Resurrecting the Role of the Product Market Wedge in Recessions

Mark Bils
University of Rochester and NBER

Peter J. Klenow
Stanford University and NBER

Benjamin A. Malin*
Federal Reserve Bank of Minneapolis

September 3, 2015

Abstract

Employment and hours appear far more cyclical than dictated by the behavior of productivity and consumption. This puzzle has been labeled “the labor wedge” — a cyclical wedge between the marginal product of labor and the marginal rate of substitution of consumption for leisure. The wedge can be broken into a product market wedge (price markup) and a labor market wedge (wage markup). Based on the wages of employees, the literature has attributed the wedge almost entirely to labor market distortions. Because employee wages may be smoothed versions of the true cyclical price of labor, we instead examine the self-employed as well as intermediate inputs. Looking at the past quarter century in the U.S., we find that price markup movements are at least as important as wage markup movements — including in the Great Recession and its aftermath. Thus sticky prices and other forms of countercyclical markups deserve a central place in business cycle research, alongside sticky wages and matching frictions.

*The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve System. We are grateful to Corina Boar, Cian Ruane and Zoe Xie for excellent research assistance.
1. Introduction

Employment and hours are more cyclical than can be explained by real labor productivity under conventional preferences for consumption and leisure – see Hall (1997), Mulligan (2002), and Chari, Kehoe and McGrattan (2007), among others. More recently, there is a growing consensus that this “labor wedge” reflects labor market frictions. Examples include Gali, Gertler, and Lopez-Salido (2007), Shimer (2009), Hall (2009), and Karabarbounis (2014a). This has helped drive an explosion of work on search, matching, and wage setting in the labor market.

The consensus that the labor wedge reflects labor market frictions is based on measuring the price of labor using average hourly earnings. The gap between average hourly earnings and labor productivity is acyclical, suggesting price markup movements are not cyclical. But it is not clear whether the marginal cost of labor to firms is well-measured by average hourly earnings. Employee wages may not reflect the true marginal cost of labor to the firm. Wages may be smoothed versions of the shadow cost due to implicit contracting (e.g., for salaried workers). One obtains a very different picture of the cyclical price of labor from using the wages of new hires, as measured by Pissarides (2009) and Haefke, Sonntag, and van Rens (2013), or from the user cost of labor, as measured by Kudlyak (2014).

In this paper, we seek evidence on cyclical distortions in the product market that do not rely on measures of the shadow price of labor. First, we estimate the labor wedge for the self-employed. If we observe significant cyclicality in the labor wedge for the self-employed, it cannot be ascribed to wage rigidities or other labor market frictions. Second, we estimate the product market wedge from intermediate inputs (energy, materials, and services). Intermediate prices should provide a truer measure of that input’s cyclical price than do average hourly earnings for labor.

Our evidence is for the U.S. from 1987 onward. For the self-employed, we look at the Current Population Survey and the Consumer Expenditure Survey, both conducted by the Bureau of Labor Statistics (BLS). For intermediates we use the BLS Multifactor Productivity Database covering 60 industries. Our consistent finding is that, contrary to the emerging consensus, product market distortions are at least as important as labor
market distortions in recent recessions.

The cyclical wedge we see for all inputs is compatible with firms whose sales are constrained in recessions by a (too high) sticky price. Given the wedges’ strong persistence, it is also consistent with firms purposefully choosing a higher markup over marginal cost in recessions. As a recent example, Gilchrist, Schoenle, Sim, and Zakrajsek (2014) find that financially-constrained firms chose higher markups rather than investing in market share during the Great Recession. Any model where expanding production has a component of investment (e.g., learning-by-doing) should have similar implications. Additionally, the product market wedge could reflect greater uncertainty, or aversion to uncertainty, in recessions, e.g., as in Arellano, Bai, and Kehoe (2012).

Our findings also speak directly to the puzzle of unemployment’s high cyclicality relative to that in labor productivity — the Shimer (2005) puzzle. A highly countercyclical product-market wedge translates into strongly procyclical labor demand, beyond what might be attributed to labor productivity. It provides a rationale for firms to create less employment in recessions without a decline in productivity, and even absent important wage stickiness.


The paper proceeds as follows. Section 2 revisits the standard labor wedge calculations as a point of comparison. Section 3 looks at the self-employed. Section 4 investigates intermediate input use. Section 5 relates our work to other efforts at measuring cyclicality of markups. Section 6 concludes.
2. The Aggregate Labor Wedge

We begin by constructing the standard representative-agent labor wedge (RAW), defined as the (log) ratio of the marginal product of labor \( (mpn) \) to the tax-adjusted marginal rate of substitution of consumption for leisure \( (mrs) \). Shimer (2009) provides a thorough derivation of the RAW, starting from the maximization problems of a representative household and firm. Constructing the wedge requires assumptions on preferences and technology; our baseline case follows Hall (1997) and Gali, Gertler, and Lopez-Salido (2007). Production features a constant elasticity with respect to hours. Preferences are separable in consumption and hours, and over time. They feature a constant intertemporal elasticity for consumption, and a constant Frisch elasticity of labor supply. These assumptions yield a log-linear labor wedge:

\[
RAW_t \equiv \ln(mp_{nt}) + \ln(1 - \tau_t) - \ln(mrs_t) \\
= \ln\left(\frac{y_t}{n_t}\right) + \ln(1 - \tau_t) - \left[\frac{1}{\sigma} \ln(c_t) + \frac{1}{\eta} \ln(n_t)\right],
\]

where \( \frac{y_t}{n_t} \) is output per hour, \( c_t \) is nondurables and services consumption per adult equivalent, \( n_t \) is hours per capita, and \( \tau_t \equiv \frac{\tau_c + \tau_n}{1 + \tau_t} \) is a combination of average marginal tax rates on consumption and labor.

For our baseline case, we use an intertemporal elasticity of substitution (IES) of \( \sigma = 0.5 \) following Hall (2009), and a Frisch elasticity of labor supply of \( \eta = 1.0 \), based on Chang, Kim, Kwon, and Rogerson (2014). The latter argue, based on a heterogeneous-agent model with both intensive and extensive labor margins, that a representative-agent Frisch elasticity of 1 (or slightly higher) is reasonable.

To gauge the cyclicality of the RAW, we project it on real GDP and hours worked. (All variables in logs and HP-filtered.) We use quarterly data from 1987 through 2012. Table 1 reports the cyclical elasticity of the wedge and its components: labor

---

1We later entertain CES production in capital, labor, and intermediate inputs.
2Nonseparable utility in consumption and leisure would not alter our results significantly. Shimer (2009) and Karabarbounis (2014a) found this as well. We find this if we calibrate the nonseparability to how consumption responds to retirement (Aguiar and Hurst, 2013) or unemployment (Saporta-Eksten, 2014).
3See the Appendix for a precise description of all variables used.
productivity, hours worked, consumption, and taxes. The wedge is strongly countercyclical (elasticity with respect to GDP: -2.69), reflecting mildly countercyclical productivity (-0.10), procyclical consumption (0.61) and very procyclical hours (1.40). In recessions, the RAW increases as the $mrs$ plummets but the $mpn$ changes little. Using the results in Table 1, it is straightforward to recalculate the wedge's cyclicity for alternative calibrations of $\sigma$ and $\eta$.

Table 1: Representative Agent Wedge

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representative Agent Wedge</td>
<td>-2.69 (0.20)</td>
<td>-2.00 (0.06)</td>
</tr>
<tr>
<td>Labor Productivity</td>
<td>-0.10 (0.08)</td>
<td>-0.28 (0.04)</td>
</tr>
<tr>
<td>Hours per capita</td>
<td>1.40 (0.07)</td>
<td>0.99 (0.01)</td>
</tr>
<tr>
<td>Consumption per capita</td>
<td>0.61 (0.03)</td>
<td>0.36 (0.02)</td>
</tr>
<tr>
<td>Tax Rates</td>
<td>0.02 (0.07)</td>
<td>-0.01 (0.04)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Sample covers 1987Q1-2012Q4. All variables in logs and HP filtered. Wedge calculation uses $\sigma = 0.5$ and $\eta = 1.0$.

As shown, the contribution of marginal tax rates to the cyclicity of the RAW is small. Because our tax measures have little impact on our results, we drop them in the remainder of the paper.4

In these baseline calculations we ignore cyclical fluctuations in the quality of the workforce and a role for overhead labor.5 We calculate that the declines in average

---

4Mulligan (2012) contends that changes in effective marginal tax rates influenced labor market behavior in the Great Recession. His focus is on how lower income workers have been affected by the expansion of means-tested assistance programs.

5We also ignore the possibility of home production, which Karabarbounis (2014b) suggests can explain part of the cyclicity in the labor wedge. Essentially, time spent producing at home implies smaller cyclical movements in effective leisure, while consumption of home-produced goods implies smaller movements in effective consumption. Note, however, that our intertemporal and Frisch elasticities are calibrated to a literature that largely estimates models without home production. Capturing the same empirical moments with a model with home production, as in Karabarbounis, would therefore imply lower intertemporal and Frisch elasticities. These lower elasticities would (at least partly) offset the smaller movements in effective leisure and consumption.
workforce quality in expansions imply that we understate the cyclical elasticity of labor’s marginal product by perhaps 0.1 to 0.2 (causing cyclicality in the wedge to be overstated). Ignoring overhead labor, conversely, causes one to overstate the procyclicality of labor's marginal product (Rotemberg and Woodford, 1999). For an overhead labor component of the magnitude suggested by Nekarda and Ramey (2013), 10 to 20 percent, the impacts of composition and overhead labor on cyclicality of labor's marginal versus average product (and on the estimated labor wedge) should approximately offset.

2.1. Extensive- and Intensive-Margin Wedges

We next construct separate wedges on the extensive margin (EMW) and the intensive margin (IMW). These distinguish between the two components of hours worked, employment and hours per worker. We make this distinction for four reasons. First, we can calibrate the Frisch elasticity of labor supply to micro estimates at the hours margin. Second, we can compare the intensive margin here to the intensive margin for the self-employed (in Section 3). Third, product market distortions should impact the wedge on both margins. If the labor wedge is only important at one margin, it would suggest the product market wedge has little cyclical importance. Finally, although the EMW appears in many theoretical models, to our knowledge it has not been constructed empirically.

In order to analyze the extensive margin, we make some additional assumptions. We consider a representative household that consists of many members. Consumption is perfectly shared across members, and labor supply decisions are made on both the extensive and intensive margins. Preferences are given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{c_t^{1-1/\sigma}}{1-1/\sigma} - \nu \left( \frac{h_t^{1+1/\eta}}{1+1/\eta} + \psi \right) e_t \right\},$$

where $c_t$ denotes per-capita consumption, $e_t$ employment, and $h_t$ hours worked per employee. $\psi$ is a fixed cost of employment, which guarantees an interior solution for the choice of hours versus employment. The marginal disutility of employment is $\nu \left( \frac{h_t^{1+1/\eta}}{1+1/\eta} + \psi \right) \equiv \nu \Omega_t h_t$, while the marginal disutility of an extra hour per worker is
For firms, we assume (i) a constant output elasticity with respect to labor, and (ii) employment and hours per worker are perfect substitutes (i.e., production depends on total hours, \( n_t = c_t h_t \)). The marginal product of labor is thus proportional to output per hour \( \frac{y_t}{n_t} \), while the marginal product of employment is \( mpn_t^\text{ext} \propto \frac{w_t}{n_t} h_t \) and the marginal product of hours per worker is \( mpn_t^\text{int} \propto \frac{w_t}{n_t} e_t \).

