ADAM HALE SHAPIRO

Estimating the New Keynesian Phillips Curve: A Vertical Production Chain Approach

It has become customary to estimate the New Keynesian Phillips Curve (NKPC) with generalized method of moments using a large instrument set that includes lags of variables that are ad hoc to the firm's price-decision problem. Researchers have also conventionally used real unit labor cost (RULC) as the proxy for real marginal cost even though it is difficult to support its significance. This paper introduces a new proxy for the real marginal cost term as well as a new instrument set, both of which are based on the micro foundations of the vertical chain of production. I find that the new proxy, based on input prices as opposed to wages, provides a more robust and significant fit to the model. Instruments that are based on the vertical chain of production appear to be both more valid and relevant toward the model.

THE NEW KEYNESIAN PHILLIPS CURVE (NKPC) has become the backbone of current macroeconomic inflation research. The model's prominence resides in its simple structure yet micro-founded principles whereby firms make their pricing decisions in a setting of nominal rigidities. There has been a large volume of econometric research related to the model. Examples include attempts to estimate the parameters, justify the structural significance, and deduce the implications for inflation persistence.[1]

---

1. Roberts (2001), Eichenbaum and Fisher (2003), Galí and Gertler (1999), Galí, Gertler, and Lopez-Salido (2005), and Linde (2005), among others, estimate the parameters for the U.S. Cogley and Sbordone (2005) argue that the parameters of the NKPC are structurally invariant. Fuhrer (2006) discusses issues of persistence implied by the NKPC.

Unfortunately, econometrically the model has had limited success. Estimates of the parameters of the model are highly sensitive to the estimation procedure, the subsample over which the model is estimated, and the data used to estimate the model. Furthermore, there has been a large debate on the specification of the model, specifically concerning whether a lagged inflation term should be included in order to better fit the model to the data. Because of such controversy, ascertaining the correct parameters of the model is vital in determining the model's specification and validity. Researchers such as Linde (2005) and Rudd and Whelan (2005) have argued that estimating the model using GMM results in specification bias, and therefore leads to unreliable estimates of the parameters. In particular, Rudd and Whelan argue that GMM produces biased estimates of the parameters, owing to scarce availability of valid and relevant instruments.

Galí and Gertler (1999), Eichenbaum and Fisher (2003), and Galí, Gertler, and Lopez-Salido (2001, 2005), however, argue that generalized method of moments (GMM) estimation still provides a robust and unbiased estimation of the parameters of the model. These researchers, estimating the NKPC with GMM, have conventionally proxied real marginal cost with real unit labor cost (RULC) and instrumented with a large instrument set that includes lags of variables that *seem* relevant to the model, such as lags of the output gap, the risk-free interest rate, commodity inflation, and wage inflation. Although these variables are not found in the firm's price-decision problem, they are chosen as instruments based on the idea that they should, at least intuitively, encompass the validity and relevancy properties that any good instrument should possess. It turns out, however, that many of these variables are only weakly correlated with the endogenous variables of the model. Furthermore, a large number of variables must be included in the instrument set in order to induce a significant estimate of the coefficient on the marginal cost term. Unfortunately, as this paper shows, a trade-off emerges, and the addition of blocks of lagged variables reduces the relevancy of the instruments, exacerbates the problem of weak instruments, and leads to estimates that are biased toward ordinary least squares (OLS) estimation.

This paper provides an alternative to the RULC proxy and the instruments currently used in GMM estimation. The proxy and instruments proposed both stem from generalizing the original model into that of a vertical chain of production, studied by Huang and Liu (2001, henceforth HL). In such a model, input prices play a vital role in firms' pricing decisions as goods travel down the chain of production from the crude-goods stage to the consumer-goods or GDP stage. The novelty of this paper lies in the incorporation of input prices into both the real marginal cost proxy and the instrument set.

The new proxy I propose, the real unit input cost (RUIC), measures the cost of a unit of output in terms of the price of inputs used to produce the good. The instruments I propose, the RUICs of "upstream" firms (i.e., firms in earlier stages of production), are micro-founded within the model, in contrast to the instruments used by current researchers. I find that these new instruments are exogenous and relevant to firms downstream on the production chain, two vital properties that good instruments should possess. Also, these instruments are sufficient by themselves to induce high relevancy and exogeneity. Importantly, they do not require inflation variables to be added to the

instrument set, a requirement that had been a major drawback of previous approaches to estimating the forward-looking NKPC.

In order to compare both proxies under vertical production-chain instruments, upstream RULC terms were constructed. Estimating the NKPC using either proxy with its own upstream instruments induces high relevancy and exogeneity. However, the RUIC and its upstream instruments still provide better estimates of the model in terms of the relevancy of the instrument set, stability of the parameters, and fit of the simulated inflation series with actual inflation. Overall, I find that the RUIC and its upstream instruments generally outperform the RULC proxy under a variety of instrument sets.

The paper is organized as follows. Section 1 discusses the foundations of the stage-of-production model. Section 2 discusses the data and marginal cost proxies used in estimation. In Section 3, I estimate the forward-looking NKPC using both the RULC and RUIC proxies with instrument sets used in the prior literature. Section 4 discusses the forward-looking model with the new instrument set that is based on the stage-of-production model. I compare the new instruments to the instruments currently being used in GMM estimation of the NKPC. Section 5 looks at the hybrid-model estimation of the forward-looking model. Section 6 discusses the results of subsample and rolling regression analysis. In Section 7, I construct series of fundamental inflation based on the discounted future stream of the RULC and RUIC proxies. I conclude in Section 8.

1. THE STAGE-OF-PRODUCTION NKPC

In the model proposed by HL, the entire economy can be thought of as a single vertical chain of production where inputs flow "downstream," from crude goods to intermediate goods to consumption goods. Output produced by the crude-goods stage is used as an input for intermediate goods. Intermediate output is then used as an input for finished goods. The chain of production concludes when inputs reach the consumption-goods sector. Generalizing the number of stages to N, output of stage n is used as an input in the production schedule of the subsequent stage— stage $n + 1$. This section discusses the stage-of-production model that generalizes the NKPC into a vertical chain of production. As will be shown, this model incorporates the NKPC estimated by previous researchers as a specific stage—stage N—of the general model.

1.1 Firms

Each stage of production consists of individual firms indexed on the unit interval. Aggregate output of a certain stage n is based on a Dixit and Stiglitz (1977) composite of output produced by the index of stage n firms:

$$\bar{Y}_n = \left[\int_0^1 Y_n(z)^{\frac{\epsilon-1}{\epsilon}} dz \right]^{\frac{\epsilon}{\epsilon-1}} . \tag{1}$$

Production flows vertically down the chain of production, beginning in stage 1, in which labor, $L_1(z)$, and productivity, A_1, are used to produce output $Y_1(z)$:

$$Y_1(z) = A_1 L_1(z). \tag{2}$$

In the subsequent stages of production, $n \in \{2, \ldots, N\}$, output of firm z in stage n is based on the assumption of a Cobb–Douglas technology, where output from stage $n - 1$ is used as an input in the production schedule of stage n:

$$Y_n(z) = A_n(\bar{Y}_{n-1}(z))^{\alpha_n}(L_n(z))^{1-\alpha_n}. \tag{3}$$

Note that the vertical production-chain structure of the model implies that the consumption aggregate, C, is an aggregate of goods produced in the final stage of production, stage N, such that $C = \bar{Y}_N$.[2] It is also important to note that, even though the firms are linked via marginal cost, they compete only with other firms in the same stage of processing. Thus, the model treats each stage of processing as a distinct aggregate economy with its own price level, $\bar{P}_n = \int_0^1 P_n(z)^{1-\epsilon} \, dz$, marginal cost, and demand structure.

To derive a demand function, we assume that customers of firms in stage n will perform a cost minimization for the firms' products. That is, customers of stage n will minimize costs by choosing $Y_n(z)$ to

$$\min \int_0^1 P_n(z)Y_n(z) \, dz$$

subject to equation (1). The demand function facing firm z at stage n is

$$Y_n^d(z) = \left[\frac{P_n(z)}{\bar{P}_n}\right]^{-\epsilon} \bar{Y}_n. \tag{4}$$

Thus, each firm faces a constant elasticity of demand curve, where demand is decreasing in the good's relative price to other goods in stage n.

1.2 Price Setting

Following Calvo (1983), each firm has a constant probability, $1 - \theta_n$, of adjusting its price. Firms at different stages of production face different demand curves, but with the same elasticity. The problem of the firm is to choose a price such that its future stream of profits is maximized. That is, the firm performs the following maximization problem:

$$\max_{P_{n,t}(z)} E_t \sum_{i=0}^{\infty} (\beta\theta_n)^i \left[\frac{P_{n,t}(z)}{\bar{P}_{n,t+i}} - \frac{\Psi_{n,t}^{Nom}}{\bar{P}_{n,t+i}}\right] Y_{n,t+i}^d(z),$$

---

2. HL show that in a model with no government spending or capital, real GDP corresponds to aggregate consumption. C can therefore be interpreted as aggregate output or aggregate consumption.

where

$$\Psi_{n,t}^{Nom} = \left(\frac{1}{A_{n,t}}\right)(W_{n,t})^{1-\alpha_n}(\bar{P}_{n-1,t})^{\alpha_n}\widetilde{\alpha}_n \tag{5}$$

is the nominal marginal cost of firms in stage n, where $\widetilde{\alpha}_n = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}$. Marginal cost is obtained by firms that minimize cost, $\bar{P}_{n-1,t}\bar{Y}_{n-1,t}(z) + W_t L_{n,t}(z)$, subject to (3)

Log-linearizing the first-order condition to problem (5) gives the price-setting function of the firm, where lowercase letters denote log-deviation from the steady state:

$$p_{n,t}^*(z) = (1 - \beta\theta_n)\sum_{i=0}^{\infty}(\beta\theta_n)^i E_t\left(\psi_{n,t+i}^{Nom}\right). \tag{6}$$

In setting its price, the firm is purely forward looking. Its optimal reset price, $p_{n,t}^*$, is chosen based on the sum of expected future nominal marginal costs, $\psi_{n,t}^{Nom}$. Since the model implies that all firms in the same stage reset to the same optimal price, the aggregate price level can be written in log-linear form as a convex combination of last period's price level and the optimal reset price, p_n^*.

$$\bar{p}_{t,n} = \theta_n\bar{p}_{n,t-1} + (1 - \theta_n)p_n^*. \tag{7}$$

Combining equations (6) and (7), we obtain N distinct NKPCs, one for each stage of production in the model.

$$\pi_{n,t} = \lambda_n\psi_{n,t} + \beta E_t(\pi_{n,t+1}) \quad \text{for } i \in \{1, \dots, N\}, \tag{8}$$

where $\lambda = \frac{(1-\theta_n)(1-\beta\theta_n)}{\theta_n}$ and $\psi_{n,t}$ is the log deviation of real marginal cost from its steady state. This equation is analogous to the general NKPC; however, here the marginal cost, inflation series, and parameters are indexed by stage of processing.

An important characteristic of the NKPC is that the inflation term inherits the purely forward-looking behavior of the firm's price setting. This becomes more apparent by iterating equation (8) forward to obtain the closed-form solution of the NKPC:

$$\pi_{n,t} = \lambda_n\sum_{i=0}^{\infty}\beta^i E_t\psi_{n,t+i}. \tag{9}$$

Equation (9) implies that inflation is completely forward looking with respect to the marginal cost term. Combined with equation (5), it also points to a linkage that resides between the different stages of processing. This linkage resides in the aggregate price level of firms in stage $n - 1$. Inflation in stage n is therefore dependent upon the pricing decisions of firms "upstream" to its own stage through the marginal cost term.

As equation (8) shows, the HL model proposes N distinct NKPCs. The aim of this paper, however, is not to estimate the parameters of the individual stage-of-processing

NKPCs but rather to utilize variables that stem from the stage-of-processing NKPCs in order to better estimate the stage N, or GDP stage, NKPC:

$$\pi_{N,t} = \lambda_{N,t}\psi_{N,t} + \beta E_t(\pi_{N,t+1}). \tag{10}$$

A nice feature of this model is that equation (10) is the same NKPC that has been studied in the previous literature. The vertical production-chain model is therefore a more generalized version of the NKPC, incorporating the "original" NKPC.

2. THE DATA

An obvious abstraction of the vertical production-chain model is that it restricts goods to flow in one direction—down the stages of processing. An alternative structure is to assume that goods travel in a "roundabout" fashion, as in Basu (1995), where all goods are inputs in the production of other goods. This paper assumes a vertical structure for two reasons. First, the available aggregate data are consistent with the stage-of-processing structure. I assume that at least *some* of the material inputs used by firms in stage n lie in the stage $n-1$ aggregate data. Second, the vertical structure of the model introduces variables that will be used to better estimate equation (10) using GMM. These variables are the RUIC of firms in stage N, as well as the RUIC of firms "upstream" to the GDP stage (i.e., $\psi_n^{(RUIC)}$ for $n < N$).

Specifically, this section introduces two different measures of the real marginal cost term that will be compared in the remainder of the paper: the RULC and the RUIC. I discuss the data used to construct these terms and then compare the visual fit of these terms to actual inflation.

2.1 Measuring Real Marginal Cost

The stage-of-production model gives three distinct algebraic representations of the real marginal cost term, ψ_n. All result from the same cost minimization problem facing the firm:[3]

$$\Psi_{n,t} \approx \frac{\bar{P}_{n-1,t}\bar{Y}_{n-1,t}}{\alpha_n \bar{P}_{n,t}\bar{Y}_{n,t}} \tag{11}$$

$$\approx \frac{(W_{n,t})^{1-\alpha_n}(\bar{P}_{n-1,t})^{\alpha_n}}{\bar{P}_{n,t}A_{n,t}}\widetilde{\alpha}_n \tag{12}$$

$$\approx \frac{W_{n,t}L_{n,t}}{(1-\alpha_n)\bar{P}_{n,t}\bar{Y}_{n,t}}. \tag{13}$$

---

3. Proxy (11) results from the first-order condition for minimizing cost with respect to aggregate input, proxy (13) results from the first-order condition with respect to labor, and proxy (12) results from utilizing both first-order conditions.

It is worth noting again that the model implies that each stage's marginal cost is deflated by its *own* price level, $\bar{P}_{n,t}$. Only the stage N marginal cost is deflated by the GDP deflator price level, $\bar{P}_{N,t}$. The model's intuition for this result is that firms in stage n compete *only* with firms within the same stage and not with firms upstream or downstream on the production chain. Thus, the general price level against which they should compare their own price is the price level of their own stage, $\bar{P}_{n,t}$.

Also noteworthy is the fact that proxy (11) includes input prices but no wages, proxy (13) includes wages but no input prices, and proxy (12) lies somewhere in the middle of this spectrum, depending on the values of the cost shares, α_n. Of the three potential proxies, proxy (12) is most susceptible to measurement error because cost shares are not available for the consumer sector.[4] Since proxies (11) and (13) shed enough light on the respective roles played by input prices and wages in the marginal cost term, we can discard proxy (12) for all intents and purposes.

