# Comment on "Optimal Inflation and the Identification of the Phillips Curve"
# (McLeay and Tenreyro, NBER Macro Annual 2019)

Matthew Rognlie

August 2019

Recently I had a dream where I was trying to explain supply and demand to an audience of intransigent economists. How could I say that demand curves sloped downward, they asked, when for so many goods, a simple plot of quantity demanded against price showed the opposite? How could these curves be useful concepts when, even by the most generous account, their parameters shifted from year to year?

I knew the answers. The incorrect slope of demand was no surprise; a plot of quantities against prices would only reveal the demand curve if the variation was caused by supply shocks. And yes, supply and demand did move around from year to year, but this didn't invalidate the concepts. If I'm selling oil, I can predict that a sudden decline in supply (embargo) will increase the price; I can predict that a new source of supply (fracking) will decrease the price; and I can predict that a decline in demand (fuel efficiency) will decrease the price. All of this is useful information.

The dream audience was unmoved. But waking up, I was relieved to find myself in a world where economists are not so silly. We understand that supply and demand is a useful framework, even if there's an identification problem, and even if there's no consensus on the parameters for a particular market. In fact, we persist even in the face of more profound theoretical complications, like imperfect competition or increasing returns.

Then I started reading commentary on the Phillips curve, and economists started seeming awfully silly again. Somehow, a weak reduced-form relationship in the aggregate data has led many people to deny the Phillips curve as a *structural* relationship. In the face of this criticism, McLeay and Tenreyro's paper is a vitally important rejoinder.

## 1 McLeay and Tenreyro's argument: the identification problem

The Phillips curve is a supply curve. It embeds the supply side of the economy, and it captures how in the aggregate, when greater pressure is put on the most important factor of production (labor), firms will want to set higher prices.

As with any supply curve, we need demand shocks for identification. A naive look at the data—if it doesn't take the source of variation into account—may find that the curve slopes in the wrong direction, or fail to detect any pattern at all.

In fact, it's worse than this: the "divine coincidence" in the basic New Keynesian model implies that when feasible, monetary policy should target zero output gap and zero inflation.[1] The only shocks that break this result are those that shift the relationship between the output gap and inflation—in other words, that shift the Phillips curve itself. But the optimal response to a "cost-push shock" to the Phillips curve that pushes up inflation is to engineer a negative output gap. As McLeay and Tenreyro powerfully point out, such shocks imply a negative relationship between the output gap and inflation in the data.

Effectively, under optimal monetary policy, there are *only* supply shocks, not demand shocks. This is the worst possible situation for identifying a supply curve.

One can illustrate the situation with an analogy. Suppose that a driver is on a long, hilly highway, and we want to study the relationship between the gas pedal and acceleration. If the driver is trying to maintain a constant speed, she'll only press harder on the pedal when trying to climb a hill. The raw data would indicate a clear negative relationship between the pedal and acceleration—the opposite of the causal relationship. And adding hills as a control variable wouldn't help: without hills, there would be no variation at all.

**Solving the identification puzzle: aggregate data.** To get a better result, we want drivers to make mistakes, ideally random ones: pushing down or letting up on the pedal for no good reason, so that the pedal-acceleration relationship is not caused by hills alone. This is related to an idea often used to study monetary policy: to look at an identified monetary shock. The response to these shocks can tell us about the Phillips curve, and indeed many other macro relationships of interest—see, for instance, Christiano, Eichenbaum and Evans (2005) or Barnichon and Mesters (2019).

The problem is that these shocks only exist if monetary policy makes random mistakes. Monetary policy in the most recent few decades, even if it has made plenty of mistakes, may have made these mistakes in a systematic enough fashion that true shocks are few and far between. This is the critique of Ramey (2016), and as McLeay and Tenreyro point out, it makes identifying the Phillips curve increasingly difficult.

One ideal situation for the econometrician would be a new, aggressive driver experimenting with radical new driving policies. And indeed, McLeay and Tenreyro find that the first part of Volcker's term has by far the most negative unemployment-inflation relationship of all periods grouped by Fed Chair, with a slope of -2.27.