There are frictions in finding employment. Firms post vacancies at the beginning of the period, and matches form and produce during the period.\(^6\) The matching technology is \( mt = v_t f(u_t) \), where \( m_t \) are matches, \( v_t \) vacancies, and \( u_t \) unemployment. \( \kappa \) denotes the opportunity cost of creating a vacancy, expressed in labor input as the fraction of the steady-state workweek \( h \). \( \delta \) is the exogenous per-period separation rate; and \( \gamma \) is the fraction of the initial period of employment devoted to training.

In this environment, the intensive margin wedge is given by

\[
IMW_t \equiv \ln(mp(n_t)^\text{int}) - \ln(mrs(n_t)^\text{int}) = \ln\left(\frac{y_t}{n_t}\right) - \left[\frac{1}{\sigma}\ln(c_t) + \frac{1}{\eta}\ln(h_t)\right].
\]

The IMW, equation (2), differs from the standard RAW, equation (1), in two ways: hours per worker \( h_t \) replaces hours per capita \( n_t \), and we calibrate \( \eta = 0.5 \). A lower \( \eta \) is appropriate given it now reflects the Frisch elasticity only at the intensive (hours) margin (e.g., Chetty et al., 2013).\(^7\)

On the extensive margin, consider creating one more vacancy in period \( t \) and reducing vacancies in \( t + 1 \) just enough to keep employment unaffected in \( t + 1 \) forward. Spending \( \kappa h_t \frac{w_t}{n_t} \) to create an additional vacancy generates \( \phi m_t / v_t \) additional matches, of which \( (1 - \delta) \) survive to \( t + 1 \). The perturbation thus requires lower

\(^6\)Blanchard and Gali (2010) and Gali (2011) use this timing, although it is more conventional for matches to start producing in the following period. The former timing gives cleaner results, but it could be altered without changing our analysis substantially.

\(^7\)Pescatori and Tasci (2012) point out that the labor wedge is less variable when calculated using hours per worker rather than hours per capita. They, however, hold the Frisch elasticity fixed across the workweek and representative agent calculations.
spending on vacancies at $t + 1$ by $(1 - \delta)\kappa h \frac{n_{t+1}}{n_t} \frac{m_t/v_t}{m_{t+1}/v_{t+1}}$. A social planner would set:

$$\frac{\phi m_t}{v_t} \left[ u'(c_t) \left( 1 - \gamma \frac{h_t}{h_t} \right) h_t \frac{y_t}{n_t} - \Omega h_t \right] - u'(c_t)\kappa h \frac{y_t}{n_t}$$

$$+ \beta (1 - \delta) E_t \left\{ u'(c_{t+1}) \left( \kappa h \frac{y_{t+1}}{n_{t+1}} \frac{m_t/v_t}{m_{t+1}/v_{t+1}} + \frac{\phi m_t}{v_t} \gamma h \frac{y_{t+1}}{n_{t+1}} \right) \right\} = 0. \quad (3)$$

I.e., the marginal benefit of an extra vacancy (utility from consuming increased output today) equals its marginal cost (the disutility of employment plus the resource cost of creating an added vacancy today less the resource savings from creating fewer future vacancies). We then rearrange equation (3) to get a (log) ratio of the marginal benefit to the marginal cost of an additional unit of labor on the extensive margin. We get

$$EMW_t = \ln \left( \frac{y_t}{n_t} \right) - \left[ \frac{1}{\sigma} \ln(c_t) + \ln(\Omega_t) \right] - S_t,$$

$$S_t \approx \frac{h}{h_t} \left( \kappa \frac{\phi m}{\phi m + \gamma} \left[ 1 - \beta (1 - \delta) E_t \left\{ u'(c_{t+1}) \left( \frac{y_{t+1}}{n_{t+1}} \frac{m_t/v_t}{m_{t+1}/v_{t+1}} \right) \right\} \right] + \gamma \left[ 1 - \beta (1 - \delta) E_t \left\{ \frac{u'(c_{t+1})}{u'(c_t)} \frac{y_{t+1}}{n_{t+1}} \frac{m_t/v_t}{m_{t+1}/v_{t+1}} \right\} \right] \right), \quad (4)$$

where $\Omega_t$ is the marginal disutility of employment (per hour worked).\(^8\)

The EMW, like the IMW, reflects movements in labor productivity, $\ln \left( \frac{y_t}{n_t} \right)$, and the marginal utility of consumption, $\frac{1}{\sigma} \ln(c_t)$. But there are differences from the IMW. Whereas the IMW reflects the marginal disutility of an extra hour, which is highly procyclical for reasonable Frisch elasticities, the extensive margin reflects the average disutility of adding a worker. We find this average disutility to be nearly acyclical. In addition, the term $S_t$, which is specific to the EMW, reflects the efficacy of spending on vacancies. In recessions $S_t$ declines as vacancies are more likely to yield a match. This lends a countercyclical component to the EMW. The cyclicality of the EMW vis-a-vis the IMW essentially reduces to whether cyclicality in the hiring term $S_t$ exceeds that in the marginal disutility of working a longer workweek.

It is well established that cyclical movements in total hours are primarily driven by

\(^8\)Our derivation makes use of the approximation $\ln(1 + x_t) \approx constant + \frac{x_t}{1 + x_t}$. The Appendix provides more details on the EMW (and IMW) construction.
employment fluctuations. That holds true for our 1987 to 2012 period, with cyclical employment fluctuation (with respect to real GDP or total hours) that are about four times that in the workweek. One might jump to the conclusion that cyclicity in the extensive (employment) margin would similarly dominate that in the intensive (workweek) margin. That jump would be unwarranted. Important components of the wedge—i.e., labor’s marginal product and consumption’s marginal utility—display the same cyclicity with respect to both margins. Any wedge differences, as just discussed, reduce to cyclicity in the hiring term $S_t$ versus that in the marginal disutility of working a longer workweek. While employment fluctuations are larger than those in the workweek, the elasticity of the marginal disutility in response to the workweek may exceed that of the hiring cost $S_t$ to employment.

To construct the EMW, we use the same variables required by the IMW, plus vacancies ($v_t$), matches ($m_t$), and additional parameters. A quarterly separation rate of $\delta = 0.105$ matches the average rate of quits, layoffs, and discharges in JOLTS. $\beta = 0.996$ implies a steady-state annual real interest rate of 1.6%, the average of the 3-month T-bill rate less core PCE inflation over our sample. Hiring costs per match, $\kappa$, are set to 0.4 quarters of output, and training costs to $\gamma = 0.16$, consistent with estimates by Barron et al. (1999). Finally, the elasticity of matches to vacancies is $\phi = 0.5$. These parameters imply a steady-state ratio of $mrs$ to $mpn$ on the extensive margin of about 0.90.

Figure 1 shows the unfiltered extensive and intensive margin wedges from 1987–2012. Table 2 reports their cyclical elasticities with respect to real GDP and hours worked. The EMW and IMW elasticities are similar, and smaller than for the RAW. An aggregate Frisch elasticity of 2.3 would make the RAW behave similarly to the EMW and IMW.

### 2.2. Decomposing the Wedge

We now empirically decompose the labor wedge into product-market (i.e., price markup) and labor-market (i.e., wage markup) components. This requires a measure of the marginal cost of labor to firms. As stressed by Gali, Gertler, and Lopez-Salido (2007) and others, the assumption that any particular wage measure reflects labor's
true marginal cost is controversial. We will show that alternative wage measures lead
to vastly different conclusions about the relative importance of the product- and
labor-market wedges. This motivates our subsequent analysis, which decomposes the
labor wedge without using wage data.

The IMW decomposition is standard and given by

\[
IMW_t = \left[ \ln \left( \frac{y_t}{n_t} \right) - \ln \left( \frac{w_t}{p_t} \right) \right] + \left[ \ln \left( \frac{w_t}{p_t} \right) - \frac{1}{\sigma} \ln(c_t) - \frac{1}{\eta} \ln(h_t) \right]
\]

(5)

where \( \frac{w_t}{p_t} \) is the (real) marginal cost of labor to firms. The intensive product market
wedge (\( \mu^{p,\text{int}} \)) is the gap between the firm’s marginal product and marginal cost of labor.
The intensive labor market wedge (\( \mu^{w,\text{int}} \)) is the gap between the firm’s marginal cost
and the household’s cost of providing an additional hour. The EMW decomposition is

\[
EMW_t = \left[ \ln \left( \frac{y_t}{n_t} \right) - S_t - \ln \left( \frac{w_t}{p_t} \right) \right] + \left[ \ln \left( \frac{w_t}{p_t} \right) + \tilde{S}_t - S_t - \frac{1}{\sigma} \ln(c_t) - \ln(\Omega_t) \right]
\]

\[
= \mu^{p,\text{ext}} + \mu^{w,\text{ext}},
\]

(6)
Table 2: Extensive & Intensive Margin Wedges

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensive Margin Wedge</td>
<td>-1.99 (0.26)</td>
<td>-1.55 (0.14)</td>
</tr>
<tr>
<td>Intensive Margin Wedge</td>
<td>-1.91 (0.13)</td>
<td>-1.38 (0.05)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Sample covers 1987Q1-2012Q4. All variables in logs and HP-filtered. Wedge calculations use $\sigma = 0.5$ and $\eta = 0.5$, and EMW expectation terms constructed by a VAR.

where $\tilde{S}_t$ takes the same form as $S_t$ (see equation (4)) but with $\phi = 1$. For intuition, temporarily let $\tilde{S}_t = S_t$ in equation (6). Doing so, it’s apparent that the extensive labor-market wedge ($\mu^{w,ext}$) mirrors the intensive ($\mu^{w,int}$), but with the household’s $mrs$ measured along the employment margin. For the extensive product-market wedge ($\mu^{p,ext}$), an additional employee produces $\ln\left(\frac{w}{m}\right) - S_t$ for the firm; i.e., in our decomposition firms pay the vacancy cost ($S_t$). Finally, using $\tilde{S}_t$ rather than $S_t$ reflects the fact that firms do not internalize the congestion effects of their decision to post another vacancy; each firm views the probability of filling a vacancy as $\frac{m}{v}$, whereas the social planner knows one more vacancy generates $\frac{\phi m}{v}$ additional matches.

Table 3 decomposes the EMW and IMW into product- and labor-market wedges using average hourly earnings (AHE) as the (nominal) measure of the firm’s marginal cost of labor ($w_t$). This wage measure is conventional and would reflect the true marginal cost if all workers were employed in spot markets. In this case, $\ln\left(\frac{w}{m}\right) - \ln\left(\frac{w}{pv}\right)$ is the (log) inverse labor share. The product-market wedge accounts for between 2 and 6% of the wedge cyclicity on the intensive margin and between 17 and 23% on the extensive margin. Thus, the results are in line with Karabarbounis’s (2014a) conclusion that the product-market wedge is relatively unimportant.