Denote percent deviations from the steady state by lowercase letters; there are two potential proxies used in estimation: the RUIC and the RULC:

$$\text{RUIC:} \quad \psi_{n,t}^{(RUIC)} = \bar{p}_{n-1,t} + \bar{y}_{n-1,t} - \bar{p}_{n,t} - \bar{y}_{n,t} \tag{14}$$

$$\text{RULC:} \quad \psi_{n,t}^{(RULC)} = w_{n,t} + l_{n,t} - \bar{p}_{n,t} - \bar{y}_{n,t}. \tag{15}$$

Like the RULC, the RUIC measures the cost of producing one unit of output. The RUIC, however, takes this measure in terms of the cost of inputs produced in the previous stage of production, whereas the RULC takes this measure in terms of the cost of labor used in the same stage. The RUIC implicitly assumes substitutability between labor and capital, and also allows for the flow of goods between stages of processing.

The Bureau of Labor Statistics (BLS) creates a producer price index (PPI) for three stages of production: the crude-goods sector, the intermediate-goods sector, and the finished-goods sector. Note that the PPI measures producer prices—essentially whole-sale prices that producers pay for goods. Including the GDP chain-type price deflator (i.e., the price level pertinent to consumers) gives $N = 4$ different price series that pertain to the stage-of-production model. Real output for the crude-, intermediate-, and finished-goods stages is measured by the stage-of-processing industrial production index published by the Board of Governors. The RUIC proxy and its upstream counterparts are therefore relatively easy to construct.

The RULC for the GDP stage of production is also easy to construct as it is obtained directly from the Bureau of Economic Analysis (BEA). However, in order to perform a comparable estimation of the RULC in the NKPC to that of the RUIC, it is necessary to construct upstream marginal cost in terms of RULC (i.e., $\psi_{n,t}^{(RULC)}$ for $n < N$). Thus, we need wage levels, $W_{n,t}$, and employment hours, $L_{n,t}$, for the same stages of processing as we have for the PPI stages. These values are not directly available from

---

4. Cost shares are available for the PPI stages in the Census Department's Annual Survey of Manufacturers. However, the Service Annual Survey does not contain cost share data.

any outside source, and they must be obtained by matching the relative importance weights[5] (RIWs) from Current Employment Survey (CES) data with those of the corresponding goods in the same stage-of-processing price index.[6] The data therefore allow us to create real marginal cost variables for three stages of production: stage N, or the GDP stage; stage $N − 1$, or the finished-goods stage; and stage $N − 2$, or the intermediate-goods stage. Figure 1 illustrates the stages of production that are relevant for this paper.

Figure 2 plots GDP chain-type price inflation against the log deviations of the two proxies of the GDP stage, $\psi_{N,t}^{(RULC)}$ and $\psi_{N,t}^{(RUIC)}$, for the full sample period 1972:2–2002:4. Note the periods in the early 1970s and early 1980s in which the RUIC shows a lengthy lag behind inflation. Researchers of the "Volcker disinflation," such as Erceg and Levin (2003) and Goodfriend and King (2005), find that inflation expectations during this period were quite slow to adapt to monetary policy actions. This interpretation would suggest that RUICs were even slower to respond to monetary policy than prices during this era. Frictions such as inventory or capital adjustment costs could cause such a delay in the response of the variable.

On the other hand, the lag of the RUIC with respect to inflation could instead be interpreted as the *lead* of inflation with respect to the RUIC. That is, the sluggishness of the RUIC could be construed as the rapid adaptation of current inflation to future RUICs, indicating the forward-looking nature of price setters as implied by the closed-form solution of the NKPC, equation (9). Both interpretations seem to be viable explanations for the lead–lag relationship between the two variables, and this paper will not attempt to resolve this issue directly.[7] Overall, comparing the RUIC with the RULC, Figure 2 shows that the RUIC is much more volatile and seems to capture better the low-frequency components of inflation. Figures 3 and 4 plot the RUIC and RULC upstream marginal cost proxies (i.e., $\psi_{N-1,t}$ and $\psi_{N-2,t}$) against inflation. Visual inspection of the upstream variables indicates that the upstream RUIC proxies follow inflation more closely than do the upstream RULC proxies.

3. GG INSTRUMENTS ESTIMATION

Galí and Gertler (1999, henceforth GG), propose estimating the NKPC under the RULC proxy (i.e., $\psi_{N,t}^{RULC}$) using an instrument set consisting of four lags of inflation,

5. The BLS does not publish the actual values used as weights in making the index. The RIW represents the value weight multiplied by the relative price change from the weight date to the RIW calculation.

6. All data were matched using four-digit SIC industry codes. If an industry at the four-digit Standard Industrial Classification (SIC) code level was not found in the CES, then a three-digit SIC match was made. CES four-digit hourly wage data are available back to 1972. For more information on the construction of each variable, please refer to Appendix A.

7. One simple measure that could help to disentangle the relationship is a Granger causality test. This would test the hypothesis that inflation "Granger causes" the RUIC, even in the presence of other factors such as the output gap, the federal funds rate, and the relative price of oil. When four lags of each variable are included in the VAR, a p-value of 0.029 is obtained on the inflation block of the RUIC equation and a p-value of 0.037 is obtained on the RUIC block of the inflation equation. When lags 2 through 5 are used, the respective p-values are 0.001 and 0.928. When lags 3 through 6 are used, the respective p-values are 0.000 and 0.602. A low p-value on the former and large p-value on the latter indicate that inflation is leading the RUIC proxy.

```
┌──────────────────────────┐
│   Crude-Goods Firms       │ ◀─── L_{N-3}
│   (Stage N-3)             │
└──────────────────────────┘
        │         ▲
      Y_{N-3}    ψ_{N-2}
        ▼         │
┌──────────────────────────┐
│   Intermediate-           │ ◀─── L_{N-2}
│   Goods Firms             │
│   (Stage N-2)             │
└──────────────────────────┘
        │         ▲
      Y_{N-2}    ψ_{N-1}
        ▼         │
┌──────────────────────────┐
│   Finished-Goods Firms    │ ◀─── L_{N-1}
│   (Stage N-1)             │
└──────────────────────────┘
        │         ▲
      Y_{N-1}     ψ_N
        ▼         │
┌──────────────────────────┐
│ Consumption-Goods Firms   │ ◀─── L_N
│ (Stage N or GDP Stage)    │
└──────────────────────────┘
        │         ▲
        C        P_N
        ▼         │
┌──────────────────────────┐
│      Households           │
└──────────────────────────┘
          Labor
```
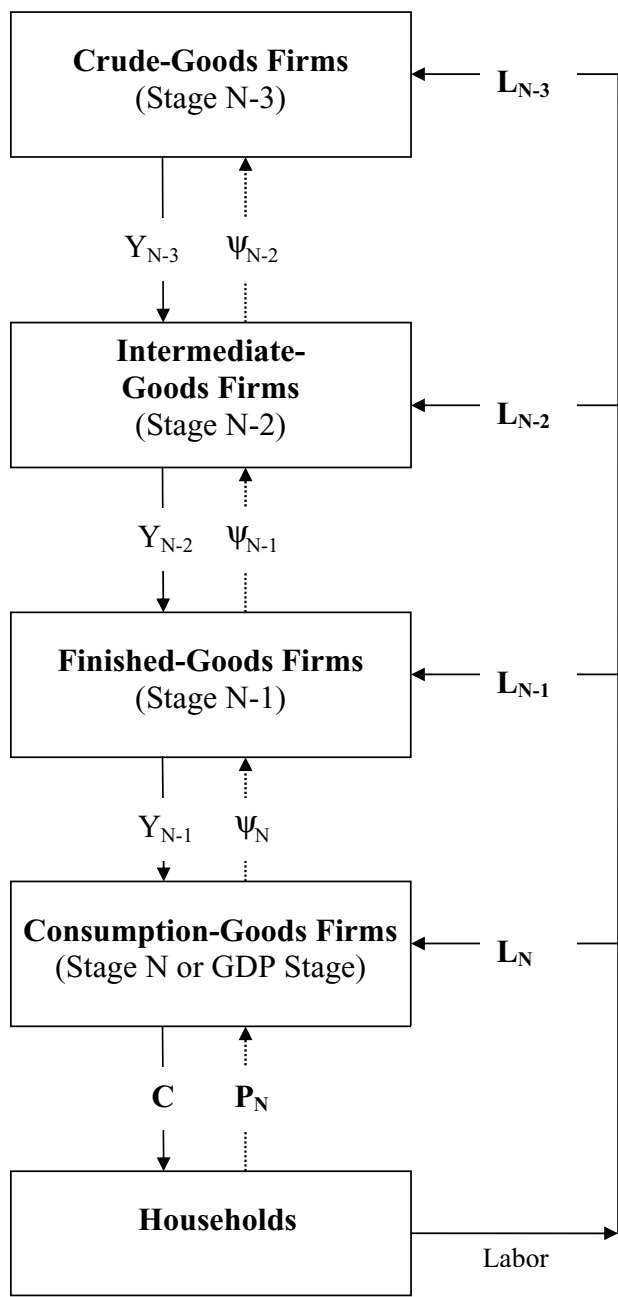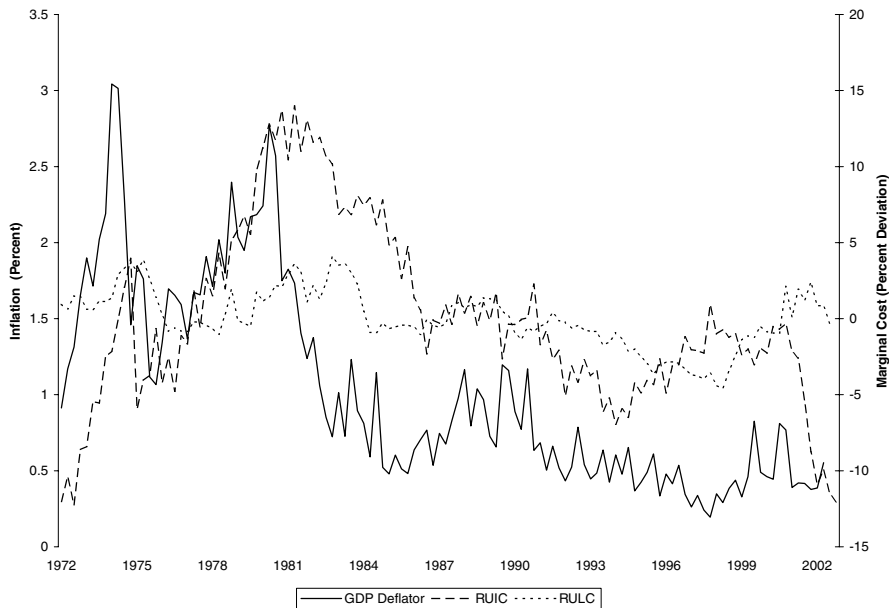
FIG. 1. Vertical Chain of Production.

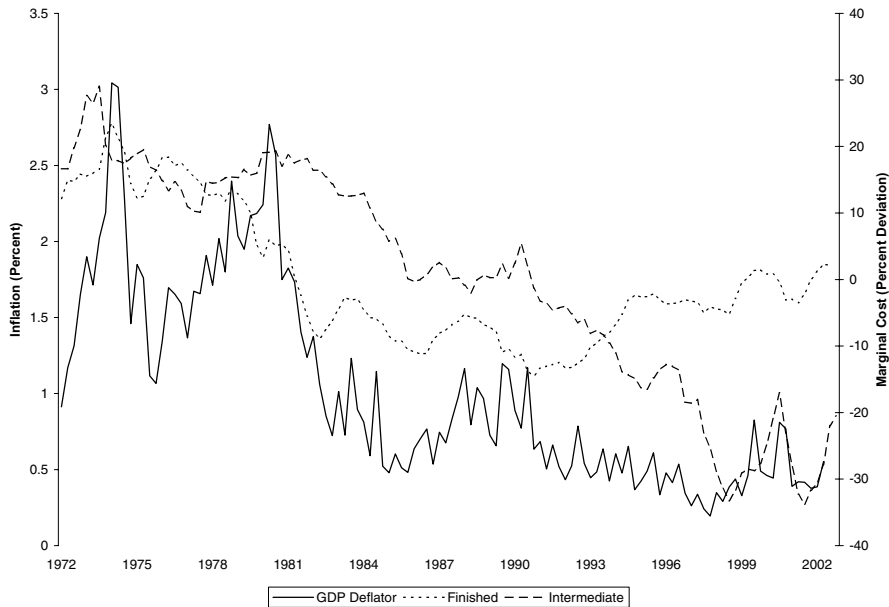FIG. 2. Inflation and Marginal Cost Proxies.

FIG. 3. Inflation and Upstream RUIC Proxies.

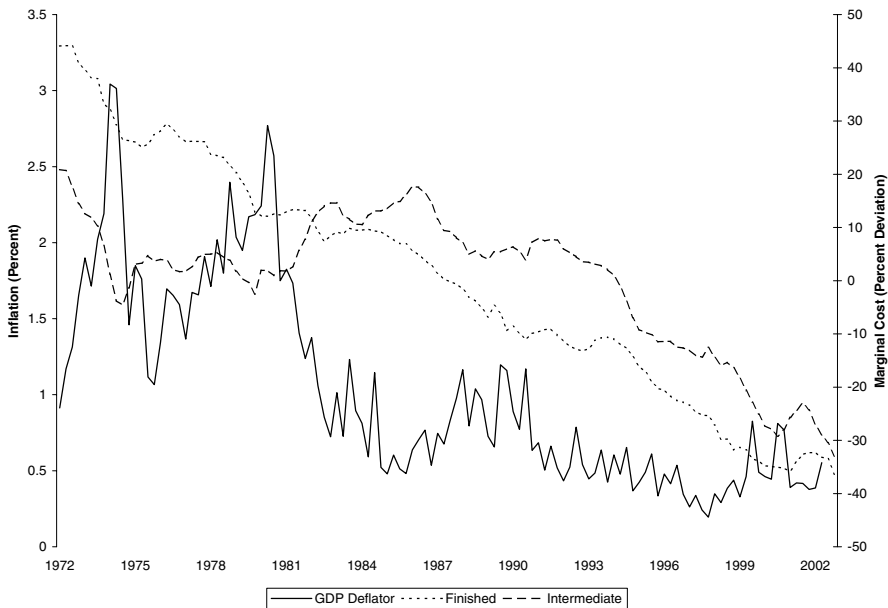NOTE: Intermediate RUIC was scaled by one-half for comparison purposes.

Fɪɢ. 4. Inflation and Upstream RULC Proxies.

the RULC proxy, the output gap, the risk-free rate, commodity inflation, and nominal wage inflation.[8] As a preliminary exercise, I compare the robustness of the RULC and RUIC proxies with respect to alterations in this instrument set.[9] Specifically, the exercise consists of estimating the model with only two blocks of instruments (three lags of inflation and three lags of real marginal cost) and then adding blocks of instruments until the GG instrument set is reached. Doing so should tell us how each additional block of instruments affects the value and significance of the estimated coefficients of the model, and also how these added blocks affect the overall relevancy and exogeneity of the entire instrument set. We can then get a clear picture of how the GG instrument set performs with the RULC and RUIC proxies.

3.1 Estimates

Table 1 reports GMM estimates under six different instrument sets using the RULC and RUIC proxies from the full sample available, 1973:2–2002:4.[10] Instrument set 1 includes only three lags of inflation and real marginal cost. Instrument set 2 adds a

---

8. The NKPC that GG estimate is analogous to the stage N, or GDP stage, NKPC in the vertical chain of production, equation (10).

