Retail inventories and inflation dynamics: The price margin channel*

Neil Mehrotra† Hyunseung Oh‡ Julio Ortiz§

October 8, 2025

Abstract

Using industry-level panel data and plausibly exogenous variation in supply conditions, we estimate the elasticity of retail price margins with respect to inventories along the retailer's optimal pricing curve. We find that this elasticity is negative and statistically significant, implying that lower finished-good inventories lead to higher price margins. We assess the implications of this channel for inflation dynamics within a New Keynesian Phillips curve (NKPC) framework that links inventories to retailers' markup behavior. Incorporating the inventory-sales ratio into the NKPC markedly improves the model's empirical fit and helps account for two notable recent inflation episodes: the missing disinflation of 2009–2011 and the COVID-era surge.

Keywords: Inflation; inventories; supply disruptions; Phillips curve.

JEL Classification Numbers: E31, E32, E22.

*We thank Chris Machol for excellent research assistance and Olivier Coibion, François de Soyres, Giuseppe Fiori, Aaron Flaaen, Nils Gornemann, Oleksiy Kryvtsov, Nitya Pandalai-Nayar, and Daniel Villar for helpful comments and suggestions. The views expressed are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System and the Federal Reserve Bank of Minneapolis, or of any other person associated with the Federal Reserve System. All errors are our own. First draft: August 2025.

†Federal Reserve Bank of Minneapolis. E-mail: neil.mehrotra@mpls.frb.org

‡Federal Reserve Board. E-mail: hyunseung.oh@frb.gov

§Federal Reserve Board. E-mail: julio.l.ortiz@frb.gov

"Retail markups in a number of sectors have seen material increases in what could be described as a price-price spiral, whereby final prices have risen by more than the increases in input prices. The compression of these markups as supply constraints ease, inventories rise, and demand cools could contribute to disinflationary pressures."

—Lael Brainard (January 2023)[1]

1 Introduction

The Phillips curve has been a key pillar of the literature on inflation and the business cycle. Nevertheless, its empirical versions to this date leave a lot of room for improvement in accounting for inflation dynamics. In particular, the COVID-era inflation surge that originated in the goods sector was not well captured by the standard Phillips curve models.

In this paper, we argue that inventories—specifically, retail inventories of finished goods—play a critical role in shaping inflation dynamics through their impact on retail prices.[2] While the standard New Keynesian Phillips curve (NKPC) abstracts from inventories, we show that retail inventory fluctuations offer a key missing link: they help explain why prices rose sharply during periods of supply disruption and low inventory availability.

Our approach begins with a stylized model of a retail firm that jointly chooses prices and inventory stocks in the face of supply disruptions. We show that transitory shocks shift the firm's stocking policy without directly affecting its pricing decision, thereby enabling identification of the optimal pricing curve using instruments for inventory disruptions.

Empirically, we estimate the elasticity of retail price margins with respect to inventories using industry-level panel data spanning 2008q1 to 2025q1. Because pricing and stocking decisions are jointly determined, accurately estimating this elasticity requires an identification strategy. Motivated by our stylized model, we construct two shift-share variables that capture plausibly exogenous transitory supply shocks. We use these shift-share variables to

[1]Speech of Federal Reserve Vice Chair Brainard on January 19, 2023, available at https://www.federalreserve.gov/newsevents/speech/brainard20230119a.htm.

[2]Between 2007 and 2023, retail and wholesale trade margins made up 46.6% of the purchasers' value of consumer goods—nearly half of what consumers pay—underscoring their importance for goods inflation dynamics.

instrument for changes in retailers' inventory-sales ratios. The first instrument reflects material supply shortages among upstream manufacturers from the Census Bureau's Quarterly Survey of Plant Capacity (QSPC). The second instrument reflects global shipping delays based on the Average Congestion Rate (ACR) of Bai et al. (2024). In both cases we regard supply shifts as occurring upstream, and map them to downstream retailers using exposure shares based on a retail industry's supply use from a given manufacturing industry.

