Empirical Assessment of the Phillips Curve II

Emi Nakamura  Jón Steinsson

Columbia University

February 2018
Motivation:

- Is New Keynesian (Calvo) Phillips curve consistent with observed inflation persistence?
  - Implies disinflations can be costless
  - In practice, it seems disinflations are costly (Ball 94, 95)
    (Imperfect credibility could explain this)

- Do we need “sticky inflation” models or adaptive expectations?

- With quarterly data, hard to get statistically significant effect of real activity on inflation, when using output gap
Simple Theory

Simple theory with Calvo pricing assumption implies:

\[ \pi_t = \beta E_t \pi_{t+1} + \lambda m c_t \]

where

\[ \lambda = \frac{(1 - \theta)(1 - \beta \theta)}{\theta} \]

and \(1 - \theta\) is frequency of price change, \(\beta\) subjective discount factor

- Theory implies that \(mc_t\) is the appropriate “forcing variable” in the Phillips curve

- Yet most empirical work uses simple measures of the output gap such as detrended output
Simple Theory

- Under certain assumptions:

\[ mc_t = \kappa x_t \]

where \( x_t = y_t - y^n_t \) denotes the output gap

- Maybe Phillips curve estimation doesn’t work because:
  - These assumptions don’t hold in reality
  - Output gap is mismeasured
With rational expectations, NK Phillips curve can be written as

$$\pi_{t+1} - \pi_t = -\lambda \kappa X_t + \epsilon_{t+1}$$

where $\epsilon_{t+1} = \pi_{t+1} - E_t \pi_{t+1}$

Traditional Phillips curve with adaptive expectations:

$$\pi_t = E_{t-1} \pi_t + \lambda \kappa X_t$$

$$\pi_t - \pi_{t-1} = \lambda \kappa X_t$$

where we are assuming $E_{t-1} \pi_t = \pi_{t-1}$

Notice the difference in the sign on the output gap term!!
(and difference in timing of inflation change)
NEW KEYNESIAN VS. OLD KEYNESIAN

\[ \pi_{t+1} - \pi_t = -\lambda \kappa x_t + \epsilon_{t+1} \]

- NK Phillips curve implies tight labor market should lead inflation to fall!!

- Theoretical logic:
  - Inflation is a jump variable in this model
  - When output gaps are expected, inflation should jump up and start falling

\[ \pi_t = \lambda \kappa \sum_{k=0}^{\infty} \beta^k E_t x_{t+k} \]

- I.e., inflation should lead output gap according to NK Phillips curve
  (Fuhrer-Moore 95)
NEW KEYNESIAN VS. OLD KEYNESIAN

- Simple estimation using quadratically detrended log GDP yields:

\[ \pi_{t+1} - \pi_t = 0.081 x_t + \epsilon_{t+1} \]

- Output gap term has “wrong sign” (from NK perspective)
Output Gap Leads Inflation

One reaction: NK Phillips curve is empirically unrealistic.
- Perhaps some sluggishness messes up this jump business
- Perhaps information frictions play a role (yield $E_{t-1} \pi_t$)

Gali-Gertler argue:
- Use of output gap is the problem
- Output gap measured with error
- Marginal costs tends to lag output gap

Gali-Gertler propose to estimate Phillips curve using marginal cost as forcing variable
Measuring Marginal Cost

- But marginal costs are unobservable as well!!

- Gali-Gertler make following assumptions:
  - Production function: \( Y_t = A_t K_t^{\alpha} N_t^{\alpha} \)
  - Labor is hired on a spot market at constant wage

- Marginal cost:
  \[
  MC_t = \frac{W_t}{P_t} \frac{\partial Y_t}{\partial N_t} = \frac{W_t}{P_t} \frac{1}{\alpha} \frac{W_t N_t}{P_t Y_t} = S_t
  \]

  proportional to labor share (average cost)

- In logs, we get:
  \[
  mc_t = s_t
  \]
Measuring Marginal Cost

- Assumptions that Gali-Gertler make to derive this are strong assumptions!!