Another possibility would be a driver struggling with a broken pedal, who is not equipped to maintain a constant speed. Yet in the Bernanke/Yellen era of the zero lower bound, McLeay and Tenreyro find a much weaker unemployment-inflation relationship, at just -0.13. Does this mean that the Phillips curve is weakening? Or is "broken pedal" not an accurate depiction of life at the zero lower bound, where the Fed was able to respond to shocks at the margin with forward guidance and quantitative easing, even if it couldn't eradicate the output gap entirely?

**Solving the identification puzzle: disaggregation.** The difficulty of answering this question has led many researchers to try their luck with an alternative strategy: using disaggregated data.[2] Since different states or metro areas in the US do not have their own monetary policies, they cannot adjust nominal interest rates in response to supply shocks. This avoids the bias that is created by monetary policy at the aggregate level.

Of course, real interest rates and relative prices can still adjust, so this approach is not free from bias, as McLeay and Tenreyro readily acknowledge. For instance, if a metro area experiences an adverse supply

[1] I remember the fury of my first-year graduate macro class, when—after tediously deriving the New Keynesian Phillips curve—I showed that inflation under optimal policy was zero anyway for all the shocks we'd considered thus far.

[2] Some examples of papers in this emerging literature are Fitzgerald and Nicolini (2014) and Beraja, Hurst and Ospina (2019).

shock that decreases productivity in both the tradable and nontradable sectors, then by making tradables less competitive this may increase unemployment, while simultaneously raising the price of nontradable goods that enter into the local CPI.

Still, this is a very useful source of evidence. Since the bias from cost-push shocks is downward, the results from this strategy provide a lower bound for the steepness of the Phillips curve, which may be tighter than the lower bound from aggregate data. In addition, I suspect that due to tradable goods and factors, the slope of regional Phillips curves should in theory be less steep than the slope of aggregate curves—so that, even more so, the regional data should understate the strength of the true relation.

In this light, McLeay and Tenreyro's results are encouraging: with year and metro area fixed effects, they find an instantaneous slope coefficient of -0.379 over the period 1990-2017, as compared to -0.150 from pooled OLS. My only caution is that in the fixed effects specification they also find much smaller coefficients on expectations and lags, so that if we interpret the results literally as a Phillips curve, then the slope of the medium-run Phillips relation—which accumulates many instantaneous responses through leads and lags—is not so steep in relative terms after all. I am not too worried by this, however: again, the regional Phillips curve is likely to understate the aggregate relationship.

**Two additional directions.** In the remaining discussion, I want to look in two directions that McLeay and Tenreyro—who understandably focus on the conventional form of the price Phillips curve—deemphasize. First, I will show that the *wage* Phillips curve is alive and well in the U.S. time series, even if transmission to prices is not immediately visible. Second, I will argue that the *theory* underlying the standard Phillips curve relies on extreme assumptions about rational expectations and common knowledge, and that although some form of Phillips curve is quite likely to exist, it need not resemble the Phillips curves derived in the conventional New Keynesian model.

## 2 The remarkable wage Phillips curve

In the aftermath of the Great Inflation and Volcker disinflation, the Phillips curve hit rock bottom in professional credibility. Summers (1991)—hardly a neoclassical—asked "Should Keynesian Economics Dispense with the Phillips Curve?". The answer was yes. He commented that the "the textbook Keynesian view of aggregate supply possesses many of the attributes that Thomas Kuhn has ascribed to dying scientific paradigms."

But a funny thing has happened in the post-Volcker era: the original Phillips curve, the negative correlation between unemployment and *wage* growth, has returned in plain sight.