Identification with shift-share instruments can be achieved by assuming exogeneity of either shifts or shares (Borusyak et al., 2022; Goldsmith-Pinkham et al., 2020). Our identification rests on exogenous shifts. In our setting, the ideal experiment would be to randomly assign observed transitory supply shifts across manufacturing industries. These shifts would affect downstream retail industries differently given the degree to which retailers rely on each manufacturer for their respective supply of finished goods. In practice, supply shifts are unobservable, so we adopt a quasi-experimental framework by constructing proxies for supply shifters.

The estimated elasticity is negative and economically significant: tighter inventories are systematically associated with higher price margins, even after controlling for cost variables. When instrumenting inventory changes with observed supply disruptions, to identify the elasticity along the retailer's optimal pricing curve, we estimate an even stronger negative relationship, supporting the hypothesis that inventory scarcity contributes directly to retail price increases and providing evidence of state-dependent retail pricing.

To assess the macroeconomic implications of this inventory-pricing channel, we extend the stylized retail firm model, embedding the negative estimated elasticity into a New Keynesian framework. Retailers face both Rotemberg-style nominal rigidities and time-varying markups arising from two sources: a Kimball (1995) demand structure and deep habit formation *a la* Ravn et al. (2006). This setup yields two NKPC representations. The sales-based NKPC highlights how inflation is tied to future marginal costs and marginal sales benefits, but is difficult to estimate directly except in the special case without time-varying markups. The stock-based NKPC, in contrast, relates inflation to observable inventory-sales ratios and expected marginal cost growth, offering a tractable and empirically robust formulation. In particular, the inventory-sales ratio serves as a sufficient statistic for

understanding inflationary pressures arising from the model's endogenous retail markup behavior.

Estimating the stock-based NKPC using macroeconomic data, we find that incorporating the retail inventory-sales ratio markedly improves the model's empirical fit compared to standard NKPC specifications. In particular, the stock-based NKPC helps account for two prominent recent inflation episodes—capturing both the missing disinflation of 2009–2011 and the COVID-era surge. The underlying intuition is that detrended inventory levels were tight in both periods, generating additional inflationary pressures that bring model-implied inflation closer to the observed data. Because changes in the inventory-sales ratio orthogonal to movements in the expected marginal cost growth should reflect movements in retail sector markups within our model, the results highlight the important role that retail price markups may play in driving inflation dynamics.

Related literature. The COVID-era surge in inflation has sparked renewed interest in the role of supply-side disruptions in driving price dynamics. A growing literature documents how production and transportation frictions contributed to rising costs across sectors (Alessandria et al., 2023; Boehm and Pandalai-Nayar, 2022; Comin et al., 2023; Ferrante et al., 2023; Ortiz, 2022). These studies emphasize upstream channels—such as global supply chain bottlenecks, convex production technologies, and sectoral reallocation—as key drivers of inflationary pressures. We build on this line of research by focusing on a downstream margin that has received comparatively less attention: the pricing behavior of retailers facing inventory shortages. While prior work highlights how supply disruptions raise production costs, our contribution is to show that they can also amplify retail price margins when firms are constrained by limited inventory availability.

A paper closely related to our topic is Cavallo and Kryvtsov (2023), which uses online micro-level stockout data to study the inflationary effects of consumer product shortages during the pandemic. We complement their analysis by leveraging industry-level panel data, which offers two distinct advantages. First, our data encompasses a broader set of retailers, including the motor vehicle dealer industry—a key contributor to goods inflation in the early stages of the pandemic. Second, we use retail gross margin data—the difference between the sales price and the acquisition cost—rather than retail prices, and we control

for several industry-level cost measures, such as average hourly earnings. This allows us to analyze retailers' pricing behavior beyond what can be attributed to underlying cost pressures during this period.