- Worker-firm relationship often long-term relationship
  - Not clear that current wage is a good proxy for marginal cost
  - May just be an installment payment on a long-term contract
  - Suppose workers performs well at time $t$:
    - Wage may not reflect this at time $t$
    - Rather worker may expect a raise / promotion in the future
  - Firms may insure workers (labor hoarding)

- Wages at a given point in time complicated by overtime
  - Marginal wage may not be the same as average wage
Supply Shocks

- Gali-Gertler estimate

\[ \pi_t = \beta E_t \pi_{t+1} + \lambda s_t \]

- Advantage of using measure of marginal costs:
  - Supply shocks should be reflected in marginal costs
Expectations of Inflations

What do Gali-Gertler do about expectations of inflation?

- They assume rational expectations
- Under this assumption, Phillips curve can be written

\[ \pi_t = \beta \pi_{t+1} + \lambda s_t + \epsilon_{t+1} \]

where \( \epsilon_{t+1} = \beta E_t \pi_{t+1} - \beta \pi_{t+1} \) (i.i.d.)

- They furthermore take structural model super seriously in assuming that there is no other error term than this expectations error
- This strong structural assumption allows them to use lagged variables as instruments (any variable dated at time \( t \) or earlier)
Empirical Specification

- Maintained assumptions:
  \[
  \pi_t - \beta \pi_{t+1} - \lambda s_t = \epsilon_{t+1}
  \]

  where \( \epsilon_{t+1} \) is an i.i.d. expectations error and therefore uncorrelated with variables at time \( t \) or earlier

- Implies:
  \[
  E_t\{(\pi_t - \beta \pi_{t+1} - \lambda s_t)z_t\} = 0
  \]

  where \( z_t \) is in the time \( t \) information set of agents
Empirical Specification

- Gali-Gertler use GMM with these orthogonality conditions

\[ E_t\{(\pi_t - \beta \pi_{t+1} - \lambda s_t)z_t\} = 0 \]

- Sample period: 1960Q1-1997Q4

- Instruments: Four lags of inflation, labor income share, output gap, long-short interest rate spread, wage inflation, and commodity price inflation (24 instruments)
**Why IV and not OLS?**

\[ \pi_t = \beta \pi_{t+1} + \lambda s_t + \epsilon_{t+1} \]

- Under maintained assumption that error term is i.i.d. expectation error dated at time \( t + 1 \), instrument only needed to estimate \( \beta \)
- More generally, other omitted variables (or cost push shocks) enter the equation and are dated at time \( t \) (i.e., affect \( \pi_t \)):

\[ \pi_t = \beta \pi_{t+1} + \lambda s_t + \eta_t \]

- In this case, both \( \beta \) and \( \lambda \) potentially biased
- Why would \( \lambda \) be biased?
  - One reason is monetary policy: \( \eta_t \) affects \( \pi_t \), monetary policy will react and this will affect \( s_t \).
Empirical Concerns

- Is $\epsilon_{t+1}$ really just an i.i.d. expectations error?
  - If assumptions needed for $mc_t = s_t$ don’t hold, it’s not
  - If expectations are not rational, it is not
- If it is not, then instruments may be invalid
  - Slow moving omitted variables correlated with past stuff
  - Same as persistent “cost-push shock”
- 24 instruments raises concerns about many-weak instruments
  - Many/Weak instruments issue is an overfitting issue in small samples
  - Using 24 relatively weak instruments may lead to substantial overfitting
**Reduced Form Results**

- Estimation with labor share:

  \[
  \pi_t = 0.023 s_t + 0.942 E_t \pi_{t+1}
  \]

  Coefficients have “correct sign” and “sensible” magnitude

- Estimation with output gap (HP-filtered GDP):

  \[
  \pi_t = -0.016 s_t + 0.988 E_t \pi_{t+1}
  \]

  Coefficient on output gap has “wrong sign”
Fig. 1. Dynamic cross-correlations.