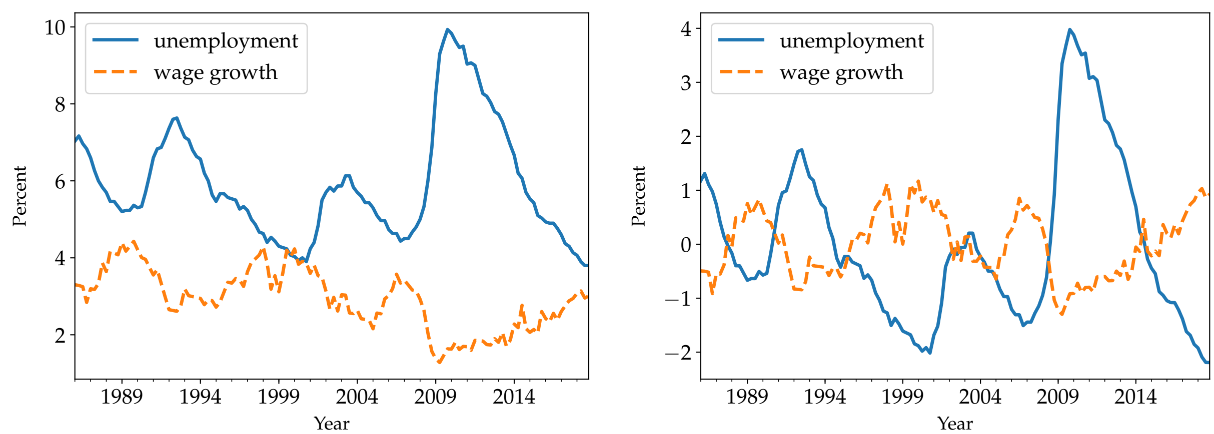
Figure 1a shows the two time series: wage growth for private industry, measured in a centered 4-quarter window around each quarter using the Employment Cost Index, and average unemployment within the quarter. The negative comovement is visually obvious: when unemployment spikes during recessions, wage growth plummets, and as unemployment creeps back down during the expansion, wage growth reasserts itself. This does not necessarily prove causality, or that the unemployment rate is the best measure of labor market slack for the Phillips curve, but as a correlation, the wage-unemployment Phillips curve appears healthy as ever.

Figure 1b detrends the data, removing secular declines in both the unemployment rate and wage growth. The correlation becomes even more apparent: every swing in unemployment has an effect on wage growth as its mirror image. A scatterplot of these observations, in figure 2a, reveals a remarkably healthy Phillips

Figure 1: Wage inflation (centered 12-month ECI growth) vs. unemployment, 1986–present

(a) Raw series

(b) Detrended series




curve. For most of the domain, the slope is roughly -1/2, with an increase in unemployment of 2 percentage points corresponding to annual wage growth that is 1 percent smaller. But it appears to flatten when unemployment reaches more than 2 percentage points above trend. The observations in this region are all from the worst period of the Great Recession, where wage growth didn't fall nearly as much as a linear extrapolation of the Phillips curve would suggest.

All this is consistent with a traditional view of the wage Phillips curve: the curve is convex, and secular trends in the natural rate of unemployment (and possibly inflation expectations) imply that the detrended data has a better fit. But does this have a causal interpretation? And what does it say about *price* inflation?

**Transmission to price inflation.** If labor market tightness causes nominal wage inflation, then transmission to price inflation is not guaranteed in the short run: although wages are an important component of overall costs, they are not 100% of costs, and markups can absorb year-to-year variation. Figure 2b, constructed analogously to figure 2a but for core PCE price inflation rather than wage inflation, reveals just how weak this short-run transmission can be: the modern price "Phillips curve" is a cloud of points with no obvious interpretation, and at best a slight negative correlation.

In the longer run, however, one-for-one transmission to price inflation seems almost inevitable. Figure 3, which plots the net labor share of corporate factor income in the US, shows why. Although there is plenty of short-to-medium term variation in the labor share, most of this is mean-reverting, and the long-run variation appears to be bounded.

A fairly large movement in the labor share would be, for instance, from 75% to 80%. If this corresponded to a decline in markups, it would imply excess wage growth over price growth of $80\%/75\% = 1.067$. Annualized over 5 years, this would be a big deal: about 1.3% per year. But over 50 years, it would only be 0.13% per year—not a major influence on long-run price inflation. There is simply no realistic *level* change in markups that can matter much, over the span of decades, for the *rate* of price inflation vs. wage inflation.