9. From this point forward I refer to the RULC and RUIC as the stage N variables, $\psi_{N,t}^{(RULC)}$ and $\psi_{N,t}^{(RUIC)}$, respectively, used to proxy real marginal cost in equation (10).

10. Estimates are based on two-stage GMM estimation using a four-lag quadratic spectral window estimate of the covariance matrix.

TABLE 1

FORWARD MODEL: GG INSTRUMENTS

Proxy	Inst. set	β	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
$\psi_N^{(RULC)}$	1	1.015	0.005 (0.661)	0.624	30.55	0.633/0.752	0.710/0.844
	2	1.007	0.006 (0.570)	0.148	23.47	0.647/0.766	0.715/0.846
	3	0.986	0.012 (0.235)	0.269	19.29	0.705/0.806	0.749/0.855
	4	0.991	0.015 (0.122)	0.393	14.77	0.716/0.810	0.761/0.861
	5	0.974	0.025 (0.001)	0.510	11.65	0.723/0.817	0.764/0.864
	6	0.953	0.031 (0.000)	0.690	10.10	0.738/0.831	0.784/0.883
$\psi_N^{(RUIC)}$	1	0.946	0.009 (0.015)	0.390	60.16	0.766/0.774	0.902/0.912
	2	0.970	0.007 (0.021)	0.095	47.85	0.783/0.787	0.909/0.915
	3	0.964	0.007 (0.008)	0.263	34.70	0.801/0.809	0.916/0.925
	4	0.952	0.009 (0.000)	0.444	25.26	0.803/0.810	0.920/0.929
	5	0.951	0.008 (0.000)	0.680	19.83	0.805/0.817	0.923/0.937
	6	0.965	0.007 (0.000)	0.714	17.03	0.815/0.830	0.924/0.941

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument sets: (1) 3 lags of ψ_N, inflation; (2) 4 lags of ψ_N, inflation; (3) 4 lags of ψ_N, inflation, output gap; (4) 4 lags of ψ_N, inflation, output gap, risk-free rate; (5) 4 lags of ψ_N, inflation, output gap, risk-free rate, commodity inflation; (6) 4 lags of ψ_N, inflation, output gap, risk-free rate, commodity inflation, wage inflation.

fourth lag of inflation and real marginal cost, and instrument sets 3 through 6 add "four-lag blocks" of variables included in the GG instrument set. Note that estimation of the NKPC using the RULC proxy in instrument set 6 is in essence a replication of GG.

The results show that estimates from the RULC proxy are more sensitive to the addition of instruments than is the RUIC proxy. Estimates of λ using the RULC proxy vary as much as six-fold, from 0.005 to 0.031, while estimates using the RUIC proxy are quite steady, varying from 0.007 to 0.009. Estimates of β using the RULC proxy range from 0.953 to 1.015, while they range from 0.946 to 0.970 using the RUIC proxy. The *p*-values of λ estimates using the RULC decrease monotonically with the addition of instrument blocks, and they are significant at the 15% level only after the risk-free rate block is added. Thus, several instrument blocks must be included in order to produce significant estimates of λ when using the RULC proxy. In contrast, estimates of λ using the RUIC proxy are all significant at the 5% level.

The fact that all estimates of λ using the RUIC proxy are statistically significant indicates that input costs are a pertinent inclusion in the NKPC model. One might expect that if there is estimable pass-through from input prices to consumer prices, this will show up in a larger estimated coefficient on the RUIC term. However, including input costs rather than wages in the marginal cost term does not increase the size of the estimated coefficient on real marginal cost. If anything, including input prices decreases the size of the coefficient, indicating that the pass-through effect is significant, but small in magnitude.

3.2 Instrument Validity and Relevance

There may be several reasons that could explain why the estimates of the RUIC are more robust to alterations in the instrument set than are the estimates of the RULC

proxy. One reason could be that the proxy itself is better in terms of measuring the actual cost faced by firms in setting prices. That is, ψ could simply fit the model better when it includes input prices and therefore be less susceptible to alterations in the instrument set. Another reason could be that the instruments themselves are more relevant and valid with regard to the model when marginal cost is proxied by RUIC. If instruments are weak or invalid, then estimates may be sensitive to the choice of instruments, as in Table 1.

There is a large literature on problems of finite-sample GMM estimation due to poor instrument choice. Instruments are considered "weak" or "irrelevant" if they do not explain enough of the variation in the variables they are supposed to instrument. Furthermore, instruments can be "invalid" if they are not exogenous to the error term. In our rational expectations scenario, good instruments are variables that are relevant, in the sense that they are used to make pricing decisions and are also in the firm's information set at time, t, and hence uncorrelated with any surprise in period $t + 1$. That is, we are looking for instruments, $Z_{n,t}$, that are correlated with the endogenous variables, $\psi_{n,t}$ and $\pi_{n,t+1}$,[11] but orthogonal to the surprise in period $t + 1$, such that:

$$E_t[(\pi_{n,t} - \lambda\psi_{n,t} - \beta\pi_{n,t+1})Z_{n,t}] = E_t[\eta_{n,t+1}Z_t] = 0. \tag{16}$$

The general choice of Z_t among researchers has been to use lagged values of variables such as those in the instrument sets shown in Table 1. We can determine whether these instruments are both relevant and valid by analyzing several statistics. Columns 5 through 8 of Table 1 report statistics that detail the exogeneity and relevance of the instrument sets. Column 5 reports the p-value of Hansen's J-statistic (1982).[12] A rejection of the null hypothesis signifies that the instruments are not satisfying the orthogonality condition, equation (14).

Table 1 shows that poor exogeneity of the instruments is not generating a problem for the NKPC with either proxy in terms of the J-statistic, as almost all of the p-values of the J-statistic (pJ-values) fall above 0.15. Beyond instrument set 2, pJ-values increase as more blocks of lagged variables are added to the instrument set.

The last three columns of Table 1 give statistics that measure the relevance of the instrument sets. Shea (1997) demonstrates that in the case of multivariate linear models with more than one endogenous regressor, the first-stage F-statistic leads to poor inference on instrument relevance. For this reason, I report the standard partial R^2 (P) as well as the Shea partial R^2 (SP) and the Cragg–Donald (1993) F-statistic (CD). The latter two statistics both take into account multiple endogenous regressors, whereas the P ignores this case.

---

11. In addition to $\pi_{n,t+1}$, we also treat $\psi_{n,t}$ as an endogenous variable. Thus, the true estimation is:

$$\pi_{n,t} = \lambda E(\psi_{n,t} \,|\, \mathcal{F}_{n,t-1}) + \beta E(\pi_{n,t+1} \,|\, \mathcal{F}_{n,t-1}).$$

12. The J-statistic is simply the GMM objective function evaluated at the second-stage estimates, distributed as χ^2 with $K_2 - K$ degrees of freedom, where K_2 is the number of instruments and K is the number of regressors.

A large SP on each endogenous variable indicates that the instruments are correlated with each distinctly.[13] However, if the SP is large on one endogenous variable but small on another, then this indicates that the instruments are *not* relevant for *both* endogenous variables, a circumstance that could ultimately bias the estimates. Table 1 shows that the SP is larger for both endogenous variables, $\pi_{N,t+1}$ and $\psi_{N,t}$, in all six instrument sets when the RUIC proxy is used instead of the RULC proxy. This reveals that the instruments used in Table 1 are more relevant to the model when real marginal cost is proxied by RUICs rather than by RULCs.

The SP works well for comparing the relevance when the number of instruments stays constant, as in comparing the same instrument sets of the upper and lower halves of Table 1. However, it is a poor indicator of the additional relevance gained by including more variables in the instrument set. That is, it is a poor judge of testing the additional relevance of moving from instrument set 1 through instrument set 6 within the use of the same proxy. Foremost, the SP does not take into account the fact that the R^2 will almost always increase with the addition of regressors. Also, it does not take into account that the estimator is more inclined to become biased as variables are added to the instrument set, as shown in Hahn and Hausman (2002). That is, the SP does not tell us whether the additional instruments are increasing the R^2 by a sufficient amount to compensate for the increase in bias that adding regressors entails.

A better statistic to measure instrument relevance when adding blocks of instruments is the CD-statistic, proposed by Stock and Yogo (2001). Stock and Yogo test for the presence of weak instruments in the case of multiple endogenous regressors by looking at the minimum eigenvalue of the matrix analog of the F-statistic from the first-stage regression.[14]

Like the SP, the CD-statistic accounts for correlation between the endogenous regressors. Unlike the SP, however, the CD-statistic takes into account the size of the instrument set and allows us to test the degree of weakness of the instrument set. A small CD-statistic implies that the instruments are weak with respect to two standards: the relative bias of the estimates with respect to their OLS counterparts and the true size of the hypothesis test $\beta = \beta_0$.[15]

---

13. The SP compares the relationship between the endogenous variables before and after they are projected on the instrument set. In the two-variable case, it measures the correlation of the orthogonal component of the projection of X_1 on X_2 with the orthogonal component of the projection of $\hat{X}_1$ and $\hat{X}_2$, where $\hat{X}_1$ and $\hat{X}_2$ are the projections of the endogenous variables on the instrument set. For analytical details see Shea (1997).

14. That is, the CD-statistic is the minimum eigenvalue of G_T:

$$G_T = \hat{\Sigma}_{vv}^{-1/2'} Y^{\perp'} P_{z^\perp} Y^\perp \hat{\Sigma}_{vv}^{-1/2} / K_2,$$

where $\hat{\Sigma}_{vv} = (Y'M_Z Y)/(T - K_1 - K_2)$. Notation is as follows: Y is the set of $T \times 1$ vectors of the endogenous regressors. Z is a $T \times K_2$ matrix of excluded instruments. $\perp$ represents the residuals from a projection on X, where X is the set of K_1 included exogenous variables. $M_X = I - P_X$, where P is the projection matrix. For more details see Stock and Yogo (2001).

15. For the table listing the critical values of the CD-statistic, refer to Appendix B.

The CD-statistics reported in Table 1 show that the added blocks of lagged variables are not increasing the relevance of the instrument set. That is, the CD-statistic of both proxies decreases monotonically with the addition of lagged blocks of variables, dropping from 30.55 to 10.10 under the RULC proxy, and from 60.16 to 17.03 under the RUIC proxy. The falling CD-statistic tells us that we are adding blocks of lagged variables that are irrelevant to the model. As Mavroeidis (2005) emphasizes, it is dangerous to throw in these irrelevant instruments, since in doing so, one causes the estimates to converge to OLS. This is just the case for the RULC proxy, as both β and λ are converging toward their OLS estimates, 0.853 and 0.039, respectively, with the addition of instruments.[16]

4. CHAIN-OF-PRODUCTION INSTRUMENTS

The above analysis reveals that, for the RULC proxy, λ becomes significant only when enough irrelevant instruments are added to the instrument set. Furthermore, λ and β change in size dramatically with the addition of these instruments. As noted by Fuhrer (2006), such differences on the estimates have great consequences for the persistence of inflation as well as for the inherent role of marginal cost in the inflation process. The fact that there is no way of telling which of the various estimates using the RULC are the true estimates poses a great problem for the econometrician. Furthermore, if we rank instrument sets by their combination of J- and CD-statistics, we should go by instrument set 1 and conclude that λ is insignificant. Thus, econometrically, the RULC proxy does not seem to fit well in the NKPC model.

On the other hand, the RUIC proxy fits for all of the instrument sets in Table 1. However, the instruments used in estimation still have problems. First, the fact that the CD-statistic still falls when we add instruments puts a damper on the use of these variables in the instrument set. If instruments are going to be added to the estimation process, we would like them to be quite relevant to the endogenous variables. If they are not, they are only clouding the true estimates of the parameters.

Second, the instruments from Table 1 include lags of inflation in the instrument set. As noted by Rudd and Whelan (2005) and Stock, Wright, and Yogo (2002), a major pitfall for the econometrician of using lags of inflation in the instrument set is that this entails describing the dynamics of the future inflation term, $\pi_{n,t+1}$, with lagged inflation terms; in essence, this means falling back on the traditional Phillips curve to help explain the forward-looking NKPC.[17] Table 2 displays the exogeneity and relevancy statistics for the six instrument sets of Table 1 excluding the inflation block. Comparing the statistics with those of Table 1, we see that the inflation block

---

16. One drawback of the CD-statistic is that it assumes spherical errors in the first-stage regression. Appendix C displays p-values of F-statistics on each instrument block for the first-stage regression after controlling for both autocorrelation and heteroskedasticity.

17. The "traditional" backward-looking Phillips curve assumes that inflation depends on lagged inflation as opposed to future expected inflation:

$$\pi_t = \beta\pi_{t-1} + \lambda\psi_t.$$

TABLE 2

FORWARD MODEL GG INSTRUMENTS: NO π BLOCK IN INSTRUMENT SET

Proxy	Inst. set	β	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
$\psi_N^{(RULC)}$	1	0.625	0.061 (0.418)	0.250	0.41	0.010/0.223	0.039/0.831
	2	0.827	0.044 (0.468)	0.120	0.32	0.011/0.222	0.041/0.831
	3	0.819	0.041 (0.031)	0.330	3.25	0.191/0.371	0.435/0.845
	4	0.924	0.025 (0.011)	0.381	4.69	0.345/0.510	0.583/0.854
	5	0.913	0.034 (0.000)	0.556	3.94	0.383/0.545	0.604/0.860
	6	0.951	0.034 (0.000)	0.454	9.79	0.675/0.774	0.762/0.873
$\psi_N^{(RUIC)}$	1	1.04	0.004 (0.691)	0.096	4.73	0.110/0.222	0.447/0.904
	2	0.895	0.012 (0.006)	0.119	6.24	0.181/0.278	0.591/0.907
	3	1.02	0.003 (0.449)	0.120	4.28	0.241/0.318	0.692/0.913
	4	0.947	0.008 (0.011)	0.309	5.60	0.391/0.460	0.784/0.923
	5	0.927	0.009 (0.000)	0.634	4.74	0.429/0.502	0.797/0.933
	6	0.945	0.008 (0.000)	0.562	12.01	0.711/0.749	0.893/0.940

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument sets: (1) 3 lags of ψ_N; (2) 4 lags of ψ_N; (3) 4 lags of ψ_N, output gap; (4) 4 lags of ψ_N, output gap, risk-free rate; (5) 4 lags of ψ_N, output gap, risk-free rate, commodity inflation; (6) 4 lags of ψ_N, output gap, risk-free rate, commodity inflation, wage inflation.

is responsible for increasing the CD-statistic, the SP, and the P for each instrument set in Table 1. Without its inclusion, most of the CD-statistics would be well below the 10% bias and 25% size thresholds of Stock and Yogo's (2001) critical values.

Finally, many of the variables used in the GG instrument set are found outside of the firm's pricing decision problem, equation (5), and could be considered an ad hoc addition to the estimation process. That is, it is not clear in the model that the instruments used by GG are the relevant state variables in which firms are basing their pricing decisions. Ultimately, this issue gives less credence to the estimates obtained through GMM.

Given these issues, we would like to find an instrument set that contains a smaller set of variables that are micro-founded within the model, contain relevant information, are theoretically exogenous to price setters, and do not necessitate inflation in the instrument set. The chain-of-production NKPC has the potential to provide such an instrument set in that it allows for a much richer set of instruments than those shown in Table 1.