Over the cycle, in the sense that a rise (decline) in current inflation should signal a subsequent rise (decline) in the output gap. Yet, exactly the opposite pattern can be found in the data. The top panel in Fig. 1 presents the cross-correlation of the current output gap (measured by detrended log GDP) with leads and lags of inflation.

As the panel indicates clearly, the current output gap co-moves positively with future inflation and negatively with lagged inflation. This lead of the output gap over inflation explains why the lagged output gap enters with a positive coefficient in Eq. (9), consistent with the old Phillips curve theory but in direct contradiction of the new.

The cross-correlations reported in Fig. 1 were computed on HP-detrended series over the period 1960:1-1997:4. We provide a more extensive discussion of Fig. 1 in the conclusion.

OUTPUT GAP VS. LABOR SHARE

- Output gap leads inflation in contradiction to theory
- Labor share strongly correlated with inflation contemporaneously
- Lagged inflation also positively correlated with inflation
- Labor share lags output gap
- Lag in response of labor share explains why it does better in Phillips curve estimation
Table 1
Estimates of the new Phillips curve

<table>
<thead>
<tr>
<th></th>
<th>(\theta)</th>
<th>(\beta)</th>
<th>(\lambda)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP deflator</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.829</td>
<td>0.926</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.024)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>(2)</td>
<td>0.884</td>
<td>0.941</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.018)</td>
<td>(0.007)</td>
</tr>
<tr>
<td><strong>Restricted (\beta)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.829</td>
<td>1.000</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td></td>
<td>(0.007)</td>
</tr>
<tr>
<td>(2)</td>
<td>0.915</td>
<td>1.000</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td></td>
<td>(0.006)</td>
</tr>
<tr>
<td><strong>NFB deflator</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.836</td>
<td>0.957</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>(2)</td>
<td>0.884</td>
<td>0.967</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.016)</td>
<td>(0.008)</td>
</tr>
</tbody>
</table>

*Notes:* This table reports GMM estimates of the structural parameters of Eq. (15). Rows (1) and (2) correspond to the two specifications of the orthogonality conditions found in Eqs. (18) and (19) in the text, respectively. Estimates are based on quarterly data and cover the sample period 1960:1–1997:4. Instruments used include four lags of inflation, labor income share, long-short interest rate spread, output gap, wage inflation, and commodity price inflation. A 12-lag Newey–West estimate of the covariance matrix was used. Standard errors are shown in brackets.

**Structural Estimates**

- Sensible estimates for $\beta$
- Estimates of $\theta$ on the high end
  - Imply price rigidity of 5-6 quarters
Inflation Inertia

- Does NK Phillips curve account for inflation inertia?
- Gali-Gertler estimate specification with fraction of rule-of-thumb agents
- Rule-of-thumb agents set

\[ p_t^b = \bar{p}_{t-1}^* + \pi_{t-1} \]

This yields

\[ \pi_t = \lambda mc_t + \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} \]

where

\[ \lambda = \frac{(1-\omega)(1-\theta)(1-\beta \theta)}{\theta + \omega[1-\theta(1-\beta)]} \]

\[ \gamma_f = \frac{\beta \theta}{\theta + \omega[1-\theta(1-\beta)]} \]

\[ \gamma_b = \frac{\omega}{\theta + \omega[1-\theta(1-\beta)]} \]

and \( \omega \) denotes the fraction of rule-of-thumb agents.
Table 2
Estimates of the new hybrid Phillips curve