**Are other interpretations possible?** If this story is right, then the medium-to-long-term price Phillips curve should be robust, and the conventional central bankers' view of inflation is correct—over a long

Figure 2: Detrended wage and price growth vs. unemployment, 1986–present

(a) Wages (centered 12-month ECI growth) (b) Prices (centered 12-month core PCE inflation)

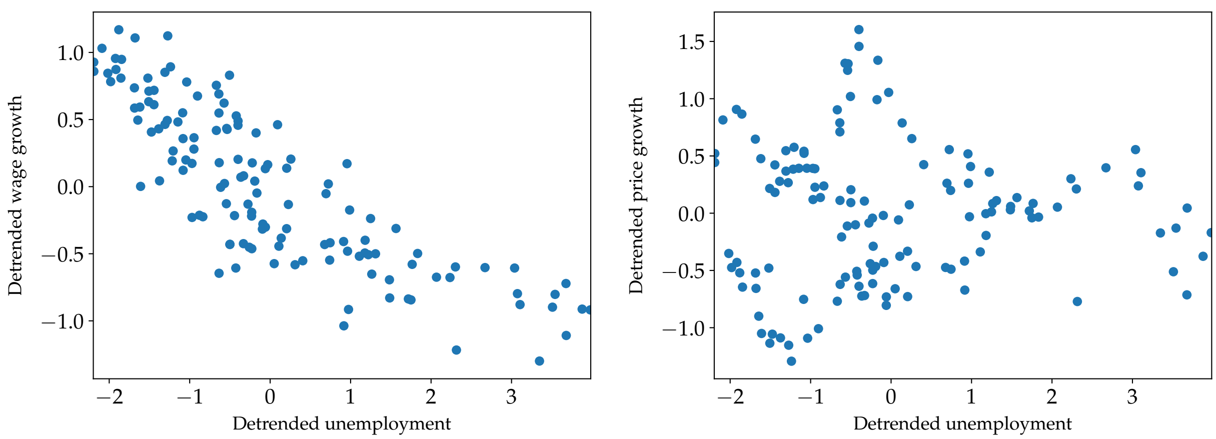
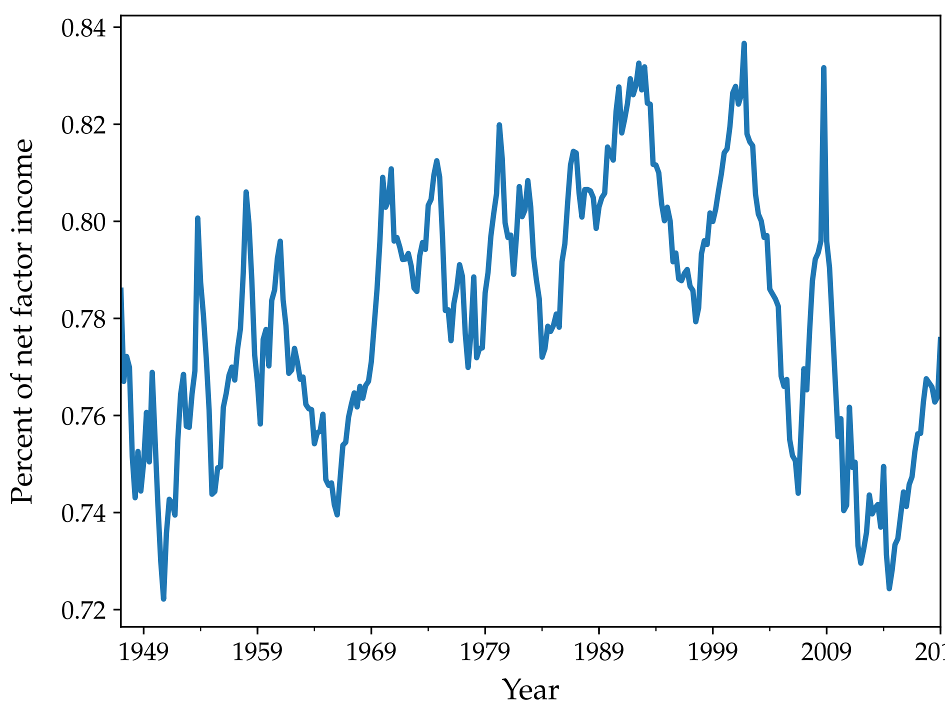



Figure 3: Net labor share of corporate factor income (NIPA)



enough time horizon.