Since the chain-of-production model incorporates early (i.e., upstream) stages of production in the downstream marginal cost term, RUICs from stages low in the production chain should be correlated with downstream marginal costs. This can be seen by combining equations (6), (7), and (14):

$$\psi_{n,t}$$

$$= \left[\overbrace{\left(\underbrace{\beta\theta_{n-1} E_t\left(p_{n-1,t+1}^*\right) + (1-\beta\theta_{n-1})(\psi_{n-1,t} + \bar{p}_{n-1,t})}_{\bar{p}_{n-1,t}} \right)(1-\theta_{n-1}) + \theta_{n-1}\bar{p}_{n-1,t-1}}^{p_{n-1,t}^*} \right] + \cdots$$

$$\cdots + \bar{y}_{n-1,t} - \bar{p}_{n,t} - \bar{y}_{n,t}, \tag{17}$$

where the marginal cost of the downstream firm, $\psi_{n,t}$, is a function of the marginal cost of the upstream firm, $\psi_{n-1,t}$. Similarly, $\psi_{n-1,t}$ is a function of $\psi_{n-2,t}$, and so on. In general, $\psi_{n,t}$ is a function of all the marginal costs of stages below it:

$$\psi_{N,t} \equiv \psi_{N,t}(\psi_{N-1,t}(\psi_{N-2,t}(\cdots(\psi_{2,t}(\psi_{1,t}), \phi_{2,t})\cdots), \phi_{N-2,t}), \phi_{N-1,t}), \phi_{N,t}), \quad (18)$$

where $\phi_{n,t}$ is the set of the other variables included in the marginal cost proxy, $\{p^*_{n-1,t+1}, \bar{p}_{n-1,t}, \bar{p}_{n-1,t+1}, \bar{p}_{n-1,t-1}, \bar{y}_{n-1,t}, \bar{p}_{n,t}, \bar{y}_{n,t}\}$. Plugging (18) into (10), we are back to the NKPC, and we see that upstream RUICs should therefore also be correlated with downstream inflation and should therefore be relevant instruments.

Furthermore, RUICs of firms in early stages of production should be exogenous to the price-setting decisions of firms in later stages. We can make this claim since early stages of production include a large proportion of crude-material and energy-producing firms—firms that are known to react quickly to exogenous factors such as geo-political, weather, and technological shocks. Since upstream RUICs are linked with these firms' pricing decisions, upstream RUICs should carry the exogenous shocks with them. An example is the airline industry, where spot prices on crude oil fluctuate independently of how commercial airlines price airline fares. Here we generalize this example and claim that upstream RUIC variables fluctuate independently with respect to the pricing decisions of downstream firms.

It is important to note that the relationship in (18) cannot be constructed through the model with the RULC proxy. This is because the model does not assume that different stages of production are interconnected through wages and labor in the production chain as are input prices. Rather, the model assumes that labor between stages is immobile and that each stage n uses its own labor force, L_n, and wage structure, W_n. Thus, the RULC proxy does not come packaged with its own instrument set as does the RUIC proxy. However, recall that we do know that the RULC and RUIC proxies are approximately equivalent since they are derived from the same cost minimization problem of the firm. If we use the approximation from (11) and (13), $\psi_n^{(RULC)} \approx \psi_n^{(RUIC)}$, we can make the same claim regarding relevance and exogeneity of the RULC proxy as we did for the RUIC proxy.[18] That is, through their equivalent approximation to the RUIC variables, we can claim that upstream RULC variables are also relevant and exogenous to downstream RULC terms. The fact that this claim requires an extra assumption already puts the RUIC and its upstream instruments in a more desirable position with regard to estimating the NKPC. However, as opposed to simply throwing out the RULC proxy, it seems reasonable to apply the assumption and provide a direct comparison to the GG instrument set as well as to the RUIC proxy and its upstream instruments.

4.1 Estimation with the Chain-of-Production Instruments

In order to compare the chain-of-production instruments to the GG instruments, I perform the same exercise as in Table 1, where I add blocks of four lags of variables

---

18. A regression of $\psi_n^{(RUIC)}$ on $\psi_n^{(RULC)}$ produces coefficients that are significant at the 1% level for stages N, $N-1$, and $N-2$, signifying that the proxies of the same stage are correlated.

TABLE 3

FORWARD MODEL: CHAIN INSTRUMENTS

Proxy	Inst. set	β	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
$\psi_N^{(RULC)}$	1	0.827	0.044 (0.468)	0.120	0.32	0.011/0.222	0.041/0.831
	2	0.933	0.021 (0.109)	0.510	18.27	0.573/0.676	0.720/0.849
	3	0.968	0.018 (0.049)	0.668	20.29	0.703/0.776	0.776/0.857
	4	0.965	0.022 (0.001)	0.391	17.70	0.752/0.828	0.785/0.865
	5	0.957	0.024 (0.000)	0.619	15.78	0.784/0.854	0.799/0.871
	6	0.946	0.027 (0.000)	0.734	13.08	0.790/0.857	0.805/0.873
	7	0.937	0.035 (0.000)	0.826	11.09	0.795/0.861	0.809/0.876
	8	0.932	0.035 (0.000)	0.922	10.49	0.816/0.881	0.829/0.895
$\psi_N^{(RUIC)}$	1	0.895	0.012 (0.006)	0.119	6.24	0.181/0.278	0.591/0.907
	2	0.927	0.010 (0.000)	0.507	27.91	0.647/0.704	0.874/0.915
	3	0.927	0.010 (0.000)	0.818	28.82	0.766/0.792	0.892/0.923
	4	0.943	0.009 (0.000)	0.527	30.55	0.829/0.840	0.916/0.928
	5	0.940	0.009 (0.000)	0.624	32.12	0.868/0.878	0.924/0.935
	6	0.941	0.009 (0.000)	0.815	26.58	0.872/0.882	0.928/0.939
	7	0.943	0.009 (0.000)	0.826	23.23	0.879/0.892	0.938/0.952
	8	0.938	0.009 (0.000)	0.923	21.02	0.889/0.903	0.939/0.956

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. p-values are in parentheses. Instrument sets: (1) 4 lags of ψ_N; (2) 4 lags of ψ_N, ψ_{N-1}; (3) 4 lags of ψ_N, ψ_{N-1}, ψ_{N-2}; (4) 4 lags of ψ_N, ψ_{N-1}, ψ_{N-2}, inflation; (5) 4 lags of ψ_N, ψ_{N-1}, ψ_{N-2}, inflation, output gap; (6) 4 lags of ψ_N, ψ_{N-1}, ψ_{N-2}, inflation, output gap, risk-free rate; (7) 4 lags of ψ_N, ψ_{N-1}, ψ_{N-2}, inflation, output gap, risk-free rate, commodity inflation; (8) 4 lags of ψ_N, ψ_{N-1}, ψ_{N-2}, inflation, output gap, risk-free rate, commodity inflation, wage inflation.

to the instrument set. Table 3 displays the results of the estimation of the NKPC, utilizing the chain-of-production instruments (i.e., ψ_{N-1}, ψ_{N-2}). Instrument set 1 includes only four lags of the real marginal cost proxy and serves as a base for comparison. Instrument sets 2 and 3 add blocks of real marginal cost proxies from stages $N-1$ and $N-2$, respectively. Beyond instrument set 3, we add instrument blocks from Table 1.

Estimates of λ using the RULC proxy are significant at the 11% level for all instrument sets beyond instrument set 1. Estimates of λ still lie in a wide range, from 0.018 to 0.035, when upstream real marginal costs are included in the instrument set, but there is now a two-fold difference as opposed to a six-fold difference. Similarly, estimates of β are now more stable as well, ranging from 0.932 to 0.968. All estimates of β and λ are significant at the 1% level with the RUIC proxy, and they are still stable with respect to the instrument set, with λ ranging from 0.009 to 0.010, in instrument sets 2 through 8.

Unlike the case shown in Table 1, where adding blocks of instruments decreased the CD-statistic, here the addition of upstream real marginal costs to the instrument set (instrument sets 2 and 3) increases the CD-statistic for both proxies. Actually, both the CD-statistic and the pJ-value increase with the addition of the ψ_{N-1} and ψ_{N-2} blocks, showing that the upstream marginal cost terms are relevant as well as valid instruments for both proxies.

A noteworthy feature of the chain instruments is that they are producing relatively high CD- and low J-statistics (high pJ-values), *without* the use of the inflation block. The fact that CD-statistics are still high, without resorting to the use of lags of inflation

TABLE 4
FORWARD MODEL—GGL INSTRUMENT SET COMPARISON

Proxy	Inst. set	β	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
$\psi_N^{(RULC)}$	1	1.05	0.004 (0.728)	0.184	18.85	0.502/0.642	0.663/0.848
	2	1.00	0.012 (0.193)	0.137	20.99	0.683/0.804	0.731/0.860
	3	0.941	0.018 (0.109)	0.468	36.02	0.661/0.740	0.740/0.829
	4	0.979	0.016 (0.110)	0.174	27.05	0.737/0.821	0.759/0.845
$\psi_N^{(RUIC)}$	1	1.00	0.006 (0.155)	0.272	19.68	0.512/0.553	0.840/0.907
	2	0.982	0.008 (0.025)	0.184	40.29	0.791/0.803	0.898/0.913
	3	0.902	0.012 (0.000)	0.421	58.36	0.754/0.784	0.875/0.909
	4	0.902	0.012 (0.000)	0.447	48.04	0.817/0.834	0.898/0.916

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. p-values are in parentheses. Instrument sets: (1) 2 lags ψ_N, output gap, wage inflation; (2) 4 lags of inflation, 2 lags ψ_N, output gap, wage inflation; (3) 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}; (4) 4 lags of inflation, 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}.

as instruments, tells us that the upstream marginal costs are taking over the relevance responsibilities of the inflation block. In fact, they are doing so while simultaneously increasing the exogeneity of the instrument set. This is revealed by the high CD-statistic and pJ-value on instrument set 3 in Table 3.[19] Once we add four lags of inflation to the instrument set (moving from instrument sets 3 to 4), there is a fall in the pJ-value for both proxies. Similar to the case in Table 1, each additional block beyond the inflation block lowers the CD-statistic for the RULC proxy. For the RUIC proxy, the CD-statistic reaches a peak with addition of the output gap block and then begins to fall back to a lower value than it reached at instrument set 3.[20]

4.2 Smaller Instrument Sets

As researchers using GMM have conceded, use of a large instrument set induces bias. Therefore, smaller instrument sets have been devised and are being used in current GMM estimation. Galí, Gertler, and Lopez-Salido (2005, henceforth GGL), for instance, propose using four lags of inflation, two lags of real marginal cost, the output gap, and wage inflation.

Table 4 reports estimates of the forward-looking model, comparing estimates produced using the GGL instruments with estimates produced using the upstream production-chain instruments. Instrument set 1 does not include the inflation block, while instrument set 2 replicates the instrument set used by GGL. Instrument sets 3 and 4 replace the output gap and the wage inflation blocks with the upstream instruments. This replacement leaves all instrument sets with the same number of total instruments, making it easy to discern the relevancy and exogeneity on a per instrument basis.

19. Note that instrument set 3 in Table 1 and instrument set 3 in Table 3 use the same number of instruments, and we can therefore compare the two instruments sets on a per instrument basis. The chain instruments increase the pJ-value from 0.269 to 0.668 for the RULC proxy, and from 0.263 to 0.818 for the RUIC proxy.

20. For more detail on this issue, Appendix D displays pF-values on each instrument block for the first-stage regression after controlling for both heteroskedasticity and autocorrelation.

The results show that in all cases, the instrument sets that include the upstream variables instead of the output gap and wage inflation variables lead to larger SP, CD, and pJ-values. That is, comparing instrument set 1 with instrument set 3, and comparing instrument set 2 with instrument set 4, the production-chain instruments are outperforming the output gap and wage inflation variables in terms of relevancy with regard to the SP and CD values in terms of exogeneity with regard to the J-statistic.

The RUIC proxy is highly significant when upstream instruments are used, with both p-values at 0.000. On the flip side, the p-values of the coefficient on the RULC proxy never dip below 10%, even with the use of the upstream instruments. Furthermore, the coefficients of both β and λ on the RUIC proxy do not change with the addition of the inflation block when upstream instruments are used. These results reiterate both the higher significance of the RUIC proxy over the RULC proxy and the stability of the upstream instruments with respect to the addition of the inflation block.[21]

5. THE HYBRID MODEL

Due to some shortcomings of the purely forward-looking NKPC, it has become conventional to include a lagged inflation term in the model, creating the Hybrid Philips Curve:[22]

$$\pi_{N,t} = \lambda_n \psi_{N,t} + \gamma_f E_t(\pi_{N,t+1}) + \gamma_l \pi_{N,t-1}. \tag{19}$$

Although a lagged inflation term is considered an ad hoc addition to the model, many researchers such as Fuhrer and Moore (1995) and Blanchard and Galí (2005) have searched for micro-foundations for including it in the NKPC.[23] Because of the prominence of the hybrid model in the literature, it is important that the chain of production, and the instruments and proxy that stem from it, apply to the hybrid model as well as to the purely forward-looking model.

5.1 GG Instrument Estimation of Hybrid Model

Table 5 reports estimates of the hybrid model, repeating the same exercise performed on the purely forward-looking model displayed in Table 1. At first glance, we

---

21. Interested readers can find results of the nonlinear GMM estimates of β and θ_N in Appendix E. Under the upstream instrument set, θ_N varies from 0.894 to 0.936 under the RUIC, depending on the subsample, while θ_N varies from 0.805 to 0.969 for the RULC. The implied results on λ do not differ from linear GMM estimation.

22. Most of the shortcomings occur when real marginal cost is proxied with the output gap, which is the conventional measure of economic activity used in the traditional Phillips curve. In such a case, the purely forward-looking model implies no trade-off between disinflation and output. Also, in such a case, the output gap proxy implies a *negative* coefficient on real marginal cost, contradicting the model. For more on this issue see GG.

23. Also see Gumbau-Brisá (2005) who incorporates beliefs and imperfect knowledge into the model in order to generate a backward-looking term.

TABLE 5

HYBRID MODEL—GG INSTRUMENTS

Proxy	Inst. set	γ_l	γ_f	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
$\psi_N^{(RULC)}$	1	0.432	0.556	0.002 (0.651)	0.350	1.82	0.075/0.075	0.760/0.762
	2	0.630	0.335	0.007 (0.225)	0.272	2.04	0.115/0.115	0.763/0.765
	3	0.497	0.487	0.004 (0.229)	0.235	3.48	0.267/0.267	0.779/0.779
	4	0.471	0.515	0.005 (0.078)	0.509	2.66	0.283/0.283	0.787/0.787
	5	0.459	0.524	0.007 (0.012)	0.510	2.30	0.310/0.309	0.793/0.793
	6	0.425	0.548	0.012 (0.000)	0.687	2.32	0.363/0.362	0.823/0.822
$\psi_N^{(RUIC)}$	1	0.172	0.789	0.007 (0.018)	0.249	3.63	0.143/0.159	0.800/0.881
	2	0.449	0.538	0.002 (0.197)	0.131	3.46	0.187/0.196	0.840/0.884
	3	0.485	0.503	0.002 (0.036)	0.261	3.50	0.274/0.279	0.881/0.897
	4	0.446	0.540	0.002 (0.000)	0.391	2.53	0.278/0.283	0.888/0.902
	5	0.447	0.539	0.002 (0.001)	0.634	2.20	0.305/0.307	0.907/0.914
	6	0.443	0.543	0.002 (0.000)	0.669	2.20	0.355/0.355	0.918/0.919

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument sets: (1) 3 lags of ψ_N, inflation; (2) 4 lags of ψ_N, inflation; (3) 4 lags of ψ_N, inflation, output gap; (4) 4 lags of ψ_N, inflation, output gap, risk-free rate; (5) 4 lags of ψ_N, inflation, output gap, risk-free rate, commodity inflation; (6) 4 lags of ψ_N, inflation, output gap, risk-free rate, commodity inflation, wage inflation.

see that estimates of λ are much smaller than those in the forward model, with the largest estimates reaching 0.012 and 0.007 for the RULC and RUIC proxies, respectively. Like estimates obtained with the purely forward-looking model, estimates of λ using the RULC proxy gain significance with the addition of the instruments. That is, *p*-values dip below 0.100 only after instrument set 4 is reached. Meanwhile, all of the estimates of λ using the RUIC proxy are significant at the 5% level, except for instrument set 2. Also, as with the forward model, estimates obtained using the RULC proxy vary as much as six-fold, from 0.002 using instrument set 1 to 0.012 using instrument set 6, while estimates of λ using the RUIC proxy are much less sensitive to the instrument set, varying from 0.007 with instrument set 1 to a steady 0.002 with all of instrument sets 2 through 6. Estimates of γ_l are significant using both proxies for all instrument sets. However, estimates of γ_f are larger than estimates of γ_l for all instrument sets with the RUIC proxy.

