& \underset{(0.13)}{0.98} + \underset{(0.07)}{0.22} y_{t-1} + \underset{(0.07)}{0.10} y_{t-2} - \underset{(0.07)}{0.08} y_{t-3} - \underset{(0.07)}{0.15} y_{t-4} \\ & - \underset{(0.014)}{0.024} x_{t-1}^{\#} - \underset{(0.014)}{0.021} x_{t-2}^{\#} - \underset{(0.014)}{0.018} x_{t-3}^{\#} - \underset{(0.014)}{0.042} x_{t-4}^{\#}. \end{aligned} \tag{5}$$

If one adds the linear terms $\{x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}\}$ to this regression and calculates the OLS $\chi^2$ test of the hypothesis that the coefficients on $\{x_{t-1}^{\#}, x_{t-2}^{\#}, x_{t-3}^{\#}, x_{t-4}^{\#}\}$ are zero using the original data set, the result is a $\chi^2(4)$ statistic of 16.93, with a $p$-value of 0.002. The last entry of Kilian and Vigfusson's Table 4 reports the OLS $\chi^2$ test on a similar specification for their data set which results in a $p$-value of 0.046. Clearly it must be differences in the specification and data set between the two papers, rather than differences in the testing methodology, that accounts for the different findings. There are a number of differences that could explain the higher $p$-value obtained by Kilian and Vigfusson.

**Different data sets.** In my earlier analysis, $t$ in (5) ran from 1949:Q2 to 2001:Q3 (or 210 total observations), whereas in Kilian and Vigfusson's analysis, $t$ runs from 1974:Q4 to 2007:Q4 (or 133 total observations). One would expect to see a higher $p$-value in a shorter sample, since fewer observations make it harder to reject any hypothesis. In addition, it is possible that there has been a structural change since 2001, so that the earlier proposed nonlinear relation (5) does a poorer job with more recent data.

**Different measure of oil prices.** In my original analysis, $x_t$ was based on the producer price index for crude petroleum, whereas Kilian and Vigfusson use the refiner acquisition cost for imported oil. The values of these two measures are compared in the top two panels of Figure 1. The RAC is not available prior to 1974, and Kilian and Vigfusson imputed values back to 1971. The two oil price measures are very similar after 1983, but are somewhat different in the 1970s. Most notably, according to RAC, the first oil shock of 1974:Q1 was three times the size of that seen in any other quarter of the 1970s, and there was very little change in oil prices in 1981:Q1. By contrast, the PPI registers the shocks of 1974:Q1, 1979:Q2-Q3, and 1981:Q1 as similar events. If one thought that a key factor in the transmission mechanism to the U.S. economy involved the price consumers paid for gasoline, the PPI may provide a better measure, since the CPI also represents these three shocks as having similar magnitude (see the bottom panel of Figure 1). In any case, it is certainly possible that for such different measures of oil prices, the functional form of the optimal forecast could differ.

**Different price adjustment.** Another difference is that (5) used for $x_t$ the nominal change in the price of oil, whereas Kilian and Vigfusson subtract the percentage change in the consumer price index in their definition of $x_t$. They argue correctly that most economic theories would involve the real rather than the nominal price of oil. On the other hand, if the nonlinearity represents threshold responses based on consumer sentiment, it is possible that these thresholds are defined in nominal terms. I would also note that empirical measurement of "the" aggregate price level is problematic, and deflating by a particular number such as the CPI introduces a new source of measurement error, which could lead to a deterioration in the forecasting performance. In any case, it is again quite possible that there are differences in the functional form of forecasts based on nominal instead of real prices.

**Inclusion of contemporaneous regressors.** The $\chi^2$ statistics in Kilian and Vigfusson's (2009) Table 4 are in fact not based on the forecasting regression (4), but instead come from testing $\gamma_0 = \gamma_1 = \cdots = \gamma_p$ in

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{p} \beta_i x_{t-i} + \sum_{i=0}^{p} \gamma_i \tilde{x}_{t-i} + \varepsilon_t^{(0)}. \tag{6}$$

Kilian and Vigfusson (2010) claim that (4) is a special case of (6), but I would disagree. Insofar as it is proposed that either equation could be estimated by OLS, the residual $\varepsilon_t^{(0)}$ in