<table>
<thead>
<tr>
<th></th>
<th>(\omega)</th>
<th>(\theta)</th>
<th>(\beta)</th>
<th>(\gamma_b)</th>
<th>(\gamma_r)</th>
<th>(\lambda)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP deflator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.265</td>
<td>0.808</td>
<td>0.885</td>
<td>0.252</td>
<td>0.682</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.015)</td>
<td>(0.030)</td>
<td>(0.023)</td>
<td>(0.020)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>(2)</td>
<td>0.486</td>
<td>0.834</td>
<td>0.909</td>
<td>0.378</td>
<td>0.591</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.020)</td>
<td>(0.031)</td>
<td>(0.020)</td>
<td>(0.016)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Restricted (\beta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.244</td>
<td>0.803</td>
<td>1.000</td>
<td>0.233</td>
<td>0.766</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.017)</td>
<td>(0.023)</td>
<td>(0.015)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>0.522</td>
<td>0.838</td>
<td>1.000</td>
<td>0.383</td>
<td>0.616</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.027)</td>
<td>(0.020)</td>
<td>(0.016)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>NFB deflator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>0.077</td>
<td>0.830</td>
<td>0.949</td>
<td>0.085</td>
<td>0.871</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.016)</td>
<td>(0.019)</td>
<td>(0.031)</td>
<td>(0.018)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>(2)</td>
<td>0.239</td>
<td>0.866</td>
<td>0.957</td>
<td>0.218</td>
<td>0.755</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.025)</td>
<td>(0.021)</td>
<td>(0.031)</td>
<td>(0.016)</td>
<td>(0.006)</td>
</tr>
</tbody>
</table>

Notes: This table reports GMM estimates of parameters of Eq. (26). Rows (1) and (2) correspond to the two specifications of the orthogonality conditions found in Eqs. (27) and (28) in the text, respectively. Estimates are based on quarterly data and cover the sample period 1960:1–1997:4. Instruments used include four lags of inflation, labor income share, long-short interest rate spread, output gap, wage inflation, and commodity price inflation. A 12-lag Newey–West estimate of the covariance matrix was used. Standard errors are shown in brackets.

Estimation of Hybrid Phillips Curve

- Estimate of $\omega$ statistically significant
  - Normalization 1 yields: $\omega = 0.265(0.031)$
  - Normalization 2 yields: $\omega = 0.486(0.040)$

- A quarter to half of agents are rule-of-thumb

- Gali-Gertler conclusion:
  - Forward-looking behavior more important than backward-looking behavior

- Estimates of $\beta$ on the low side at around 0.9
Model Fit

- Hybrid Phillips curve has following solution:

\[ \pi_t = \delta_1 \pi_{t-1} + \left( \frac{\lambda}{\delta_2 \gamma_t} \right) \sum_{k=0}^{\infty} \left( \frac{1}{\delta_2} \right)^k E_t s_{t+k} \]

- Gali-Gertler forecast \( s_{t+k} \) (with VAR?) to construct “fundamental inflation” (i.e., RHS of above equation)

- Compare actual inflation and fundamental inflation
Actual vs. Fundamental Inflation

Overall fundamental inflation tracks the behavior of actual inflation very well. It is particularly interesting to observe that it does a good job of explaining the recent behavior of inflation. During the past several years, inflation has been below trend. Output growth has been above trend, on the other hand, making standard measures of the output gap highly positive. As a consequence, traditional Phillips curve equations have been overpredicting recent inflation.

However, because real unit labor costs have been quite moderate recently despite rapid output growth, our model of fundamental inflation is close to target.

Sbordone (1998) similarly finds that inflation is well explained by a discounted stream of future real marginal costs, though using a quite different methodology to parametrize the model.

As exception is Lown and Rich (1997). Because they augment a traditional Phillips curve with the growth in nominal unit labor costs, their equation fares much better than the standard formulation. Though the way unit labor costs enters our formulation is quite different, it is similarly the sluggish behavior of unit labor costs that helps the model explain recent inflation.

Source: Gali and Gertler (1999)
CRITIQUES OF GALI-GERTLER

- Subsequent work has found Gali-Gertler’s results to be highly sensitive to instruments used, vintage of data, model specification.

- Mavroeidis-Plagborg-Moller-Stock 14 argue fundamental problem is weak instruments:
  - Inflation is notoriously difficult to forecast.
  - Lagged variables weak instruments for future inflation.

- More recent literature has used many fewer instruments to avoid many-instruments problem.
and Gertler also develop the now-standard hybrid NKPC, whose lagged inflation terms introduce intrinsic persistence of the inflation rate on top of the extrinsic persistence imparted by the forcing variable.