Alternative frameworks for understanding the labor market, however, emphasize the durable nature of the firm-worker relationship and imply the contemporaneous
Table 3: **Wedge Decomposition: Average Wage**

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensive Margin Wedge</td>
<td>-1.99 (0.26)</td>
<td>-1.55 (0.14)</td>
</tr>
<tr>
<td>Product Market (AHE)</td>
<td>-0.32 (0.14)</td>
<td>-0.35 (0.09)</td>
</tr>
<tr>
<td>Intensive Margin Wedge</td>
<td>-1.91 (0.13)</td>
<td>-1.38 (0.05)</td>
</tr>
<tr>
<td>Product Market (AHE)</td>
<td>-0.04 (0.08)</td>
<td>-0.08 (0.05)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Sample is from 1987Q1 through 2012Q4. All variables in logs and HP filtered. Expectation terms in EMW constructed using a VAR. The extensive product market wedge ($\mu_{p,ext}$) follows equation (6), and the intensive product market wedge ($\mu_{p,int}$) follows equation (5).

Wage plays no allocative role. For example, in matching models with search frictions, what matters is the (expected) economic surplus generated over the life of the match and not the wage payment at any one time. Implicit contracting models similarly imply that the flow wage payment is not allocative. Barro (1977) and Rosen (1985) both forcefully drive home that an acyclical, or even countercyclical wage, can coincide with a highly procyclical true price of labor. (See Basu and House, 2015, as well.) Recent examples of empirical support for implicit contracting include Ham and Reilly (2013) and Belliou and Kaymak (2012).

In light of this, Table 4 shows how alternative measures of firms’ marginal cost of labor affect the EMW decomposition. We consider the wages of new hires (NH) and Kudlyak’s (2014) user cost of labor (UC), which have been argued to be more relevant for job formation in search frameworks. We find that, depending on the wage measure...
used, the product market wedge can account for almost none or essentially all of the EMW (i.e., from 16% to 110%). We presume the product market wedge is similar on both the intensive and extensive margins — so this implies it could be anywhere from nearly zero to nearly all of the intensive margin as well.

Table 4: **Wedge Decomposition: Alternative Wage Measures**

<table>
<thead>
<tr>
<th>Elasticity wrt</th>
<th>GDP</th>
<th>Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensive Margin Wedge</td>
<td>-1.99 (0.26)</td>
<td>-1.55 (0.14)</td>
</tr>
<tr>
<td>Product Market (AHE)</td>
<td>-0.32 (0.14)</td>
<td>-0.35 (0.09)</td>
</tr>
<tr>
<td>Product Market (NH)</td>
<td>-0.98 (0.16)</td>
<td>-0.81 (0.09)</td>
</tr>
<tr>
<td>Product Market (UC)</td>
<td>-2.17 (0.21)</td>
<td>-1.65 (0.09)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Sample is from 1987Q1 through 2012Q4. All variables in logs and HP filtered. The product market wedge follows equation (6).

3. **The Self-Employed Wedge**

We now consider the cyclicality of hours and the labor wedge specifically for self-employed workers. If we observe significant cyclicality in the wedge for these workers, presumably it cannot be attributed to wage rigidities or other labor market distortions.

As a starting point, we note that self-employment has been as cyclical as total employment. The share of self-employed in nonagricultural industries declined slightly during each of the past two NBER-defined recessions: from 10.1 to 10.0 percent during 2001 and from 10.5 to 10.4 percent from 2007 to 2009. The unemployment rate, while the relative fall in the user cost of labor is 3.4%. The results in Table 4 reflect adjusting the time series of average hourly earnings by these cyclical factors.

---

self-employment share exhibits lower-frequency fluctuations, but, if we HP-filter, the resulting cyclical series is completely acyclical with respect to GDP or aggregate hours. Becoming self-employed requires starting a business; so fluctuations in self-employment could be affected by financing costs and constraints. In particular, the decline in self-employment during the Great Recession may partly reflect financing constraints. Going forward, we thus focus on the intensive (hours) margin for examining the self-employed.

We base our analysis on the Annual Social and Economic supplements to the CPS, typically referred to as the March CPS. In the March supplement household members report their hours and earnings for the previous calendar year. They also report the earnings and class of worker at their primary job – the job held longest during the prior year. The class-of-worker variables allow us to distinguish the self-employed separately for agriculture and non-agriculture. We begin our sample in 1987, the first year that data on primary-job earnings are available. Advantages of the March supplement are: (i) it is large; (ii) its top-coding of earnings is less extreme than in the monthly surveys; and (iii) some households can be matched across two consecutive March surveys, allowing us to examine year-over-year changes for a given set of workers. Our unmatched sample contains 197,723 self-employed workers for 1987 to 2012 (1,901,936 wage earners).

Figure 2 reports usual weekly hours and total annual hours worked separately for self-employed (nonagricultural) and those earning wages and salaries for 1987 to 2012. The intensive margin is clearly more cyclical for the self-employed. During the Great Recession (2007-2009) the workweek for self-employed declined by 4.9 percent (2 full

---

12These numbers are for nonagriculture, which represents 94 percent of the self-employed. For agriculture, self-employment (again from Hipple) is acyclical. We focus on nonagriculture workers as top and bottom coding of earnings in the CPS is extreme for farmers. For farmers it is also implausible to treat realized income as known at the time labor input is chosen, an assumption implicit in calculations of the labor wedge. Finally, we presume farmers face a competitive market.

13We require that workers be between ages 20 and 70 and worked at least 10 hours per week and at least 10 weeks during the year. Some earnings and hours responses are top coded. For each survey we trim the top and bottom 9.6% of workers by earnings on primary job. 9.6% is just large enough to remove top-coding of earnings for self-employed in all 26 years of surveys. We trim the bottom for symmetry; this also serves to remove all negative earnings. Usual hours are top coded at 99 per week. We trim the top 1.2 percent of workers by weekly hours. This is the minimal trimming that removes top-coded hours for all years.
hours) compared to only 1.7 percent for wage earners.\footnote{If we regress hours per week on real GDP (both series in logs and hp-filtered), the impact for self-employed at 0.37 (standard error 0.14) is nearly twice that of 0.20 (s.e. 0.02) for wage earners.} Similarly, annual hours declined by 6.9 percent for self-employed compared to 3.2 percent for wage earners.

Figure 2 might be influenced by composition bias. For example, if workers becoming self-employed in expansions work more hours than the typical self-employed worker, then hours in Figure 2 will have a procyclical bias. For this reason, we match self-employed workers across consecutive March supplements, constructing growth rates for their hours and earnings.\footnote{We follow standard matching procedures for the March CPS. Respondents are matched across years based on household and person identifiers and conformity of respondent’s sex, race, and age.} Using these growth rates, we express hours and earnings relative to 1987. We are not able to match workers across 1994 and 1995 calendar years due to a CPS sample redesign. For 1994-1995 we impute to each series its mean growth rate. We then create a level series indexed to 1987. In all subsequent statistics, we exclude years 1994 and 1995. Our matched sample includes 39,306 self-employed workers, prior to trimming to deal with top coding.\footnote{The March CPS responses for weeks worked and usual hours per week are for all prior-year jobs, whereas class of worker and earnings refer to the primary (longest-held) job. To achieve an earnings-compatible measure of hours growth, we restrict our self-employed sample to those who received 95
Comparing these hours indices for self-employed and wage-earners, the workweek is more cyclical for the self-employed. The elasticity of the workweek with respect to real GDP (both variables HP-filtered) is 0.28 (s.e. 0.07) for self-employed versus 0.17 (s.e. 0.03) for wage-earners. For annual hours are slightly more cyclical for wage-earners, with an elasticity with respect to real GDP of 0.57 (s.e. 0.07), compared to 0.54 (s.e. 0.13) for the self-employed. (Similar remarks apply if we measure the cycle by aggregate hours.)

In Table 5 we report the cyclicality of the intensive-margin labor wedge. The first column is estimated for all workers, not just the self-employed. It repeats the analysis from Section 2, but uses workweek fluctuations constructed from the matched-CPS surveys. It is also annual, rather than quarterly, and excludes years 1994 and 1995 because we are unable to match those years in the CPS. It dispenses with the tax wedge, which we found to have little impact. As in Section 2, we find a strongly countercyclical wedge. The elasticities of the wedge with respect to real GDP and aggregate hours, -1.87 and -1.20, are modestly smaller than in Section 2 (estimates there being -1.91 and -1.38), with the difference reflecting a slightly less procyclical workweek.

Columns 2-4 of Table 5 show how the wedge’s cyclicality changes as we sequentially replace the estimates of cyclicality in hours, productivity, and consumption for all workers with estimates for the self-employed. Column 2 constructs the wedge using fluctuations in the workweek just for self-employed workers, maintaining the same aggregate series for productivity and consumption. Not surprising, given the greater cyclical nature of the workweek for self-employed described previously, this results in a slightly more cyclical labor wedge. The elasticity with respect to real GDP goes from -1.87 (s.e. 0.10) to -2.06 (s.e. 0.17).

We next replace aggregate labor productivity with a measure of productivity percent of earnings from their primary self-employed job. (The average of that earnings share across the two years must be at least 0.95.)

17 Figures A2-A5 in the Appendix plot the series that underlie these estimates.
18 If we used fluctuations in annual hours, rather than weekly hours, then the wedge in column 1 with respect to real GDP, maintaining a Frisch elasticity of 0.5, would become -2.66 (s.e. 0.15), while that in column 2 would become -2.60 (s.e. 0.24).
Table 5: **Cyclicality of Labor Wedge, All Workers vs. Self-Employed**

<table>
<thead>
<tr>
<th>Elasticity wrt</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real GDP</td>
<td>-1.87 (0.10)</td>
<td>-2.06 (0.17)</td>
<td>-1.97 (0.25)</td>
<td>-3.23 (1.00)</td>
</tr>
<tr>
<td>Total Hours</td>
<td>-1.20 (0.05)</td>
<td>-1.41 (0.10)</td>
<td>-1.29 (0.16)</td>
<td>-1.93 (0.61)</td>
</tr>
</tbody>
</table>

**Hours**

<table>
<thead>
<tr>
<th></th>
<th>All workers</th>
<th>SE</th>
<th>SE</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPN</td>
<td>Agg. y/n</td>
<td>Agg. y/n</td>
<td>SE earn/hr</td>
<td>SE earn/hr</td>
</tr>
<tr>
<td>Consumption</td>
<td>NIPA PCE</td>
<td>NIPA PCE</td>
<td>NIPA PCE</td>
<td>+ CE Adj.</td>
</tr>
</tbody>
</table>

Notes: Sample is based on matched-March CPS self-employed outside government and agriculture. CPS observations are weighted. Each cell represents a separate regression. Regressions have 24 annual observations, 1987-1993 and 1996-2012. Newey-West standard errors are in parentheses. Hours are weekly. NIPA PCE refers to aggregate real expenditures on nondurables and services. CE adjustment incorporates consumption for the self-employed vs. all persons from the Consumer Expenditure Surveys.

specific to the self-employed.\textsuperscript{19} Our measure equals a self-employed worker's annual earnings divided by annual hours. Movements in self-employed earnings per hour correspond to movements in marginal product assuming: (i) a constant elasticity of output with respect to the self-employed worker's labor, as with Cobb-Douglas production in Section 2; and (ii) earnings of the self-employed worker are proportional to their output. Self-employed earnings per hour could overstate the procyclicality of labor's marginal product by ignoring the overhead component of self-employed labor. That could be especially important for self-employed production, given its small scale.