(6) must be uncorrelated with $x_t$ and $\tilde{x}_t$, whereas there is no such requirement on the residual $\varepsilon_t$ in (4). If one interprets both (4) and (6) as population linear projections, in general the values of $\{\phi_i, \beta_i, \gamma_i\}_{i=1}^p$ in (4) would not be the same as the values of $\{\phi_i, \beta_i, \gamma_i\}_{i=1}^p$ in (6). The former are values that give optimal one-quarter-ahead forecasts of GDP, while the latter coefficients represent the answer to a different question.[2]

Even if we did believe that there is an implicit assumption that the error $\varepsilon_t$ in (4) associated with an optimal one-quarter-ahead forecast also happens to be uncorrelated with $x_t$ and $\tilde{x}_t$, in which case (4) would be a special case of (6) with $\beta_0 = \gamma_0 = 0$, one would still expect there to be a loss in power in testing the hypothesis $\gamma_0 = \cdots = \gamma_p = 0$ on the basis of (6) (which requires estimating the nuisance parameters $\beta_0$ and $\gamma_0$) relative to testing the hypothesis ($\gamma_1 = \cdots = \gamma_p = 0$) on the basis of (4) (which imposes the maintained true values for $\beta_0$ and $\gamma_0$).

**Number of lags.** My original regression (5) used $p = 4$ lags, whereas Kilian and Vigfusson have used $p = 6$ lags throughout. If the truth is $p = 4$, estimating and testing the additional lags will result in a reduction in power. On the other hand, it might be argued that an optimal linear forecast of $y_t$ requires more than 4 lags of $y_{t-i}$ and $x_{t-i}$, and that omitting the extra lags accounts for the apparent success of a nonlinear specification (since $x_{t-4}^{\#}$ incorporates some additional information about $x_{t-i}$ for $i > 4$).

**The contribution of each factor.** Table 1 identifies the role of each of these differences in turn, by changing one element of the specification at a time and seeing what effect it

[2] For example, if $y_t = x_t + \varepsilon_t^{(0)}$ and $x_t = \rho x_{t-1} + u_t$, then in (4) we would have $\beta_1 = \rho$ whereas in (6) the value of $\beta_1$ is 0.

has on the results. The first row gives the $p$-value for the last entry reported in Kilian and Vigfusson's Table 4, while the second row gives the $p$-value for my specification on the original data set. The third row isolates the effect of the choice of sample period alone, by estimating my original specification using the sample period adopted by Kilian and Vigfusson. Instead of a $p$-value of 0.002 obtained for the original sample, the $p$-value is only 0.013 on the new data set. Is this because the sample is shorter, or because the relation has changed? One can test for the latter possibility by using data for 1949:Q2-2007:Q4 to re-estimate equation (5) and test the hypothesis that the coefficients on $\{x_{t-1}^{\#}, x_{t-2}^{\#}, x_{t-3}^{\#}, x_{t-4}^{\#}\}$ were different subsequent to 2001:Q4 compared with those prior to 2001:Q4. One fails to reject the null hypothesis of no change in these coefficients ($\chi^2(4) = 5.03$, $p = 0.284$). Thus a key explanation for why Kilian and Vigfusson find weaker evidence of nonlinearity is that they have used a shorter sample.

Subsequent rows of Table 1 use Kilian and Vigfusson's 1974:Q4-2007:Q4 sample, but change other elements of their specification one at a time. Row 4 uses the real change in the producer price index of crude petroleum in place of the real change in the refiner acquisition cost of imported oil, but otherwise follows Kilian and Vigfusson in all the other details. Using the PPI instead of the RAC would reduce Kilian and Vigfusson's reported $p$-value from 0.046 to 0.024. Row 5 keeps the RAC, but uses the nominal price rather than the real. This change alone would again have reduced the $p$-value from 0.046 to 0.028. Row 6 simply omits the contemporaneous term, basing the test on (4) rather than (6), and would be another way to reduce the $p$-value to 0.027. Finally, row 7 shows that using $p = 4$

instead of $p = 6$ would also increase the evidence of nonlinearity. Furthermore, a test of the null hypothesis that $\beta_5 = \beta_6 = \gamma_5 = \gamma_6 = 0$ in Kilian-Vigfusson's original (6) fails to reject ($\chi^2(4) = 3.90$, $p = 0.420$), suggesting that this again is another factor in reducing the power of their tests.

To summarize, Kilian and Vigfusson make a number of changes from previous research, including a shorter sample, different oil price measure, different price adjustment, inclusion of contemporaneous terms, and longer lags. Each of these changes, taken by itself, would lead them to find weaker evidence of nonlinearity than previous research. Taken together, they explain why Kilian and Vigfusson find more support for linear specifications than have previous researchers.