Is there any alternative? Since the wage Phillips curve in figure 2a is just a correlation, the case is not quite settled. I find it difficult, however, to think of a different story that matches both the data and our own understanding of the labor market. By and large, firms set wages in nominal terms, and in response to tighter hiring conditions and the threat of turnover, they allow nominal wages to grow a little more. If this wage growth causes markups to shrink but tight labor market conditions persist, these firms are likely to restore their markups by raising prices, not trimming wages (which would be suicidal in a tight labor market).

Perhaps reduced markups exert some downward pressure on wages—so that more wage growth this year, by reducing markups, means less growth next year, as firms adjust wages to bring markups back into line. If this effect was strong enough, in principle it could stop transmission to price inflation entirely: firms would keep prices constant and allow markups to absorb wage growth fluctuations in the short run, then eventually reverse course to avoid any long-term change in markups. But in this case, I'd imagine that the unemployment-wage relationship would be disrupted, to the point where figure 2a would be far less clean.

**Why is the wage Phillips curve so much clearer than the price Phillips curve?** I have suggested a story for why figure 2a is so much nicer than figure 2b: the link between labor market tightness and wage growth is much more immediate than the link with price growth. First, markups provide a short-run buffer between wages and prices. Additionally, several other shocks may affect prices more than nominal wages—for instance, energy shocks, or even some productivity shocks—further muddying the price Phillips curve.

McLeay and Tenreyro offer a distinct and complementary story: since central banks target price rather than wage inflation, we should expect the identification problem to be worse for the price Phillips curve. This is a nice corollary to their main argument, and I think it is almost certainly part of the explanation for why the wage curve performs so much better. One can even interpret the Great Recession in this light: one reason why central banks were not more aggressive to combat unemployment in the early 2010s was that they were looking at price inflation (which, thanks in part to energy prices, was surprisingly high) rather than wage inflation (which was not).

More generally, McLeay and Tenreyro's identification argument is compelling enough that it's hard to imagine figure 2a could exist in a world where central banks systematically targeted wage inflation.

# 3 The subtle theory of Phillips curves

Most of us remember the standard formulation of the New Keynesian Phillips curve for price inflation. McLeay and Tenreyro use this formulation in their theoretical discussion—understandably, since the point of their paper is to address the mainstream New Keynesian literature.

But where does this equation come from? I will argue that the usual specification is surprisingly subtle and assumption-contingent. Some form of Phillips curve likely exists, but it is probably not the one in the textbooks.

To start, consider the most basic log-linearized Calvo model as in Galí (2015), with no firm-specific cost shocks. This implies that firm $i$ setting its price in period $t$ chooses an optimal price

$$p_{it}^* = (1-\beta\theta)\mathbb{E}_{it}\sum_{s=0}^{\infty}(\beta\theta)^s(\Theta\widehat{mc}_{t+s}+p_{t+s}) \tag{1}$$

where $p_{t+s}$ is the aggregate price level at time $s$, $\widehat{mc}_{t+s}$ is the deviation of aggregate real marginal cost from steady state, $\beta$ is the discount rate, $1-\theta$ is the likelihood of price adjustment, and $\Theta \in [0,1]$ is closer to zero when there is more strategic complementarity.

Equation (1) is straightforward, and not hard to derive from the basic firm problem. In the case $\Theta = 1$ with no strategic complementarity, it simply states that the firm wants to move its price in proportion to the discounted average of expected nominal marginal cost, over the lifetime of that price. The discount rate $\beta$ and price survival probability $\theta$ combine into the factor $\beta\theta$ in (1) that discounts future costs.

Applying the law of iterated expectations, we can consolidate (1) to the simpler-looking expression

$$p_{it}^* = \mathbb{E}_{it}[(1-\beta\theta)(\Theta\widehat{mc}_t + p_t) + \beta\theta p_{it+1}^*] \tag{2}$$

which replaces terms from $t+1$ onward with firm $i$'s expectation of its own optimal reset price at $t+1$. This is possible because of the geometric discounting in (1), under which tomorrow's expected choice enters today as a sufficient statistic for the future.