\textsuperscript{19}We deflate earnings by the nondurables and services PCE deflator. We use the midpoint formula to calculate the percentage change in a worker's earnings across matched years. This avoids extreme values for workers with very low earnings in one of the matched years.
of operations. That would cause us to understate the cyclicality in their labor wedge. Another concern is that reported earnings could misstate actual earnings. The self-employed tend to understate earnings. Hurst, Li, and Pugsley (2014) show that the ratio of consumption to income is higher in survey data for the self-employed, consistent with the self-employed understating income. The concern for us would be if these workers underreport at a lower rate in recessions.

Going from Column 2 to 3 of Table 5, we replace aggregate labor productivity with self-employed earnings per hour. Aggregate labor productivity has been modestly countercyclical since 1987, with an elasticity with respect to real GDP of -0.21 (s.e. 0.07). Self-employed earnings per hour have been less cyclical (elasticity -0.13, s.e. 0.19). Thus, the estimated labor wedge becomes slightly more procyclical, with an elasticity of -1.97 (s.e. 0.25) with respect to real GDP. In summary, the wedge calculated with measured productivity and hours for the self-employed is just as cyclical as that for all workers. Figure 3 plots the time series of these two wedges.

We have assumed the cyclicality of consumption for self-employed workers is the same as for consumption per capita. For robustness, we estimate self-employed consumption relative to aggregate consumption based on quarterly growth rates in
household spending on nondurables and services in the Consumer Expenditure Surveys (CE). We add these estimates of relative consumption to aggregate consumption to obtain an estimate of consumption for the self-employed.

The elasticity of aggregate consumption with respect to real GDP is 0.64 (s.e. 0.04). Self-employed consumption is even more procyclical, with an elasticity of 1.27. But the standard error is too large, at 0.56, to reliably infer that the self-employed have more procyclical consumption. The big standard error reflects the small number of self-employed observations in the CE. If we do use this measure of consumption, however, we get an even more cyclical wedge for the self-employed. This is illustrated in the last column of Table 5. The self-employed labor wedge now exhibits an elasticity of -3.23 (s.e. 1.00) with respect to real GDP. In this section’s subsequent exercises, we revert to measuring self-employed consumption by aggregate consumption, rather than adopting such a noisy measure.\textsuperscript{20}

Table 6 presents two robustness exercises. First, self-employed who are incorporated might take income in the form of corporate profits rather than business earnings. It is not obvious how incorporated self-employed treat these profits in answering the CPS earnings question. For this reason, as an alternative measure of labor productivity, we consider earnings per hour excluding the incorporated self-employed. This series is more procyclical than earnings per hour for all self-employed. Its elasticity with respect to real GDP is 0.28 (s.e. 0.26), whereas the measure for all self-employed is slightly countercyclical. The first two columns of Table 6 repeat Columns 1 and 3 from Table 5, while Table 6, Column 3 measures productivity by earnings per hour for those not incorporated. The wedge becomes less countercyclical, with an elasticity with respect to real GDP of -1.57 (s.e. 0.24). Nevertheless, the self-employed labor wedge remains extremely cyclical, and still nearly as cyclical as that estimated for all workers (Column 1).\textsuperscript{21}

\textsuperscript{20}We also examined cyclicalit y of consumer expenditure for the self employed in the Panel Study of Income Dynamics (PSID). The PSID has reasonably broad expenditure measures starting in 1999; but only biannually. In the PSID relative self-employed consumption, relative to all households, dropped by 2.8% from 2007 to 2009, corresponding to the most recent recession. (Unfortunately, biannual observations miss the timing of the 8-month recession during 2001.) But, overall, the PSID numbers suggest similar cyclicalit y of consumption for the self-employed as for all households.

\textsuperscript{21}A wedge constructed solely for those not incorporated is slightly less cyclical than in Column 3 (elasticity of -1.39 with respect to real GDP -1.39).
### Table 6: Cyclicality of the Wedge, All vs. Self-Employed, Alternatives

<table>
<thead>
<tr>
<th>Elasticity wrt</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real GDP</td>
<td>-1.87 (0.10)</td>
<td>-1.97 (0.25)</td>
<td>-1.57 (0.24)</td>
<td>-1.64 (0.32)</td>
</tr>
<tr>
<td>Total Hours</td>
<td>-1.20 (0.05)</td>
<td>-1.29 (0.16)</td>
<td>-1.03 (0.20)</td>
<td>-1.03 (0.19)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hours</th>
<th>All workers</th>
<th>SE</th>
<th>SE</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPN Agg. y/n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPS weighting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPS weights</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ Ind. shares</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Sample is based on matched-March CPS self-employed outside government and agriculture. CPS observations are weighted. Each cell represents a separate regression. Regressions have 24 annual observations, 1987-1993 and 1996-2012. Newey-West standard errors are in parentheses. Hours are weekly. NIPA PCE consumption.

A second robustness exercise considers that self-employed workers are distributed differently across industries than are wage earners. For instance: Self-employment is about twice as frequent among workers in construction, a highly cyclical industry, than overall; it is considerably less common in durable manufacturing, which is also highly cyclical. We constructed a self-employed labor wedge reweighting observations by industry so that the weighted shares of self-employed workers by industry mimics that for all workers. (We do this for a breakdown of 12 major industries.) For example, if self-employment is twice as frequent in construction, then self-employed workers in construction are down-weighted by a factor of one-half. The results are given in Table 6, Column 4. The cyclicality of the self-employed wedge is modestly reduced. The elasticity is now -1.64 (s.e. 0.32) with respect to real GDP. Again, however, it remains
extremely cyclical, nearly as cyclical as that calculated for all workers.\textsuperscript{22}

Thus we conclude that self-employed workers exhibit a highly countercyclical labor wedge. Depending on specification choices, it is either as cyclical as the wedge calculated for all workers or nearly as cyclical. Because this wedge is presumably not driven by wage or other labor market distortions, it is evident of a highly countercyclical product market wedge. By extension, we find it suggestive of a countercyclical product market wedge for the overall economy.

4. Intermediate Inputs

The conventional way to estimate the Product Market Wedge (PMW) is based on the cyclicity of labor's share of income, e.g., Karabarbounis (2014a). But, in principle, any input with a well measured marginal product and marginal price can be used to infer marginal cost and therefore price markups. Here we investigate the cyclicity of spending on intermediate inputs – materials, energy, and services – relative to gross output.

Intermediate inputs are promising for several reasons. First, intermediates are used by all industries, so evidence on them goes beyond just manufacturing or the self-employed. Second, adjustment costs for intermediates are believed to be low relative to adjustment costs for capital or even labor. See, for example, Basu (1995) or Levinsohn and Petrin (2003). Third, the wedge calculations based on cyclicity in a factor's share of output assumes that factor has no overhead component. This assumption seems much more defensible for intermediates than for labor.

A key question is whether intermediate prices reflect the marginal cost of intermediate inputs. Long term relationships between firms and their suppliers could raise the same implicit contracting issues that arise with labor. Still, intermediates offer an independent piece of evidence \textit{vis a vis} labor. And, as with labor, one would

\textsuperscript{22}As discussed in Section 2, if the calibration is misspecified by ignoring countercyclical home production, this can impart some cyclicity to the labor wedge (Karabarbounis, 2014b). Could this misspecification fall disproportionately on the self-employed? This would require that the self-employed exhibit less comparative advantage in the market versus home production. While certainly possible, we do not judge this true \textit{a priori}. The self-employed exhibit higher average market earnings. This would seem to require that the self-employed are even more able at home production so as to overwhelm their market advantage.
expect price smoothing relative to true input costs to impart a procyclical bias to the estimated PMW.

4.1. Technology for Gross Output

We assume a C.E.S. production function for gross output in an industry:

$$y = \left[ \theta m^{1-\frac{1}{\varepsilon}} + (1 - \theta) \left[ z_v \left( (1 - \alpha)k^{1-\frac{1}{\omega}} + \alpha(z_n n)^{1-\frac{1}{\omega}} \right) \frac{1}{\varepsilon - 1} \right]^{1-\frac{1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon - 1}}$$

where $y$ denotes gross output, $m$ intermediate inputs, $k$ capital, and $n$ labor. Technology shocks can be specific to value added ($z_v$) or labor ($z_n$). The elasticity of substitution between intermediates and value added is $\varepsilon$ and between capital and labor (within value added) is $\omega$.

This technology implies the marginal product of output with respect to intermediate inputs is

$$\frac{\partial y}{\partial m} = \theta \left( \frac{y}{m} \right)^{\frac{1}{\varepsilon}}.$$ 

Based on this marginal product, we can estimate the PMW as

$$\mu^p = \frac{p}{p_m \frac{\partial y}{\partial m}} = \frac{p \theta \left( \frac{y}{m} \right)^{\frac{1}{\varepsilon}}}{p_m}.$$ 

In the special case of Cobb-Douglas aggregation of intermediates and value added ($\varepsilon = 1$), the PMW is the inverse of intermediates’ share:

$$\mu^p = \frac{\theta \frac{p y}{p_m m}}{p_m m}.$$ 

A higher price-cost markup boosts gross output relative to spending on intermediates in an industry. This is analogous to using inverse labor’s share to measure price markup movements. A countercyclical markup would show up as a procyclical intermediate inputs share.
4.2. Evidence on the Cyclicality of Intermediate Inputs

We use the Multifactor Productivity Database from the U.S. Bureau of Labor Statistics on industry gross output and KLEMS inputs (capital, labor, energy, materials and services). It contains annual data from 1987–2012 and covers 60 industries (18 in manufacturing).23

Figure 4 plots the weighted-average industry intermediate share against GDP, where both variables are in logs and HP-filtered. As shown, spending on intermediates relative to gross output is highly procyclical. This is also true if we define the cycle in terms of hours worked.

We next run regressions of the inverse intermediate share on the cycle. The specification is

\[
\log \left( \frac{p_{it} y_{it}}{p_{mit} m_{it}} \right) = \alpha_i + \beta p \log(cyc_t) + \epsilon_{it}
\]

\footnote{The Appendix provides more details. KLEMS intermediate inputs come from BEA annual input-output accounts. These reflect purchases during the year minus inventory accumulation. (For details, see www.bea.gov/papers/pdf/IOmanual_092906.pdf.)}
where \( cyc_t \) is either real GDP or hours worked, and all variables are HP filtered. The industry fixed effects \( (\alpha_i) \) should take out changes in the aggregate share due to shifting industry composition over the cycle. We weight industries by the average share of their value added in all industry value added from 1987–2012. Standard errors are clustered by year.

