The Cyclicality of the Price-Cost Markup

Christopher J. Nekarda
Board of Governors of the Federal Reserve System

Valerie A. Ramey
University of California, San Diego and NBER

May 27, 2019

Abstract

A countercyclical markup of price over marginal cost is a key transmission mechanism for demand shocks in New Keynesian (NK) models. This paper re-examines the foundation of those models by studying the cyclicality of the price-cost markup in the private economy. We find that how the markup is measured matters for its unconditional cyclicality. Measures of the markup based on the inverse of the labor share are moderately procyclical, but are moderately countercyclical for some generalizations of the production function. NK models predict that the cyclicality of the markup should vary depending on the nature of the shock. Consistent with the NK model, we find that the markup is procyclical conditional on TFP shocks and countercyclical conditional on investment-specific technology shocks. In contrast, we find that the markup increases in response to a positive demand shock. Thus, the transmission mechanism for the effects of demand shocks in sticky-price NK models is not consistent with the data. Our results suggest that NK models might benefit from a renewed focus on wage rigidities.

This paper updates Nekarda and Ramey (2013). The views in this paper are those of the authors and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff. We are grateful to Susanto Basu, Mark Bils, Larry Christiano, Olivier Coibion, Steven Davis, Davide Debortoli, Martin Eichenbaum, Robert Hall, Garey Ramey, Sergio Rebelo, Matthew Shapiro, Harald Uhlig, Raf Wouters, and participants at numerous seminars for helpful comments and suggestions. Ben Backes and Myungkyu Shim provided excellent research assistance. Valerie Ramey gratefully acknowledges financial support from National Science Foundation grant SES-0617219 through the National Bureau of Economic Research.
1 Introduction

The markup of price over marginal cost plays a key role in New Keynesian (NK) macroeconomic models. In these models, a demand shock raises output by lowering the price markup. As pointed out by Broer et al. (2019), a lower markup leads to higher output during booms largely through its effect on profits. In particular, lower markups reduce profits, generating a negative wealth effect that induces households to raise their labor supply. The two-agent New Keynesian and heterogeneous-agent New Keynesian (HANK) models also rely heavily on countercyclical price markups to amplify shocks. As Debortoli and Galí (2018) point out, “in both models the amplification/dampening of aggregate shocks depends critically on the cyclical properties of markups (or equivalently the labor share)...”\(^1\) Indeed, price markups are required to be so countercyclical in the leading HANK models that expansionary monetary shocks cause profits to fall.\(^2\) Even in the medium-scale NK models that also incorporate sticky wages, countercyclical movements in the price markup play a key role in the transmission of monetary and fiscal policy shocks. For example, in the estimated dynamic stochastic general equilibrium model from Smets and Wouters (2007), the price markup is countercyclical in response to a monetary shock.

The dependence of Keynesian models on a countercyclical price markup is a feature only of the models formulated since the early 1980s. From the 1930s through the 1970s, the Keynesian model was founded on the assumption of sticky wages.\(^3\) Some researchers believed that the implications of this model were at odds with the cyclical properties of real wages, leading to a debate known as the “Dunlop-Tarshis” controversy.\(^4\) In response to the perceived disparity between the data and predictions of the traditional Keynesian model, the literature shifted in the early 1980s to relying on sticky prices rather than sticky wages for the transmission of shocks.\(^5\) While the medium-scale NK models add wage stickiness, virtually all current NK models rely on countercyclical price markups in response to demand shifts.

---

1. p. 31.
3. Such as Keynes (1936); Phelps (1968); Taylor (1980).
4. In fact, Dunlop (1938) and Tarshis (1939) were repeatedly misquoted by the literature as showing that real wages were procyclical. Neither of them showed this. Both authors showed that money wages and real wages were positively correlated, and Tarshis went on to show that real wages were in fact negatively correlated with aggregate employment. Dunlop (1998) discusses the debate in his retrospective article.
5. Gordon (1981); Rotemberg (1982).
Is the price markup countercyclical in the data? There is no consensus, because estimating the cyclicity of the markup is one of the more challenging tasks in macroeconomics. Theory prescribes a comparison of price and marginal cost; however, available data typically include only average cost. As we will discuss, researchers have used a variety of techniques to measure the markup directly, or have inferred its movements using indirect evidence. Some researchers have estimated the markup to be procyclical while others have estimated it to be countercyclical.

In this paper, we assess how various measures of the aggregate markup move over the business cycle and how they respond to leading macroeconomic shocks. We find that how the markup is measured matters for its unconditional cyclicity — that is, its relationship with an indicator of the business cycle. Markups measured as the inverse of the labor share are moderately procyclical, but markups based on more general production functions are procyclical or countercyclical depending on the details of the empirical implementation.

Our main focus is on the conditional cyclicity of the markup — that is, how the markup responds to a particular type of shock. Studying the conditional cyclicity of the markup is the appropriate way to test the predictions of NK models, since they predict the markup should behave differently in response to different shocks. Unlike the unconditional cyclicity, our estimates of the sign of the conditional cyclicity do not depend on the empirical measure of the markup. Consistent with the NK model, we find that the markup is procyclical conditional on total factor productivity (TFP) shocks and countercyclical conditional on investment-specific technology (IST) shocks. In contrast to the sticky-price NK model predictions, we find that the markup increases in response to expansionary monetary policy shocks and government spending shocks. Thus, the transmission mechanism for the effects of these policy shocks in sticky-price NK models is not consistent with the data.

Based on this evidence, we join a rising chorus recommending that the NK model return to the traditional Keynesian emphasis on wage rigidities rather than price rigidities. Broer et al. (2019) advocate a shift from price stickiness to wage stickiness based on insights from heterogeneous agent models. Auclert and Rognlie (2017) recommend the same shift based on undesirable interactions between Greenwood, Hercowitz and Huffman (1988) preferences and flexible wages. Based on our findings of the conditional cyclicity of the price markup, we too recommend this shift. As discussed earlier, the macroeconomics literature originally abandoned sticky wage models because real
wages were found to be mildly procyclical, which was believed to be at odds with the predictions of the model. What that pre–Kydland and Prescott (1982) literature failed to realize was that business cycles are driven by a multitude of shocks, not just demand shocks, so simple analyses of the unconditional cyclicality of variables such as real wages and price markups are not dispositive.

2 Relationship to the literature

Industrial organization economists have a long history of studying the cyclicality of price-cost margins. Macroeconomists only began studying this issue in the mid-1980s when macro models started to emphasize price setting behavior of firms. Four principal methods have been used to measure the markup directly and two additional methods have been used to assess the cyclicality of the markup indirectly.

The first of the direct methods uses the standard industrial organization concept of a price-cost margin constructed from revenues and variable costs. Domowitz, Hubbard and Petersen (1986) use this method in a panel of four-digit standard industrial classification (SIC) manufacturing industries and find that margins are significantly procyclical. Anderson, Rebelo and Wong (2018) use confidential detailed data from the retail industry and measure markups by comparing well-measured individual product prices to the replacement cost of the good. This latter cost measure should be a very good proxy for marginal cost. They find that markups are acyclical or mildly procyclical.

The second method builds on Hall’s (1986) generalization of the Solow residual to estimate the cyclicality of markups. For example, Haskel, Martin and Small (1995) extend Hall’s framework to allow for time-varying markups and apply it to a panel of two-digit U.K. manufacturing industries. They find that markups are markedly procyclical. Marchetti (2002) applies a similar framework to two-digit manufacturing industries in Italy. He finds no clear pattern of cyclicality of markups; in only 2 of 13 industries does he find consistent evidence across specifications of countercyclical markups.

The third method uses generalized production functions with quasi-fixed factors to estimate markups relative to marginal cost estimated from stochastic Euler equations. Using this type of approach, Morrison (1994) finds weakly procyclical markups in Canadian manufacturing, and Chirinko and Fazzari (1994) find acyclical or procyclical markups in firm-level data. Galeotti and Schianterelli (1998) test the Rotemberg and Saloner (1986) game-theoretic hypothesis and find that, consistent with this hy-
hypothesis, markups depend negatively on the current level of output but positively on the growth of output.

The fourth method is the only one that aligns with the measured markup in NK models. This method uses the labor input margin to estimate marginal cost. Under standard assumptions, such as Cobb-Douglas (C-D) production functions and no overhead labor, this method implies that the markup is inversely proportional to the labor share. Since the labor share is countercyclical during the post-WWII period, this measure of markups implies that markups are on average procyclical. Most of the papers using reduced form methods to measure the cyclicity of markups have applied adjustments to the standard model to account for reasons why marginal labor costs might be more procyclical than average labor costs. For example, Bils (1987) argues that the marginal hourly wage paid to workers should be more procyclical than the average wage. He constructs a measure of marginal cost based on estimates of the marginal wage and finds that his markup series has a negative correlation with industry employment in a panel of two-digit industries, suggesting countercyclicality. Rotemberg and Woodford (1991), Rotemberg and Woodford (1999), Oliveira Martins and Scarpetta (2002), and Galí, Gertler and López-Salido (2007) apply additional adjustments to the standard model, such as substituting a constant elasticity of substitution (CES) production function for C-D and allowing for overhead labor. Their applications of these adjustments typically convert procyclical markups (based on standard assumptions) into countercyclical markups. Bils, Klenow and Malin (2018) (BKM) argue that wages are not allocative in typical employment relationships and formulate other measures of markups using either self-employed workers or intermediate goods. Their measures for the period after 1987 suggest countercyclical markups.

The two indirect methods for assessing the cyclicity of the markup use entirely different frameworks. Bils and Kahn (2000) present a model of inventories and stockouts in which the joint cyclicity of the ratio of sales to inventories and the discounted growth rate of output prices reveals the cyclicity of markups. They use this framework to conclude that markups are countercyclical in several two-digit U.S. manufacturing industries. Hall (2012) exploits standard advertising theory to show that countercyclical markups imply that advertising should also be highly countercyclical. He shows, in fact, that advertising is somewhat procyclical.

More recently, several papers have studied the decomposition of the labor wedge—that is, the log difference between the marginal product of labor and the representative household’s marginal rate of substitution between leisure and consumption. The labor wedge can be decomposed into the log sum of the price markup and the wage markup. Since the labor wedge is markedly countercyclical, either the price markup or the wage markup or both must be countercyclical. Karabarbounis (2014) works through numerous cases, employing a variety of data sets and assumptions about preferences and concludes that almost all of the countercyclicality of the labor wedge is due to the wage markup. In contrast, BKM argue that wages are not allocative in typical employment relationships. They instead study self-employed workers and argue that their labor wedge must be due entirely to a price markup. Their estimate of the labor wedge for self-employed workers is as countercyclical as the one for the aggregate economy. On this evidence they conclude that price markups are countercyclical.

Finally, a relatively recent literature has documented and offered explanations for the global decline in the labor share, such as Karabarbounis and Neiman (2014). Some of the papers in this literature have introduced additional methods for estimating the markup. Barkai (2017), Gutiérrez (2017), and Gutiérrez and Philippon (2017) assume constant returns to scale and infer markups from measured profit rates. Karabarbounis and Neiman (2018), however, offer arguments against the profit rate approach. De Loecker, Eeckhout and Unger (2018) follow the cost-minimization approach but instead of focusing on only labor, they use Compustat’s variable on the cost of goods sold as a measure of the variable input in order to infer markups. Since the focus of this literature is on trends, none of these papers has analyzed the cyclicality of their measures of markups.

Overall, this literature has used a host of innovative and clever ways to measure markups. Given the mixed results of the literature, it is surprising that the countercyclicality of markups is often treated as an uncontroversial stylized fact by the NK literature.