**Post-sample performance.** Kilian and Vigfusson (2010) also conduct some out-of-sample exercises, though they do not report these for the specification that forms the basis of their 2009 analysis, which I now examine. For the $h = 1$ forecasting horizon, one can implement their approach by estimating equation (4) using the real RAC measure for the price of oil, $p = 6$, and $t =$ 1974:Q4 to $T$, which is the exact specification adopted in their 2009 paper. I used these estimated coefficients and the value for $y_T, x_T$, and $\tilde{x}_T$ to predict the value of $y_{T+1}$, and repeated the exercise for $T =$ 1990:Q2 to 2010:Q1. The average out-of-sample mean squared error for this specification is 0.33, compared with a value of 0.39 for a specification that omits the nonlinear terms ($\gamma_i = 0$ for $i = 1, ..., 6$). This improvement in MSE that results from including nonlinear terms is statistically significant on the basis of the Clark and West (2007) statistic. Furthermore, when one drops the linear terms (that is, set $\beta_i = 0$ for $i = 1, ..., 6$) the MSE improves further to a value of 0.29. Either nonlinear specification also offers an improvement over the pure autoregression ($\beta_i = \gamma_i = 0$ for $i = 1, ..., 6$) which has an out-of-sample MSE of 0.34. Thus this exercise suggests that nonlinear terms are helpful, and that a parsimonious model that uses only the nonlinear terms does better than Kilian and Vigfusson's (2009) proposal of including both linear and nonlinear terms.

Some may feel that a true out-of-sample evaluation should refer to how well a relationship holds up after it has been published. Kilian and Vigfusson's (2009) original paper was estimated on data for $t =$ 1974:Q4 to 2007:Q4. I used those exact coefficients and specification to construct one-quarter-ahead forecasts of real GDP growth over 2008:Q1 to 2010:Q1. The true out–of-sample MSE for that model is 0.42, which is a 55% improvement over the MSE of 0.95 for a specification estimated over the same original sample that excludes the nonlinear terms.

Interestingly, if one uses the specification (5) exactly as it was published in equation (3.8) in Hamilton (2003), which used the nominal PPI rather than real RAC and whose coefficients were estimated 1949:Q2 to 2001:Q3, the MSE over 2008:Q1 to 20010:Q1 is 0.32, a 76% improvement over Kilian and Vigfusson's (2009) preferred nonlinear representation.

# 3 Censoring bias.

In Section 2 of their paper, Kilian and Vigfusson (2009) demonstrate that if the true relation is linear and one mistakenly estimates a nonlinear specification, the resulting estimates are asymptotically biased. These results parallel the demonstration in Hamilton (2003) that if the true relation is nonlinear and one mistakenly estimates a linear specification, the resulting estimates are asymptotically biased. Both statements are of course true, and are illustrations of the broader theme that one runs into problems whenever one tries to estimate a misspecified model.

Kilian and Vigfusson suggest one should take the high road of including both linear and nonlinear terms as a general strategy to avoid either problem. While that would indeed work if one had an infinite sample, in practice it is not always better to add more parameters, particularly in a sample as small as that used by Kilian and Vigfusson. After all, the same principle would suggest we include *both* the RAC and PPI as the oil price measure on the right-hand side, since there is disagreement as to which is the better measure, and nonlinear transformations of both the real and nominal magnitudes. Nobody would do that, and nobody should. All empirical research necessarily faces a trade-off between parsimony and generality, and one is forced to choose some point on that trade-off in literally every empirical study that has ever been done. My personal belief is that there are very strong arguments for trying to keep the estimated relations parsimonious. I note for example the results from the preceding section that more parsimonious representations have better out-of-sample performance.

# 4 On calculating impulse-response functions.

Kilian and Vigfusson (2010) seem to agree that a one-quarter-ahead forecast of real GDP should make use of a nonlinear rather than a linear functional form, though a reader of both their (2009) and (2010) papers could be forgiven for failing to come away with the understanding that this was in fact their conclusion. Kilian and Vigfusson (2010) nonetheless assert "all of these results are tangential to the question at hand because the results of slope-based tests are not informative about the degreee of asymmetry in the response functions of real GDP." This strikes me as an odd position to take for a number of reasons, several of which are in fact nicely articulated in Kilian and Vigfusson's (2009) original analysis.