Table 7 presents the results. The inverse intermediate share, a proxy for the price-cost markup, is systematically countercyclical. This is true for both measures of the cycle, for all industries together, and separately for manufacturing and non-manufacturing industries. It is also true if we weight industry-years by Tornqvist value added shares rather than industry shares over the entire sample (not reported in the table).  

### Table 7: Cyclicality of Inverse Intermediate Share

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Industries</td>
<td>-0.94 (0.24)</td>
<td>-0.59 (0.15)</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.95 (0.32)</td>
<td>-0.65 (0.20)</td>
</tr>
<tr>
<td>Non-Manufacturing</td>
<td>-0.94 (0.24)</td>
<td>-0.57 (0.16)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Annual data from 1987-2012 for 60 industries (1560 industry-year observations), 18 manufacturing and 42 non-manufacturing. All variables in logs and HP filtered. Regressions include industry fixed effects and use industry average value added shares as weights. Standard errors are clustered by year.

What share of the total wedge might be accounted for by the PMW? To answer this, we construct an industry-specific total wedge that is consistent with the gross-output production function we consider. We replace aggregate labor productivity \( \left( \frac{v}{n} \right) \) with nominal gross output per hour in each BLS industry (relative to the consumption

\[24\] For manufacturing, we can break intermediate inputs into materials, energy, and services. The inverse shares for materials and energy are both countercyclical, and significantly so. The inverse share of spending on services, in contrast, is acyclical. Perhaps services are contracted less in spot markets than materials, or especially, energy.
deflator). We also consider preferences that allow for an industry-specific marginal rate of substitution. The industry-\(i\) (intensive-margin) total wedge is thus

\[
\ln \left( \mu_{int}^i \right) = \ln \left( \frac{p_i m_{i}^{int}}{p_m r_{s}^{int}} \right) = \ln \left( \frac{p_i v_i}{p m_{s}^{int}} \right) + \ln \left( \frac{y_i}{v_i} \right) - \frac{1}{\eta} \ln \left( \frac{h_i}{h} \right) + \ln \left( \frac{m_{i}^{int}}{m r_{s}^{int}} \right),
\]

which differs from the aggregate labor wedge in three possible ways. Value-added per hour could be more or less cyclical in the BLS industries than for the average industry (the first term on the right-side). The cyclicality of gross output may differ from value added (the second term). Finally, hours worked per worker, and thus the \(m r_{s}\), could be more (or less) cyclical for the BLS industries (the third term).

Table 8 presents the cyclical elasticities. The all-industry wedge has a smaller elasticity (-0.89 wrt GDP) than the aggregate IMW (-1.91) from Section 2. Why? Gross output is more procyclical than value-added (cyclical elasticity wrt GDP of 0.49) and nominal value-added labor productivity is more procyclical in our BLS industries than in the aggregate (cyclical elasticity wrt GDP of 0.33). Workweeks – which are industry-specific and thus affect the all-industry, manufacturing and non-manufacturing wedges in different ways – account for the remainder.

Comparing Tables 7 and 8, the intermediates-based PMW accounts for essentially all of the cyclical labor wedge. Figure 5 provides visual corroboration by plotting the weighted-average industry (inverse) intermediate share against the total wedge.

It is often argued that is tough to substitute between intermediates and value added. Bruno (1984) and Rotemberg and Woodford (1996) estimate elasticities of 0.45 and 0.69, respectively, for U.S. manufacturing. Oberfield and Raval (2014) obtain

\footnote{Preferences are \(E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{1}{1-\sigma} - \sum_{i} \left[ \nu \left( \frac{h_{it}^{1+1/\eta} + \psi}{h_{it}} \right) \epsilon_{it} \right] \right\} \) and thus \(\ln \left( \frac{m r_{s}^{int}}{m r_{s}^{int}} \right) = \frac{1}{\eta} \ln \left( \frac{h_{it}}{h_{it}} \right) + \ln \left( \frac{e_{it}}{e_{it}} \right) \). In the Appendix we consider alternative preferences, the extensive margin, and \(\varepsilon \neq 1\).}

\footnote{Because some of our industries only have workweek data starting in 1990, we use the aggregate average workweek from 1987 through 2012, which had an elasticity with respect to GDP of 0.32 (s.e. 0.03), adjusted by the relative elasticity of industry-specific workweeks to the aggregate from 1990 through 2012. These elasticities are 0.32 (s.e. 0.03) for the aggregate, 0.22 (s.e. 0.04) for all KLEMS industries, 0.41 (0.08) for manufacturing, and 0.16 (0.04) for non-manufacturing. Appendix Table A1 reports the cyclicality of the total wedge using a common workweek for all industries (i.e., omitting the industry-specific adjustments).}

\footnote{The total wedge in Figure 5 is constructed using the aggregate average workweek. Using industry-specific workweeks produces similar plots, just with three fewer years.}
Table 8: Cyclicality of Intensive-Margin Total Wedge

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Industries</td>
<td>-0.89 (0.26)</td>
<td>-0.59 (0.13)</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.72 (0.39)</td>
<td>-0.39 (0.20)</td>
</tr>
<tr>
<td>Non-Manufacturing</td>
<td>-0.93 (0.24)</td>
<td>-0.65 (0.12)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Annual data from 1987-2012 for 60 industries (1560 industry-year observations), 18 manufacturing and 42 non-manufacturing. All variables in logs and HP filtered. Regressions include industry fixed effects and use industry average value added shares as weights. Standard errors are clustered by year.

estimates ranging from 0.63 to 0.90 by looking across regions in U.S. manufacturing. Atalay (2014) estimates even smaller elasticities (below 0.1). A smaller elasticity makes the PMW based on intermediates more countercyclical. Because firms are shifting toward intermediates in booms, the marginal product of intermediates will fall faster if substitutability is more limited, making marginal cost more procyclical. Thus the price-cost markups implied by intermediate inputs becomes more countercyclical.\(^{28}\)

5. Discussion and Relationship to Literature

How does our work relate to other attempts at measuring the cyclicality of price markups? The challenge is capturing cyclicality in marginal costs. Researchers must make assumptions about firms’ production functions in order to infer marginal cost based on quantities and prices of inputs and output. Assuming firms minimize costs, marginal cost is equated across input margins; so one can consider the cost of marginally increasing output via any input. Many studies have focused on labor – e.g., Bils (1987), Rotemberg and Woodford (1999), and Nekarda and Ramey (2013). Labor’s share of income – i.e., the average price of labor divided by its average product – often

\(^{28}\)A smaller \(\varepsilon\) simultaneously makes the total wedge less countercyclical because firms shift away from value-added (labor) in booms.
serves as the baseline measure of marginal cost, with corrections made to address concerns that average prices and products may not equal marginal ones.

We argued (Section 2) that using wage data to infer the marginal price of labor may be especially fraught with difficulty, and subsequently eschewed wage data altogether. We did not, however, entirely avoid the labor input margin. Our analysis of the self-employed focused precisely on their hours worked. We observed that the marginal revenue of a self-employed person’s labor should equal that person’s marginal rate of substitution (MRS) between consumption and leisure. We then measured the MRS for the self-employed using data on their consumption and hours worked.

Our second approach focuses instead on intermediate inputs. Related, Rotemberg and Woodford (1999) point out that one could use intermediate inputs to infer markups. Basu (1995) finds that quantities of intermediate inputs rise relative to real output in expansions, but he does not explore relative price movements. On the other hand, the work of Murphy, Shleifer, and Vishny (1989), which documents relative price movements for broad categories, suggests the relative price of intermediates is likely to be procyclical. By looking at the intermediate share of income, we combine
quantities and prices to obtain a measure of marginal cost.\textsuperscript{29}

Other researchers have also used approaches that do not require wage data. We briefly discuss these approaches, noting their key assumptions and their conclusions on the cyclicality of markups.

Galeotti and Schiantarelli (1998) measure marginal cost using capital inputs. This requires, like our approach, making assumptions on the production function to infer a marginal product (of capital, in this case). In addition, they must also take a stand on the stochastic discount factor, because the marginal price (i.e., rental rate) of capital is not directly observable but depends on both today’s acquisition price of capital and tomorrow’s expected discounted price.\textsuperscript{30} They find evidence of countercyclical markups: specifically, markups fall when the level of output is high and its expected growth is low.

Hall (2014) considers a simple model of advertising, in which increased advertising shifts the (static) demand curve faced by a firm. If the markup is high, then the firm will want to advertise more. Indeed, the model implies that the ratio of advertising expenditures to revenue is proportional to the markup. And, since the advertising expenditure share is acyclical, this model suggests markups are also acyclical.

We note, however, that this conclusion is not robust to reasonable model alterations. For example, we show, in the appendix, that if advertising affects consumers’ reservation price rather than shifting quantity demanded, then changes to the price elasticity of demand have no effect on the advertising expenditure share, but do cause markup changes. A second alteration that breaks the tight contemporaneous link between advertising expenditures and markups posits that advertising also affect future demand (e.g., Bagwell, 2007). Some evidence that advertising is, at least partially, an investment is provided by Campella, Graham, and Harvey (2010), who

\textsuperscript{29}We also consider a more recent sample and more disaggregated industries than these prior studies. Kim (2015) estimates markup fluctuations for disaggregated manufacturing for years 1958 to 2009 based on spending just on energy intermediates. Energy is a fairly small component of our intermediates; for 2010, it constituted only 4.4% of total intermediate spending in manufacturing, 6.2% in other industries. For his sample, Kim estimates that markups increase in response to financial shocks, but otherwise are procyclical. If we similarly consider only energy spending in manufacturing, for our 1987-2012 sample period, we find energy’s share of gross output is highly procyclical and estimate a highly countercyclical price markup across a wide range of elasticities of substitution.

\textsuperscript{30}They also consider adjustment costs; doing so requires a functional form assumption.
found that planned marketing expenditures fell much more for financially constrained firms than unconstrained firms during the Great Recession.\footnote{Hall does consider a dynamic model, with an annual depreciation rate of 60 percent for advertising’s impact. But, because he maintains a constant discount factor, there is no channel from high discounting to reduced advertising investment during recessions.}

Another approach, pursued by Bils and Kahn (2000) and Kryvtsov and Midrigan (2012), exploits the tight theoretical relationship between markups and (finished) inventories. Consider a firm’s decision to add a unit of inventories. The expected (net) benefit has two components: if the additional unit is sold in the current period, the benefit is the markup of price over marginal cost; if not, it is the expected discounted ratio of future to current (real) marginal cost (i.e., the firm will not have to acquire the inventory in the future). If the current markup declines, holding all else equal, both components are lower – as a lower markup implies a higher real marginal cost – and the firm will reduce its inventories relative to sales.