In this paper, we first revisit the arguments from the literature and then proceed to construct markups based on new data and new methods to implement the theoretical measures. We argue that the cyclicality of markups sheds light on the NK model only when analyzed conditional on the type of shock. Our results suggest that markups do not behave as predicted in response to traditional NK demand shocks, such as monetary shocks or government spending shocks.
3 Theoretical framework

This section lays out the theoretical framework that forms the basis of our main estimates of the markup. We first derive general expressions for the markup based on the firm's cost minimization problem. Next, we explain why we think that labor hours is the best margin for measuring the markup. Finally, we show several possible measures of the markup based on assumptions about the production function.

3.1 Deriving the price markup from cost minimization

The theoretical markup, $M$, is defined as

$$M = \frac{P}{MC},$$

where $P$ is the price of output and $MC$ is the nominal marginal cost of increasing output. The inverse of the right hand side of equation 1, $MC/P$, is also known as the real marginal cost in the NK literature.

A cost-minimizing firm should equalize the marginal cost of increasing output across all possible margins for varying production. Thus, it is valid to construct the marginal cost of varying output based on changing any one input. Most of the literature has considered variable inputs in order to avoid the challenges involved in estimating adjustment costs.

Focusing on the cost-minimization problem for variable inputs, consider the problem of a firm that chooses variable inputs $x_i$, $i = 1, \ldots, N$ to minimize

$$\text{Cost} = \sum_{i=1,\ldots,N} (w_i \cdot x_i) + \text{terms not involving } x,$$

subject to

$$\bar{Y} = F(x_1, x_2, \ldots, x_N),$$

where $w_i$ is the factor price, $x_i$ is the variable input, $Y$ is output, and $F(\ldots)$ is the production function. Letting $\lambda$ be the Lagrange multiplier on the constraint, we obtain
the first-order condition for \( x_i \) as:

\[
(4) \quad w_i = \lambda \cdot \frac{\partial Y}{\partial x_i}.
\]

Since the multiplier \( \lambda \) is equal to the marginal cost of raising output, we can substitute equation 4 into equation 1 to derive the markup based on using input \( x_i \) to raise output:

\[
(5) \quad M_{x_i} = \frac{1}{s_{x_i}} \cdot \left( \frac{\partial Y}{\partial x_i} \cdot \frac{x_i}{Y} \right)
\]

where

\[
(6) \quad s_{x_i} = \left( \frac{w_i \cdot x_i}{P \cdot Y} \right)
\]

is \( x_i \)'s factor share of output. The term in parentheses in equation 5 is the elasticity of output with respect to \( x_i \). Thus, the markup can in theory be measured as the product of the inverse of any variable input’s share and the output elasticity with respect to that input.

### 3.2 Why we measure the markup using the labor margin

In principle, one can choose the first-order condition for any variable input as the basis for the markup measure. The traditional variable input studied is the labor input margin. Hamermesh and Pfann’s (1996) summary of the literature suggests that adjustment costs on the number of employees are relatively small and that adjustment costs on hours per worker are essentially zero. Bils (1987) and Rotemberg and Woodford (1999) study markups based on the hours per worker margin, and virtually all modern NK models use the total labor hours margin.

Some have criticized the use of the labor margin, arguing that a key part of the marginal cost measure—average hourly wages—may not be a good indicator of the true marginal cost of an extra hour of work. This critique takes two forms. The first is Bils's (1987) argument that the ratio of the marginal wage to the average wage for a particular worker may be procyclical because of an overtime premium or other costs associated with inducing extra hours worked. Using approximations, various simplifying assumptions, and annual industry data, he found that his adjustments to average wages transformed the markup from being procyclical to countercyclical. Appendix D
revisits Bils's (1987) argument. It derives a ratio of marginal to average wages based on observables and shows that once we replace Bils's (1987) approximations and simplifying assumptions with exact expressions and the richer data that is now available, there is very little cyclical variation in the marginal to average wage ratio. Thus, we conclude that the average wage is a good indicator of the marginal wage for a worker.

The second critique of the labor hours margin and its reliance on average hourly wages is based on ideas from the implicit contract literature. This critique argues that because wages may be payments on an implicit contract in an ongoing employment relationship, they are smoothed relative to the true marginal cost of increasing hours. There are two leading pieces of evidence offered in support of this view. The first is Bils's (1985) finding that new hire wages are more procyclical than existing worker wages. The second is Beaudry and DiNardo's (1991) finding that workers’ long-term wages depend on the state of the economy at the time they were first hired.

Two recent papers call into question the interpretation of this evidence. First, Hagedorn and Manovskii (2013) show theoretically that changing quality of job matches over the business cycle can lead to composition effects. They demonstrate empirically that the results of Bils (1985) and Beaudry and DiNardo (1991) disappear after controlling for job match quality. Second, Gertler, Huckfeldt and Trigari (2019) document that new hire worker wages are no longer more procyclical than existing worker wages once one adjusts for the cyclicality of the composition of new hires. Based on their new findings, Gertler, Huckfeldt and Trigari (2019) conclude that “the sluggish behavior of wages for existing workers is a better guide to the cyclicality of the marginal cost of labor than is the high measured cyclicality of new hires wages unadjusted for composition effects.” Thus, the latest findings support the use of average wages as a measure of marginal cost of labor hours.

Recent work has provided evidence that, even if wages are sticky, they appear to be allocative. Olivei and Tenreyro (2007), Olivei and Tenreyro (2010), and Bjöklund, Carlsson and Skans (2019) present evidence that monetary shocks have larger effects when they occur just after annual labor contracts are signed, suggesting that wage stickiness has real effects.

Based on the latest evidence, as well as its prominent role in NK models, we therefore choose to measure the markup using the labor input margin.

---

7. Such as Barro (1977) and Hall (1980).
Of course, one could consider several margins. Two other possibilities are offered by BKM, who argued against the standard labor measure based on implicit contract arguments. One of their alternatives is based on self-employed individuals. They measure the markup as the log difference between the current business income per hour worked and the marginal rate of substitution for self-employed workers. We believe this measure has some weaknesses. For example, it relies on the assumption that the returns to self-employed labor are adequately reflected in current business income. Recent work by Bhandari and McGrattan (2019) estimates that the self-employed spend many hours building up customer bases and other intangible capital and that intangible capital is worth 60 percent of total assets of these businesses. The returns to those intangible assets are realized over a long period of time and are not adequately reflected in current business income. If investment in intangible capital is procyclical, then BKM’s markup measure will be biased toward countercyclicality.\(^9\)

BKM’s second set of alternative measures relies on variations in intermediate inputs, such as materials, energy, and business services. The elasticity of substitution between the various intermediate goods and value added is a key part of the construction of their markups. There are far fewer estimates of those elasticities of substitution than of the elasticity of substitution between capital and labor. Thus, we believe that the labor margin implementation relies on better-established estimates.

### 3.3 Production function assumptions

The markup using the labor input margin can be expressed as

\[ M = \frac{1}{s} \cdot \left( \frac{\partial Y}{\partial L} \cdot \frac{L}{Y} \right) \]

where \( s \) is the labor share of output and the term in parentheses is the elasticity of output with respect to labor hours \( L \).

The formula for the markup above requires an estimate of the marginal product of labor, necessitating assumptions about the production function. Under the standard assumptions that the production function is Cobb-Douglas (denoted by a superscript

\(^9\) In addition, we show in appendix C that the unconditional markup became more countercyclical after 1995. Because BKM’s sample starts only in 1987, their study is more likely to find a countercyclical markup.
“CD”) in total hours, the markup is given by

\[ M_{CD} = \frac{\alpha}{s}, \]  

where \( \alpha \) is the exponent on labor input in the production function and \( s \) is the labor share.

Rotemberg and Woodford (1999) note several reasons why the standard assumption of a production function that is C-D in total hours may lead to estimates of the markup that are biased toward being procyclical. We now consider the most plausible generalizations.

The first generalization is overhead labor. In this generalization, the labor term in the production function is instead \( (L - \bar{L})^\alpha \) where \( \bar{L} \) represents overhead labor hours. With a C-D production function and overhead labor (denoted by “CD, OH”), the markup is given by:

\[ M_{CD, OH} = \frac{\alpha}{s'}, \]  

where

\[ s' = \frac{W (L - \bar{L})}{PY}, \]  

is the labor share of non-overhead labor.

A second generalization allows the elasticity of substitution between inputs to deviate from unity. For example, consider the following CES production function:

\[ Y = \left[ \alpha_L (ZL)^{\sigma-1} + \alpha_K (uK)^{\sigma-1} \right]^{\frac{\sigma}{\sigma-1}}, \]  

where \( Z \) is labor-augmenting technology, \( u \) is capital utilization, \( K \) is the stock of capital, \( \sigma \) is the elasticity of substitution between labor and capital, and \( \alpha_L \) and \( \alpha_K \) are distribution parameters.\(^{10}\)

Computing the elasticity with respect to hours \( L \) and substituting into equation 7

---

\(^{10}\) We will discuss later how the distribution parameters are normalized.
yields the markup in the CES case:

\[
\mathcal{M}_{CES} = \frac{1}{s} \cdot \alpha_L \cdot \left( \frac{Y}{ZL} \right)^{\frac{1}{\sigma} - 1} = \frac{1}{s} \cdot \left[ 1 - \alpha_K \cdot \left( \frac{Y}{uK} \right)^{\frac{1}{\sigma} - 1} \right].
\]

This equation shows two ways of writing the CES markup. The first expression, \( \mathcal{M}_{CES}^L \), uses the elasticity with respect to hours and the second expression, \( \mathcal{M}_{CES}^K \), uses Euler’s theorem to re-express it as a function of the output-capital ratio. It is important to note that the second expression is based on the labor margin even though capital appears there. In both cases, the first term, \( \frac{1}{s} \), is the C-D markup. The impact of the CES generalization depends on the value of \( \sigma \) and the cyclicality of output per effective hour, \( \frac{Y}{ZL} \), or equivalently, the ratio of output to capital input, \( \frac{Y}{uK} \). We consider markups based on both versions since measurement of each of the ratios is not straightforward.

We can also combine overhead labor and CES production by substituting \( Z(L - \bar{L}) \) for labor input. Working through this substitution, we derive the markup for both CES production and overhead labor:

\[
\mathcal{M}_{CES, OH} = \frac{1}{s'} \cdot \alpha_L \cdot \left( \frac{Y}{Z(L - \bar{L})} \right)^{\frac{1}{\sigma} - 1} = \frac{1}{s'} \cdot \left[ 1 - \alpha_K \cdot \left( \frac{Y}{uK} \right)^{\frac{1}{\sigma} - 1} \right].
\]

Rotemberg and Woodford (1999) and Galí, Gertler and López-Salido (2007) implement these two generalizations using log-linear approximations around a steady-state and then calibrating parameters based on zero profit conditions and assumptions on steady-state markups. We will use more direct measures that do not rely on zero profit conditions or the assumption of a constant steady state markup, an assumption recently called into question by De Loecker, Eeckhout and Unger (2018) and others.

4 Empirical measures of the markup

The remainder of the paper uses the theory from the previous section to derive new measures of the aggregate price markup and assesses their cyclicality. This section describes how we constructed our measures of the markup. The next two sections report our results for the unconditional and conditional cyclicality. The unconditional analysis updates and expands the previous literature on the markup cyclicality. However, as we
emphasized in the introduction, what matters for assessing economic models is how the markup moves in response to shocks. The conditional analysis assesses how our markup measures respond to identified demand and supply shocks.

### 4.1 Baseline markup

As discussed in section 3, the markup is proportional to the inverse of the labor share when the production function is Cobb-Douglas and there is no overhead labor (equation 8). The logarithm of the markup for this case is given by:

\[
\mu_{CD}^{t} = -\ln s_{t},
\]

where \( s_{t} \) is the labor share. We use the labor share in the private business sector for our baseline measure.\(^{11}\) The markup is computed from Bureau of Labor Statistics (BLS) data as value added divided by total labor compensation. We use quarterly data from 1947:Q1 through 2017:Q4, the last period for which the output data have been through at least one annual revision.