The issues that Kilian and Vigufsson (2009) address have not to do with whether the equation for forecasting real GDP (4) is a nonlinear function of lagged oil prices, but instead with the specification of a separate equation for forecasting the oil prices themselves. They argue, quite correctly, that one cannot use an equation such as

$$x_t^{\#} = c + \sum_{i=1}^{p} b_i x_{t-i}^{\#} + \sum_{i=1}^{p} d_i y_{t-i} + u_t, \tag{7}$$

as an element for constructing multi-period forecasts that go into generating impulse-response functions. The problem comes from the fact that while (5) in such a two-variable VAR might be exactly the correct equation to use to forecast GDP, the fitted values of (7) cannot possibly represent the conditional expectation of oil prices

$$E(x_t^{\#} | x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, ...), \tag{8}$$

since (7) could generate a negative predicted value for $x_t^{\#}$, which an optimal forecast (8) would never allow. This point was first noted by Balke, Brown and Yücel (2002), though most researchers have ignored the concern.

There is certainly a problem with applying mechanically the standard linear impulse-response tools in such a setting as a result of the difference between (7) and (8). But also at a more fundamental level, researchers need to reflect on the underlying question that they are intending such calculations to answer. A variable such as $x_t^{\#}$ is nonnegative by definition, and therefore the conditional expectation (8) must always be a positive number. Thus if one defines an "oil shock" as a deviation from the conditional expectation,

$$u_t = x_t^{\#} - E(x_t^{\#}|x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, ...), \tag{9}$$

then there is a range of positive realizations of $x_t^{\#}$ that are defined to be a "negative oil shock". More generally, insofar as an impulse-response function is intended to summarize the revision in expectations of future variables associated with a particular realization of (9), as Gallant, Rossi, and Tauchen (1993), Koop, Pesaran and Potter (1996), and Potter (2000) emphasized, such an object is in principle different for every different information set $\{x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, ...\}$ and size of the shock $u_t$. There are an infinite number of questions one could ask about dynamic response functions in a nonlinear system, with a potentially different answer for each history and each size shock. Which of these is "the" impulse-response function of interest? For small shocks, one would expect from Taylor's Theorem that a linear representation of the function would be a good approximation around the point of linearization. In most of their analysis, Kilian and Vigfusson seem to assume that the object of interest is a one-standard deviation shock averaged across the dates in the sample. Given this decision as to the question they propose to answer, and particularly given the underlying weak evidence of nonlinearity for their data set and specification, Kilian and Vigfusson find limited evidence of nonlinearity in the impulse-response function. On the other hand, by "oil shock," many of us may instead have in mind the consequences of extraordinary events. I note that, even with their favored specification and data set, Kilian and Vigfusson find statistically significant evidence of nonlinearity when they examine the effects of two-standard-deviation shocks..

In any case, there is a much simpler and direct way to get at this question. Any answer from the infinite set of possible impulse-response functions in a nonlinear system is nothing more than an answer to a particular conditional forecasting question plotted as a function of the horizon. Jordá (2005) notes that it is possible to estimate the latter directly as primitive objects independent of the equation used to forecast oil prices themselves, by simple OLS estimation of the equation for forecasting GDP $h$ periods ahead directly,

$$y_{t+h-1} = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{p} \beta_i x_{t-i} + \sum_{i=1}^{p} \gamma_i \tilde{x}_{t-i} + \varepsilon_t, \tag{10}$$

on which one can readily test the null hypothesis of linearity in the form of the restriction $\gamma_1 = \gamma_2 = \cdots = \gamma_p = 0$. For $h > 1$, the errors in (10) are serially correlated, for which one could correct using the regression-coefficient covariance matrix proposed by either Hansen and Hodrick (1980) with $h - 1$ lags or by Newey and West (1987) using $L > h$ lags.[3]

The Jordá (2005) approach is also perfectly valid when $x_t$ exhibits discrete dynamics, a

[3] The Hansen-Hodrick results reported in Table 2 for $h = 1$ differ slightly from those reported in Table 1 because for $h = 1$, the Hansen-Hodrick formula becomes White's (1980) heteroskasticity-consistent estimate rather than the usual OLS, and the White standard errors turn out to be smaller than OLS standard errors for this application. The Newey-West results used $L = 5$ lags.

case for which Kilian and Vigfusson (2010) note that their impulse-response analysis could be problematic.