The relationship between inventories and inflation has been explored in several studies, particularly in the context of business cycle models. Jung and Yun (2005), Lubik and Teo (2012), and Kryvtsov and Midrigan (2010, 2013) examine how inventory dynamics interact with pricing, often within frameworks that assume constant elasticity of substitution (CES) in demand. These papers typically find either a limited role for inventories in explaining inflation dynamics or a need for additional mechanisms—such as time-varying, countercyclical markups—to reconcile theory with the data, especially in response to monetary policy shocks. Our empirical analysis using industry-level data confirms that retail price margins respond in ways that are inconsistent with CES demand, supporting the need for richer pricing structures. In contrast to the prior literature, our model explicitly incorporates standard features that generate time-varying price markups, including habit formation and Kimball-style demand. We derive a representation of the NKPC in which the inventory-sales ratio serves as a sufficient statistic for retail markups. This approach allows us to empirically assess the role of time-varying markups in inflation dynamics without relying on direct and often noisy measures of price margins.

Our paper also contributes to the empirical Phillips curve literature, which has grappled with the declining explanatory power of traditional NKPC specifications in recent decades (Mavroeidis et al., 2014; Stock and Watson, 2020; Coibion and Gorodnichenko, 2025). A central debate in this literature concerns the inclusion of forward-looking variables that capture the information and beliefs shaping firms' pricing decisions. For example, Coibion and Gorodnichenko (2015) show that replacing professional inflation forecasts with household survey expectations can help resolve the 2009–2011 missing disinflation puzzle, highlighting the value of expectation measures that better reflect firms' information set. Reis (2023) similarly argues for combining multiple indicators of inflation expectations to obtain early and more accurate signals of inflation dynamics. In this spirit, we propose the inventory-sales ratio as a complementary forward-looking variable that reflects retailers' near-term beliefs and strategic pricing behavior. Incorporating this measure directly into

the NKPC significantly improves its empirical fit, allowing it to account for key episodes such as the 2009–2011 missing disinflation and the surge in goods inflation following the COVID-era supply disruptions.

The remainder of the paper is organized as follows. Section 2 presents a stylized retail firm model that motivates our identification strategy for estimating the inventory stock elasticity of the retail price margin. Section 3 implements this strategy using industry-level panel data. Section 4 extends the stylized model and derives a NKPC with inventories that is consistent with the empirically observed negative inventory elasticity of retail markups. Section 5 estimates the inventory-augmented Phillips curve to quantify the macroeconomic importance of retail pricing behavior, as reflected in the inventory-sales ratio. Section 6 concludes.

2 A stylized retail firm model

In this section, we present a stylized model of a retail firm to illustrate how temporary supply disruptions can be used to identify the firm's optimal pricing curve as a function of its inventory stock. The key insight is that while the firm's optimal *stocking* decision depends on both current and expected future costs of acquiring goods, its *pricing* decision depends only on the expected future cost. Therefore, changes in the current cost of acquiring goods—so long as they do not affect expectations of future costs—will shift the firm's stocking curve but leave its pricing curve unchanged. In particular, if supply disruptions are perceived to be transitory relative to other cost shocks, they can serve as valid instruments to trace out the firm's pricing curve. We first show this identification logic in a static one-period model, and then extend the insight to a dynamic two-period setting that makes explicit the role of intertemporal inventory decisions in separating pricing from stocking behavior.

2.1 One-period model

To build intuition about the retail firms' joint decisions over stocking and pricing, we consider a one-period model in which the firm purchases a stock of finished goods and sets a price at which to sell them. The firm simultaneously chooses the stock level and the

selling price, taking as given the demand function and the unit cost of stocking. Demand for the good, denoted s, is given as $s = \boldsymbol{s}(a, p)$ where a is the stock and p is the price.[3] We assume the demand function satisfies $\partial \boldsymbol{s}/\partial a > 0$, $\partial^2 \boldsymbol{s}/\partial a^2 < 0$, and $\partial \boldsymbol{s}/\partial p < 0$. The first two conditions are standard in the inventory literature. For example, in stockout avoidance models, an additional unit of stock raises sales up to the point where stockouts are eliminated; in stock-elastic demand models, sales increase with an additional unit of stock but at a diminishing rate.