For reference, we compare our baseline markup with a new measure from De Loecker, Eeckhout and Unger (2018), who use Compustat data to estimate firm-level markups, from which they create an aggregate markup. They use the same cost minimization problem outlined in section 3, but choose cost of goods sold (COGS) as their variable input. COGS is an accounting concept that includes all expenses that can be directly related to the production of goods, combining the costs of labor, materials, energy, and other intermediate goods. For their baseline measure, they use sector-year specific C-D production functions to estimate the output elasticity.

### 4.2 Overhead labor

As shown in equation 9, the generalization of the markup to allow for overhead labor requires actual estimates of overhead labor. Despite macroeconomists’ fondness for relying on overhead labor to explain a variety of phenomenon, few researchers have made the effort to try to measure overhead labor directly, either in the macro literature

---

11. Galí, Gertler and López-Salido (2007) use the nonfarm business version of this measure, while Rotemberg and Woodford (1999) favor the nonfinancial corporate business sector. We did not find any major difference in the cyclicity among these different measures; see appendix B for details.
or the labor literature. Most macroeconomists have used very indirect ways to estimate the ratio of overhead to variable labor. For example, Rotemberg and Woodford (1999) use zero-profit conditions and assumptions on the steady-state markup to estimate that the ratio of overhead labor to variable labor is 0.4. This high value is key to converting the procyclical baseline markup to being countercyclical.

Ramey (1991) argued that the number of nonproduction or supervisory workers is probably an upper bound on the number of overhead workers. It is an upper bound because even nonproduction and supervisory workers shows significant cyclicity of employment. For example, using Hodrick-Prescott (HP) filtered data, we find that the elasticity of the log of employment of nonproduction workers to GDP is positive and statistically significant and is about half of the elasticity of production workers with respect to GDP.

We verify the use of nonproduction workers as an upper bound on the number of overhead workers by comparing a direct measure of overhead workers to the fraction of nonproduction workers. Our direct measure is computed from the number of workers at automobile assembly plants when they are running one shift versus two shifts. Our notion is that employment should go up by less than 100 percent when a second shift is added if part of employment under one shift consists of overhead labor. According to Levitt, List and Syverson (2013), adding a second shift increased employment at the automobile plant in their study by 80 percent. This implies that overhead labor is 11 percent of total employment when two shifts are running and 20 percent of total employment when one shift is running. This ratio is consistent with narrative evidence from automobile industry periodicals during the 1970s and 1980s. Since automobile assembly plants run two or more shifts 80 percent of the time, the steady-state ratio of overhead to total employment at plants should be closer to 11 percent.

We then compare this direct measure to the fraction of nonproduction workers at automobile assembly plants. Detailed industry-level manufacturing data from the Manufacturing Industries Database, published by the National Bureau of Economic Research (NBER) and the Census Bureau’s Center for Economic Studies, shows that in the 4-digit automobile assembly industry (SIC 3711) over the period from 1958 to 2009, the ratio of nonproduction worker to total employment varied between 16 and 22 percent, with

12. p. 675.
13. These data were collected as part of the Bresnahan and Ramey (1994) project. There were a number of articles that mentioned how many workers were laid off when the second shift was eliminated and how many workers were left.
a mean of 18 percent. Thus, the evidence on employment by shifts in the automobile industry supports our contention that nonproduction workers are an upper bound on overhead labor.

We therefore construct the markup with overhead labor using the production worker share:

\[
\mu_{t}^{CD, OVH} = -\ln s',
\]

where \(s'\) is defined in equation 10. Specifically, the markup is given by current dollar output in private business divided by the wage bill for production workers. The wage bill is calculated from data on employment, average weekly hours, and average hourly wages of production workers in the private sector.\(^{14}\) This markup measure starts in 1964, when estimates of production worker employment for the entire economy begin.

### 4.3 CES production function

From equation 12, the logarithm of the CES measure of the markup can be measured by either of the following equivalent methods:

\[
\mu_{t}^{CES} = -\ln s_{t} + \ln \left[ \alpha_{L} \cdot \left( \frac{Y_{t}}{Z_{t}L_{t}} \right)^{\frac{1}{\sigma} - 1} \right]
\]

or

\[
\mu_{t}^{CES} = -\ln s_{t} + \ln \left[ 1 - \alpha_{K} \cdot \left( \frac{Y_{t}}{u_{t}K_{t}} \right)^{\frac{1}{\sigma} - 1} \right].
\]

Both variations use the labor adjustment margin, but each expresses the elasticity of output to labor in a different way.

Both expressions require a value of the elasticity of substitution (\(\sigma\)). Chirinko (2008) surveys the substantial literature that estimates the elasticity of substitution.

---

\(^{14}\) In addition to production workers covered by the BLS’s survey of establishments, we add the self-employed and workers in agriculture, who are part of the private business sector but are not covered by the establishment survey. As Krueger (1999) notes, “it is unclear how the income of the [unincorporated self-employed] should be categorized in the labor-capital dichotomy, because some of the income earned by self-employed workers clearly represents labor income, while some represents a return on investment or economic profit” (p. 45). We add 62.7 percent of their income to the wage bill, attributing the rest to capital income. See appendix A for additional details.
between capital and labor and concludes that it is in the range of 0.4 to 0.6. Karabar-bounis and Neiman (2014) use differential long-run trends in labor shares and the relative price of investment goods across countries and estimate a much higher value, around 1.25. More recently, Chirinko and Mallick (2017) use a low-pass filter on U.S. panel data and find an estimate around 0.4. Since we study capital-labor interactions at a higher frequency, i.e. the business cycle frequency, we believe an elasticity below 1 is more likely than one above 1. Thus, we use the midpoint, 0.5, of Chirinko’s (2008) survey as our elasticity of substitution. This particular value has the additional advantage that it gives the two terms in parenthesis in equation 16 and equation 17 an exponent of 1.

How the CES generalization changes the cyclicality of the markup relative to the C-D case depends on the cyclicality of these second terms. Consider first the version in equation 16. If labor-augmenting technological change, $Z_t$, is acyclical then the cyclicality of the term in brackets depends on the cyclicality of labor productivity. If capital is slow to adjust, then diminishing returns to labor implies that labor productivity should be countercyclical. Thus, the second term resulting from the CES generalization adds a countercyclical term to the C-D log markup. On the other hand, procyclical labor-augmenting technological change $Z_t$ would make this second term less countercyclical, or even procyclical. Thus, the relative cyclicality of $Z_t$ is important.

Unfortunately, there are no direct measures of $Z_t$. We consider two measures of the markup using the expression in equation 16 based on two measures of $Z_t$. The first assumes that $Z_t$ follows a trend but that there is no cyclical variation after we detrend the series. The second uses Galí’s (1999) structural vector autoregression (SVAR) method to estimate technology shocks that can be used to create a technology level series. This SVAR identifies technology shocks as those shocks that have permanent effects on labor productivity in the long-run; thus any movements in labor productivity due to cyclical variations in utilization of factors are excluded from this series. We use a simple bivariate SVAR in productivity growth and per capital hours growth, allowing for four lags.

An alternative approach is expressed in equation 17. In this case, the cyclicality of the CES adjustment depends on the cyclicality of the ratio of output to utilized capital. If this ratio is procyclical, as one would expect with slow-moving capital stocks, then it imparts some countercyclicality to the markup since it enters with a negative sign.

We measure the output-capital ratio using real private business output in the numer-
ator and the productive real capital stock for private business in the denominator. The measure of the capital stock (which excludes consumer durables) is derived from the U.S. Department of Commerce, Bureau of Economic Analysis (BEA)’s fixed asset tables, which are annual. The annual data are interpolated to quarterly frequency using the Denton method, with quarterly real private fixed investment as our indicator series.\footnote{See appendix A for additional details.}

We consider three alternatives based on different estimates of capital utilization since there is no readily available series on aggregate capital utilization ($u_t$).\footnote{The Board of Governors of the Federal Reserve System publishes a measure of capacity utilization for the industrial sector, but as Shapiro (1986) notes, this concept is distinct from capital utilization.} The first assumes that utilization is constant. In practice, capital utilization is procyclical, so assuming constant utilization will make $\frac{Y_t}{u_t K_t}$ appear to be more procyclical than it actually is, resulting in an estimated markup that is more countercyclical than it actually is.

The second alternative is based on a utilization series we construct from available estimates of the workweek of capital. Our method proceeds in several steps. First, we estimate the elasticity of the workweek of capital in manufacturing to output in manufacturing at a business-cycle frequency. Shapiro (1986) constructs a quarterly series on the workweek of capital in manufacturing from 1952 to 1982 based on data on shiftwork from the Area Wage Survey of the BLS. Gorodnichenko and Shapiro (2011) construct an annual series from 1974 to 2004 on the workweek of capital in manufacturing based on the U.S. Census Bureau’s Survey of Plant Capacity. For each of these series, we regress the HP–filtered log of the workweek of capital in manufacturing on the HP–filtered log of industrial production in manufacturing. For both series, we estimate an elasticity around 0.3.

The second step involves a decision on how to use that information. Even if Shapiro’s (1986) quarterly series extended over our entire sample, it would be incorrect to use it as our utilization measure for all of private business. This is because manufacturing output is much more procyclical than private business output. Indeed, a regression of the cyclical component of either manufacturing output or the workweek of capital on the cyclical component real private business output yields estimated elasticities above 1.7. To create a capital workweek series suitable for the entire private business sector, we assume that the elasticity of the workweek of capital in private business to the cyclical component of output in private business is also 0.3, as estimated for manufacturing.
Thus, we assume that the cyclical variation of $\frac{Y_t}{u_t}$ is the same in private business as it is in manufacturing.

The third alternative takes Fernald’s (2012) utilization series that he derives in order to estimate utilization-adjusted TFP. This measure is calculated using hours per worker as a proxy for unobserved capital utilization and effort. Note that this measure may over-correct for capital utilization, since it may also include variation in labor effort, and thus make $\frac{Y_t}{u_t K_t}$ less procyclical than it actually is. In sum, the constant utilization measure likely induces a countercyclical bias to the markup and Fernald’s (2012) utilization measure likely induces a procyclical bias.

For all measures based on equation 17, units of $\frac{Y_t}{u_t K_t}$ matter. Therefore, we normalize using one of the options recommended by Klump, McAdam and Willman (2012) and Cantore and Levine (2012). In particular, we set $\alpha_K$ in equation 17 based on the average labor share and the average of $\frac{Y_t}{u_t K_t}$ over the sample.

To summarize, we derive five potential measures of the markup based on CES production functions. Two measures are based on equation 16 and differ according to how labor-augmenting technological progress $Z$ is estimated. Three measures are based on equation 17 and differ according to how utilization $u$ is estimated. In later sections, we emphasize the measure based on the output-capital ratio and with utilization estimated from the workweek of capital, because we think it has the least cyclical bias of these measures.

### 5 Unconditional cyclicality of the markup

#### 5.1 Cobb-Douglas production function

Figure 1 plots our baseline markup, described in section 4.1. For comparison, we also show De Loecker, Eeckhout and Unger (2018) (DLEU)'s measure, which is annual and covers from 1950 to 2014. Both measures appear to peak near the middle of expansions, to decline going into a recession, and then to rise coming out a recession. That said, the cyclicality is somewhat masked by clear trends in the two series. Both markup measures have upward trends, but the timing and magnitude is somewhat different. The inverse of the labor share displays little trend until the early 2000s and then rises from around 1.6 to almost 1.8. In contrast, the DLEU markup shows a pronounced trend starting in the early 1980s, rising from a trough around 1.2 in 1980 to over 1.6 in 2014.
The downward trend in the labor share— or upward trend in the markup— has attracted considerable attention in recent years.\textsuperscript{17} To abstract from these substantial low-frequency movements for assessing the cyclicality, we detrend using the HP filter with a standard smoothing parameter.\textsuperscript{18} Figure 2 plots the detrended markup series, together with detrended real GDP. The cyclical components of the two markup measures are broadly similar, typically reaching a cyclical peak mid-way to late in an expansion and reaching a cyclical trough early in a recession.