Table 2 reports results of these tests using both the Kilian-Vigfusson data set and specification (that is, $x_t$ the real RAC, $p = 6$, and $t + h - 1$ running from 1974:Q4 to 2007:Q4) and the original Hamilton (2003) data set and specification ($x_t$ the nominal PPI, $p = 4$, and $t + h - 1$ running from 1949:Q2 to 2001:Q3). Interestingly, for every specification and every horizon one finds quite strong evidence of nonlinearity.

The evident reconciliation of results is that, although there is not much evidence of a nonlinear response to small changes in the Kilian-Vigfusson data set and specification, the results are quite consistently indicating nonlinear consequences of larger movements in oil prices.

# 5 Conclusion.

The evidence is convincing that the relation between GDP growth and oil prices is nonlinear. The recent paper by Kilian and Vigfusson (2009) does not challenge that conclusion, but does offer a useful reminder that we need to think carefully about what question we want to ask with an impulse-response function in such a system and cannot rely on off-the-shelf linear methods for an answer.

# References


Atkeson, Andrew, and Patrick J. Kehoe (1999) "Models of energy use: putty-putty versus putty-clay," *American Economic Review*, 89, 1028-1043.

Balke, Nathan S., Stephen P.A. Brown, and Mine Yücel (2002) "Oil Price Shocks and the U.S. Economy: Where Does the Asymmetry Originate?" *Energy Journal,* 23, pp. 27-52.

Bernanke, Ben S. (1983) "Irreversibility, Uncertainty, and Cyclical Investment," *Quarterly Journal of Economics, 98*, pp. 85-106.

Bresnahan, Timothy F., and Valerie A. Ramey (1993) "Segment Shifts and Capacity Utilization in the U.S. Automobile Industry," *American Economic Review Papers and Proceedings, 83,* no. 2, pp. 213-218.

Carlton, Amelie Benear (2010) "Oil Prices and Real-Time Output Growth." Working paper, University of Houston.

Clark, Todd E., and Kenneth D. West (2007) "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics,* 138, 291-311.

Cuñado, Juncal, and Fernando Pérez de Gracia (2003) "Do Oil Price Shocks Matter? Evidence for Some European Countries," *Energy Economics,* 25, pp. 137–154.

Dahl, Carol A. (1993) "A Survey of Oil Demand Elasticities for Developing Countries," *OPEC Review* 17(Winter), pp. 399-419.

Davis, Steven J. (1987a) "Fluctuations in the Pace of Labor Reallocation," in K. Brunner and A. H. Meltzer, eds., *Empirical Studies of Velocity, Real Exchange Rates, Unemployment and Productivity*, Carnegie-Rochester Conference Series on Public Policy, 24, Amsterdam:

North Holland.

_____ (1987b) "Allocative Disturbances and Specific Capital in Real Business Cycle theories," *American Economic Review Papers and Proceedings, 77*, no. 2, pp. 326-332.

_____, and John Haltiwanger (2001) "Sectoral Job Creation and Destruction Responses to Oil Price Changes," *Journal of Monetary Economics,* 48, pp. 465-512.

Edelstein, Paul, and Lutz Kilian (2009) "How Sensitive are Consumer Expenditures to Retail Energy Prices?" *Journal of Monetary Economics*, 56(6), pp. 766-779.

Elder, John, and Apostolos Serletis (2010) "Oil Price Uncertainty," *Journal of Money, Credit and Banking*, forthcoming.

Engemann, Kristie M., Kevin L. Kliesen, and Michael T. Owyang (2010) "Do Oil Shocks Drive Business Cycles? Some U.S. and International Evidence," working paper, Federal Reserve Bank of St. Louis.

Ferderer, J. Peter (1996) "Oil Price Volatility and the Macroeconomy: A Solution to the Asymmetry Puzzle," *Journal of Macroeconomics, 18*, pp. 1-16.

Finn, Mary G. (2000) "Perfect Competition and the Effects of Energy Price Increases on Economic Activity," *Journal of Money, Credit and Banking* 32, pp. 400-416.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen (1993) "Nonlinear Dynamic Structures," *Econometrica* 61, pp. 871-908.

Hamilton, James D. (1983) "Oil and the Macroeconomy Since World War II," *Journal of Political Economy, 91*, pp. 228-248.

_____ (1988) "A Neoclassical Model of Unemployment and the Business Cycle," *Jour-*

*nal of Political Economy, 96*, pp. 593-617.