It is apparent from Table 5 that relevancy is a problem for the GG instrument set in the hybrid model, as all of the relevancy statistics are much smaller than they were with the forward model. Specifically, for both proxies, CD-statistics are all very low, lying below 3.63. Note that this is below the 4.32 critical CD value that allows 30% bias. SP- and P-values on the forward inflation term, $\pi_{N,t+1}$, are also very small, never exceeding 0.363. These low relevancy statistics are due to the fact that $\pi_{N,t-1}$ is considered an exogenous variable and is therefore included as an instrument but not taken into account in measuring instrument relevancy.[24] The excluded instruments therefore need to be correlated with the endogenous variables, over and above $\pi_{N,t-1}$. Thus, the smaller relevancy statistics are in large part a result of the fact that much of

24. Recall that CD, SP, and P measure the *extra* relevance of excluded instruments over the included instruments.

TABLE 6

HYBRID MODEL—GGL AND CHAIN INSTRUMENTS

Proxy	Inst. set	γ_l	γ_f	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$.
$\psi_N^{(RULC)}$	1	0.266	0.727	0.008 (0.256)	0.047	2.80	0.130/0.130	0.774/0.772
	2	0.403	0.590	0.006 (0.256)	0.108	4.21	0.260/0.260	0.786/0.786
	3	0.343	0.645	0.004 (0.551)	0.517	6.81	0.270/0.271	0.746/0.749
	4	0.412	0.588	0.002 (0.665)	0.225	5.58	0.321/0.323	0.761/0.764
$\psi_N^{(RUIC)}$	1	0.143	0.832	0.006 (0.110)	0.073	1.68	0.084/0.102	0.722/0.877
	2	0.345	0.648	0.003 (0.009)	0.131	3.90	0.252/0.257	0.862/0.880
	3	0.256	0.717	0.006 (0.002)	0.422	8.52	0.318/0.326	0.855/0.877
	4	0.381	0.605	0.003 (0.004)	0.313	6.87	0.371/0.372	0.883/0.885

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. p-values are in parentheses. Instrument sets: (1) 2 lags ψ_N, output gap, wage inflation; (2) 4 lags of inflation, 2 lags ψ_N, output gap, wage inflation; (3) 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}; (4) 4 lags of inflation, 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}.

the relevancy of the instrument set is coming from the first lag of inflation, and not from the excluded instruments.

5.2 Estimation of the Hybrid Model Using Chain-of-Production Instruments

Table 6 replicates the GGL instruments and production-chain instruments used in Table 4. Two similarities emerge between Tables 5 and 6 concerning the values of the estimates of γ_l, γ_f, and λ. Like the GG instruments in Table 5, the chain instruments produce estimates of γ_f that are larger than γ_l. Also, estimates of λ using the RUIC proxy are all significant, whereas the estimates of λ using the RULC proxy are all insignificant.[25]

Like the purely forward-looking model estimates, the chain instruments produce better relevancy and validity statistics than do the GGL instruments. All CD-statistics are larger when production-chain instruments are used. It reaches its highest value when the upstream instruments are used without the inflation block, instrument set 3. Furthermore, the upstream real marginal costs increase the relevancy of the instrument set toward the future inflation term according to the SP and P for both proxies. For the RUIC proxy, the upstream variables increase these values on the future inflation term and the marginal cost term. The upstream instruments also improve the validity of the instrument set for the hybrid model, as seen by the large pJ-values on instrument sets 3 and 4.

6. SUBSAMPLE ANALYSIS

Many researchers have argued that estimates of the NKPC are sensitive to the change in the Federal Reserve's operating policy during the Volcker years of the early

---

25. Imposing the homogeneity condition $\gamma_f + \gamma_l = 1$, as advocated by Mavroeidis (2005), did not change the results.

TABLE 7

FORWARD MODEL: SUBSAMPLES

Proxy	Inst. set	β	λ	pJ	CD	$SP/P\pi_{t+1}$	$SP/P\psi_n^{(i)}$.
			1973:2–1989:4				
$\psi_N^{(RULC)}$	1	1.03	0.052 (0.030)	0.599	17.51	0.630/0.631	0.740/0.741
	2	0.998	0.047 (0.013)	0.477	13.48	0.715/0.731	0.771/0.789
	3	0.994	0.044 (0.045)	0.577	20.16	0.672/0.664	0.718/0.710
	4	0.921	0.045 (0.002)	0.485	14.12	0.722/0.733	0.765/0.777
$\psi_N^{(RUIC)}$	1	0.964	0.015 (0.001)	0.557	16.71	0.630/0.623	0.901/0.891
	2	0.977	0.013 (0.000)	0.686	14.49	0.744/0.734	0.903/0.890
	3	1.00	0.013 (0.001)	0.366	22.84	0.705/0.697	0.894/0.884
	4	0.978	0.013 (0.000)	0.636	15.51	0.762/0.751	0.898/0.885
			1980:1–2002:4				
$\psi_N^{(RULC)}$	1	1.26	−0.026 (0.106)	0.136	6.53	0.317/0.548	0.496/0.856
	2	1.01	0.008 (0.367)	0.173	13.29	0.644/0.807	0.692/0.866
	3	1.07	−0.003 (0.738)	0.691	14.14	0.503/0.659	0.643/0.843
	4	1.01	0.006 (0.438)	0.327	13.74	0.648/0.794	0.701/0.860
$\psi_N^{(RUIC)}$	1	1.12	0.000 (0.996)	0.079	3.62	0.203/0.557	0.545/0.942
	2	0.951	0.008 (0.000)	0.285	13.38	0.481/0.767	0.602/0.960
	3	0.903	0.011 (0.006)	0.053	8.92	0.387/0.721	0.510/0.949
	4	0.964	0.008 (0.001)	0.264	12.10	0.600/0.817	0.704/0.958

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument sets: (1) 2 lags ψ_N, output gap, wage inflation; (2) 4 lags of inflation, 2 lags ψ_N, output gap, wage inflation; (3) 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}; (4) 4 lags of inflation, 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}.

1980s. Specifically, many have noted a fall in the magnitude of the coefficient on real marginal cost in the period after the Volcker disinflation, when using the RULC proxy.[26] Although it is very likely that the coefficients are not perfectly stable over time, large movements in the magnitude or significance of the coefficients when the subsample is altered do not bode well for the NKPC. I therefore check the stability of the forward-looking and hybrid model under two subsamples and examine a 70-quarter-window rolling regression of the forward-looking model.

6.1 Forward Model Subsamples

Table 7 reports estimates of the forward NKPC over two overlapping subsamples, 1973:2–1989:4 and 1980:1–2002:4, using the same instrument sets as used in the estimates in Tables 4 and 6. These subsample periods mimic the GG sample periods, making it easy to compare with their results, and they also roughly coincide with Paul Volcker's arrival in 1979 and departure in 1987 as chairman of the Federal Reserve. Similar to results found in the prior literature using the RULC proxy, these results show a fall in λ in the later subsample. Instrument set 1, for instance, shows a drop from 0.052 to −0.026. Similarly, the GGL instrument set and the upstream instruments suggest that λ falls to an insignificant value in the second subsample.

---

26. Eichenbaum and Fisher (2003) produce an estimate of λ of 0.064 in the sample period 1959:1–1979:2 and 0.041 in the sample period 1982:3–2001:4, using nonlinear GMM on the forward model. GG give an estimate of λ of 0.07 in the sample period 1970:1–1989:4, and 0.051 in the sample period 1980:1–1997:4, using nonlinear GMM on the hybrid model.

The RUIC proxy does not produce such a large fall in λ. Excluding instrument set 1, the largest drop in the magnitude of λ is 0.005, and this decline occurs under instrument set 4. The upstream instruments alone (instrument set 3) show a change of only 0.002. The estimates of λ obtained using the RUIC proxy are all significant and positive at the 1% level in both samples. Estimates of λ obtained using the RULC proxy, however, are significant and positive only in the first subsample. In the second subsample, estimates of λ are highly insignificant in all cases except with instrument set 1, where the estimate is negative and significant at the 11% level.

The RUIC and RULC proxies tell different stories about the structural significance of the marginal cost term in the two subsamples. The RULC proxy suggests that the marginal cost term was once a significant component in the NKPC but has become insignificant. The RUIC proxy, however, implies that the relationship between inflation and marginal cost has not changed much between time periods.

It is apparent from Table 7 that the model holds up better in the earlier period 1973:2–1989:4 than in the later period 1980:1–2002:4. That is, comparing the instrument sets of the same proxy in each subsample, we see higher pJ-values, CD-statistics, and SPs in the first subsample. A large reason for the lackluster relevancy performance of the instruments in the second subsample, particularly the RUIC instruments, is related to the large discrepancy between the SP and the P of the first-stage regression of the endogenous variables, $\pi_{n,t+1}$ and $\psi_{n,t}$. This high discrepancy is due in large part to high correlation between the endogenous regressors, $\pi_{n,t+1}$ and $\psi_{n,t}$, in the second subsample.[27] A higher correlation of the endogenous regressors will lower the CD-statistic and SP-value, as these two statistics factor in the correlation among the regressors in measuring relevancy.[28] The P does not factor this into account, hence, the relatively high values in the second subsample.

Furthermore, the low pJ-values in the second subsample are due entirely to poor exogeneity on behalf of lagged RUIC.[29] The main point here is that it is not necessarily the case that the upstream instruments are less relevant nor less exogenous in the second subsample than in the first. Rather, the high correlation of the endogenous variables and the poor exogeneity of lagged values of the RUIC proxy in the instrument set are causing the CD-statistic, and the SP- and pJ-values to fall in the second subsample.

27. A regression of $\pi_{n,t+1}$ on $\psi_{n,t}$ in the first subsample produces R^2 values of 0.047 and 0.020 for the RULC and RUIC proxies, respectively. In the second subsample, the R^2 values are 0.333 and 0.532, respectively.

28. The SP for instance, does a "before and after instrumenting" analysis of the relationship between the explanatory variables. In our case, it compares the relationship between $\psi_{n,t}$ and $\pi_{n,t+1}$ with the relationship between $\hat{\psi}_{n,t}$ and $\hat{\pi}_{n,t+1}$, the projections of the endogenous variables on the instrument set. If the relationship between the endogenous regressors changes after they are both projected on the instrument set, the SP will fall. If there is no correlation between the endogenous variables to begin with, the SP will not differ much from the P. However, if the two endogenous regressors are correlated, as they are in the second subsample, then the SP will be more sensitive to alterations in the instrument set and will differ substantially from the P.

29. Refer to Appendix F for a more detailed demonstration.

TABLE 8

HYBRID MODEL—SUBSAMPLES

Proxy	Inst. set	γ_l	γ_f	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
				1973:2–1989:4				
$\psi_N^{(RULC)}$	1	0.288	0.724	0.023 (0.176)	0.467	2.00	0.170/0.227	0.530/0.706
	2	0.452	0.547	0.007 (0.500)	0.536	2.38	0.280/0.316	0.668/0.752
	3	0.247	0.771	0.021 (0.249)	0.470	2.29	0.195/0.269	0.493/0.680
	4	0.362	0.654	0.014 (0.221)	0.533	2.30	0.276/0.321	0.634/0.740
$\psi_N^{(RUIC)}$	1	0.365	0.623	0.004 (0.208)	0.398	2.03	0.173/0.236	0.650/0.885
	2	0.443	0.560	0.002 (0.030)	0.606	2.34	0.283/0.317	0.780/0.885
	3	0.140	0.859	0.010 (0.005)	0.400	3.38	0.259/0.317	0.716/0.879
	4	0.336	0.671	0.005 (0.033)	0.580	2.83	0.353/0.366	0.783/0.880
				1980:1–2002:4				
$\psi_N^{(RULC)}$	1	0.396	0.546	0.011 (0.094)	0.022	1.45	0.094/0.106	0.693/0.779
	2	0.383	0.587	0.007 (0.177)	0.100	3.53	0.282/0.305	0.733/0.794
	3	0.151	0.904	−0.003 (0.587)	0.662	1.80	0.115/0.120	0.728/0.762
	4	0.404	0.600	0.000 (0.980)	0.335	2.94	0.246/0.260	0.744/0.782
$\psi_N^{(RUIC)}$	1	0.257	0.698	0.006 (0.004)	0.045	1.52	0.097/0.101	0.844/0.861
	2	0.333	0.614	0.005 (0.001)	0.143	3.63	0.291/0.291	0.880/0.882
	3	0.352	0.621	0.004 (0.094)	0.061	4.08	0.226/0.244	0.803/0.868
	4	0.354	0.609	0.004 (0.024)	0.211	4.33	0.325/0.341	0.853/0.893

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument sets: (1) 2 lags ψ_N, output gap, wage inflation; (2) 4 lags of inflation, 2 lags ψ_N, output gap, wage inflation; (3) 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}; (4) 4 lags of inflation, 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}.

6.2 Hybrid Model Subsamples

Table 8 reports estimates of the hybrid model over the same two overlapping subsamples used in the forward-model estimation. Again, we look at the same four instrument sets.

Estimates of γ_f are again all larger than estimates of γ_l, indicating that forward-looking behavior is important in both subsamples. Also, it is clear that the RUIC proxy outperforms the RULC proxy in terms of significance of λ. The RULC proxy is insignificant in both subsamples when upstream instruments are used in the instrument set. The RUIC proxy is insignificant only in the first subsample using the GGL set without the inflation block.

Also, similar to the forward-looking model subsamples, estimates of λ using the RUIC proxy remain relatively stable between subsamples. That is, there is little variation between the estimates of λ in the first subsample compared with the estimates of λ in the second when marginal cost is proxied by RUIC. There is clearly more variation between subsamples using the RULC proxy. Also, like the forward-looking model estimates, the hybrid model subsample estimates show a uniform fall in the value of λ when we move from the first subsample to the second.