Note, however, that empirically all else might not be equal. In particular, the appropriate discount rate may vary over the cycle, as may the way in which inventories affect sales. So one must account for both in order to infer markups from inventory data. Importantly, both studies assume \( \frac{\partial \text{sales}}{\partial \text{inventories}} \) is a time-invariant function of the sales-inventory ratio.\footnote{Bils and Kahn (2000) assume the elasticity of sales to inventories is constant, while in Kryvtsov and Midrigan’s (2012) model, which features demand uncertainty, the probability of a stock-out is a constant function of the ratio of inventories to expected demand. The authors then derive a relationship between the discount factor, sales-to-inventory ratio and markup. If, say, the elasticity of sales to inventories varied over the cycle or there were scale effects of holding finished inventories, then an additional time-varying variable would enter that relationship, and markups could no longer be inferred from observable measures of the discount factor and sales-to-inventory ratio.} Given their assumptions and a highly procyclical sales-to-inventory ratio, both studies conclude that markups are countercyclical.

Relatedly, in the appendix we consider work-in-process (WIP) inventories and again infer countercyclical markups. The intuition for the relationship between inventories and markups is similar to that described above; if markups are high (i.e., real marginal costs low) relative to the future, a firm should shift production from tomorrow to today and increase its stock of WIP inventories. The WIP framework is somewhat simpler than that for finished goods: only the relative (inter-temporal) markup appears rather than both the relative markup and the level of the markup. Also, one does not need to
take a stand on how inventories affect sales, but instead on how WIP inventories enter the production function.

In summary, these non-wage approaches to measuring price markups yield results broadly consistent with our own: namely, countercyclical markups. However, they all involve dynamics, requiring one to measure any adjustment costs and the stochastic discount factor. Our self-employed and intermediates approaches, on the other hand, only require static measurements.

6. Conclusion

Hours worked fall more in recessions than can be explained by optimal changes in labor supply in response to real labor productivity. This “labor wedge” could reflect frictions in the labor market (e.g., sticky wages and matching problems), frictions in the product market (e.g., sticky prices), or some combination.

Research has increasingly focused on problems in labor markets, in particular for firms hiring workers. Using average hourly earnings, the labor wedge seems to arise between the cost of labor to firms and the value of jobs to workers. But this inference could be mistaken if the true cost of labor to firms is more cyclical than average hourly earnings. If labor's price is measured by the wages of new hires or a user cost of labor, instead of by average hourly earnings, the labor wedge arises as much between the cost of labor and real labor productivity.

To bring new evidence to bear on this debate, we estimate the product market component of the labor wedge without relying on workers’ wages. First, we look at the self-employed. The labor wedge appears nearly as cyclical for the self-employed as for wage earners, even though sticky wages and matching frictions should not be barriers to the self-employed working more hours. The hours of the self-employed appear to fall in recessions in no small part because of difficulty, or reluctance, in selling their output (for example due to sticky prices). The other primary evidence we present is for intermediate inputs. We find that output prices rise relative to the marginal cost of producing by buying more intermediates in recessions. Again, this suggests that firms face difficulty converting production into revenue in recessions.
We estimate that at least three-quarters of the labor wedge's cyclical variation reflects product market distortions as opposed to labor market distortions. While labor market distortions matter, they are less important than has been inferred using average hourly earnings. Our evidence is consistent with a price of labor that is at least as cyclical as the new hire wage.

Our evidence does not determine the exact nature of these product market distortions, which is critical for informing stabilization policy. Our results do suggest, however, that in recessions firms have trouble selling their output — as if their shadow value of output is low relative to its market price. Why would firms’ shadow value of output be pushed down in recessions? One explanation is price stickiness that constrains production from translating into added sales. Output's shadow value also falls relative to price in models of countercyclical desired markups. (See Rotemberg and Woodford, 1999, for a review.) And in any setting where producing at the margin has an investment component (e.g., the customer base model in Gilchrist et al, 2014). Our evidence is also consistent with models where expanding production puts firms in a riskier position, and risk (or risk avoidance) heightens during recessions (e.g., Arellano et al, 2012).
APPENDIX

Resurrecting the Role of the Product Market Wedge in Recessions

For Online Publication

Mark Bils
University of Rochester and NBER

Peter J. Klenow
Stanford University and NBER

Benjamin A. Malin
Federal Reserve Bank of Minneapolis

September 3, 2015
This appendix provides a detailed description of our data, a more thorough explanation of some calculations, and some robustness results.

**A1. Representative-Agent Wedge**

Variables include:

- $y_t$: (Real) Output per hour; BLS, Labor Productivity & Costs, Business Sector.

- $c_t$: (Real) Nondurables and services consumption per adult equivalent; NIPA consumption data, adjusted for indirect taxes following Prescott (2004) and McDaniel (2007). Adult-equivalent population = Population $\geq 16 + 0.5 \cdot (\text{Population} \leq 15)$.

- $n_t$: Hours worked per capita; Hours worked from BLS (LPC, Business Sector), and population is Civilian Pop 16+.

- $\tau_t = (\tau^c_t + \tau^n_t)/(1+\tau^c_t)$, where $\tau^c_t$ is the average tax rate on consumption, following McDaniel (2007), and $\tau^n_t$ is the average marginal labor tax rate, using NBER TaxSim to extend Barro-Redlick (2011) through 2012.

**A2. Extensive- and Intensive-Margin Wedges**

To construct the IMW and EMW, some variables (e.g., $y_t/n_t$ and $c_t$) are the same as used for the RAW. Additional variables include:

- $h_t$: Average weekly hours worked (per worker); BLS, LPC, Business Sector.

- $v_t$: Vacancies (per capita); Pre-1995 is help-wanted index, and post-1995 is Barnichon’s (2010) spliced series of help-wanted and JOLTS. Population is 16+.

- $m_t$: Matches (per capita); Post-1994 from Fallick-Fleischman (2004), and pre-1994 is backcast using data on unemployment and vacancies, following Blanchard and Diamond (1989).
All variables are seasonally adjusted.

The calibration is described in the text with the exception of $\psi$, the fixed (utility) cost of employment. One can derive an expression for $\psi$ by combining the steady-state optimality conditions for the extensive and intensive margins and assuming the EMW and IMW are the same in steady state. The result is $\psi \equiv \frac{h^{1+1/\eta}}{\eta+1} \left[ 1 - (\eta + 1) [1 - \beta(1 - \delta)] \left[ \frac{\sigma w}{\phi m} + \gamma \right] \right].$

The EMW includes expectational terms in $S_t,$ e.g., $E_t \left\{ \frac{w'(c_{t+1}) w_{t+1} w_{t+1}}{w'(c_t) w_t w_t} \right\}.$ We construct these using 3-variable, 4-lag VARs consisting of real GDP growth, aggregate (log) hours worked, and the respective expectational term. We estimate the VAR using data over the entire sample period, and then use the estimated coefficients to construct time series of the expectational terms.

Finally, we constructed the EMW using alternative data. We assumed $\frac{\beta w'(c_{t+1})}{w'(c_t)} = \frac{1}{1+r_{t+1}}$ and measured the (ex post) real interest rate, $r_{t+1},$ as the 3-month T-bill rate less (realized) core PCE inflation at $t + 1$. Our results change little. The cyclical elasticity of the EMW with respect to GDP was -1.89 (s.e. 0.28) and with respect to aggregate hours was -1.54 (0.15).

**A3. Aggregate Wedge Decomposition**

The decomposition requires wage measures. For our baseline (labeled AHE), we assume $\frac{w_{t} n_t}{p_{t} y_t}$ is the labor share of income as measured in the BLS’s LPC Business Sector. Because we also have a series for labor productivity $\frac{w_{t}}{n_{t}},$ we can back out the average real wage in the economy.

Kudlyak (2014) estimated the semi-elasticities of average hourly earnings, new hire wages, and the user cost of labor, respectively, to the unemployment rate. We use these estimated elasticities, along with the time series of unemployment and our (baseline) average wage measure, to construct time series for new hire wages and the user cost of labor.
A4. Self-Employed

Figures A1–A4 display time series of the variables that underlie our estimates of the cyclicality of the all-worker and self-employed wedges, respectively. Figure A1 displays HP-filtered (log) indices for hours per week for both the self-employed and wage-earners, while Figure A2 presents the same comparison for annual hours. Figure A3 presents HP-filtered aggregate labor productivity, self-employed earnings per hour, and earnings per hour for the unincorporated self-employed. We construct the time series in these three figures using data from the March CPS, as described in Section 3 of the paper.

Figure A4 presents our consumption series for the self-employed together with aggregate consumption, both HP-filtered. To construct consumption for the self-employed, we use the Consumer Expenditure Surveys (CE) from 1987 through 2012 to get a quarterly series for the growth rate of consumption of self-employed workers relative to that for a representative sample of households in the CE. The relative growth rate, in turn, is integrated to obtain a series for relative consumption of the self-employed, indexed to the beginning of 1987. We add this relative estimate to
Figure A2: Annual Hours: Self-employed vs. Wage-earners

Figure A3: Alternative Productivity Measures
NIPA aggregate consumption to arrive at an estimate of the cyclicality of consumption for the self-employed. The following paragraphs describe the construction of the quarterly growth rates of consumption for self-employed workers.

The CE has been an ongoing quarterly survey since 1980, with about 5000 households interviewed each quarter. Households are surveyed about their detailed expenditures for the previous three months. Each household is surveyed for up to four consecutive quarters, allowing for construction of up to three observations on quarterly growth. We focus on expenditures on non-durables and services, which we construct by aggregating those individual categories that are clearly not durables by NIPA standards. We include expenditures on housing: for renters this is captured by household rent; for home-owners it reflects the owner’s estimate of its rental value (rental equivalence). The categories we can classify as nondurables and services constitute about two thirds of household expenditures. We deflate these expenditures by the GDP deflator for nondurables and services. Individual growth rates across any two quarters are calculated by the midpoint formula to reduce the impact of extreme values.

During each of the first and fourth quarterly interviews on expenditures, households are surveyed about the work experience of its household members during
the past 12 months. We focus on the work history in the latter survey, as the work history over the prior 12 months conforms to the time-frame for reported expenditures. (For a small number of households, we fill in for missing employment information from responses collected in earlier quarters.) We create a sample of workers from the CE households, including all members that meet our sample requirements. These requirements are chosen to mimic our treatment of the CPS data: (i) Individuals must be between ages 20-70; (ii) They must report working at least 10 weeks during the year, at a workweek of 10 hours or more when working; (iii) We exclude workers in the top or bottom 9.6 percent of the earnings distribution and the top 1.2 percent of hours per week. These last exclusions are chosen to match those we made on the CPS data, dictated by its top coding of earnings and hours. We make two other sample restrictions in order to measure quarterly growth rates of household consumption. We exclude households in the top and bottom one percent of expenditures in any quarter in order to eliminate top-coded expenditures and outliers. We exclude households that exhibited a change in household size across the quarters that are the basis for the growth rate. In all calculations we employ the CE sampling weight that is designed to make the sample representative of the U.S. civilian non-institutionalized population.