_____ (1996) "This is What Happened to the Oil Price Macroeconomy Relation," *Journal of Monetary Economics*, 38, pp. 215-220.

_____ (2003) "What Is an Oil Shock?", *Journal of Econometrics,* 113, pp. 363-398.

_____ (2005) "What's Real About the Business Cycle?" *Federal Reserve Bank of St. Louis Review,* July/August (87, no. 4), pp. 435-452.

_____ (2009) "Causes and Consequences of the Oil Shock of 2007-08," *Brookings Papers on Economic Activity,* Spring 2009, pp. 215-261.

Hansen, Lars P., and Robert J. Hodrick (1980) "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy* 88, pp. 829-853.

Herrera, Ana María, Latika Gupta Lagalo, and Tatsuma Wada (2010), "Oil Price Shocks and Industrial Production: Is the Relationship Linear?," working paper, Wayne State University.

Hooker, Mark A. (1996) "What Happened to the Oil Price-Macroeconomy Relationship?," *Journal of Monetary Economics,* 38, pp. 195-213.

Jiménez-Rodrígueza, Rebeca, and Marcelo Sánchez (2005) "Oil Price Shocks and Real GDP Growth: Empirical Evidence for Some OECD Countries," *Applied Economics*, 37, pp. 201–228.

Jordá, Òscar (2005) "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95(1): 161-182.

Kilian, Lutz, and Robert J. Vigfusson (2009) "Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases?" Working Paper, University of Michigan (http://www-personal.umich.edu/~lkilian/kvsubmission.pdf).

Kilian, Lutz, and Robert J. Vigfusson (2010) "Nonlinearities in the Oil-Output Relationship," Working Paper, University of Michigan.

Kim, Dong Heon (2009) "What is an Oil Shock? Panel Data Evidence," working paper, Korea University.

Kim, In-Moo, and Prakash Loungani (1992) "The Role of Energy in Real Business Cycle Models," *Journal of Monetary Economics, 29*, no. 2, pp. 173-189.

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter (1996) "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics,* 74, pp. 19-147.

Leduc, Sylvain, and Keith Sill (2004) "A Quantitative Analysis of Oil-Price Shocks, Systematic Monetary Policy, and Economic Downturns," *Journal of Monetary Economics,* 51, pp. 781-808.

Lee, Kiseok, Shawn Ni and Ronald A. Ratti, (1995) "Oil Shocks and the Macroeconomy: The Role of Price Variability," *Energy Journal, 16,* pp. 39-56.

Loungani, Prakash (1986) "Oil Price Shocks and the Dispersion Hypothesis," *Review of Economics and Statistics,* 68, pp. 536-539.

Mork, Knut A. (1989) "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results," *Journal of Political Economy, 91*, pp. 740-744.

Newey, Whitney K., and Kenneth D. West (1987) "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, pp. 703-708.

Potter, Simon M. (2000) "Nonlinear impulse response functions," *Journal of Economic Dynamics and Control,* 24, pp. 1425-1446.

Ramey, Valerie A., and Daniel J. Vine (2010) "Oil, Automobiles, and the U.S. Economy: How Much Have Things Really Changed?", working paper, UCSD.

Ravazzolo, Francesco, and Philip Rothman (2010) "Oil and US GDP: A Real-Time Out-of-Sample Examination," working paper, East Carolina University.

West, Kenneth D., and Whitney K. Newey (1987) "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, pp. 703-708.

White, Halbert L. (1980) H. White: "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48, pp. 817-838.

Table 1
*P*-values for test of null hypothesis of linearity for alternative specifications

| | sample | oil measure | price adjustment | contemporaneous | # lags | *p*-value |
|---|---|---|---|---|---|---|
| (1) | 1974:Q4-2007:Q4 | RAC | real | include | 6 | 0.046 |
| (2) | **1949:Q2-2001:Q3** | **PPI** | **nominal** | **exclude** | **4** | 0.002 |
| (3) | 1974:Q4-2007:Q4 | **PPI** | **nominal** | **exclude** | **4** | 0.013 |
| (4) | 1974:Q4-2007:Q4 | **PPI** | real | include | 6 | 0.024 |
| (5) | 1974:Q4-2007:Q4 | RAC | **nominal** | include | 6 | 0.028 |
| (6) | 1974:Q4-2007:Q4 | RAC | real | **exclude** | 6 | 0.027 |
| (7) | 1974:Q4-2007:Q4 | RAC | real | include | **4** | 0.036 |

Notes to Table 1: *P*-values for test that $\gamma_0 = \gamma_1 = \cdots = \gamma_p = 0$ in equation (6) (for rows with "include" in contemporaneous column) or test that $\gamma_1 = \cdots = \gamma_p = 0$ in equation (4) (for rows with "exclude" in contemporaneous column). Boldface entries in each row indicate those details of the specification that differ from the first row.