To assess the unconditional cyclicality more systematically, we estimate the elasticity of the detrended markup with respect to detrended real GDP using the following regression:
\[ \mu_t = \beta y_t + \epsilon_t, \]
where $\mu$ is the cyclical component of the log markup and $y$ is the cyclical component of log real GDP.\textsuperscript{19} To account for serial correlation, we report Newey and West (1987)

\textsuperscript{17} See, Elsby, Hobijn and Şahin (2013), Karabarbounis and Neiman (2014), and Gutiérrez and Piton (2019), among others.
\textsuperscript{18} We also explored other detrending methods, including the Baxter-King (BK) filter, a first-difference filter, and Hamilton’s (2018) two-year-difference filter. We found that the HP and BK filters gave very similar results, whereas the first difference filter implied higher elasticities; see appendix B. We found the two-year-difference filter to be sensitive to low frequency movements.
\textsuperscript{19} Hall (2012) assesses cyclicality with respect to labor market variables rather than GDP.
standard errors. We prefer the elasticity over the correlation because it describes the magnitude of the response as well as the cyclical.

Line 1 of table 1 reports the cyclicality of our baseline markup measure calculated from 1947 to 2017. The markup is mildly procyclical, with an estimated elasticity of 0.2. That is, when real GDP is 1 percent above its trend, this markup measure is 0.2 percent above its trend, on average.

Because some parts of labor compensation might be considered more a fixed cost per worker than a payment per hour, we also consider a measure of the labor share that includes only wages and salaries. As shown on line 2, the elasticity of this markup measure is 0.1, somewhat smaller than for compensation. In addition, although the baseline elasticity is statistically significant, we cannot reject that the elasticity of the markup based on wages and salaries is zero.

We next consider alternative measures of the markup that allows for overhead labor. Line 3 of table 1 shows the results for the markup over wages and salaries (e.g., line 2) for the sample starting in 1964 and line 4 shows the markup assuming all nonproduction and supervisory workers are overhead labor. Although the estimated elasticity declines, the cyclical behavior of productivity changed dramatically in the mid-1980s and because some shocks, such as technology shocks, are often found to drive output and labor in opposite directions, we chose GDP as the best measure of cyclical.
### Table 1. Unconditional Cyclicality of the Price-Cost Markup

<table>
<thead>
<tr>
<th>Measure</th>
<th>Elasticity</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CD production function, 1947–2017</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Labor compensation</td>
<td>.20**</td>
<td>(.07)</td>
</tr>
<tr>
<td>2. Wages and salaries</td>
<td>.12</td>
<td>(.09)</td>
</tr>
<tr>
<td><strong>CD production function, overhead labor, 1964–2017</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. All worker wages and salaries</td>
<td>.14</td>
<td>(.10)</td>
</tr>
<tr>
<td>4. Prod. worker wages and salaries</td>
<td>.12</td>
<td>(.19)</td>
</tr>
<tr>
<td><strong>CES production function, 1947–2017</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. $\mu_L$, naive technology trend</td>
<td>.46***</td>
<td>(.12)</td>
</tr>
<tr>
<td>6. $\mu_L$, SVAR technology trend</td>
<td>.39***</td>
<td>(.09)</td>
</tr>
<tr>
<td>7. $\mu_K$, constant capital utilization</td>
<td>−.37***</td>
<td>(.07)</td>
</tr>
<tr>
<td>8. $\mu_K$, variable utilization (Shapiro)</td>
<td>−.22**</td>
<td>(.07)</td>
</tr>
<tr>
<td>9. $\mu_K$, variable utilization (Fernald)</td>
<td>−.02</td>
<td>(.08)</td>
</tr>
<tr>
<td><strong>CES production function, overhead labor, 1964–2017</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. $\mu_L$, naive technology trend</td>
<td>.36</td>
<td>(.23)</td>
</tr>
<tr>
<td>11. $\mu_L$, SVAR technology trend</td>
<td>.30</td>
<td>(.21)</td>
</tr>
<tr>
<td>12. $\mu_K$, constant capital utilization</td>
<td>−.52**</td>
<td>(.18)</td>
</tr>
<tr>
<td>13. $\mu_K$, variable utilization (Shapiro)</td>
<td>−.35</td>
<td>(.18)</td>
</tr>
<tr>
<td>14. $\mu_K$, variable utilization (Fernald)</td>
<td>−.18</td>
<td>(.19)</td>
</tr>
</tbody>
</table>

**Annual, 1950–2014**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Elasticity</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. Labor compensation</td>
<td>.12</td>
<td>(.09)</td>
</tr>
<tr>
<td>16. Cost of goods sold</td>
<td>.17</td>
<td>(.13)</td>
</tr>
</tbody>
</table>

Notes: Elasticity of detrended log markup with respect to detrended log real GDP; series detrended using the HP filter. Standard errors that are robust to serial correlation are reported in parentheses; ‘***’, ‘**’, ‘*’ indicates significance at the 0.1-, 1-, and 5-percent level. For CES production function, elasticity of substitution between capital and labor $\sigma = 0.5$. See section 4.3 for a description of the CES markup measures. Markup based on cost of goods sold comes from De Loecker, Eeckhout and Unger (2018).
Figure 3. Estimates of the Price-Cost Markup with CES Production Function

The next generalization we consider is a CES production function and a lower elasticity of substitution between capital and labor. Figure 3 plots the cyclical components of the 5 measures of the markup based on a CES production function discussed in section 4.3, together with the baseline C-D markup for comparison. Focusing first on the markup measures based on the output-labor ratio (equation 16), denoted by $\mu_L$, these measures have noticeably larger cyclical swings than the baseline C-D markup, the blue line. This is particularly true prior to the mid-1980s, when labor productivity switched from being procyclical to acyclical. Indeed, these measures move much more in line with the baseline after the mid-1980s. The other three measures, denoted by $\mu_K$, are based on the output-capital ratio (equation 17). These measures of the markup tend to reach a cyclical peak just after a recession and reach a trough prior to or at the start of a recession.

The third panel of table 1 reports estimates of the unconditional cyclicality of the

as expected, both estimates are small positive numbers that are not statistically different from zero. Thus, for the aggregate private business sector, our estimates of overhead labor do not support the idea of significant countercyclicality.
CES markup measures. Lines 5 and 6 show that the CES markups based on the output-labor ratio are more procyclical than the baseline, with estimated elasticities of 0.4 to 0.5. These measures are more procyclical because they add to the baseline detrended labor productivity, which is procyclical, on average, over the post-war period. Even when the labor productivity term is divided by the potentially procyclical labor-augmenting progress $Z$ estimated using long-run restrictions, the CES–based markup is even more procyclical than the C-D markup. This result is surprising.

As shown by lines 7 through 9, CES markups based on the output-capital ratio are countercyclical or acyclical. When we assume constant capital utilization (line 7), the elasticity of the markup with respect to real GDP is $-0.4$. As seen in equation 17, a procyclical $\frac{Y_t}{u_tK_t}$ will make the markup less procyclical (or more countercyclical) than the C-D markup. Because capital stocks are slow to adjust, $\frac{Y_t}{K_t}$ has an elasticity near one with output. Line 8 shows that the markup based on the cyclicality of the workweek of capital is countercyclical (elasticity of $-0.2$), but less so than under the assumption of constant capital utilization. As shown on line 9, the markup based on Fernald’s (2012) estimate of factor utilization is essentially acyclical, with an estimated elasticity of zero. The relative cyclicality of these three measures lines up with what we would expect, given the differing cyclicality of the utilization measures.

Finally, lines 10 through 14 of table 1 show the results when we combine the two generalizations, allowing for both overhead labor and CES production functions. Not surprisingly, we find the markup to be somewhat more countercyclical than when we do not assume overhead labor. That being said, the results are imprecisely estimated and are not statistically different from zero apart from the case with constant capital utilization.20

To summarize our unconditional results, we find that the markup estimate based on C-D production functions are slightly procyclical or acyclical, even allowing for overhead labor. In contrast, the markup estimates based on a CES production function have estimated elasticities ranging from 0.5 to $-0.5$. Our preferred measure, based on output-capital ratio and with capital utilization estimated from the workweek of capital, is modestly countercyclical, with an elasticity of $-0.2$ or $-0.4$, depending on whether when we also allow for overhead labor.

---

20. One source of the increased imprecision might be the shorter sample, which is due to data availability.
5.3 Discussion

The results for our baseline measure should not be a surprise to anyone who has studied the cyclicality of labor share. In fact, table 1 of Galí, Gertler and López-Salido (2007) report a correlation of the price-cost markup with GDP of 0.28 for their sample and data.

Our finding of a procyclical markup after accounting for overhead labor (line 4) is at odds with those in Rotemberg and Woodford (1999) and Galí, Gertler and López-Salido (2007). These papers use steady-state approximations and calibrations that result in adjusting the standard markup by subtracting 0.4 times the cyclical variation in labor from the measure. Since labor input is strongly procyclical, it is easy to see how their adjustment would make the markup look much more countercyclical.

The role of capital utilization also sheds light on one of the reasons that Rotemberg and Woodford (1999) and Galí, Gertler and López-Salido (2007) found very countercyclical markups when they used a CES generalization. As Appendix B of Galí, Gertler and López-Salido (2007) outlines, both sets of authors operationalize the CES production function assumption differently. They take approximations around the steady state and specify the log markup as

\[ \mu \approx -\ln s + \theta (\ln Y - \ln K), \]

They calibrate the value of \( \theta \) based on a nonlinear combination of several steady-state elasticities. In addition to an elasticity of substitution between capital and labor of 0.5, their value of \( \theta \) also depends on their calibrations of steady-state labor share of 0.7 and a steady-state average gross markup near unity. These values imply a value of \( \theta \) equal to \(-0.4\). Since Rotemberg and Woodford’s (1999) and Galí, Gertler and López-Salido’s (2007) do not allow for variable capital utilization, the term in parenthesis is very procyclical. Since that terms is multiplied by \(-0.4\), they find a very countercyclical markup.

Cyclical capital utilization is generally believed to be important empirically and is a key part of the leading medium scale NK models. For example, Christiano, Eichenbaum and Evans (2005) find that variable capital utilization is crucial for matching their data. In their model, the elasticity of capital utilization to a monetary policy shock is about 80 percent of the elasticity of output. Their empirical work, however, implies a higher elasticity of capital utilization. In particular, they find that two of their three empirical
indicators of utilization imply that the elasticity of capital utilization with respect to a monetary policy shock is greater than the elasticity of output.

6 Conditional cyclicality of the markup

The unconditional cyclicality estimates presented in the last section are useful for describing the patterns in the data, but they are not useful for assessing how well the behavior of the markup fits the predictions of NK models. In both NK models with only sticky prices and in medium-scale models with both sticky prices and sticky wages, the cyclicality of the price markup depends crucially on the source of the shock. For example, demand shocks, such as monetary policy shocks and government spending shocks, should lead to countercyclical movements in the markup since an expansionary shock raises output and marginal cost, but firms cannot immediately adjust their prices. In the Smets and Wouters (2007) model, investment-specific technology (IST) shocks also lead to a countercyclical markup because these shocks do not raise productivity in the short run. Conversely, as pointed out by Galí (1999), a labor-augmenting or neutral technology shock should lead to procyclical movements in the markup since a positive technology shock raises output and reduces marginal cost, but prices do not adjust.

Estimated medium scale NK models identify parameters and shocks using data along with assumptions about the structure of the model and the time series process driving the unobserved shocks.21 Virtually all of those models assume C-D production functions. Here we present independent evidence on the cyclicality of the markup based on our production function generalizations and on shocks identified using time series methods.