A disconcerting feature of the hybrid-model subsample estimation is the inability of the upstream instruments to replicate the relatively large relevancy statistics they produce in the full sample hybrid estimation. That is, the CD-statistics stay at the same low levels achieved in the full sample using the GG instruments. The upstream
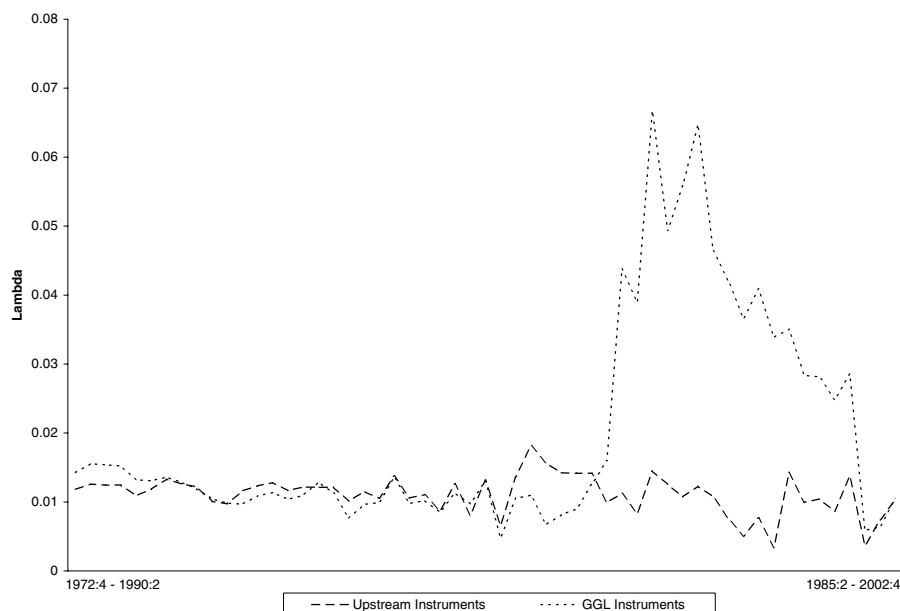
FIG. 5. Subsample Rolling Regression: RUIC Proxy.

NOTE: Intervals of 70 quarters were estimated. GGL instruments: Instrument set 1 of Table 4. Upstream instruments: Instrument set 3 of Table 4.

instruments still generally outperform the GGL instrument sets, but poor relevancy is still a problem for subsample hybrid NKPC estimation. The largest CD-statistic found in Table 8 is 4.33, which lies below the 30% allowed bias and well below the 25% max size in Stock and Yogo's (2001) critical values table. It is therefore hard to get reliable subsample statistics for the hybrid model, even when using the upstream instruments.

The low relevancy statistics and numerous insignificant estimates of λ obtained using the RULC proxy show that it is very difficult to uphold the hybrid model in the subsample setting. Relatively large relevancy statistics are associated with significant values of λ when proxying with RUIC; however, this alone is not enough to lead us to conclude that the model is holding well for either proxy in the hybrid subsample case.

6.3 Subsample Rolling Regression

In order to get a broader picture of the stability of λ with respect to subsample estimation, rolling estimates were performed on the forward model in spans of 70 quarters.[30] Figure 5 displays rolling estimates of λ, proxying marginal cost with

30. Neither altering the window size nor looking at estimates of the hybrid model changed the relative results.
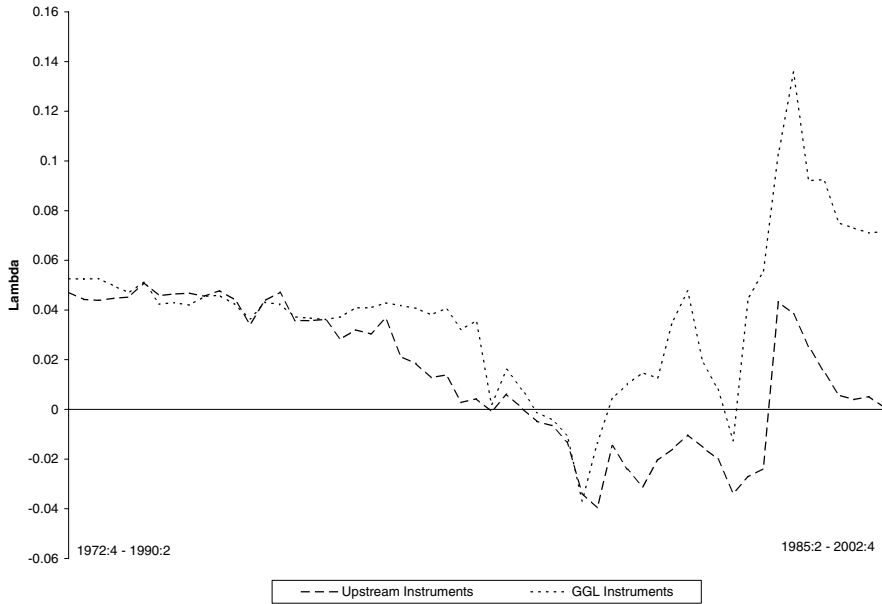
FIG. 6. Subsample Rolling Regression: RULC Proxy.

NOTES: Intervals of 70 quarters were estimated. GGL instruments: Instrument set 1 of Table 4. Upstream instruments: Instrument set 3 of Table 4.

RUIC under two different instrument sets: the GGL instrument set without inflation and the upstream instrument set consisting of two lags of stage N, $N-1$, and $N-2$ RUIC terms (instrument sets 1 and 3 of Table 7). Similarly, Figure 6 displays rolling estimates of λ under the same two instrument sets but proxying marginal cost with RULC.

Figures 5 and 6 show that the estimated coefficient of λ is generally more stable under the upstream instruments than under the GGL set. This is particularly true for the RUIC proxy. The RULC estimates are relatively unstable under both instrument sets. Of all four plots in these two figures, λ displays the least variation under the GMM estimation that uses the RUIC proxy and its upstream instruments.[31]

Figure 7 displays 95% confidence bands of the rolling estimates for both proxies under the upstream instruments. The band under the RUIC proxy is relatively stable and generally lies above zero. In contrast, the size of the band under the RULC proxy is much more volatile and dips below zero for most of the estimates.

These three figures reiterate the subsample results obtained in the preceding subsections. That is, estimation of λ under the RULC proxy is very sensitive to the sample

31. A measure of the stability of λ with respect to the subsample can be obtained by comparing the variance of the estimates obtained in Figures 5 and 6, which are as follows: GGL RULC: 9.2 e-4, Upstream RULC: 7.7 e-4, GGL RUIC: 2.5 e-4, Upstream RUIC: 7.8 e-6.
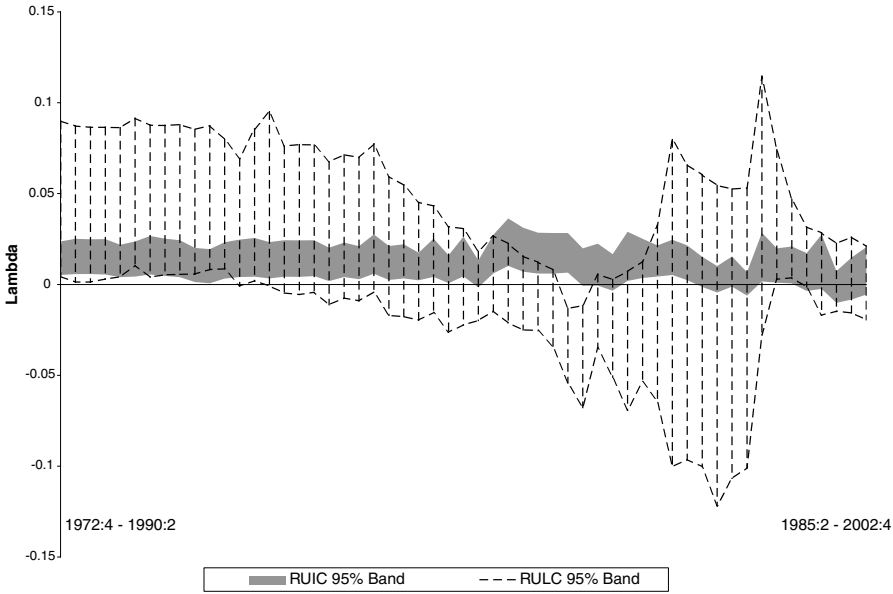
FIG. 7. Subsample Rolling Regression: Upstream Instruments—95% Confidence Intervals.

NOTES: Intervals of 70 quarters were estimated. Both estimates were taken using the upstream instrument set 3 of Table 4.

period, whereas estimates of λ under the RUIC proxy are quite robust. Furthermore, this result is bolstered by using the upstream instruments, which help to stabilize the estimates with respect to variations in the sample period.[32]

7. FUNDAMENTAL INFLATION SIMULATION

As a final exercise, I compare the goodness of fit of the RULC and RUIC proxies. That is, I construct the "fundamental" inflation series based on the model's solution:

$$\pi^*_{N,t} = \lambda_N \sum_{i=0}^{\infty} \beta^i E(\psi_{N,t+i} \mid \mathcal{F}_{N,t}), \tag{20}$$

where expectations are taken as conditional on $\mathcal{F}_{N,t}$, the information of the firm at time t.

---

32. In addition to subsample analysis, I performed two other robustness checks. First, I ran the same GMM regressions with inflation measured as both the core CPI and the core PCE. The results found in the paper do not change. Second, I ran regressions that included current and lagged upstream RUICs as dependent variables. The coefficients on these variables show up as insignificant. Refer to Appendices G and H for the results of these robustness checks.
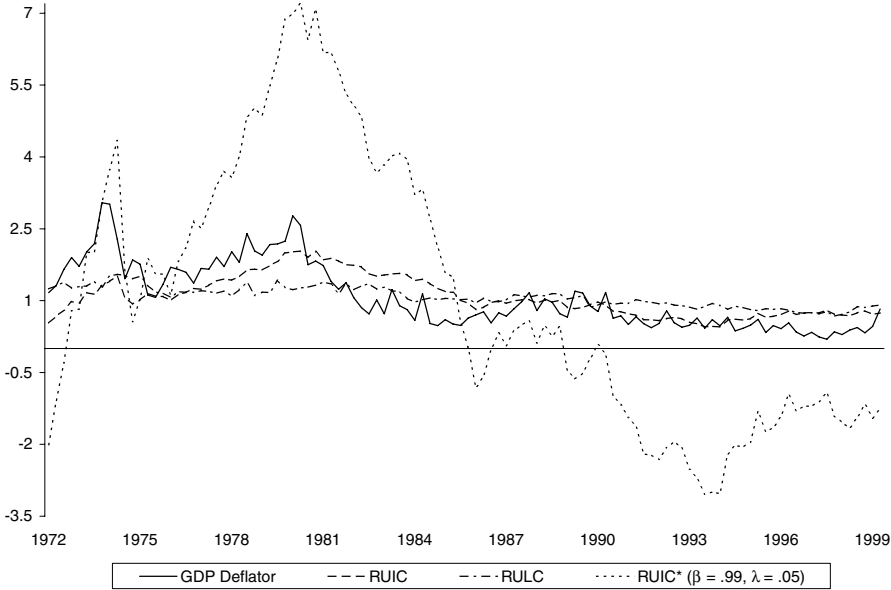
Fig. 8. Inflation: Actual versus Fundamental with RULC and RUIC.

NOTES: Fundamental inflation is constructed using instrument set 3 of Table 4 as a proxy for the information set. K is set to 12 and the parameters β and λ are taken from the estimates in Table 4. RUIC and RULC indicate the series constructed using real unit input cost and real unit labor cost as the marginal cost proxy, respectively. RUIC* represents fundamental inflation constructed from RUIC as the marginal cost proxy and setting β to 0.99 and λ to 0.05. Intercepts and root mean square errors (RMSEs) were measured by regressing the fundamental inflation series on actual inflation, restricting the slope parameter to 1. RMSEs for the three series are as follows:

RUIC: RMSE $= 0.507$
RULC: RMSE $= 0.529$
RUIC*:   RMSE $= 2.360$.

Specifically, present values of inflation are constructed from

$$\pi_{N,t}^* = \lambda_N \sum_{i=0}^{K} \beta^i E(\psi_{N,t+i} \mid Z_{N,t}), \tag{21}$$

where $Z_{N,t}$ is a matrix of variables representing information of stage N, $E(\psi_{N,t+i} \mid Z_{N,t})$ are the sample moments (i.e., the predicted values from a regression of $\psi_{N,t+i}$ on the instruments) and K is set to 12 quarters.

Figure 8 plots GDP deflator inflation against fundamental inflation implied by the RUIC and the RULC marginal cost proxies. The upstream instrument set without inflation (two lags of stage N, $N - 1$, and $N - 2$ marginal cost proxies) is used as a proxy for information of stage N firms at time t, $Z_{N,t}$. The parameters β and λ_N are taken directly from Table 4. The figure also depicts fundamental inflation using the RUIC proxy with values for β and λ standard in the NKPC literature. That is, for the third series, labeled RUIC*, I fix β to 0.99 and λ to 0.05. Visually, the fundamental

inflation series based on the RUIC proxy using the parameter estimates found in Table 4 fits actual inflation best, especially its low-frequency components of the early 1970s and 1980s. This fundamental inflation series also fits inflation best in terms of the root mean square error (RMSE).[33] Altering β and λ away from the estimates found in this paper acts to move fundamental inflation away from actual inflation. This can be seen visually and also by the large increase in RMSE. Furthermore, taking expectations on upstream instruments fits actual inflation better than taking expectations on the GGL instruments when using the respective estimated parameters from Table 4. Specifically, when switching from the GGL instruments to the upstream instruments as a proxy for the information set, the RMSE falls from 0.542 to 0.507 under the RUIC proxy, and from 0.604 to 0.529 under the RULC proxy.[34] This exercise demonstrates that fundamental inflation constructed with the RUIC proxy and its upstream instruments provides a relatively nice fit with actual inflation.

8. CONCLUSION

Finding an estimation method that gives unbiased and efficient estimates of the NKPC is quite important if the NKPC is to be utilized as either a monetary policy tool or a correct specification of macroeconomic behavior. This paper shows that the limited success of estimating the NKPC model using GMM is correlated with both the poor choice of marginal cost proxy and the poor choice of instruments. Overall, I find that the use of the RUIC proxy and its upstream instruments yields better estimates of the NKPC when using GMM.

In comparison with the RULC proxy, the RUIC proxy, based on the vertical chain-of-production model, induces a better fit with the model in terms of relevancy with respect to the instruments, significance of the coefficient on the marginal cost term, and stability of this coefficient with respect to changes in the instrument set. The evidence for this conclusion is displayed in Tables 1 and 5, which compare the different proxies using the GG instrument sets.

This paper shows that generalizing the NKPC model to a vertical chain of production allows for the creation of instruments that are micro-founded. These variables, the RUICs of firms in stages below the GDP stage (i.e., upstream to the GDP stage), keep both relevancy and exogeneity high without relying on a large number of overidentifying restrictions. Furthermore, they perform these duties without requiring the inclusion of lags of inflation, a requirement that has been a major drawback of GMM estimation of the NKPC. Generally, the combination of the RUIC and its upstream instruments results in a rather robust and stable estimation of the NKPC. In terms of rolling subsample estimates of λ and the measure of fundamental inflation, this combination performs better than the RULC proxy with either the GGL instrument set or even its own upstream instruments.

33. The RMSE is found by regressing fundamental inflation on actual inflation, restricting the slope parameter to equal 1 but allowing the intercept to be estimated freely.