We classify workers as self-employed, as opposed to wage-earners, if they report that the job for which they received most earnings was self-employment and, in fact, at least 95 percent of their reported earnings over the past 12 months is from (non-farming) self-employment. This conforms well to our definition in Section 3 based on CPS data. We do not observe consumption at the individual level (e.g., for a self-employed member versus a wage-earning member). Thus we have to make the simplifying assumption that households equate consumption across members. For example, if a household has one self-employed worker and one wage earner, then that household contributes two members to our overall sample, and one member to our self-employed sample. But the growth rate in consumption in any quarter will be the same for both members of that household. We have 11,849 quarterly observations on consumption growth that apply for self-employed workers, which equals 115 per quarter on average.
A5. Intermediates

We first derive an industry-level, intensive-margin total wedge using more general technology and preferences than we used for our baseline results. We then derive the industry-level, extensive-margin total wedge.

The gross-output production function implies a marginal product of labor on the intensive margin of

$$m_{\text{int}}^{\text{int}} = \alpha(1-\theta) \left( \frac{y_{it}}{v_{it}} \right)^{\frac{\epsilon}{1-\theta}} \left( \frac{v_{it}}{n_{it}} \right)^{\frac{\omega}{1-\theta}} (z_{v, it}^e z_{n, it}^e)^{\frac{\omega-1}{\omega}} e_{it}.$$ 

For our baseline, $\epsilon = \omega = 1$, this simplifies to $m_{\text{int}}^{\text{int}} = \alpha(1-\theta) \frac{y_{it}}{n_{it}} e_{it}$.

Our baseline used the following preferences:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{c_{i, t}^{1-1/\sigma}}{1-1/\sigma} - \nu \sum_i \left[ \left( \frac{h_{i, t}^{1+1/\eta}}{1+1/\eta} \right) + \psi \right] e_{it} \right\},$$

so the marginal rate of substitution of consumption for an extra hour per worker in industry $i$ is $m_{\text{int}}^{\text{int}} = \nu h_{i, t}^{1/\eta} c_{i, t}^{1/\sigma} e_{it}$.

Thus, our baseline industry-$i$ (intensive-margin) labor wedge is (up to an additive constant)

$$\ln(\mu_{i, \text{int}}) = \ln \left( \frac{p_{\text{int}}^{\text{int}}}{p m_{\text{int}}^{\text{int}}} \right) = \ln \left( \frac{p_{i, y_{i}}}{p n_{i}} \right) - \left[ \frac{1}{\sigma} \ln(c) + \frac{1}{\eta} \ln(h_{i}) \right]$$

$$= \ln \left( \frac{p_{i, n_{i}}}{p_{n}} \right) + \ln \left( \frac{y_{i}}{v_{i}} \right) - \frac{1}{\eta} \ln \left( \frac{h_{i}}{h} \right) + \ln \left( \frac{m_{\text{int}}^{\text{int}}}{m_{\text{int}}^{\text{int}}} \right), \quad (1)$$

where $m_{\text{int}}^{\text{int}} \equiv \alpha(1-\theta) \frac{y_{it}}{n_{it}} e_{it}$ and $m_{\text{int}}^{\text{int}} \equiv \nu h_{i, t}^{1/\eta} c_{i, t}^{1/\sigma} e_{it}$ are based on aggregate data. For $\epsilon, \omega \neq 1$, it’s straightforward to see how the total wedge would be altered. Specifically, for $\epsilon < 1$, the total wedge becomes less countercyclical if gross output is more procyclical than value-added.

Note that our preferences assume separability across labor supply in different industries. This seems reasonable because workweeks are person-specific. But, we
Table A1: Cyclicality of (Common-MRS) Intensive-Margin Total Wedge

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Industries</td>
<td>-1.10 (0.26)</td>
<td>-0.72 (0.13)</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.55 (0.39)</td>
<td>-0.35 (0.20)</td>
</tr>
<tr>
<td>Non-Manufacturing</td>
<td>-1.25 (0.24)</td>
<td>-0.82 (0.12)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Annual data from 1987 - 2012 for 60 industries (1560 industry-year observations), 18 manufacturing and 42 non-manufacturing. All variables in logs and HP filtered. Regressions include industry fixed effects and use industry average value added shares as weights. Standard errors are clustered by year.

could consider alternative preferences, say

\[
\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{c_{it}^{1-1/\sigma}}{1 - 1/\sigma} - \nu \left( \frac{h_{it}^{1+1/\eta}}{1 + 1/\eta} + \psi \right) e_{it} \right\},
\]

where \( h_t \equiv \frac{\sum_i h_{it} e_{it}}{\sum_i e_{it}} \) and \( e_t \equiv \sum_i e_{it} \). In this case, \( mrs_{int} = \nu h_{it}^{1/\eta} c_{it}^{1/\sigma} e_{it} \). The industry-\( i \) labor wedge is thus (for baseline technology, \( \varepsilon = \omega = 1 \))

\[
\ln(\mu_{it}^{int}) = \ln \left( \frac{p_{n,lt}}{p_{int}} \right) + \ln \left( \frac{y_{it}}{v_{it}} \right) + \ln \left( \frac{m_{int}}{m_{ns}^{int}} \right).
\]

Under these preferences, labor supply is perfectly substitutable across industries and only aggregate labor supply, \( e \) and \( h \), matters. The labor wedge no longer needs an adjustment for industry-specific workweeks. Table A1 shows how replacing industry-specific workweeks with aggregate average weekly hours worked affects our results (compare to Table 8 in the paper). Manufacturing industries exhibit more procyclical workweeks, and the labor wedge is thus less countercyclical for manufacturing. On the other hand, it’s more countercyclical for non-manufacturing and all industries. (For the latter, recall that the 60 industries covered by KLEMS are not necessarily representative of the entire economy.)
Moving to the extensive margin, $mpn_{it}^{\text{ext}} = mpn_{it}^{\text{int}} \frac{h_{it}}{e_{it}}$ and $mrs_{it}^{\text{ext}} = mrs_{it}^{\text{int}} \frac{\Omega_{it}}{h_{it}/\eta}$. The industry-$i$ extensive-margin labor wedge is

$$\ln(\mu_i^{\text{ext}}) = \ln\left(\frac{p_i mpn_{it}^{\text{ext}}}{p mrs_{it}^{\text{ext}}}\right) - S_i = \ln(\mu_i^{\text{int}}) - \ln\left(\frac{\Omega_i}{h_i/\eta}\right) - S_i,$$

or, for our baseline case, is

$$\ln(\mu_i^{\text{ext}}) = \ln\left(\frac{p_i}{p_n} \frac{\ln\left(\frac{v_i}{v_n}\right)}{\ln\left(\frac{\Omega_i}{\Omega}\right)} + \ln\left(\frac{\mu_i^{\text{ext}}}{\mu_i}\right) - (S_i - S). \quad (2)$$

Due to data limitations (i.e., vacancies and matches are not available at the industry-level), we assume $S_{it}$ differs across industry only because of industry-specific workweek movements. That is, $S_{it} = \left[h_i/h_n\right] S_t$. Table A2 displays the cyclicality of the extensive-margin total wedge.

**Table A2: Cyclicality of (Common-MRS) Extensive-Margin Total Wedge**

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Industries</td>
<td>-1.14 (0.52)</td>
<td>-0.86 (0.28)</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.59 (0.66)</td>
<td>-0.49 (0.35)</td>
</tr>
<tr>
<td>Non-Manufacturing</td>
<td>-1.28 (0.49)</td>
<td>-0.96 (0.26)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Annual data from 1987-2012 for 60 industries (1560 industry-year observations), 18 manufacturing and 42 non-manufacturing. All variables in logs and HP filtered. Regressions include industry fixed effects and use industry average value added shares as weights. Standard errors are clustered by year.

To construct the industry-level total wedge and the intermediates-based product market wedge, some variables (e.g., $c_t$) are the same as used earlier in the paper. Additional variables include:

- $p_t$: Price deflator for nondurables and services consumption; Tornqvist index of NIPA implicit price deflators for nondurables consumption and services
consumption.

- $p_{it}, y_{it}, n_{it}, p_{mit} m_{it}$: Respectively, the gross output deflator, real gross output, hours worked, and expenditures on intermediates (tornqvist index of materials, energy and services) by industry from BLS KLEMS.

- $h_{it}$: Average weekly hours worked (per worker); Ratio of hours worked (from BLS KLEMS) to industry-specific employment (calculated with data underlying BLS LPC dataset).

A6. Advertising

Hall (2014) considers a simple theory of advertising (further simplified here), in which a firm’s objective is

$$max_{p,A} (p - mc) \frac{A^\alpha}{p^\epsilon} - \kappa A,$$

where $p$ is the firm’s price, $A$ its advertising volume, $mc$ the marginal cost of production, $\kappa$ the cost of a unit of advertising, and $-\epsilon$ and $\alpha$ are the elasticities of demand with respect to price and advertising. The first-order condition for advertising yields an expression for the ratio of advertising expenditure to revenue:

$$\frac{\kappa A}{pQ} = \alpha \left[ 1 - \frac{1}{p/mc} \right].$$ (3)

Hall’s finding that the advertising expenditure share of revenue is acyclical, combined with eq’n (3), suggests markups are also acyclical.

But this tight link between the advertising expenditure share and markup is fairly tenuous. For example, if we slightly alter how advertising affects demand, say it lowers the reservation price of consumers rather than shifting the demand curve out,\(^1\) then the firm’s objective becomes

$$max_{p,A} (p - mc) \left( \frac{p}{A^\alpha} \right)^{-\epsilon} - \kappa A,$$

\(^1\)Assume a fixed population and that individual $i$’s willingness to pay for a good is given by $x_i = Z A^\alpha \Omega_i$, where $Z$ is an aggregate shifter, $A$ is advertising for the good, and $\Omega_i$ is the individual preference. If $\Omega_i$ is distributed basic Pareto, $f(\Omega_i) = \epsilon \Omega_i^{-(1+\epsilon)}$ for $\Omega_i \geq 1$, then demand for the good is $Z^\epsilon A^{\alpha \epsilon} p^{-\epsilon}$, where $p$ is the price of the good.
and optimal advertising requires
\[
\frac{\kappa A}{pQ} = \alpha \epsilon \left[ 1 - \frac{1}{p/mc} \right].
\]
(4)

In this case, an increase in the price elasticity of demand lowers the price markup – i.e.,
\[
p/mc = \epsilon / (\epsilon - 1)
\]
– but has no affect on the advertising share.

A7. Inventories

Here, we show how data on work-in-process (WIP) inventories can be used to infer a price markup. Following Christiano (1988), we assume a production function that uses WIP inventories as one of its inputs. For a firm in industry \( i \),
\[
y_{it} = g(z_{it}, n_{it}, k_{it})q_{it}^{\varphi_{it}},
\]
where \( y_{it} \) denotes output, \( q_{it} \) is beginning-of-period inventories, and \( z_{it}, n_{it}, \) and \( k_{it} \) are TFP, hours worked and capital, respectively. The elasticity of output with respect to inventories, \( \varphi_{it} \), is allowed to vary both across industry and time. The law of motion for inventories is assumed to be
\[
q_{i,t+1} = (1 - \delta_q)q_{it} + y_{it} - y^f_{it},
\]
where \( \delta_q \) is the depreciation rate of inventories, \( y^f_{it} \geq 0 \) is output of finished goods, and \( q_{i,t+1} \geq 0 \). That is, total output \( y_{it} \) is the sum of gross investment in WIP inventories and finished-good output. The latter includes both final sales and (gross) investment in finished-goods inventories, but it is not necessary to separate these two for our purposes.