6.1 Identification of shocks

We study the response of our markup measures to four types of shocks: monetary policy, government spending, TFP, and investment-specific technology (IST). We use standard SVARs to identify the shocks and estimate the responses. All four SVARs are estimated on quarterly data, include four lags, as well as a quadratic time trend. We plot bootstrapped standard errors.

The monetary SVAR includes log real GDP per capita, the log of the GDP price deflator, the log of commodity prices, the federal funds rate, and a measure of the log

As in Christiano, Eichenbaum and Evans (1999), the monetary policy shock is identified as a shock to the federal funds rate using a Choleski decomposition. We order the federal funds rate second to last, with the markup being the last variable. We do not allow contemporaneous effects of the markup on the federal funds rate so that changes in the markup variable across specifications have little effect on the estimated federal funds shock.

The government spending SVAR includes the updated version of Ramey’s (2011) military news variable, divided by nominal GDP, along with log real GDP per capita, the log of the GDP price deflator, the three-month Treasury bill rate, and the log of the markup. Government spending news shocks are identified as the shocks to the military news variable, ordered first in the Choleski decomposition.

The TFP SVAR includes the log level of Fernald’s (2012) utilization-adjusted measure of TFP, log real GDP per capita, log of the GDP price deflator, the three-month Treasury bill rate, and the log of the markup. TFP shocks are identified as the shocks to Fernald’s TFP variable, ordered first in the Choleski decomposition.

Finally, to identify the IST shock we use Fisher’s (2006) identifying assumption that only IST shocks can have a long-run effect on the relative price of investment goods. We first estimate the shock in a system with long-run restrictions. That system includes the log difference of the deflator for equipment investment relative to the deflator for non-durable plus services consumption, log difference in real GDP per capita, log difference of the GDP price deflator (i.e. inflation), and the level of the three-month Treasury bill rate. We then incorporate that shock into an SVAR, ordered first, along with log real GDP per capita, log of the GDP price deflator, the three-month Treasury bill rate, and the log of the markup.

Because the federal funds rate only became available in 1954, the monetary SVAR is estimated from 1954:Q3 through 2017:Q4. The other three SVARs are estimated from 1947:Q1 through 2017:Q4.

### 6.2 Estimates of conditional cyclicality

Figure 4 shows the estimated impulse responses for log real GDP and the log of two measures of the markup in response to each of the four identified shocks. For ease of

---

22. We use Krippner’s (2013) estimate of the shadow federal funds rate in place of the actual funds rate from 2009:Q1 to 2016:Q3. We also estimated a version that ends estimation in 2008 and found very similar results.
comparison, we consider expansionary shocks in all four cases. The baseline measure, in which production function is C-D, is the inverse of the labor share. The second measure is the markup assuming a CES production function, measured by the output-capital ratio, with variable capital utilization based on the workweek of capital.\textsuperscript{23}

Because we are interested in how the estimated conditional responses compare to NK models, we also plot simulations from the Smets and Wouters (2007) (SW) model, estimated using their data and sample.\textsuperscript{24} We normalize the simulations so that the peak effect on output is the same as in our estimated SVARs.

Consider the effects of a monetary policy shock, shown in figure 4a. Output rises in both our SVAR estimates and in the simulation from the SW model, though the response of output occurs more quickly in the model. Both of our markup measures rise, meaning they are procyclical, whereas the SW simulations show a countercyclical response of their markup.\textsuperscript{25}

Figure 4b shows the responses to a positive government spending shock. Our SVAR estimates imply that output and the markup rise robustly in response. In contrast, the SW’s simulations imply an increase in output but a small decline in the markup. It is important to note, though, that SW’s government spending shock is actually a mix of shocks to government spending plus net exports, which mute the countercyclicality of the markup to this shock.

Thus, our SVARs estimates imply procyclical price markup movements in response to the two demand shocks we study. This result is at odds with SW’s model estimates, as well as those of all other NK models with which we are familiar. Also interesting is that even our markup measures based on a CES production function, which ranged from procyclical to countercyclical in the unconditional analysis, are procyclical conditional on the demand shocks.

Figure 4c shows the responses to a positive TFP shock. In this case, the SVAR responses line up very well with the SW responses. All of the estimates show an increase

\textsuperscript{23} We do not use the version with overhead labor, since the necessary data on production workers is available only starting in 1964. Cutting the early sample tends to increase the standard error bands without changing the qualitative results.

\textsuperscript{24} Recall that the SW model also has sticky wages, so they must rely less on the movement of the price markup than a NK model with just sticky prices. In those models, price markup movements are much more pronounced.

\textsuperscript{25} Smets and Wouters (2007) graph the log deviation of real marginal cost rather than the log deviation of the price markup. However, log real marginal cost is just the negative of the log of the price markup.
Figure 4. Conditional Cyclicality of the Price Markup

(a) Monetary Policy Shock

(b) Government Spending Shock

Continued on next page.
Figure 4. Conditional Cyclicality of the Price Markup (continued)

(c) Technology Shock

(d) Investment-Specific Technology Shock

Notes: Impulse response of log real GDP and log markup to a shock to variable indicated in heading; shaded areas indicate 90-percent confidence interval around estimate. CES markup measure based on output-capital ratio and workweek of capital. Estimation of monetary SVAR begins in 1954:Q3; all others start in 1947:Q1.
in output and markups, all with quite similar dynamics.

At this point, the reader may wonder how we could have found countercyclicality of the CES–based markup in the unconditional analysis when we are finding procyclical CES–based markups in response to demand and TFP shocks. The answer to this apparent puzzle is provided in figure 4d, which shows the responses to a positive IST shock. Output rises in both of our SVAR specifications, as well as in SW’s simulations. Our output responses are more persistent because we identify our shocks as those having permanent effects on the relative price of investment goods whereas SW assume stationary processes. In contrast to the responses to the three previous shocks, the markup response from our SVARs is significantly countercyclical. The estimated response from SW is also countercyclical but is muted compared to ours.

Table 2 summarizes these results as well as those for our other measures of the markup by calculating the implied elasticity with respect to real GDP. In order to summarize the entire dynamic pattern succinctly, we extend the method introduced in the government spending multiplier literature that calculates multipliers as ratios of integrals under impulse response functions (IRFs). In our case, we are interested in elasticities, which we calculate as the ratio of the cumulative IRF of the log markup (that is, the integral under the impulse response curve) over a 20-quarter horizon to the cumulative IRF of log output over the same horizon.

The main take-away from table 2 is that the estimated elasticities are positive for the monetary policy shock, the government spending shock, and the TFP shock and negative for the IST shock. Moreover, as can be seen by looking down the columns, our estimates have the same sign across all measures of the markups we consider. This stands in contrast to the unconditional elasticities, where some measures were procyclical while others were countercyclical.

Finally, on line 15 we report the comparable elasticities from the SW model. Their model estimates imply markups decrease in response to monetary policy and government spending shocks, which is not consistent with our findings. The responses in their model to the TFP shock and the IST shock are qualitatively consistent with our estimates.

---

26. See, for example, Ramey (2016), pp. 116 and 119.
### Table 2. Conditional Cyclicality of the Price-Cost Markup

<table>
<thead>
<tr>
<th>Measure</th>
<th>Monetary policy</th>
<th>Govt. spending</th>
<th>TFP</th>
<th>IST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CD production function, 1947–2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Labor compensation</td>
<td>.72</td>
<td>.67</td>
<td>.49</td>
<td>−.42</td>
</tr>
<tr>
<td>2. Wages and salaries</td>
<td>1.14</td>
<td>.82</td>
<td>1.30</td>
<td>−.70</td>
</tr>
<tr>
<td><strong>CD production function, overhead labor, 1964–2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. All worker wages and salaries</td>
<td>1.44</td>
<td>.22</td>
<td>.64</td>
<td>−1.02</td>
</tr>
<tr>
<td>4. Prod. worker wages and salaries</td>
<td>3.03</td>
<td>.98</td>
<td>1.44</td>
<td>−.84</td>
</tr>
<tr>
<td><strong>CES production function, 1947–2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. $\mu_L$, naive technology trend</td>
<td>1.08</td>
<td>1.24</td>
<td>2.74</td>
<td>−.29</td>
</tr>
<tr>
<td>6. $\mu_L$, SVAR technology trend</td>
<td>.78</td>
<td>.63</td>
<td>.47</td>
<td>−.39</td>
</tr>
<tr>
<td>7. $\mu_K$, constant capital utilization</td>
<td>.18</td>
<td>.89</td>
<td>.41</td>
<td>−.79</td>
</tr>
<tr>
<td>8. $\mu_K$, variable utilization (Shapiro)</td>
<td>.34</td>
<td>1.13</td>
<td>.63</td>
<td>−.80</td>
</tr>
<tr>
<td>9. $\mu_K$, variable utilization (Fernald)</td>
<td>.52</td>
<td>1.02</td>
<td>.22</td>
<td>−.79</td>
</tr>
<tr>
<td><strong>CES production function, overhead labor, 1964–2017</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. $\mu_L$, naive technology trend</td>
<td>3.20</td>
<td>3.94</td>
<td>3.60</td>
<td>−.90</td>
</tr>
<tr>
<td>11. $\mu_L$, SVAR technology trend</td>
<td>3.13</td>
<td>1.00</td>
<td>1.50</td>
<td>−.89</td>
</tr>
<tr>
<td>12. $\mu_K$, constant capital utilization</td>
<td>2.41</td>
<td>.73</td>
<td>1.40</td>
<td>−1.23</td>
</tr>
<tr>
<td>13. $\mu_K$, variable utilization (Shapiro)</td>
<td>2.51</td>
<td>.80</td>
<td>1.52</td>
<td>−1.27</td>
</tr>
<tr>
<td>14. $\mu_K$, variable utilization (Fernald)</td>
<td>2.69</td>
<td>.50</td>
<td>1.28</td>
<td>−1.23</td>
</tr>
<tr>
<td><strong>Memo:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Smets and Wouters (2007)</td>
<td>−.51</td>
<td>−.10</td>
<td>.18</td>
<td>−.25</td>
</tr>
</tbody>
</table>

Notes: Implied elasticity of markup with respect to real GDP based on ratio of cumulative IRFs over 20-quarter horizon. For CES production function, elasticity of substitution between capital and labor $\sigma = 0.5$. See section 4.3 for a description of the CES markup measures. Smets and Wouters (2007) results are from our calculations.
6.3 Discussion

These conditional results shed light on the cyclical behavior of the markup and how well NK models can capture that behavior. While our four estimated shocks do not exhaust the list of possible shocks, they nonetheless provide some insight into the unconditional cyclicality we estimated in the previous section. Recall that some of our CES–based markup measures suggested that the markup is mildly countercyclical. The unconditional elasticities depend on both the individual elasticities to each shock and on the variance of each shock in the sample. Interestingly, a large literature, surveyed in Ramey (2016), finds that IST shocks are some of the most important shocks driving output and hours at business cycle frequencies. Thus, even if the markup is procyclical in response to monetary policy, government spending, and TFP, the markup can, in principle, be countercyclical overall if IST shocks are the dominant shocks driving business cycles.

It is also interesting to note that long-run trends in investment-specific technological change also play a central role in Karabarbounis and Neiman’s (2014) explanation for the global decline in the labor share. In particular, they argue that labor shares declined globally since 1980 because of the acceleration of the pace of investment-specific technological change coupled with an elasticity of substitution between capital and labor above 1. We also find a central role for IST shocks as the only one of our measured shocks that produces a countercyclical markup. However, the shock appears to be so important that it leads the unconditional estimate of the markup to be countercyclical for some of the CES–based measures. Recall that our results are based both on the C-D specification and the CES specification assuming an elasticity of substitution between capital and labor of 0.5. Further results (not shown) indicate that the unconditional cyclicality of the CES–based markup becomes procyclical if we measure the markup using Karabarbounis and Neiman’s (2014) value of the elasticity of 1.25. The response of the CES–based markup using their assumed elasticity continues to be countercyclical conditional on IST shocks, but noticeably less so than when we use our assumption of an elasticity of 0.5.
7 Conclusion

This paper has presented new evidence on the cyclicality of aggregate price markup, and in particular on the cyclicality conditional on leading macroeconomic shocks. We began by arguing that the labor input margin continues to be the best way to measure the markup, citing new evidence that measured wages are a good indication of the marginal cost of an extra hour of labor. Even focusing on that measure, though, we derived a range of measures of the markup by varying assumptions about elasticities of substitution between capital and labor, whether there is overhead labor, and how key inputs are measured.