34. For a visual demonstration, please refer to Appendix I.

An important result obtained in this paper is the smaller size of the coefficient λ obtained using the RUIC proxy than obtained using the RULC proxy. Even with the inclusion of input prices in the marginal cost term, λ never rises above 0.02. This low measure of λ implies a very low sensitivity of inflation to marginal cost, and hence a high degree of price rigidity in the NKPC. Specifically, the estimates of λ with the RUIC imply that the price stickiness parameter θ_N lies in the range of 0.894 to 0.936, whereas estimates using the RULC imply a range of θ_N from 0.805 to 0.969. The estimates of θ_N with RUIC therefore imply a tight range of price stickiness that lies in the somewhat high end of the price stickiness spectrum implied by the RULC. Thus, the estimates taken with the RUIC suggest that a large proportion of the GDP deflator consists of very rigidly priced goods, such as service sector goods discussed in Dyne et al. (2004).

Finally, the lagged response of the RUIC to inflation seen in Figure 2 has implications for frictions on the input side of the marginal cost term. Specifically, frictions dealing with stages lower in the production chain, such as capital adjustment costs, presumably play a large role in explaining why disinflations are associated with large falls in output. Incorporating investment and inventory frictions in the NKPC, as in Woodford (2005) and Christiano, Eichenbaum, and Evans (2005), therefore seems to be an area deserving more attention.

APPENDIX A: DATA APPENDIX

Inflation (π_N): 100 times the log change in the gross domestic product (GDP) chain-type price deflator.

Output Gap: 100 times the log difference between chain-weighted real GDP and the Congressional Budget Office's (CBO) estimate of potential GDP.

Wage Inflation: 100 times the log change in nominal nonfarm compensation per hour.

Commodity Inflation: 100 times the log change in the crude-stage producer price index (PPI).

Risk-Free Rate: The federal funds rate.

RUIC Stage N: 100 times the log of the PPI finished-goods price level plus the log of the Federal Reserve Board of Governors (FRBG) finished-goods (stage $N - 1$) industrial production (IP) index minus the log of nominal GDP.

RUIC Stage $N - 1$: 100 times the log of the PPI intermediate-goods (stage $N - 2$) price level plus the log of the FRBG intermediate IP index minus the log of the PPI finished-goods price level minus the log of the FRBG IP finished-goods index.

RUIC Stage $N - 2$: 100 times the log of the PPI crude-goods price level plus the log of the FRBG crude IP index minus the log of the PPI intermediate-goods price level minus the log of the FRBG IP intermediate-goods (i.e., primary and semi-finished processing) index.

RULC Stage N: 100 times nominal unit labor costs provided by the Bureau of Economic Analysis (BEA) (i.e., the log of non-farm compensation less the log

of non-farm output per hour) less the log of the non-farm implicit price deflator for the non-farm business sector.

Wages Stage $N - 1$: A weighted average of average hourly earnings of selected goods-producing industries from the Current Employment Survey (CES), provided by the Bureau of Labor Statistics (BLS). The weights were matched with the relative importance weights (RIWs) of the finished-goods PPI stage provided by the BLS using four-digit Standard Industrial Classification (SIC) codes. If a four-digit SIC good from the PPI was not found in the CES, a three-digit SIC match was made.

Wages Stage $N - 2$: A weighted average of average hourly earnings of selected goods-producing industries from the CES, provided by the BLS. The weights were matched with the RIWs of the intermediate-goods PPI stage provided by the BLS using four-digit SIC codes. If a four-digit SIC good from the PPI was not found in the CES, a three-digit SIC match was made.

Labor Hours Stage $N - 1$: A weighted average of hours times employees of selected goods-producing industries from the CES, provided by the BLS. The weights were matched with the RIWs of the finished-goods PPI stage provided by the BLS using four-digit SIC codes. If a four-digit SIC good from the PPI was not found in the CES, a three-digit SIC match was made.

Labor Hours Stage $N - 2$: A weighted average of hours times employees of selected goods-producing industries from the CES, provided by the BLS. The weights were matched with the RIWs of the intermediate-goods PPI stage provided by the BLS using four-digit SIC codes. If a four-digit SIC good from the PPI was not found in the CES, a three-digit SIC match was made.

RULC Stage $N - 1$: 100 times the log of finished-goods (stage $N - 1$) wages plus the log of finished-goods labor hours minus the log of the PPI finished-goods price level minus the log of the FRBG IP finished-goods index.

RULC Stage $N - 2$: 100 times the log of intermediate-goods (stage $N - 2$) wages plus the log of intermediate-goods labor hours minus the log of the PPI of intermediate-goods price level minus the log of the FRBG IP intermediate-goods (primary and semi-finished processing) index.

APPENDIX B: STOCK AND YOGO CRITICAL VALUES FOR CRAGG–DONALD F-STATISTIC

K	TSLS allowed bias				TSLS desired max size			
	0.05	0.10	0.20	0.30	0.10	0.15	0.20	0.25
5	13.97	8.78	5.91	4.79	19.45	11.22	8.38	6.89
6	15.72	9.48	6.08	4.78	21.68	12.33	9.10	7.42
7	16.88	9.92	6.16	4.76	23.72	13.34	9.77	7.91
9	18.30	10.43	6.22	4.69	27.51	15.24	11.03	8.85
13	19.64	10.84	6.21	4.56	34.62	18.84	13.45	10.68
17	20.23	10.99	6.16	4.45	41.51	22.35	15.83	12.49
21	20.54	11.04	6.10	4.37	48.31	25.82	18.20	14.29
25	20.73	11.06	6.05	4.32	55.07	29.29	20.56	16.10

NOTES: The critical values are a function of the number of endogenous regressors and the number of instruments. The significance level is 5% and the number of endogenous regressors is set at two. K is the number of instrumental variables. The test rejects the allowed bias or desired maximum size with its TSLS counterpart if the Cragg–Donald statistic is larger than the critical value.

APPENDIX C: GG HETEROSKEDASTICITY AND AUTOCORRELATION CHECK

		RULC		RUIC	
Inst. set	Regressors	π_{t+1}	$\psi^{(RULC)}$	π_{t+1}	$\psi^{(RUIC)}$
1	3 Lags ψ_N	0.467	0.000	0.028	000
	3 Lags π_N	0.000	0.000	0.000	0.000
2	4 Lags ψ_N	0.417	0.000	0.057	0.000
	4 Lags π_N	0.000	0.001	0.000	0.000
3	4 Lags ψ_N	0.441	0.000	0.274	0.000
	4 Lags π_N	0.000	0.022	0.000	0.000
	4 Lags O.G.	0.003	0.200	0.050	0.065
4	4 Lags ψ_N	0.313	0.000	0.563	0.000
	4 Lags π_N	0.000	0.570	0.000	0.001
	4 Lags O.G.	0.004	0.010	0.016	0.163
	4 Lags RFF	0.618	0.190	0.935	0.046
5	4 Lags ψ_N	0.772	0.000	0.966	0.000
	4 Lags π_N	0.000	0.669	0.000	0.003
	4 Lags O.G.	0.008	0.033	0.031	0.185
	4 Lags RFF	0.646	0.190	0.930	0.073
	4 Lags CInf	0.150	0.295	0.159	0.056
6	4 Lags ψ_N	0.659	0.000	0.936	0.000
	4 Lags π_N	0.000	0.277	0.000	0.838
	4 Lags O.G.	0.008	0.010	0.061	0.466
	4 Lags RFF	0.390	0.021	0.977	0.039
	4 Lags CInf	0.131	0.467	0.156	0.057
	4 Lags WInf	0.040	0.001	0.076	0.159

NOTES: p-values of F-statistics are shown. A four-lag quadratic spectral window estimate of the covariance matrix was used.

The above table shows that for the RULC, the newly added blocks have both p-values of the F-statistic (pF-values) smaller than 0.15 only with the addition of the wage inflation block. The RUIC has such an occurence only for the output gap block. We can see that because of collinearity, pF-values rise and fall with the addition of new instruments, rendering once-relevant instruments irrelevant. There is no case in which all pF-values fall with the addition of new instruments and that the new instruments themselves generally have large pF-values.

APPENDIX D: CHAIN-OF-PRODUCTION INSTRUMENTS HETEROSKEDASTICITY AND AUTO-CORRELATION CHECK

		RULC		RUIC	
Inst. set	Regressors	π_{t+1}	$\psi^{(RULC)}$	π_{t+1}	$\psi^{(RUIC)}$
1	4 Lags ψ_N	0.000	0.000	0.015	0.000
2	4 Lags ψ_N	0.045	0.000	0.000	0.000
	4 Lags π_{N-1}	0.000	0.000	0.000	0.001
3	4 Lags ψ_N	0.215	0.000	0.215	0.000
	4 Lags ψ_{N-1}	0.000	0.030	0.000	0.015
	4 Lags ψ_{N-2}	0.000	0.133	0.000	0.080

(*Continued*)

APPENDIX D

CONTINUED

Inst. set	Regressors	RULC		RUIC	
		π_{t+1}	$\psi^{(RULC)}$	π_{t+1}	$\psi^{(RUIC)}$
4	4 Lags ψ_N	0.262	0.000	0.478	0.000
	4 Lags ψ_{N-1}	0.000	0.039	0.000	0.003
	4 Lags ψ_{N-2}	0.018	0.033	0.020	0.053
	4 Lags π_N	0.000	0.011	0.000	0.179
8	4 Lags ψ_N	0.131	0.000	0.529	0.000
	4 Lags ψ_{N-1}	0.075	0.379	0.003	0.144
	4 Lags ψ_{N-2}	0.612	0.320	0.000	0.000
	4 Lags π_N	0.007	0.333	0.000	0.094
	4 Lags O.G.	0.002	0.025	0.000	0.383
	4 Lags RFF	0.430	0.118	0.481	0.114
	4 Lags C Inf	0.585	0.478	0.080	0.000
	4 Lags W Inf	0.003	0.013	0.050	0.038

NOTES: p-values of F-statistics are shown. A four-lag quadratic spectral window estimate of the covariance matrix was used.

The above table reports pF-values on the block of instruments in question for the first-stage regression after controlling for both autocorrelation and heteroskedasticity. The p-values of the F-statistics on all of the upstream real marginal costs in instrument sets 2, 3, and 4 are significant at the 15% level, confirming their relevancy. pF-values of the upstream real marginal costs rise significantly after the addition of the blocks of "non-chain" instruments used in instrument set 8 of Table 3. This is similar to what happened when adding the inflation block in Appendix B.

APPENDIX E: NONLINEAR GMM ESTIMATION

Proxy	Inst. set	β	θ_N	Implied λ
		Full sample		
$\psi_N^{(RULC)}$	1	1.05 (0.000)	0.999 (0.030)	0.000 (0.999)
	2	1.00 (0.000)	0.899 (0.000)	0.011 (0.350)
	3	0.933 (0.000)	0.895 (0.000)	0.019 (0.163)
	4	0.976 (0.000)	0.896 (0.000)	0.015 (0.198)
$\psi_N^{(RUIC)}$	1	0.978 (0.000)	0.925 (0.000)	0.007 (0.122)
	2	0.968 (0.000)	0.928 (0.000)	0.008 (0.069)
	3	0.904 (0.000)	0.929 (0.000)	0.012 (0.023)
	4	0.906 (0.000)	0.932 (0.000)	0.011 (0.013)
		1973:2–1989:4		
$\psi_N^{(RULC)}$	1	1.03 (0.000)	0.782 (0.000)	0.054 (0.092)
	2	1.01 (0.000)	0.810 (0.000)	0.043 (0.068)
	3	0.989 (0.000)	0.805 (0.000)	0.049 (0.110)
	4	0.999 (0.000)	0.794 (0.000)	0.053 (0.013)
$\psi_N^{(RUIC)}$	1	0.968 (0.000)	0.896 (0.000)	0.015 (0.037)
	2	0.975 (0.000)	0.904 (0.000)	0.013 (0.038)
	3	1.00 (0.000)	0.894 (0.000)	0.012 (0.040)
	4	0.969 (0.000)	0.909 (0.000)	0.012 (0.018)

APPENDIX E

CONTINUED

Proxy	Inst. set	β	θ_N	Implied λ
		Full sample		
		1980:1–2002:4		
$\psi_N^{(RULC)}$	1	1.13 (0.000)	0.942 (0.999)	−0.003 (0.999)
	2	1.00 (0.000)	0.918 (0.000)	0.007 (0.517)
	3	1.06 (0.000)	0.969 (0.999)	−0.001 (0.999)
	4	0.992 (0.000)	0.922 (0.000)	0.007 (0.465)
$\psi_N^{(RUIC)}$	1	1.13 (0.000)	0.988 (0.000)	0.000 (0.999)
	2	0.947 (0.000)	0.934 (0.000)	0.008 (0.032)
	3	0.903 (0.000)	0.936 (0.000)	0.010 (0.050)
	4	0.964 (0.000)	0.934 (0.000)	0.007 (0.055)

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. p-values are in parentheses. Instrument sets: (1) 2 lags ψ_N, output gap, wage inflation; (2) 4 lags of inflation, 2 lags ψ_N, output gap, wage inflation; (3) 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}; (4) 4 lags of inflation, 2 lags ψ_N, ψ_{N-1}, ψ_{N-2}.

APPENDIX F: p-VALUES OF C-STATISTICS OF FORWARD-MODEL SUBSAMPLES

Inst set	Block	RULC	RUIC
	1973:2–1989:4		
3	2 Lags ψ_N	0.487	0.196
	2 Lags ψ_{N-1}	0.287	0.305
	2 Lags ψ_{N-2}	0.394	0.300
	1980:1–2002:4		
3	2 Lags ψ_N	0.497	0.021
	2 Lags ψ_{N-1}	0.696	0.721
	2 Lags ψ_{N-2}	0.843	0.351

NOTES: p-values of C-statistics are shown. A large p-value of the C-statistic indicates that we do not reject the null hypothesis that the suspect instruments are valid. A four-lag quadratic spectral window estimate of the covariance matrix was used.

The above table reports p-values of the C-statistic on the block of instruments in question. The C-statistic, or "difference-in-Hansen" statistic, tests the orthogonality of a subset of instruments. A large C-statistic (a low p-value of the C-statistic) indicates that we reject the null hypothesis that the instruments are orthogonal to the expectation error. In the second subsample, the low p-value on the $\psi_{t,n}$ block indicates the poor exogeneity of this variable. The upstream variables do, in fact, produce higher p-values on the C-statistic in the second subsample.

APPENDIX G: INFLATION SPECIFICATION

An incorrect measure of the relevant inflation term, π_N, in the final stage of production could lead to biased estimates of the coefficients obtained in the paper thus far. To demonstrate, suppose the GDP chain-type price deflator is actually composed of "sticky-price" goods (goods produced by firms that have probability of no price

adjustment greater than zero: $\theta_N > 0$) and "flexible-price" goods (goods produced by firms that have probability of no price adjustment approximately equal to zero: $\theta_N \approx 0$).

Since the NKPC captures only the dynamics of sticky-price goods inflation (i.e., the model assumes $\theta_N > 0$), a large proportion of flexible-price goods in the GDP deflator would imply they show up in the error term, η_{t+1}. As crude goods are usually considered flexible-price goods, the upstream RUIC terms could very well be correlated with η_{t+1}. This hypothesis would in turn cause a violation of the exogeneity of the instruments and bias our estimates.