After sales, the firm may liquidate unsold inventory, $a - s$, incurring a cost. The total cost of stocking, net of liquidation, is given by:

$$(c + \xi)a - (1 - \gamma)c(a - s),$$

where c is the unit cost of stocking, ξ captures an additional shadow cost faced by the retailer (e.g., supply disruptions not embedded in c), and γ governs the liquidation cost. The parameter ξ is a stand-in for the difficulty in stocking by the retail firm not captured in the unit cost of stocking c.

The firm's profit is given by:

$$\phi = p\boldsymbol{s}(a, p) - (c + \xi)a + (1 - \gamma)c(a - \boldsymbol{s}(a, p)).$$

The firm chooses a and p to maximize this profit.

The first order condition with respect to a (stocking) is:

$$p\frac{\partial \boldsymbol{s}}{\partial a} - (c + \xi) + (1 - \gamma)c\left(1 - \frac{\partial \boldsymbol{s}}{\partial a}\right) = 0,$$

which simplifies to:

$$(p - (1 - \gamma)c)\frac{\partial \boldsymbol{s}}{\partial a} = \gamma c + \xi. \tag{1}$$

Absent both a shadow cost of stocking ($\xi = 0$) and a liquidation cost ($\gamma = 0$), the right-hand side of (1) becomes zero. In that case, as long as $p > (1 - \gamma)c$, the firm would optimally

[3] For simplicity, we normalize aggregate demand to 1.

stock an infinite quantity ($a \to \infty$), driving $\partial \boldsymbol{s}/\partial a \to 0$. Therefore, at least one of ξ or γ must be strictly positive to yield a finite interior solution. The resulting stocking curve determines the optimal level of the stock as a function of the price, conditional on the cost parameters c, ξ, and γ.

The first order condition with respect to price is:

$$s = -(p - (1-\gamma)c)\frac{\partial \boldsymbol{s}}{\partial p}. \tag{2}$$

This pricing equation pins down the optimal price as a function of the level of the stock, taking c and γ as given. For example, suppose the demand function has constant price elasticity:

$$\boldsymbol{s}(a, p) = u(a) \times p^{-\eta},$$

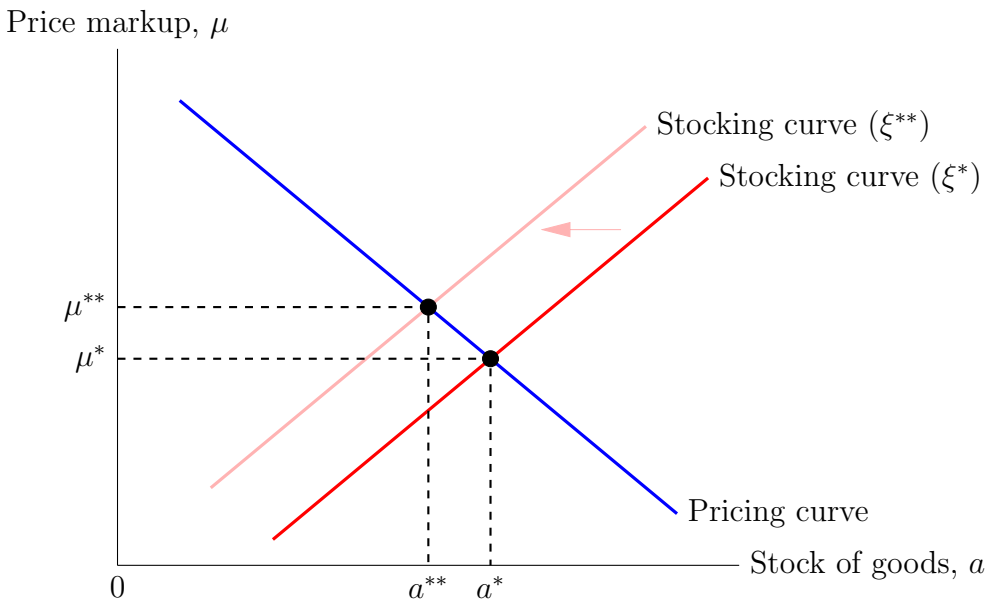
where $u(a)$ is increasing in a and $\eta > 1$. Then the price markup is constant and given by:

$$\mu \equiv \frac{p}{(1-\gamma)c} = \frac{\eta}{\eta - 1},$$

which is the standard result under CES demand. In this case, the pricing curve is flat, and changes in the stock have no effect on the markup.