An optimizing firm minimizes the expected present discounted cost of producing a given path of finished goods. One perturbation on its cost-minimizing strategy would be to produce an additional unit of output in the form of WIP inventories at time \( t \) and then reduce production just enough at \( t + 1 \) — i.e., by \( (1 - \delta_q + \varphi_{i,t+1} \frac{y_{i,t+1}}{q_{i,t+1}}) \) — to keep
inventories unaffected at \( t + 2 \) forward. At an optimum,

\[
\frac{mc_{it}}{pt} = \mathbb{E}_t \left[ M_{t,t+1} \frac{mc_{i,t+1}}{pt+1} \left( 1 - \delta_q + \varphi_{i,t+1} \frac{y_{i,t+1}}{q_{i,t+1}} \right) \right],
\]

(5)

where \( \frac{mc_{it}}{pt} \) is the (real) marginal cost of production and \( M_{t,t+1} \) its discount factor. In words, the firm equates the marginal cost of output to its marginal benefit, which is reduced future production costs.\(^2\) Because the industry product-market wedge is \( \mu_{it}^p \equiv \frac{pt}{mc_{it}} \), we can write (5) as

\[
\frac{pt/pt^*}{\mu_{it}^p} = \mathbb{E}_t \left[ M_{t,t+1} \frac{pt_{i,t+1}/pt_{t+1}}{\mu_{i,t+1}^p} \left( 1 - \delta_q + \varphi_{i,t+1} \frac{y_{i,t+1}}{q_{i,t+1}} \right) \right].
\]

(6)

We assume the stochastic discount factor is given by \( M_{t,t+1} \equiv \beta \frac{u'(c_{t+1})}{u'(c_t)} \) and that the joint conditional distribution of \( u'(c_{t+1}) / u'(c_t) \), \( \frac{pt_{i,t+1}}{pt_{t+1}} \), \( \frac{pt_{i,t+1}}{pt} \), and \( 1 - \delta_q + \varphi_{i,t+1} \frac{y_{i,t+1}}{q_{i,t+1}} \) is log-normal and homoskedastic.\(^3\) We then take logs of equation (6) and get (up to a constant)

\[
ln(\mu_{it}^p) \approx ln \left( \frac{pt}{pt^*} \right) + \mathbb{E}_t \left\{ -ln \left( \frac{u'(c_{t+1})}{u'(c_t)} \right) + ln \left( \mu_{i,t+1}^p \right) - ln \left( \frac{pt_{i,t+1}}{pt_{t+1}} \right) - ln \left( \frac{\varphi_{i,t+1} y_{i,t+1}}{1 - \delta_q q_{i,t+1}} \right) \right\},
\]

where \( \frac{\varphi_{i,t+1} y_{i,t+1}}{1 - \delta_q q_{i,t+1}} \) \( \approx ln \left( 1 + \frac{\varphi_{i,t+1} y_{i,t+1}}{1 - \delta_q q_{i,t+1}} \right) \). Iterating forward for \( ln(\mu_{i,t+s}^p) \) and using \( u'(c_t) = c_t^{-1/\sigma} \) yields the inventory-based PMW:

\[
ln(\mu_{it}^p) \approx -\frac{1}{\sigma} ln(c_t) + ln \left( \frac{pt}{pt^*} \right) - \mathbb{E}_t \sum_{s=1}^{\infty} \frac{\varphi_{i,t+s} y_{i,t+s}}{1 - \delta_q q_{i,t+s}} + \text{constant terms}.
\]

(7)

The intuition for equation (7) is as follows. Suppose the economy is in a recession in period \( t \), so the log marginal utility of consumption, \( -ln(c_t) / \sigma \), is high. If the firm’s price markup and relative price are not cyclical, then (7) says the path of future

\(^2\)Note that an optimizing firm will always produce to the point that the marginal value of an extra unit of output equals its marginal cost. In a model in which the firm can adjust sales at the margin, the marginal value of output is simply marginal revenue. If the firm cannot adjust sales, the additional unit of output is held as an inventory and valued accordingly. The value of a finished-good inventory is the expected discounted revenue it generates when it is eventually sold. The value of a WIP inventory, on the other hand, is that the firm enters the next period with a larger stock of WIP inventories.

\(^3\)As explained by Campbell (2003), log-normality implies the log of an expectation can be expressed as an expectation of the log plus a variance term. The conditional homoskedasticity means the variance term is not time-varying.
output-to-inventory ratios must be high. That is, the firm should be depleting future WIP inventories in order to push output out the door today and boost consumption.

Alternatively, if the expected path of output-to-inventory ratios is not cyclical, then for equation (7) to hold, the firm's real marginal cost ($mc_{it}/p_t$) must be low in recessions. In turn, either the product-market wedge ($\mu_{it}^p$) is high or the firm's relative price ($p_{it}/p_t$) is low in recessions. That is, if firms do not deplete inventory investment in recessions, one explanation is that product market distortions keep the firm's price high relative to its marginal cost.

To measure the PMW according to equation (7), we turn to NIPA, which provides quarterly and monthly measures of inventories, sales, and sales price deflators by industry. We define industry output as sales plus the change in (total) inventories, and we use quarterly data from 1987-2012 for comparison to previous sections. WIP inventories are available for 22 (roughly 2-digit) manufacturing industries, but the industry classification changed from the SIC to NAICS system in 1997. To bridge that year and create consistent industry definitions, we aggregate some industries, leaving 14 sectors.

To calibrate the parameters in equation (7), we first note that inventory-to-output ratios exhibited significant low-frequency movement over our sample period. We thus let $\varphi_{it}$ vary over time and set $\varphi_{it} = \left[ \frac{1}{\beta} - (1 - \delta_q) \right] \frac{\bar{q}_{it}}{\bar{y}_{it}}$, where $\frac{\bar{q}_{it}}{\bar{y}_{it}}$ is a quadratic trend fitted to the inventory-output ratio. Our quarterly calibration sets $\beta = 0.996$ and $\delta_q = 0.01$. As a result $\varphi_{it}$, which measures the share of output attributable to inventories, is quite low, about 0.2%, on average.

Constructing the inventory-based wedge requires computing, at each point in time, the sum of expected future output-to-inventory ratios. We estimate

---

4Specifically, the Output-to-WIP-Inventory ratio, $\frac{q_{it}}{y_{it}}$, and price deflator for (industry) sales, $p_{it}$, are taken from the NIPA Underlying Detail Tables, Real Inventories and Sales.

5We use a Tornqvist index to construct chain-weighted growth rates of real sales, real inventories, and price deflators for the combined industries. For bridging across the 1996-97 break, we made two assumptions. For inventories, we assume the industry shares of nominal inventories don't change between Dec 96 and Jan 97. (This is feasible since the inventory data is reported for both classifications in 1997, but there is no such overlap for the sales.) For sales, we assume the growth rate in the nominal inventory-to-shipments ratio is the same as that of the real inventory-to-shipments ratio (in Jan 97). The former is constructed using data from the Census M3 survey, which has a consistent NAICS industry classification across 96-97.

6This specification for $\varphi_{it}$ assures equation (6) holds in (detrended) steady-state.
industry-specific, 3-variable, 12-(monthly)-lag VARs consisting of real GDP growth, aggregate (log) hours worked, and the industry-specific output-to-inventory ratio. The latter two variables are quadratically detrended. We estimate the VARs using data over the entire sample period, and then use the estimated coefficients to produce a time series for the expected sum of future output-to-inventory ratios.\footnote{We considered a second approach to calculating the expected sum of future output-to-input ratios, which involved truncating the sum at either 4 or 8 quarters and calculating the (ex-post) realized sum. Because we project the constructed wedge on the time-\(t\) business cycle, using the (ex-post) realized values is valid for our purposes. It does require using a 1-sided hp-filter for the business cycle, so the difference between expected and realized values of the output-to-inventory ratios is orthogonal to the time-\(t\) cycle. This second approach produced results for the wedge that were very similar to the VAR approach.}

Figure A5 plots the weighted-average industry PMW against GDP. As shown, the wedge is quite countercyclical. This is also true if we define the cycle in terms of hours worked. Figure A6 plots the PMW again, but now aggregated to an annual frequency and plotted against the weighted-average manufacturing-industry total wedge constructed in Section 4 of the paper rather than against GDP. The PMW accounts for most of the cyclical variation in the total wedge.
We next run regressions of the industry-level wedge on the cycle

$$\log (\mu_{it}^p) = \alpha_i + \beta^p \log (\text{cyc}_t) + \epsilon_{it},$$

where the weights are the industry’s average share of output and standard errors are clustered by period. Table A3 displays the results at an annual frequency for comparison to the total wedge.\(^8\) The strongly countercyclical PMW (-0.70 elasticity with respect to GDP) accounts for nearly all of the cyclicality in the total wedge (-0.73).

Finally, we have used WIP inventories for our calculations because these align most closely with the theory, which posits a role for inventories in production. Christiano (1988) argues for total inventories (i.e., including materials, WIP, and final goods inventories), noting that labor inputs can be conserved by transporting materials in bulk and holding finished inventories. For robustness, we redo our calculations using total inventories instead of WIP inventories. The results are fairly similar to those reported in Table A3: the cyclical elasticity of the product-market wedge is -0.56 wrt

---

\(^8\)The quarterly elasticities are more precisely estimated: -0.80 (s.e. 0.12) with respect to GDP and -0.33 (0.08) with respect to Hours.
Table A3: Cyclicality of Inventory-based PMW

<table>
<thead>
<tr>
<th></th>
<th>Elasticity wrt GDP</th>
<th>Elasticity wrt Total Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Market Wedge</td>
<td>-0.70 (0.26)</td>
<td>-0.26 (0.15)</td>
</tr>
<tr>
<td>Marginal Utility of Consumption</td>
<td>-1.33 (0.06)</td>
<td>-0.76 (0.08)</td>
</tr>
<tr>
<td>Relative Price</td>
<td>0.83 (0.20)</td>
<td>0.60 (0.11)</td>
</tr>
<tr>
<td>Expected Output/Inventory Path</td>
<td>0.21 (0.05)</td>
<td>0.10 (0.04)</td>
</tr>
</tbody>
</table>

Note: Each entry is from a separate regression. Annual data from 1987-2012 for 14 manufacturing industries (375 industry-years). Variables in logs and HP filtered. Regressions include industry fixed effects and use industry average value added shares as weights. Standard errors clustered by year. See equation (7) for the wedge components.

GDP and -0.22 wrt Hours.⁹

---

⁹Using total inventories enables one to consider industries outside of manufacturing (e.g., wholesale and retail trade).
References


Atalay, Enghin, “How Important are Sectoral Shocks?,” June 2014.


Oberfield, Ezra and Devesh Raval, “Micro Data and Macro Technology,” May 2014.