Our analysis of the elasticity of the markup with respect to output, both filtered to focus on variation at business-cycle frequencies, yields a range of estimates from procyclical to countercyclical, depending on the measure. The baseline C-D measure is procyclical, and remains so after we account for overhead labor. Some measures of the markup based on a CES production function are procyclical whereas others are countercyclical.

Turning to the conditional analysis, we identify four macroeconomic shocks using standard time-series methods from the literature: monetary policy shocks, government spending shocks, TFP shocks, and IST shocks. The markup increases in response to expansionary monetary policy, government spending, and TFP shocks. In contrast, the markup decreases in response to the IST shock. These findings for the conditional cyclicality hold for all measures of the markup that we considered.

We compare our results to those from the Smets and Wouters (2007) model. We find that the responses of our various measures of the markup are qualitatively consistent with those from the SW model for the two technology shocks we analyze. In contrast, we find that the responses of the markup to monetary policy and government spending shocks are inconsistent with the simulations from the SW model. In particular, we find that the markup increases in response to expansionary demand shocks whereas the SW model predicts a decrease. Because this key sticky-price transmission mechanism for monetary policy and government spending shocks is at odds with the data, our results suggest that NK models might benefit from a renewed focus on wage rigidities rather than price rigidities.
References


## Appendix

### A Data sources

<table>
<thead>
<tr>
<th>Data series</th>
<th>N-R mnemonic</th>
<th>Source and mnemonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real gross domestic product</td>
<td>rgdp</td>
<td>FRED</td>
</tr>
<tr>
<td>Gross domestic product</td>
<td>ngdp</td>
<td>FRED</td>
</tr>
<tr>
<td>Implicit price deflator, gross domestic product</td>
<td>pgdp</td>
<td>FRED</td>
</tr>
<tr>
<td>Output, private business</td>
<td>output_bus</td>
<td>BLS</td>
</tr>
<tr>
<td>Implicit price deflator, private business</td>
<td>pd_bus</td>
<td>BLS</td>
</tr>
<tr>
<td>Hours worked, private business</td>
<td>hours_bus</td>
<td>BLS</td>
</tr>
<tr>
<td>Employment, private business</td>
<td>emp_bus</td>
<td>BLS</td>
</tr>
<tr>
<td>Compensation of employees, private business</td>
<td>comp_bus</td>
<td>BLS</td>
</tr>
<tr>
<td>Wages and salaries, private business</td>
<td>ws_bus</td>
<td>FRED</td>
</tr>
<tr>
<td>Employment, production workers in private business</td>
<td>emp_prodwrk_priv</td>
<td>FRED</td>
</tr>
<tr>
<td>Average weekly hours, production workers in private business</td>
<td>ah_prodwrk_priv</td>
<td>FRED</td>
</tr>
<tr>
<td>Average hourly earnings, production workers in private business</td>
<td>nw_prodwrk_priv</td>
<td>FRED</td>
</tr>
<tr>
<td>Employment, unincorporated self-employed</td>
<td>emp_selfemp</td>
<td>FRED</td>
</tr>
<tr>
<td>Employment, agriculture and related industries</td>
<td>emp_agr</td>
<td>FRED</td>
</tr>
<tr>
<td>Population</td>
<td>pop</td>
<td>FRED</td>
</tr>
<tr>
<td>Civilian noninstitutional population</td>
<td>civpop16rev</td>
<td>BLS</td>
</tr>
<tr>
<td>Effective federal funds rate</td>
<td>rff</td>
<td>FRED</td>
</tr>
<tr>
<td>Shadow federal funds rate</td>
<td>ssr</td>
<td>Krippner (2013)</td>
</tr>
<tr>
<td>Three month treasury bill, secondary market rate</td>
<td>tb3</td>
<td>FRED</td>
</tr>
</tbody>
</table>

Continued.
<table>
<thead>
<tr>
<th>Data series</th>
<th>N-R mnemonic</th>
<th>Source and mnemonic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>dtpf_util</td>
</tr>
<tr>
<td>Utilization of capital and labor</td>
<td>n.a.</td>
<td>Fernald (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dutil</td>
</tr>
<tr>
<td>Spot commodity price index</td>
<td>n.a.</td>
<td>CRB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PZALL</td>
</tr>
<tr>
<td>Personal consumption expenditures, nondurable goods</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCND</td>
</tr>
<tr>
<td>Personal consumption expenditures, services</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCESV</td>
</tr>
<tr>
<td>Real PCE, nondurable goods (chain-type quantity index)</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNDGRA3Q086SBEA</td>
</tr>
<tr>
<td>Real PCE, services (chain-type quantity index)</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DSERRA3Q086SBEA</td>
</tr>
<tr>
<td>PCE, nondurable goods (chain-type price index)</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DNDGRG3Q086SBEA</td>
</tr>
<tr>
<td>PCE, services (chain-type price index)</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DSERRG3Q086SBEA</td>
</tr>
<tr>
<td>Implicit price deflator, gross private domestic investment, equipment</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y033RD3Q086SBEA</td>
</tr>
<tr>
<td>Index of industrial production, manufacturing</td>
<td>n.a.</td>
<td>FRED</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IPMANSICS</td>
</tr>
<tr>
<td>Workweek of capital, manufacturing</td>
<td>n.a.</td>
<td>Shapiro (1986)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Table III</td>
</tr>
<tr>
<td>Plant hours per week, manufacturing</td>
<td>n.a.</td>
<td>GS (2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>phw_adj_K4</td>
</tr>
</tbody>
</table>

Notes: Unpublished BLS data can be downloaded from https://www.bls.gov/lpc/special_requests/nonfarm_business.zip. Civilian noninstitutional population (FRED series CE16OV) is adjusted by authors to smooth revisions to population controls. CRB stands for Commodity Research Bureau. GS stands for Gorodnichenko and Shapiro (2011); see http://www.umich.edu/~shapiro/data/SPC/.

We construct the baseline measure of the price-cost markup as the inverse of the labor share, which is value added divided by labor compensation in the private business sector. Both series are measured in current dollars. The adjustments to the baseline measure are described in section 4.3.
Our measure of private-sector production worker wages and salaries is calculated as follows. We start with production workers from the BLS establishment survey

\[ ws_{\text{ces}} = \frac{\text{nw}_{\text{prodwrk}} \cdot (\text{emp}_{\text{prodwrk}} / 1000) \cdot (\text{ah}_{\text{prodwrk}} \cdot 52)}{1000} \]

For the self-employed and agriculture, there are published data on employment, but not for workweeks or wages. We assume these two groups work the same hours and earn the same nominal wages as production workers in the CES private sector.

\[ ws_{\text{se}} = \frac{\text{nw}_{\text{prodwrk}} \cdot (\text{emp}_{\text{selfemp}} / 1000) \cdot (\text{ah}_{\text{prodwrk}} \cdot 52)}{1000} \]

\[ ws_{\text{agr}} = \frac{\text{nw}_{\text{prodwrk}} \cdot (\text{emp}_{\text{agr}} / 1000) \cdot (\text{ah}_{\text{prodwrk}} \cdot 52)}{1000} \]

Finally, production worker wages and salaries are the sum of

\[ ws_{\text{bus prod}} = ws_{\text{ces}} + .627 \times ws_{\text{se}} + ws_{\text{agr}}. \]

That is, we assume a bit less than \( \frac{2}{3} \) of self-employed income accrues to labor, with the remainder reflecting both returns to their work effort and returns to business assets. This value is the average from lines 6–8 in table 1 from Elsby, Hobijn and Şahin (2013).

The measure of the real productive capital stock for private business is computed as follows. We begin with annual data on the real stock of private fixed capital, from line 3 of Fixed Asset Table 1.2 from the BEA. The annual data are interpolated to quarterly frequency using the Denton method, with quarterly real private fixed investment as our indicator series.\(^{27}\) The index level of the capital stock was normalized to the value of real productive capital stock in 2012 taken from the BLS’s MFP program.\(^{28}\)

The investment-specific technology shock identification is based on the relative price of equipment investment. This relative price is measured as the ratio of the implicit price deflator for gross private domestic investment in equipment divided by the implicit price deflator for personal consumption expenditures on nondurable goods plus services. The latter is constructed from the series on each component separately using Whelan’s (2002) method for aggregating chain weighted series.

\(^{27}\) We implement the interpolation using the Denton command in Stata; see Baum and Hirstakeva (2014).

\(^{28}\) [https://www.bls.gov/mfp/special_requests/capital.xlsx](https://www.bls.gov/mfp/special_requests/capital.xlsx); sheet PG, cell B290.
B  Robustness and alternative specifications

B.1  Alternate measures and detrending methods

Table B1 reports the unconditional cyclicality of the various markup measures for two alternative detrending methods. The first two columns show results for the Baxter-King (BK) filter, while the second uses a first-difference filter. The results for the BK filter are similar to those from the HP filter that we reported in the main text. Detrending using the first-difference filter yields much more procyclical markups.

Table B2 reports the elasticities of the markup when measured in different sectors of the U.S. economy (in rows) and using the three detrending methods we consider (in columns). Across all three methods, the markup in the entire private business sector is more the most procyclical while that in the nonfinancial corporate business sector is the least procyclical. The elasticities range from 0.1 to 0.3, depending on the measure and method.
Table B1. Unconditional Cyclicality of the Price-Cost Markup - Robustness

<table>
<thead>
<tr>
<th>Measure</th>
<th>Baxter-King</th>
<th>First difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elasticity</td>
<td>Std. err.</td>
</tr>
<tr>
<td>CD production function, 1947–2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Labor compensation</td>
<td>.18*</td>
<td>(.08)</td>
</tr>
<tr>
<td>2. Wages and salaries</td>
<td>.10</td>
<td>(.10)</td>
</tr>
<tr>
<td>CD production function, overhead labor, 1964–2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. All worker wages and salaries</td>
<td>.06</td>
<td>(.12)</td>
</tr>
<tr>
<td>4. Prod. worker wages and salaries</td>
<td>−.06</td>
<td>(.20)</td>
</tr>
<tr>
<td>CES production function, 1947–2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. $\mu_L$, naive technology trend</td>
<td>.40***</td>
<td>(.12)</td>
</tr>
<tr>
<td>6. $\mu_L$, SVAR technology trend</td>
<td>.38***</td>
<td>(.10)</td>
</tr>
<tr>
<td>7. $\mu_K$, constant capital utilization</td>
<td>−.39***</td>
<td>(.08)</td>
</tr>
<tr>
<td>8. $\mu_K$, variable utilization (Shapiro)</td>
<td>−.24**</td>
<td>(.08)</td>
</tr>
<tr>
<td>9. $\mu_K$, variable utilization (Fernald)</td>
<td>−.03</td>
<td>(.09)</td>
</tr>
<tr>
<td>CES production function, overhead labor, 1964–2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. $\mu_L$, naive technology trend</td>
<td>.15</td>
<td>(.25)</td>
</tr>
<tr>
<td>11. $\mu_L$, SVAR technology trend</td>
<td>.12</td>
<td>(.23)</td>
</tr>
<tr>
<td>12. $\mu_K$, constant capital utilization</td>
<td>−.69***</td>
<td>(.19)</td>
</tr>
<tr>
<td>13. $\mu_K$, variable utilization (Shapiro)</td>
<td>−.52**</td>
<td>(.19)</td>
</tr>
<tr>
<td>14. $\mu_K$, variable utilization (Fernald)</td>
<td>−.35</td>
<td>(.20)</td>
</tr>
<tr>
<td>Annual, 1950–2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Labor compensation</td>
<td>.12</td>
<td>(.09)</td>
</tr>
<tr>
<td>16. Cost of goods sold</td>
<td>.22</td>
<td>(.12)</td>
</tr>
</tbody>
</table>