4.2 GIV Estimation

Using linear and nonlinear GIV methods, Galí and Gertler (1999) find that, while the backward-looking inflation term is significant, the forward-looking final expectations term dominates; they also obtain a significant and correctly signed coefficient on the labor share (unlike the output gap). The NKPC restrictions are not rejected by overidentification tests or by visual inspection of fitted inflation. Galí, Gertler, and López-Salido (2001) take the model to aggregate Eurozone data, largely confirming the U.S. findings. Benigno and López-Salido (2006) find some heterogeneity in estimated coefficients for major Eurozone countries. Eichenbaum and Fisher (2004, 2007) evaluate a variant of the NKPC with price indexation that was developed by Christiano, Eichenbaum, and Evans (2005).

Figure 3. Point Estimates Reported in the Literature

Notes: Point estimates of $\lambda$ (vertical axis) and $\gamma_f$ (horizontal axis) reported in the literature. Only estimates that use U.S. data and the labor share as forcing variable are plotted. For some papers the semistructural point estimates have been imputed from point estimates of deeper parameters. The dotted blue lines indicate 95 percent confidence intervals for $\lambda$ where available. We include papers with readily available estimates and more than twenty-five Google Scholar citations as of mid-September 2012: Galí and Gertler (1999); Galí, Gertler, and López-Salido (2001); Fuhrer and Olivei (2005); Gagnon and Khan (2005); Guay and Pelgrin (2005); Henzel and Wollmershäuser (2008); Jondeau and Le Bihan (2005); Roberts (2005); Sbordone (2005); Dufour, Khalaf, and Kichian (2006); Fuhrer (2006); Kiley (2007); Kurmann (2007); Rudd and Whelan (2007); Brissimis and Magginas (2008); and Adam and Padula (2011).

Source: Mavroeidis, Plagborg-Moller, Stock (2014)
Sensitivity to Data Vintage

- Rudd-Whelan 07 emphasize sensitivity to data vintage

- Mavroeidis-Plagborg-Moller-Stock 14 run Gali-Gertler 99 hybrid specification with Gali-Gertler-Lopez-Salido 01 instruments on Gali-Gertler 99 sampler period for two data vintages
  - Roughly replicate Gali-Gertler 99 results for 2008 data vintage
  - With 2012 data vintage, slope of Phillips curve 30% smaller and insignificant
The coefficient on the labor share \( \lambda \) is generally estimated to be positive, but borderline significant (using the usual strong-instrument inference). In table 3 we replicate these findings using data of the same vintage as Galí and Gertler (1999), but with the Galí, Gertler, and López-Salido (2001) instrument set. Later papers have mostly obtained insignificant \( \lambda \) estimates, and like Rudd and Whelan (2007) we find that this is even true on the Galí and Gertler (1999) sample if revised data (as of 2012) is used.

Using the output gap as forcing variable also typically yields an insignificant estimate of \( \lambda \), and early papers in the literature tended to find negative point estimates.

The estimation results reported in the literature differ in terms of the choice of data series, estimation sample, and various other aspects of the specification, such as the number of inflation lags, any additional regressors, the measurement of inflation expectations, and the identification assumptions, including the set of instruments and other identifying restrictions. As we showed in figure 3, estimates of \( \lambda \) and \( \gamma_f \) reported in various papers differ markedly, but the key message is that all highly cited papers obtain a positive slope coefficient (\( \lambda > 0 \)), and, with the exception of Fuhrer (2006) and Henzel and Wollmershäuser (2008), generally find forward-looking behavior to be dominant (\( \gamma_f > 0.5 \)). The results presented in figure 3 are a tiny subset of possible specifications. Table 4 presents various dimensions of the specification choice that have been considered in the literature. These combinations of choices produce a very large number of specifications that are not objectionable on a priori grounds.

Table 3 presents various dimensions of the specification choice that have been considered in the literature. The only components of the table that have not been explored extensively in the literature are some of the real-time data series (but see Paloviita and Mayes (2005), Dufour, Khalaf, and Kichian (2006), and Wright (2009)) and the use of survey expectations as instruments (but see Wright (2009) and Nunes (2010)). The latter is motivated by evidence that surveys typically forecast inflation better than most alternatives; see Ang, Bekaert, and Wei (2007).