Taken together, the optimal stocking and pricing curves jointly determine the firm's equilibrium stock and price (or markup). A key feature of this simple model is that the shadow cost ξ appears only in the stocking condition—not in the pricing condition. As a result, shifts in ξ move the stocking curve without affecting the pricing curve, allowing us to trace out the firm's pricing behavior. Figure 1 illustrates this identification strategy. When $\xi = \xi^*$, the firm's equilibrium stock and markup are (a^*, μ^*). When ξ increases to $\xi^{**} > \xi^*$, the equilibrium shifts to (a^{**}, μ^{**}). Supply disruptions that affect ξ not captured by c therefore serve as valid instruments to recover the pricing curve. If the pricing curve is flat—as in the case of constant price elasticity—then changes in ξ would shift the stock without altering the markup. In contrast, if the pricing curve is downward sloping, as shown in the figure, then a higher ξ leads to lower stock and a higher markup. Verifying this inverse relationship between inventory and price margins is the focus of the empirical

FIGURE 1: Illustration of a Retailer's Equilibrium



Notes: This plot is illustrative. A finished-goods inventory model with a negatively sloped stocking curve would yield qualitatively similar implications.

analysis that follows.

2.2 Two-period model

We now extend the logic of the one-period model to a dynamic, two-period setting in which inventories can be carried across periods. This extension verifies that the identification strategy remains valid when the firm makes intertemporal stocking decisions.

Assume the retail firm lives for two periods, facing the same demand function in each period as in the one-period model, $s = \boldsymbol{s}(a, p)$. In the first period, the firm purchases a stock of goods a_1, sets a price p_1, and sells $s_1 = \boldsymbol{s}(a_1, p_1)$. The remaining inventory $a_1 - s_1$ is carried over to the second period, subject to a depreciation rate δ. In the second period, the firm purchases n additional units, bringing the total available stock to:

$$a_2 = (1 - \delta)(a_1 - s_1) + n.$$

It then sets price p_2 and sells $s_2 = \boldsymbol{s}(a_2, p_2)$ units. As in the one-period model, leftover inventory at the end of the second period is liquidated at cost, governed by the parameter γ.

The firm maximizes the sum of discounted profits:

$$\pi = p_1 s_1 - (c_1 + \xi_1)a_1 + \beta \mathbb{E}_1 \left[p_2 s_2 - c_2(a_2 - (1-\delta)(a_1 - s_1)) + (1-\gamma)c_2(a_2 - s_2) \right],$$

where β is the discount factor, and $\mathbb{E}_1$ denotes expectations conditional on information available in period 1. We assume that the second-period supply disruption ξ_2 is unanticipated, so $\mathbb{E}_1 \xi_2 = 0$.

The first-order condition with respect to a_1 (initial stocking) is:

$$(p_1 - \beta(1-\delta)\mathbb{E}_1 c_2) \frac{\partial \boldsymbol{s}(a_1, p_1)}{\partial a_1} = c_1 + \xi_1 - \beta(1-\delta)\mathbb{E}_1 c_2,$$

which mirrors the one-period stocking condition in equation (1).

Similarly, the first-order condition with respect to p_1 is:

$$s_1 = -(p_1 - \beta(1-\delta)\mathbb{E}_1 c_2) \frac{\partial \boldsymbol{s}(a_1, p_1)}{\partial p_1},$$

which corresponds to the pricing condition in equation (2).

As in the one-period case, the shadow cost ξ_1 enters only the stocking equation and not the pricing equation. This separation confirms that temporary supply disruptions—captured by changes in ξ_1—remain valid instruments for identifying the slope of the pricing curve in a dynamic environment.