Notes: Elasticity of detrended log markup with respect to detrended log real GDP; detrending method listed in column heading. Standard errors that are robust to serial correlation are reported in parentheses; ‘***’, ‘**’, ‘*’ indicates significance at the 0.1-, 1-, and 5-percent level. For CES production function, elasticity of substitution between capital and labor $\sigma = 0.5$. See section 4.3 for a description of the CES markup measures. Markup based on cost of goods sold comes from De Loecker, Eeckhout and Unger (2018).
Table B2. Unconditional Cyclicality of the Price-Cost Markup - Robustness (II)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Hodrick-Prescott</th>
<th>Baxter-King</th>
<th>First-difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Private business</td>
<td>.20**</td>
<td>(.07)</td>
<td>.18*</td>
</tr>
<tr>
<td>2. Private nonfarm business</td>
<td>.13*</td>
<td>(.07)</td>
<td>.13</td>
</tr>
<tr>
<td>3. Nonfin. corp. business</td>
<td>.11</td>
<td>(.08)</td>
<td>.13</td>
</tr>
</tbody>
</table>

Notes: Elasticity of detrended log markup with respect to detrended log real GDP; detrending method listed in column heading. Markup measured using C-D production; data are from the BLS. Standard errors that are robust to serial correlation are reported in parentheses; ‘***’, ‘**’, ‘*’ indicates significance at the 0.1-, 1-, and 5-percent level.

B.2 Capital utilization

Our estimated series for the workweek of capital in private business was based on estimated elasticities of the workweek to output. As a robustness check, we also considered the alternative based on elasticities of the workweek of capital to labor hours, specifically to hours per worker and to the number of workers in case they had different relationships with the workweek of capital. To explore this alternative, we used Shapiro’s (1986) quarterly manufacturing workweek from 1952 to 1952. We first compared the elasticity of the cyclical component in the workweek of capital to the cyclical component of total hours versus output in manufacturing. In both cases, the elasticity was estimated to be 0.32.

We then regressed the workweek capital in manufacturing on employment and average hours per worker in manufacturing (all involving cyclical components and logarithms). The estimated elasticity with respect to employment was 0.164 (SE = .07) and to average hours was 0.925 (SE = .21). They are significantly different from each other. We created an alternative utilization series for private business using the estimated coefficients for the two hours margins (i.e. employment and average hours) in private business. When we studied the unconditional elasticity of the CES-based markup with this new measure of utilization, the result was virtually the same as for our baseline measure.

---

29. All cyclical components were extracted using a standard HP filter.
C Structural break in markup cyclicality

The unconditional cyclicality of the markup changed dramatically in the mid-1990s, switching from procyclical to countercyclical. This section describes this structural break and explores which components can explain the change in cyclicality. We find that the switch in the markup cyclicality is nearly all due to a decline in the cyclicality of labor productivity.

Figure C1 plots a rolling 40-quarter elasticity of the detrended markup with respect to detrended real GDP. The blue shaded area is the 90-percent confidence interval around the point estimate, based on Newey and West (1987) standard errors to account for autocorrelation in the residuals. The orange line is the full-sample elasticity. The markup is somewhat more procyclical, on average, than the full-period estimate through the late 1990s. However, the markup turns countercyclical in samples ending in the late-1990s and beyond. Indeed, by the mid-2000s the markup is significantly countercyclical.

To identify the timing of the break more accurately, we run a standard test for a structural break with an unknown break date.\(^{30}\) The break test finds overwhelming evidence of a structural break, with the maximum value for the test statistic (19.7)

\(^{30}\) Andrews (1993).
Table C1. Cyclicality of Markup in Private Business and Its Components

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Markup</td>
<td>.20**</td>
<td>.26***</td>
<td>−.20</td>
<td>−.46</td>
</tr>
<tr>
<td>2. Real output</td>
<td>1.22***</td>
<td>1.20***</td>
<td>1.34***</td>
<td>.14</td>
</tr>
<tr>
<td>3. Hours</td>
<td>.96***</td>
<td>.89***</td>
<td>1.39***</td>
<td>.50</td>
</tr>
<tr>
<td>4. Output per hour</td>
<td>.26***</td>
<td>.32***</td>
<td>−.05</td>
<td>−.37</td>
</tr>
<tr>
<td>5. Real compensation</td>
<td>1.03***</td>
<td>.94***</td>
<td>1.54***</td>
<td>.60</td>
</tr>
<tr>
<td>6. Hours</td>
<td>.96***</td>
<td>.89***</td>
<td>1.39***</td>
<td>.50</td>
</tr>
<tr>
<td>7. Comp. per hour</td>
<td>.07</td>
<td>.05</td>
<td>.15</td>
<td>.10</td>
</tr>
</tbody>
</table>

Notes: Elasticity of detrended log markup with respect to detrended log real GDP; series detrended using the HP filter. Standard errors that are robust to serial correlation are reported in parentheses; ‘***’, ‘**’, ‘*’ indicates significance at the 0.1-, 1-, and 5-percent level.

occurring in 1995:Q1. We take this date as our structural break and explore the unconditional cyclicality of the markup in two subsamples.

Line 1 of table C1 reports the elasticity of the markup with respect to real GDP. The first column reports the estimate for the full sample (1847–2017), while the second and third columns report estimates from the two subsamples. The final column reports the change in the elasticity from the early period to the later period. As expected given figure C1, the elasticity switches from procyclical (0.3) in the sample from 1947–94 to countercyclical (−0.2) in the sample from 1995–2017.

Recall that for the baseline markup measure, the markup is equal to output divided by labor compensation. The ½ percentage point decrease in markup elasticity could come from a change in the cyclicality of the numerator (real business output), the denominator (real compensation), or both.31 To explore this, figure C2 plots the cyclical components of the markup together with its components. Because they are expressed in logs, the markup (the blue line) is the difference between the orange line and the gold line. Both output and labor compensation are more cyclical than the markup. In addition, it is apparent from the figure that compensation is a bit less cyclical than output prior to 1995 (the dashed vertical line) and noticeably more cyclical than output

31. The markup is generally calculated using current dollars. However, to compare these components with real GDP, we convert both business sector output and labor compensation to constant dollars using the price deflator for business output. This leaves the markup unaffected, but removes the effects of inflation.
Figure C2. Detrended Markup and Its Components

![Graph showing detrended markup and its components with percent deviation from trend on the y-axis and years 1950 to 2015 on the x-axis.]

Notes: Dashed series is detrended real GDP after 1995.

Lines 2 and 5 in table C1 report the elasticity of the numerator and denominator of the markup with respect to real GDP. Although real business output became slightly more procyclical after 1995, real compensation became significantly more procyclical and this can explain the markup’s move from procyclical to countercyclical between the two periods.

Why did compensation become more procyclical? To answer this, note the following accounting identities:

\[
\text{output} = \text{productivity} \times \text{hours} \\
\text{compensation} = \text{hours} \times \text{compensation per hour}
\]

and so

\[
\text{markup} = \frac{\text{productivity} \times \text{hours}}{\text{hours} \times \text{compensation per hour}}
\]

Since hours appears in the numerator and the denominator, those terms offset. Thus, the changing cyclicality of the markup reflects changes in the cyclicality of productivity and compensation per hour.

Figure C3 plots the cyclical components of productivity and compensation per hour. As shown by line 6 of table C1, real compensation became more procyclical because
hours became more procyclical. However, since hours also appears in the numerator, this component can’t explain the change in cyclicality of the markup.

As shown by line 4 of table C1, productivity moved from moderately procyclical to weakly countercyclical, while real compensation per hour, line 7, was little changed. Thus, the decline in cyclicality in the markup from the early period to the later period is mostly due to a decline in the cyclicality of labor productivity. Labor productivity was more procyclical in the earlier period — for example, firms may have kept more workers on payrolls during downturns, making compensation less cyclical than output. In the later period, when productivity is roughly acyclical, firms adjusted labor input more closely with demand and so compensation is more procyclical.

D The marginal wage versus the average wage

This section revisits Bils’s (1987) argument that the marginal hourly wage is more procyclical than the average hourly wage because of the additional cost of overtime hours. The first section begins by generalizing the theory we presented in the paper to distinguish hours per worker from the number of workers. It then develops a relationship between marginal and average wages based on parameters and variables that can be measured. The second section uses that relationship to measure the ratio of marginal to average wages in the aggregate data and to assess its cyclicality.
D.1 Theory

We generalize the labor input by decomposing total labor hours, $L$, into hours per worker, $h$, and the number of workers, $N$ — that is, $L = hN$. The firm chooses $h$ to minimize

\[ \text{Cost} = W_A(h) \cdot hN + \text{other terms not involving } h, \]

subject to $\tilde{Y} = F(Z hN, \ldots)$. $W_A$ is the average hourly wage (which is potentially a function of average hours), $N$ is the number of workers, $Y$ is output, and $Z$ is the level of labor-augmenting technology. Letting $\lambda$ be the Lagrange multiplier on the constraint, we obtain the first-order condition for $h$ as:

\[ W_A'(h) \cdot h + W_A(h) = \lambda \cdot F_1(Z hN, \ldots) \cdot Z, \]

where $W_A'$ is the derivative of the average wage with respect to $h$ and $F_1$ is the derivative of the production function with respect to effective labor, $Z hN$. The multiplier $\lambda$ is equal to marginal cost, so the marginal cost of increasing output by raising hours per worker is given by:

\[ MC = \lambda = \frac{W_A' \cdot h + W_A}{Z \cdot F_1(Z hN, \ldots)}. \]

The denominator of equation D.3 is the marginal product of increasing hours per worker; the numerator is the marginal increase in the wage bill (per worker). The markup is the price divided by marginal cost.

Following Bils, we specify the average wage function as:

\[ W_A(h) = W_S \left[ 1 + \rho \cdot \theta \cdot \frac{v}{h} \right]. \]

where $W_S$ is the straight-time wage, $\rho$ is the premium for overtime hours, $\theta$ is the fraction of overtime hours that command a premium, and $v/h$ is the ratio of average overtime hours to total hours. The term $\rho \cdot \theta \cdot \frac{v}{h}$ captures the idea that firms may have to pay a premium for hours worked beyond the standard workweek.\(^{32}\) Bils did not include

\(^{32}\) It would also be possible to distinguish wages paid for part-time work versus full-time work. However, Hirsch (2005) finds that nearly all of the difference in hourly wages between part-time and full-time workers can be attributed to worker heterogeneity rather than to a premium for full-time work.
the $\theta$ term in his specification because he used data for manufacturing from the BLS’s establishment survey, in which overtime hours are defined as those hours commanding a premium (that is, $\theta = 1$). In our data, we define overtime hours as those hours in excess of 40 hours per week. Because overtime premium regulations do not apply to all workers, we must allow for the possibility that $\theta$ is less than unity.