Table 2
*P*-values for test of null hypothesis of linearity of *h*-quarter-ahead forecasts of real GDP using alternative data sets and specifications.

| Forecast horizon | Kilian and Vigfusson (2009) | | Hamilton (2003) | |
|---|---|---|---|---|
| | Hansen_Hodrick | Newey-West | Hansen-Hodrick | Newey-West |
| $h = 1$ | 0.002 | 0.000 | 0.001 | 0.000 |
| $h = 2$ | 0.000 | 0.000 | 0.000 | 0.000 |
| $h = 3$ | 0.000 | 0.037 | 0.000 | 0.000 |
| $h = 4$ | 0.000 | 0.000 | 0.001 | 0.001 |

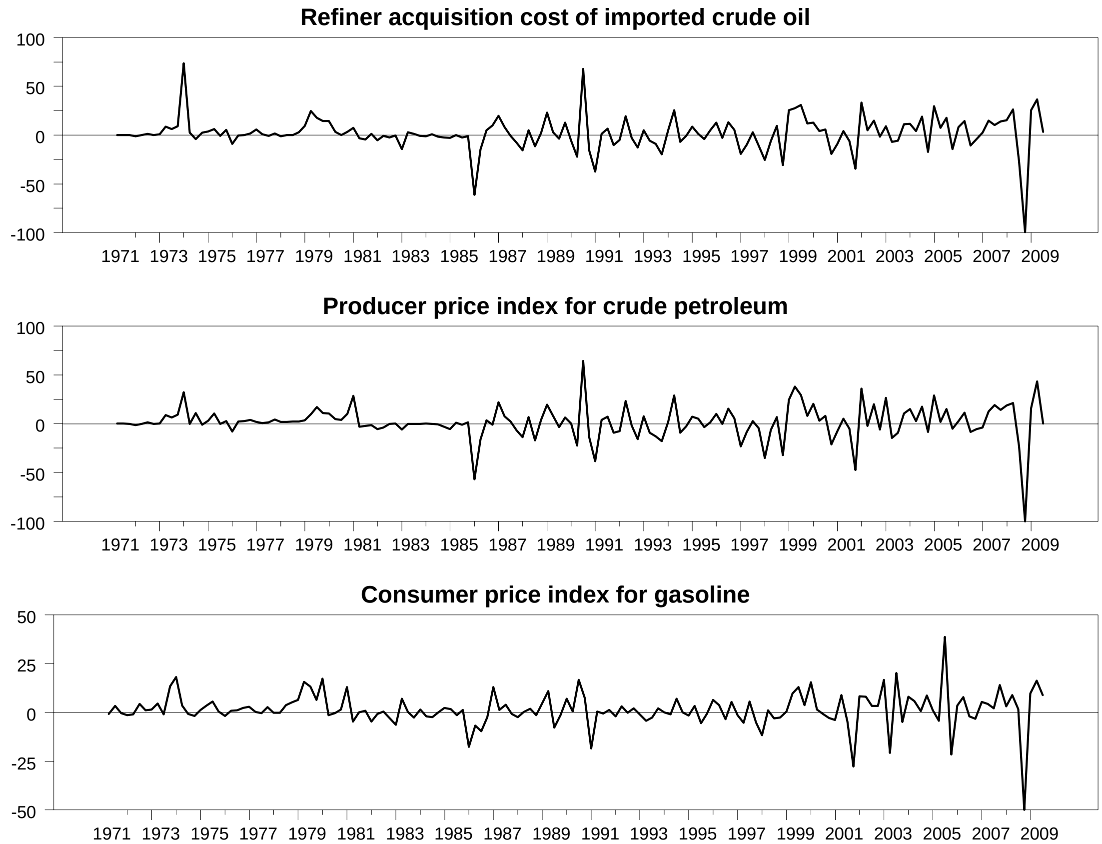


Figure 1. Quarterly percent changes in PPI, RAC, and gasoline CPI. Third series is seasonally adjusted, first two are not.