I would argue that (1) and (2) are somewhat robust, in the sense that many different formulations of a forward-looking pricesetting problem will lead to similar expressions. In the absence of exact geometric discounting (like that produced by Calvo or Rotemberg pricing), we would not be able to consolidate to the form (2), but it might not be *too* far off. Perhaps the biggest weakness is the implicit invocation of rational expectations when applying the law of iterated expectations between (1) and (2).

The much more doubtful features of the New Keynesian Phillips curve emerge *after* (2). How, indeed, do we go from discounting by $\beta\theta$ in (1) and (2) to discounting by just $\beta$ in the traditional NKPC? The answer is a near-miracle of common knowledge and rational expectations.

To see this, we can rewrite (2) as

$$\pi_{it}^* = (1-\beta\theta)\Theta\mathbb{E}_{it}\widehat{mc}_t + \beta\theta\mathbb{E}_{it}\pi_{it+1}^* + \mathbb{E}_{it}\pi_t \tag{3}$$

where $\pi_t \equiv p_t - p_{t-1}$ and $\pi_{it}^* \equiv p_{it}^* - p_{t-1}$

Now, if we impose common knowledge, we can—since there are no firm-specific cost shocks—drop the $i$ subscripts and also drop the expectations operator for date-$t$ outcomes. Then, if we multiply by $(1-\theta)$ and use the log-linearized aggregate price law of motion $\pi_t = (1-\theta)\pi_t^*$, we can rewrite this as

$$\begin{aligned}(1-\theta)\pi_t^* &= (1-\beta\theta)(1-\theta)\Theta\widehat{mc}_t + \beta\theta\mathbb{E}_t(1-\theta)\pi_{t+1}^* + (1-\theta)\pi_t \\ \pi_t &= (1-\beta\theta)(1-\theta)\Theta\widehat{mc}_t + \beta\theta\mathbb{E}_t\pi_{t+1} + (1-\theta)\pi_t \end{aligned} \tag{4}$$

The final step is to subtract the $(1-\theta)\pi_t$ on the far right of (4) from both sides and divide by $\theta$, giving the standard form of the marginal cost NKPC:

$$\pi_t = \frac{(1-\beta\theta)(1-\theta)}{\theta}\Theta\widehat{mc}_t + \beta\mathbb{E}_t\pi_{t+1} \tag{5}$$

Now we know how the discount factor of $\beta\theta$ in (2) becomes just $\beta$ in (5): the term $(1-\theta)\pi_t$ on the right of (4) added an extra factor of $\theta^{-1}$. But we should also be suspicious: originally we had $\mathbb{E}_{it}\pi_t$ in (3), and we were able to replace by $\pi_t$ to get the term on the right only by imposing common knowledge. Away from common knowledge, we will generally have $\mathbb{E}_{it}\pi_t \neq \pi_t$; and if variation in higher-order expectations is muted, this will result in much more equilibrium discounting than in (5). If expectations are formed adaptively, or higher-order expectations coalesce around recent events or historical averages, then the true Phillips curve will have lags as well.[3]

And, of course, it is an even greater leap to write a Phillips curve in terms of the output gap or unemployment directly, rather than with marginal cost as in (1). If the relationship between the output gap and marginal cost is more complex than in the simplest New Keynesian model—as it surely is—then swapping the two without further modifying the NKPC is incorrect.[4]

**Lessons from the theory.** At a conceptual level, the existence of *some* kind of Phillips curve seems quite robust. Prices are set in response to costs, and if there are nominal rigidities impeding immediate adjustment, then the rate of price *change* will respond to costs. Costs, in turn, are affected by demand pressure. The same is true for wages and the labor market.

But to go from a result like (3)—which requires "only" firm optimality and rational expectations—to the specific, standard form of the New Keynesian Phillips curve requires extraordinary assumptions, assumptions that are sure to be false. We should not expect such a curve to appear robustly in the data. If we try to estimate it, the numbers we'll get back are at best reduced-form correlations, which need not carry over between different policy regimes or the impulses to different shocks. For instance, if inflation targeting has been successful at anchoring expectations—especially higher-order expectations—at a certain level, then the Phillips curve (5) is likely to be shallower and feature much more discounting.[5]

Both McLeay and Tenreyro, and the applied literature more broadly, are sophisticated about this. They contemplate specifications with, for instance, different mixes of leads and lags (and also none at all). They run analyses separately by subperiod to account for the possibility that inflation targeting or other changes in policy have altered the curve.