To be certain that the upstream RUIC variables are valid instruments with regard to the NKPC, we would ideally like to estimate the model under an inflation measure that is free of flexible-price goods:

$$\pi_{N,t}^{sticky} = \lambda_N \psi_{N,t} + E_t\left(\pi_{N,t+1}^{sticky}\right). \tag{G1}$$

Constructing such a price index, where inflation is completely purged of flexible-price goods, would be quite problematic given the ambiguous distinction between sticky- and flexible-price goods.[35] Nonetheless, there do exist other inflation measures that might do a better job of capturing the sticky component of inflation than does the GDP deflator. As a robustness exercise, I estimate the NKPC with the RUIC and its upstream instruments under the core CPI and the core PCE. These inflation measures do not include food or energy, which are conventionally regarded as flexible-price goods. These inflations may therefore better resemble π^{sticky}.

Under the joint hypothesis that (i) these inflation series better capture sticky-price inflation and (ii) the upstream instruments are correlated with the flexible component of the GDP deflator, the pJ-value under these series should be significantly *higher* than that under the GDP deflator. The new error under these price series, η_{t+1}^*, theoretically does not include flexible-price goods, and thus it should be *less* correlated with the instruments and therefore raise the p-value of the J-statistic.

The table below displays results of the forward-looking model, using the upstream instruments under the core CPI and the core PCE.[36] The table shows that for the RUIC proxy, the pJ-value *falls* when moving from the GDP deflator to either the core CPI or the core PCE, from 0.421 to 0.394 and 0.179, respectively. These statistics indicate that the upstream RUIC instruments are more valid with regard to the GDP deflator than with either the core CPI or the core PCE. Thus, the above joint hypothesis is rejected. We can conclude that either the deflator contains a smaller share of flexible-price goods than the core CPI or core PCE, or that it contains a greater share of flexible-price goods but that the upstream RUIC terms are not correlated with them.

---

35. Bils and Klenow (2004) rank goods by their frequency of price adjustment. Based on this criterion, a new aggregate price series could conceivably be constructed that leaves out goods with price adjustments lasting less than one quarter. However, ambiguity problems still arise because there are many goods whose prices still lie in the margin between flexible and sticky (within 2.5 to 3.5 months since the last price change). Furthermore, given that these frequencies were taken over 2 years, 1995–97, the durations are subject to change over subsamples.

36. The same procedure was performed with the hybrid model; the relative magnitudes of the p-values of the J-statistic did not change.

In any case, we can be certain that food and energy prices in the GDP deflator do not cause an exogeneity problem with the upstream RUIC instruments.

When estimating under the RULC proxy, the pJ-value falls when switching to the core PCE, but *rises* when using the core CPI. Thus, compared with ψ_{N-1}^{RUIC} and ψ_{N-2}^{RUIC}, there is more evidence under our joint hypothesis that ψ_{N-1}^{RULC} and ψ_{N-2}^{RULC} are correlated with the flexible component of inflation. This gives more credence to the use of the RUIC proxy and its upstream variables as instruments.

Overall, this exercise demonstrates that changing the inflation series does not invalidate the results obtained earlier. That is, the use of the RUIC proxy and its upstream instruments in estimating the NKPC is robust to the inflation measure used. Specifically, under the core CPI and core PCE measures, the RUIC proxy and its upstream instruments produce high relevancy and exogeneity statistics as well as estimates of λ that are significant at the 5% level.

Proxy	Inflation measure	β	λ	pJ	CD	SP/Pπ_{t+1}	SP/P$\psi_n^{(i)}$
$\psi_N^{(RULC)}$	GDP deflator	0.941	0.018 (0.109)	0.468	36.02	0.661/0.740	0.740/0.829
	Core CPI	0.889	0.023 (0.184)	0.674	13.34	0.414/0.519	0.662/0.829
	Core PCE	0.933	0.021 (0.037)	0.142	24.54	0.570/0.701	0.675/0.829
$\psi_N^{(RUIC)}$	GDP deflator	0.902	0.012 (0.000)	0.421	58.36	0.754/0.784	0.875/0.909
	Core CPI	0.884	0.006 (0.031)	0.394	23.74	0.555/0.660	0.765/0.909
	Core PCE	0.917	0.009 (0.048)	0.179	42.30	0.693/0.788	0.799/0.909

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument set: 2 lags $\psi_N, \psi_{N-1}, \psi_{N-2}$.

APPENDIX H: MARGINAL COST SPECIFICATION

A direct way to test whether the upstream instruments are included in η_{t+1} is to estimate the model with the inclusion of upstream marginal costs. If the coefficients on the upstream instruments show up as significant, then this gives some evidence that η_{t+1} in previous estimations included upstream RUIC terms.

The table below displays GMM results of two different specifications of the marginal cost term in the NKPC. Specification A adds the *current* $\psi_{N-1,t}$ and $\psi_{N-2,t}$ terms to equation (10), and specification B adds two lags of ψ_{N-1} and ψ_{N-2}:[37]

$$(A)\, \pi_{N,t} = \lambda\psi_{N,t} + \lambda_{1,0}\psi_{N-1,t} + \lambda_{2,0}\psi_{N-2,t} + \beta E_t(\pi_{N,t+1})$$
$$(B)\, \pi_{N,t} = \lambda\psi_{N,t} + \lambda_{1,1}\psi_{N-1,t-1} + \lambda_{1,2}\psi_{N-1,t-2} + \lambda_{2,1}\psi_{N-2,t-1}$$
$$+ \cdots\lambda_{2,2}\psi_{N-2,t-2} + \beta E_t(\pi_{N,t+1}).$$

The results show that when either the current lag or two lags of the RUIC term are added to the NKPC, their coefficients show up as insignificant, and the overall effect on λ is quite small. Although current RULC terms are an insignificant inclusion in

___

37. These estimations were performed using two lags of inflation, the output gap, wage inflation, and real marginal cost in the instrument set. The upstream terms were treated as exogenous. Using upstream instruments in these specifications yields poor relevancy statistics.

the model, lags of the RULC terms are quite significant. Overall, this specification test provides evidence against the case that η_{t+1} includes upstream RUIC terms and justifies the use of upstream RUIC terms over RULC terms in the instrument set of the NKPC.[38]

Proxy	Specification	β	λ	$\lambda_{1,0}$	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{2,0}$	$\lambda_{2,1}$	$\lambda_{2,2}$
$\psi_N^{(RULC)}$	A	1.04	0.026 (0.013)	−0.003 (0.272)	-	-	0.002 (0.369)	-	
	B	1.04	0.032 (0.016)	-	−0.042 (0.034)	0.039 (0.061)	-	0.018 (0.314)	−0.016 (0.348)
$\psi_N^{(RUIC)}$	A	0.885	0.011 (0.002)	0.001 (0.627)	-	-	3 e-4 (0.675)	-	
	B	0.942	0.007 (0.031)	-	−0.008 (0.671)	0.007 (0.709)	-	−0.006 (0.315)	0.006 (0.291)

NOTES: A four-lag quadratic spectral window estimate of the covariance matrix was used. *p*-values are in parentheses. Instrument set: 2 lags inflation, output gap, wage inflation, ψ_N.

APPENDIX I: FUNDAMENTAL INFLATION: ALTERING THE INFORMATION SET
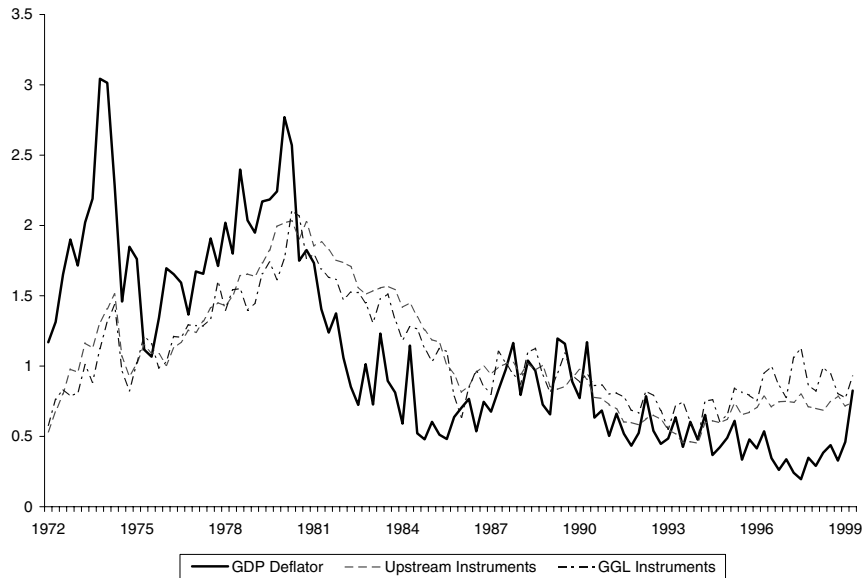

FIG. I1. Inflation: Actual versus Fundamental with RUIC Proxy.

NOTES: GGL instruments (instrument set 1 of Table 4): RMSE = 0.542. Upstream instruments (instrument set 3 of Table 4): RMSE = 0.507.

38. *F*-tests were performed to test the hypothesis that the added variables are jointly significant. The results did not change.
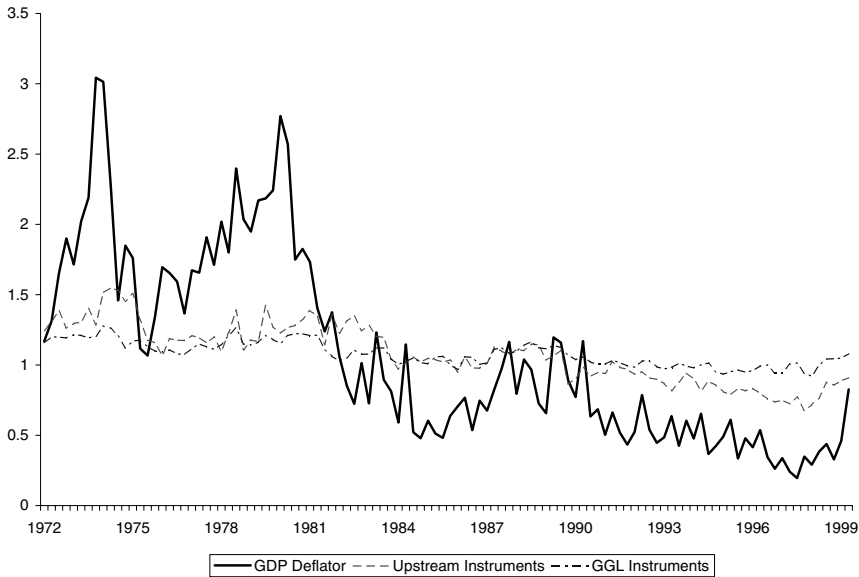
Fig. I2. Inflation: Actual versus Fundamental with RULC Proxy.

NOTES: GGL instruments (instrument set 1 of Table 4): RMSE = 0.604. Upstream instruments (instrument set 3 of Table 4): RMSE = 0.529.

LITERATURE CITED

Basu, Susanto. (1995) "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare." *The American Economic Review,* 85, 512–31.

Bils, Mark, and Pete Klenow. (2004) "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy,* 112, 947–85.

Blanchard, Olivier, and Jordi Galí. (2005) "Real Wage Rigidities and the New Keynesian Model." Federal Reserve Bank of Boston, Working Paper 05-14.

Calvo, Guillermo. (1983) "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics,* 12, 383–98.

Christiano, Lawrence, Martin Eichenbaum, and Charles L. Evans. (2005) "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy,* 113, 1–45.

Cogley, Timothy, and Argia M. Sbordone. (2005) "A Search for a Structural Phillips Curve." Federal Reserve Bank of New York, Staff Reports 203.

Cragg, John G.,, and Stephen G. Donald. (1993) "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory,* 9, 222–40.

Dhyne, Emmanuel, Luis J. Alvarez, Herve Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffman, Nicole Jonker, Patrick Lunnermann, Fabio Rumier, and Jouko Vilmunen. (2004) "Price Setting in the Euro Area: Some Stylised Facts from Individual Consumer Price Data." ECB Working Paper No. 524.

Dixit, Avanish, and Joseph. E. Stiglitz. (1977) "Monopolistic Competition and Optimum Product Diversity." *American Economic Review,* 67, 297–308.

Eichenbaum, Martin, and Jonas D. M. Fisher. (2003) "Testing the Calvo Model of Sticky Prices. 2003. Federal Reserve Bank of Chicago." *Economic Perspectives,* Q II, 40-53.

Erceg, Christopher, and Andrew T. Levin. (2003) "Imperfect Credibility and Inflation Persistence." *Journal of Monetary Economics,* 50, 915–44.

Fuhrer, Jeffrey. (2006) "Intrinsic and Inherited Inflation Persistence." *International Journal of Central Banking,* 2, 49–86.

Fuhrer, Jeffrey, and Jeffery Moore. (1995) "Inflation Persistence." *Quarterly Journal of Economics,* 110, 127–59.

Galí, Jordi, and Mark Gertler. (1999) "Inflation Dynamics: A Structural Econometric Approach." *Journal of Monetary Economics,* 44, 195–222.

Galí, Jordi, Mark Gertler, and David Lopez-Salido. (2001) "European Inflation Dynamics." *European Economic Review,* 45, 1237–70.

Galí, Jordi, Mark Gertler, and David Lopez-Salido. (2005) "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve." *Journal of Monetary Economics,* 52, 1107–18.

Goodfriend, Marvin, and Robert G. King. (2005) "The Incredible Volcker Disinflation." *Journal of Monetary Economics,* 52, 981–1015.

Gumbau-Brisá, Fabia. (2005) "Heterogeneous Beliefs and Inflation Dynamics: A General Equilibrium Approach." Working Paper, Federal Reserve Bank of Boston 05-16.

Hahn, Jinyong, and Jerry Hausman. (2002) "IV Estimation with Valid and Invalid Instruments." Mimeo, Massachusetts Institute of Technology.

Hansen, Lars P. (1982) "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica,* 50, 1029–54.

Huang, Kevin, and Zheng Liu. (2001) "Production Chains and General Equilibrium Dynamics." *Journal of Monetary Economics,* 48, 437–62.

Linde, Jesper. (2005) "Estimating New Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach." *Journal of Monetary Economics,* 52, 1135–49.

Mavroeidis, Sophocles. (2005) "Identification Issues in Forward-Looking Models Estimated by GMM, with an Application to the Phillips Curve." *Journal of Money, Credit, and Banking,* 37, 421–48.

Roberts, John M. (2001) "How Well Does the New Keynesian Sticky-Price Model Fit the Data?" Federal Reserve Board Finance and Economics Discussion Series: 2001-13.

Rudd, Jeremy, and Karl Whelan. (2005) "New Tests of the New Keynesian Phillips Curve." *Journal of Monetary Economics,* 52, 1167–81.

Shea, John. (1997) "Instrument Relevance in Multivariate Linear Models: A Simple Measure." *Review of Economics and Statistics,* 79, 348–52.

Stock, James, Jonathan W. Wright, and Motohiro Yogo. (2002) "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics,* 20, 518–29.

Stock, James H., and Motohiro Yogo. (2001) "Testing for Weak Instruments in Linear IV Regression." NBER Technical Working Papers 0284.

Woodford, Michael. (2005) "Firm Specific Capital and the New-Keynesian Phillips Curve." *International Journal of Central Banking,* 1, 1–46.