We assume that the firm takes the straight-time wage, the overtime premium, and the fraction of workers receiving premium pay as given, but recognizes the potential effect of raising $h$ on overtime hours $v$. With this functional form, the marginal cost of increasing output by raising hours per worker is given by:

$$MC = \lambda = \frac{WS \left( 1 + \rho \cdot \theta \cdot \frac{\partial v}{\partial h} \right)}{Z \cdot F_1 (ZhN, \ldots)}.$$  

Equation D.5 makes it clear that the marginal cost of increasing hours per worker is not necessarily equal to the average wage, as is commonly assumed. Following Bils (1987), we call the term in the numerator the “marginal wage” and denote it by $W_M$:

$$W_M = WS \left( 1 + \rho \cdot \theta \cdot \frac{\partial v}{\partial h} \right).$$  

To the extent that the marginal wage has different cyclical properties from the average wage, markup measures that use the average wage may embed cyclical biases. Bils (1987) used approximations to the marginal wage itself to substitute for marginal cost in his markup measure. We instead derive an expression that does not require approximation. In particular, we combine the expressions for the average wage and the marginal wage to obtain their ratio:

$$\frac{W_M}{W_A} = \frac{1 + \rho \cdot \theta \cdot \frac{\partial v}{\partial h}}{1 + \rho \cdot \theta \cdot \frac{v}{h}}.$$  

This ratio can be used to convert the observed average wage to the theoretically-correct marginal wage required to estimate the markup. We show below that the ratio of overtime hours to average hours, $v/h$, is procyclical. Thus, the denominator in equation D.7 is procyclical. How $W_M/W_A$ evolves over the business cycle depends on the relative cyclicality of $\partial v/\partial h$.

In the case where the wage is increasing in average hours, the markup in any of the previous formulations can be adjusted by multiplying $W_A$ by $\frac{W_M}{W_A}$. For example in the
C-D case, the markup is given by:

\[(D.8) \quad \mathcal{M}_{\text{CD}}^M = \frac{P}{W_M \left[ \alpha \left( \frac{Y}{\alpha N} \right) \right]} = \frac{\alpha}{s \left( \frac{W_M}{W_A} \right)},\]

where we use equation D.7 to convert average wages to marginal wages.

**D.2 Cyclicality of the markup using marginal wages**

We now consider the cyclicality of the marginal-average wage factor and how it affects the cyclicality of the markup. In this case, the measured markup for the C-D case (in natural logarithms) is given by

\[(D.9) \quad \mu_{\text{CD}}^M = -\ln s - \ln \left( \frac{W_M}{W_A} \right),\]

where \(\mu \equiv \ln \mathcal{M}\). The last term is the log of the wage factor used in the average-marginal wage adjustment factor (equation D.7).

To construct the ratio of marginal to average wages, we require (1) estimates of the marginal change in overtime hours with respect to a change in average total hours, \(\partial v / \partial h\); (2) estimates of the ratio of overtime hours to average hours, \(v / h\); (3) the fraction of overtime hours that command a premium, \(\theta\); and (4) the premium for overtime hours, \(\rho\).

The series for \(\partial v / \partial h\) is the most challenging to measure. Bils (1987) speculated that \(\partial v / \partial h\) was procyclical because a given increase in average hours would be more likely to come from an increase in overtime hours if the starting level of average hours was higher. He implemented this idea by regressing the change in average overtime hours, \(\Delta v\), on the change in average total hours, \(\Delta h\), in annual two-digit SIC manufacturing data, and allowing the coefficient in the regression to be a polynomial of average hours.

Average hours based on industry or aggregate data are not ideal for measuring this component for several reasons. As Bils pointed out, higher moments of the average hours distribution could matter because all workers do not work the same average hours. For example, it matters for the marginal wage whether average hours are increasing because more workers are moving from 38 to 39 hours per week or more workers are moving from 40 to 41 hours per week. Ideally, we want to compute the ratio of the change in overtime hours to the change in average hours at the level of the
individual worker and then average over all workers at each point in time. That is, we want to construct the “average marginal” change in overtime hours with respect to a change in average hours. The ideal way to do this is to use panel data on individual workers.\footnote{We are indebted to Steven Davis for suggesting this method for calculating $\partial v / \partial h$.}

We measure both $\partial v / \partial h$ and $v / h$ using individual-level data from Nekarda’s (2013) Longitudinal Population Database, a monthly panel data set constructed from CPS microdata that matches individuals across all months, available for 1976 to 2017. In order to match the BLS private business data, we limit the sample to private-sector workers. We calculate $v / h$ as follows. For all employed workers in each month we sum average weekly overtime hours (defined as those hours in excess of 40 per week) and average weekly hours. We seasonally adjust these two series separately (as discussed below) and then form our series as $\sum v / \sum h$.

To calculate $\partial v / \partial h$, for each matched individual $i$ who is employed in two consecutive months we calculate

\[ \left( \frac{\Delta v}{\Delta h} \right)_{it} = \frac{v_{it} - v_{it(t-1)}}{h_{it} - h_{it(t-1)}}. \]

Then for each month $t$ we take the average over all individuals $N_t$:

\[ \left( \frac{\Delta v}{\Delta h} \right)_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{\Delta v}{\Delta h} \right)_{it}. \]

Ideally, we would limit the matches to individuals employed in the same job over the two consecutive months, but the same-job measure does not exist prior to 1994. However, we found that the matched same-job measure was nearly identical to the matched employment measure after 1994, so we used the matched measure for individuals employed in consecutive months for the entire sample.

The raw data have significant seasonal variation. The CPS asks respondents to report actual hours worked during the week of the month containing the twelfth. Two holidays, Easter and Labor day, periodically fall during the reference week. When one of these holidays occurs during the reference week, actual hours worked falls substantially.

We seasonally adjust the series we calculate from the LPD ($h$, $v$, and $dv/dh$) using the Census Bureau’s X-12-ARIMA program. We include exogenous variables for the...
Figure D1. $\partial v/\partial h$ and $v/h$

Source: Authors' calculations from Nekarda (2013).
Notes: Shaded areas represent periods of business recession as determined by the NBER.

Easter and Labor day holidays that fall during the reference period and remove the estimated effect of these holidays on each series. We then take the quarterly average of the monthly series to match our other aggregate data.

The blue line in figure D1 shows the value of $\partial v/\partial h$. The series shows obvious procyclicality: it tends to rise during expansions and fall during recessions. It also exhibits some low frequency movements, rising from the mid-1970s to late 1990s and then trending lower thereafter. Because $\partial v/\partial h$ appears in the numerator of the wage factor, its procyclicality makes the wage factor more procyclical. But because the wage factor appears in the denominator of the markup, procyclicality of $\partial v/\partial h$ has a countercyclical influence on the markup.

The orange line in figure D1 shows the fraction $v/h$. It is procyclical as well, though it tends to peak a bit before the peak of the business cycle. Like $\partial v/\partial h$, it also displays low frequency movements, although the decline since the late-1990s is more pronounced. Thus, the wage factor in equation D.7 contains a procyclical series in both the numerator and denominator. Hence, the cyclicity of the factor depends in large part on the relative cyclicality of $\partial v/\partial h$ versus $v/h$.

Two more parameters are also required to construct the marginal-average wage factor. One is the fraction of overtime hours that command a premium, $\theta$. We define as overtime hours, any hours worked greater than 40 hours per week. As some of those
hours may come from salaried workers or persons with second jobs, not all hours over 40 are paid a premium. The only direct information is from the May supplements to the CPS in 1969–81, which asked workers whether they received higher pay for hours over 40 hours per week.

We calculate the share of overtime hours that are paid a premium using data from CPS May extracts provided by the NBER.\textsuperscript{34} The overtime variable (x174) is a dummy for whether an individual receives higher pay for work exceeding 40 hours in a week. (Note that the value 0 indicates that a worker received premium pay.)

We drop all individuals that do not report total hours (variable x28). We calculate overtime hours as hours worked at primary job (variable x182) less 40 when this is reported; otherwise, overtime hours is calculated as total hours worked less 40. An individual's paid overtime hours is the product of overtime hours and the indicator for whether overtime hours are paid a premium. We aggregate overtime hours, paid overtime hours, and total hours by year using the individual sampling weights (variable x80). For a given year, the share of overtime that is paid a premium is the ratio of paid overtime hours to total overtime hours.

Unfortunately, the key question on premium pay was dropped from the May supplement after 1985. A potential alternative source of information is the BLS's Employer Costs for Employee Compensation (ECEC) survey which provides information on total compensation, straight time wages and salaries, and various benefits, such as overtime pay, annually from 1991 to 2001 and quarterly from 2002 to the present. If one assumes a particular statutory overtime premium, then one can construct an estimate of $\theta$ from these data. We assume that the statutory premium is 50 percent and construct a $\theta$ accordingly.

Figure D2 shows annual estimates of $\theta$ based on these two sources. From 1969 to 1981, $\theta$ averages 0.33, meaning that only one-third of hours over 40 command a premium. From 1991 to 2009, $\theta$ averages 0.27. Although it appears that the estimate of $\theta$ from the Current Population Survey falls during recessions, regressing $\theta$ on average hours does not yield a significant relationship.\textsuperscript{35} On the other hand, the fraction of hours paid a premium is slightly countercyclical in the ECEC data.\textsuperscript{36} It is difficult to tell whether the structure of the economy actually changed or whether the two surveys are

\textsuperscript{34} http://www.nber.org/data/cps_may.html

\textsuperscript{35} The coefficient from this regression is 0.02 and has a $t$ statistic of 1.40.

\textsuperscript{36} Regressing $\theta$ estimated from the ECEC on CPS average hours yields a coefficient of $-0.03$ with a $t$ statistic of $-3.2$. 

Page 55 of 57
Figure D2. Fraction of Overtime Hours Worked Paid a Premium

Source: Authors’ calculations using data from May CPS extracts (NBER) and the Employer Costs for Employee Compensation survey (BLS).
Notes: The implied $\theta$ for the early sample is based on individual worker reports on hours and whether they are paid a premium from the May CPS extract. The implied $\theta$ for the later sample is based on aggregated data on wages and salaries and overtime compensation from the ECEC survey, coupled with our constructed measure of $v/h$.

...simply not comparable. Because there is little cyclical variation in $\theta$ in either survey, we assume that $\theta$ is a constant equal to the average across the two surveys of 0.3.$^{37}$

Based on the information from these two data sources, we use a value of $\theta = 0.3$ for the private economy.

The final input required for the wage factor is the premium paid for overtime hours, $\rho$. The Fair Labor Standards Act requires that employers pay a 50 percent premium for hours in excess of 40 per week for covered employees. Evidence from Carr (1986) indicates that in 1985, 92 percent of those who earned premium pay received a 50 percent premium.$^{38}$ Although there is considerable evidence that the implicit premium could be closer to 0.25, we use a $\rho$ of 0.50 to reflect the statutory premium.$^{39}$ Using the

---

37. If we instead assume that $\theta$ is procyclical with the coefficient of 0.024 on average hours, our estimates of the marginal-average wage factor change little.
38. See Wetzel (1966) and Taylor and Sekscenski (1982) for other estimates.
39. Trejo (1991) has questioned whether the true cost of an extra overtime hour for those covered is actually 50 percent. He shows that the implicit cost of overtime hours is lower than 50 percent because straight-time wages are lower in industries that offer more overtime. Hamermesh (2006) updates his analysis and finds supporting results: The implicit overtime premium is 25 percent, not 50 percent. The results using a 25 percent premium lie between those using the average wage and those using the 50 percent premium.
higher overtime premium will bias the analysis toward finding countercyclical markups.

Figure D3 shows the marginal-average wage factor. Although the movements in the wage factor are procyclical, the magnitude of the variation is so small that it does not change the cyclicality of the markup. Specifically, the estimated elasticity from 1976 to 2017 of the baseline markup in private business to real GDP is 0.14. This falls to 0.07 measured using marginal wages. Neither of these estimates is statistically different from zero.

These results stand in contrast to those of Bils (1987), who found countercyclical markups in two-digit annual manufacturing data from 1956 to 1983. Our explorations suggest that Bils’ results are due to the combination of details in the implementation of his method for estimating $\partial v/\partial h$. We show that even within his framework, small adjustments in the method eliminate the finding of countercyclicality.\textsuperscript{40}

---

\textsuperscript{40} See Appendix B of Nekarda and Ramey (2013) for details.