**TABLE 3**

**Baseline GIV Estimates Using Different Data Vintages**

<table>
<thead>
<tr>
<th>Data vintage</th>
<th>Const.</th>
<th>( \lambda )</th>
<th>( \gamma_f )</th>
<th>( \gamma_b )</th>
<th>Hansen test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>0.041</td>
<td>0.026</td>
<td>0.615</td>
<td>0.340</td>
<td>5.263</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.013)</td>
<td>(0.057)</td>
<td>(0.058)</td>
<td>[0.628]</td>
</tr>
<tr>
<td>2012</td>
<td>-0.049</td>
<td>0.018</td>
<td>0.719</td>
<td>0.240</td>
<td>9.816</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.012)</td>
<td>(0.099)</td>
<td>(0.095)</td>
<td>[0.199]</td>
</tr>
</tbody>
</table>


Source: Mavroeidis, Plagborg-Moller, Stock (2014)
Run huge number of different a priori reasonable specifications with a common dataset

Main findings:
- Large amount of specification uncertainty
- Large amount of sampling uncertainty

Both conclusions due to weakness of identification
<table>
<thead>
<tr>
<th>Specification settings</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation ($\pi_t$)</td>
<td>GDP deflator, CPI, chained GDP def., GNP def., chained GNP def., NFB GDP def., PCE, core PCE, core CPI, filtered GDP def. gap, smoothed GDP def. gap, filt. CPI gap, sm. CPI gap, SPF-based CPI gap, filt. core CPI gap, sm. core CPI gap, filt. PCE gap, sm. PCE gap, filt. core PCE gap, sm. core PCE gap</td>
</tr>
<tr>
<td>Labor share (ls)</td>
<td>NFB, NFB coint. relation, HP filtered NFB gap, Baxter-King filt. NFB gap, linearly detrended NFB gap, quadratically detrended NFB gap, real-time NFB HP gap, real-time NFB BK gap, real-time NFB lin. detr. gap, real-time NFB quadr. detr. gap</td>
</tr>
<tr>
<td>Reduced form</td>
<td>Unrestricted, VAR</td>
</tr>
<tr>
<td>Survey forecasts ($\pi_{t+s}$)</td>
<td>SPF CPI, SPF GDP def., GB GDP def.</td>
</tr>
<tr>
<td>Expectations</td>
<td>$\pi_{t+1}$ (endogenous), $\pi_{t+1</td>
</tr>
<tr>
<td>$\pi_{t+s-1}$ (endog.)</td>
<td>$\pi_{t+s-1}$ (exog.)</td>
</tr>
<tr>
<td>Instruments</td>
<td>GG: 4 lags of $\pi$, ls, ygap, 10y–90d yield spread, wage infl., commodity price infl.</td>
</tr>
<tr>
<td></td>
<td>GGLS: 4 lags of $\pi$, and 2 lags of ls, ygap, wage infl.</td>
</tr>
<tr>
<td></td>
<td>small: 4 lags of $\pi$, and 3 lags of forcing variable</td>
</tr>
<tr>
<td></td>
<td>exact: 1 extra lag of each endog. reg. (just-identified)</td>
</tr>
<tr>
<td></td>
<td>RT: 2 real-time lags of GDP def. inflation, $\Delta ls$, ygap</td>
</tr>
<tr>
<td></td>
<td>survey: 2 lags of 1-quarter SPF/GB forecasts, forcing variable</td>
</tr>
<tr>
<td></td>
<td>Extra regressors (e.g., oil) added to instruments (if endog., use 2 lags)</td>
</tr>
<tr>
<td>Inflation lags</td>
<td>0 lags (pure NKPC), 1 lag, 4 lags</td>
</tr>
<tr>
<td>Parameter restrictions</td>
<td>No restrictions, $\gamma(1) = \gamma_f$ (inflation coefficients sum to 1)</td>
</tr>
<tr>
<td></td>
<td>With $\gamma(1) = \gamma_f$, use lags of $\Delta \pi$, instead of $\pi$, as instruments</td>
</tr>
<tr>
<td>Oil shocks</td>
<td>None, log change of WTI spot price divided by GDP def.</td>
</tr>
<tr>
<td>Interest rate</td>
<td>None, 90-day Treasury rate</td>
</tr>
<tr>
<td>GMM estimator</td>
<td>2-step, CUE</td>
</tr>
</tbody>
</table>