I worry, however, that as a profession we have not fully acknowledged just how hard it *should* be to estimate Phillips curves, especially price Phillips curves, even if we understand the identification conundrum posed by McLeay and Tenreyro.[6] Economists love nothing better than a single, tractable equation—officially blessed by microfoundations—that can be taken to the data. The New Keynesian Phillips curve is exactly this. Most card-carrying macroeconomists can recite it, but I suspect that very few could explain exactly how it is derived, or what assumptions are required.[7]

This has led to some unfortunate polarization of the debate. We must either, it sometimes seems, accept the Phillips curve as a quantitative relationship that should be estimated in some form, or else deny the existence of a Phillips curve altogether. Yet a much better answer is in the middle. Like supply and demand, the Phillips curve is a conceptual building block that is indispensable to thinking about the world. It can also offer rough guidance to supplement our very limited information—for instance, by telling us in 2010 or 2011 that sustained inflation was highly unlikely. The standard is not, and should not be, a stable curve that holds up in regression after regression.

This is not to deny the importance of empirical work on the relationship between slack and inflation.

[3] None of these ground is new to me: it is covered, for instance, by Angeletos and Lian (2018) and Angeletos and Huo (2018).

[4] Applying this reasoning, Sbordone (2002) estimated a price Phillips curve directly using evidence on marginal cost.

[5] This raises the unpleasant possibility that inflation targeting may be a self-defeating approach to stabilization, since by weakening the Phillips curve it cannibalizes its own signal. This would be an extreme version of the paper's mechanism: beyond just limiting the visibility of the Phillips curve to the econometrician, the central bank's behavior might drastically weaken the relationship itself.

[6] On top of the theoretical issues I emphasize, Mavroeidis, Plagborg-Møller and Stock (2014) convincingly demonstrate the formidable econometric difficulties of estimating Phillips curves.

[7] One piece of evidence for this statement: although substituting non-rational or heterogeneous expectations of $\pi_{t+1}$ directly into (5) is incoherent—given that the derivation requires rational expectations and common knowledge, and relaxing these assumptions changes other parts of the equation—many prominent papers do exactly that.

But these efforts should aim to be indicative, not authoritative. A goal of estimating "the" Phillips curve is almost certainly too ambitious, and the best work may have more of a micro flavor, building up our knowledge of price- and wage-setting along whatever dimensions we can. Empirical success is more likely when there are fewer steps in the causal chain (unemployment to wage inflation), and less likely when there are more (output gap to price inflation).

Before we can do any of this, however, we need to throw out bad arguments against Phillips curves, and understand why estimating them poses such a challenge. This makes McLeay and Tenreyro's paper an essential foundation for future work.

## References


**Angeletos, George-Marios and Chen Lian**, "Forward Guidance without Common Knowledge," *American Economic Review*, 2018, *108* (9), 2477–2512.

_ **and Zhen Huo**, "Myopia and Anchoring," 2018.

**Barnichon, Regis and Geert Mesters**, "Identifying Modern Macro Equations with Old Shocks," 2019.

**Beraja, Martin, Erik Hurst, and Juan Ospina**, "The Aggregate Implications of Regional Business Cycles," *Econometrica (forthcoming)*, 2019.

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

**Fitzgerald, Terry J and Juan Pablo Nicolini**, "Is There a Stable Relationship Between Unemployment and Future Inflation?: Evidence from US Cities," 2014.

**Galí, Jordi**, *Monetary Policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications*, Princeton University Press, 2015.

**Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H Stock**, "Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve," *Journal of Economic Literature*, 2014, *52* (1), 124–88.

**Ramey, Valerie A**, "Macroeconomic Shocks and Their Propagation," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 71–162.

**Sbordone, Argia M**, "Prices and Unit Labor Costs: a New Test of Price Stickiness," *Journal of Monetary Economics*, 2002, *49* (2), 265–292.

**Summers, Lawrence H**, "Should Keynesian Economics Dispense with the Phillips Curve?," in "Issues in Contemporary Economics," Springer, 1991, pp. 3–20.