Notes: List of the specification options that we consider when estimating the NKPC (9). The efficient GMM weight matrix is computed using the Newey and West (1987) heteroskedasticity and autocorrelation consistent estimator with automatic lag truncation, except for VAR specifications, which use the White (1980) heteroskedasticity consistent estimator.

Source: Mavroeidis, Plagborg-Moller, Stock (2014)
To gauge the sensitivity of the results about the importance of forward-looking behavior to variations in data, sample, and identification assumptions, we obtain estimates of the coefficients \((\lambda, \gamma_f)\) in the baseline NKPC (9) for various combinations of the specification choices listed in table 4. We then plot the point estimates in \((\gamma_f, \lambda)\)-space. These plots do not convey any information about sampling uncertainty, i.e., they are not confidence sets. Confidence sets for a subset of those specifications are analyzed in section 5.3 below. However, these plots, which we refer to as “clouds,” do give a useful visual impression of the specification uncertainty.

We study the specifications with the labor share and output gap as forcing variable separately, because the coefficient \(\lambda\) on the forcing variable is not comparable across these cases. As we are only able to report a limited number of results here, we invite interested readers to explore the myriad of possible clouds using our interactive Matlab plotting tool, available in the online supplement 38.

We first look at the specification settings that have been used in the literature (i.e., not using real-time data or survey expectations as instruments). Figures 4 and 5 report the results with the labor share and output gap as forcing variable, respectively. Figure 4 also contains the Galí and Gertler (1999) vintage point estimate and associated Wald confidence ellipse from table 3 for comparison. These plots contain more than 600,000 estimates combined. Observe that the plotted parameter space \((\gamma_f, \lambda)\) is much larger than that of figure 3. Table 5 reports 38

---

**Figure 4.** Point Estimates: Labor Share Specifications

*Notes:* Point estimates of \(\lambda, \gamma_f\) from the various specifications listed in table 4 that use the labor share as forcing variable, excluding real-time and survey instrument sets. The black dot and ellipse represent the point estimate and 90 percent joint Wald confidence set from the 1998 vintage results in table 3.

Source: Mavroeidis, Plagborg-Moller, Stock (2014)
Summary statistics for the point estimates in figures 4 and 5. The main messages from the figures are that (i) estimates of the coefficient on $\lambda$ and $\gamma_f$ from the various specifications listed in table 4 that use the output gap as forcing variable, excluding real-time and survey instrument sets.

**Figure 5.** Point Estimates: Output Gap Specifications

*Notes:* Point estimates of $\lambda$, $\gamma_f$ from the various specifications listed in table 4 that use the output gap as forcing variable, excluding real-time and survey instrument sets.

**Source:** Mavroeidis, Plagborg-Moller, Stock (2014)
Overall conclusion:

“Literature has reached a limit on how much can be learned about the New Keynesian Phillips curve from aggregate macroeconomic time series.”

“New identification approaches and new datasets are needed to reach an empirical consensus.”
RECENT BEHAVIOR OF THE LABOR SHARE

- Since about 2000, labor share has been trending downward.
- If labor share is a good measure of marginal costs, downward trend should create massive deflationary pressure.
  (Coibion-Gorodichenko 15)
- Doesn’t seem to line up with the evolution of inflation.
Figure 1. Labor share, payroll share, and replicated labor share in U.S. nonfarm business sector.

Source: Bureau of Labor Statistics, Bureau of Economic Analysis, and authors’ calculations

Source: Elsby, Hobijn, Sahin (2014)