2.3 Empirical proxy for the stock of goods: Inventory-sales ratio

In the next section, we turn to the data and implement this identification strategy using industry-level panel data to estimate the elasticity of retail price margins with respect to the stock of goods. In empirical work, we proxy the stock using the end-of-period inventories. Indeed, our stylized model implies a tight relationship between the stock of goods, a, and

end-of-period inventories, inv, as given by:

$$inv = a - s.$$

Letting $is = inv/s$ denote the inventory-sales ratio and $\overline{IS}$ its long-run average, we obtain the following log-linear approximation:

$$\hat{a} = \hat{s} + \frac{\overline{IS}}{\overline{IS} + 1}\widehat{is},$$

where hatted variables denote log-deviations from their respective averages. This expression implies that changes in the stock-sales ratio and the inventory-sales ratio are proportionally related. As such, conditional on sales, fluctuations in the stock of goods are directly reflected in movements of the inventory-sales ratio. Accordingly, we use the inventory-sales ratio as an empirical proxy for the stock of goods in our model—controlling for sales—as it conveys equivalent information about stock dynamics and is both directly observable and widely used in the empirical literature.

3 Inventories and prices at the industry level

In this section, we use industry-level panel data to estimate how the retail and wholesale margin price reacts to changes in the inventory stock.[4]

3.1 Industry-level data

We assemble quarterly data for eight industries across the retail and wholesale trade sectors, which we collectively refer to as the retail sector for simplicity, spanning 2008q1 to 2025q1. From the Bureau of Labor Statistics (BLS), we collect industry-specific Producer Price Index (PPI) data, average hourly earnings, and total hours worked. Note that in the trade sectors, the PPI reflects gross margins—that is, the difference between the sales price and the acquisition cost of goods sold by retailers. From the Bureau of Economic Analysis

[4]Since both retailers and wholesalers primarily manage finished-good inventories and typically do not transform the products as manufacturers do, we also include data from the wholesale trade industry for completeness of our analysis.

(BEA), we incorporate information on real inventories and real sales, which are sourced from the National Income and Product Accounts tables. We merge the BLS and BEA data at the three-digit North American Industry Classification System (NAICS) level, which represents the most granular classification available across these data sets. Further details on data construction are provided in the Appendix.

3.2 Industry-level regression framework

To formally examine the relationship between inventory holdings and changes in the industry gross margin, we estimate the following regression,

$$\Delta_k \log(PPI_{i,t}) = \beta \Delta_k \log(IS_{i,t}) + X_{i,t}\gamma + \varepsilon_{i,t}, \tag{3}$$

where Δ_k denotes the k-th difference operator. The outcome variable is the percentage change in the gross margin, measured as the log-difference in the PPI from period $t-k$ to t. The coefficient of interest, β, captures the relationship between changes in the inventory-sales ratio and changes in the gross margin. All control variables are contained in $X_{i,t}$.

Because stocking and pricing decisions are jointly determined, estimating equation (3) via ordinary least squares (OLS) would yield a biased estimated of β. In particular, as shown in the stylized model, unobserved persistent or anticipated shocks can shift both the pricing and stocking curves. Shifts in the pricing curve induce a positive correlation between stocks and margins. As a result, we expect the OLS estimate of β to be biased upward.

To estimate equation (3), we therefore adopt an identification strategy which exploits plausibly exogenous and transitory supply-side variation which we discuss in turn.

3.3 Identification with shift-share instruments

We seek to establish causality by instrumenting for the retailer's inventory-sales ratio. Motivated by our stylized model, we construct two instruments that exploit plausibly exogenous variation in upstream suppliers' ability to deliver finished-goods inventories to downstream retailers. Delivery disruptions can arise from (i) production problems—including suppli-

ers' difficulties sourcing raw materials—and (ii) delays in shipping finished goods. Our instruments capture both channels by measuring retailers' exposure to shocks that affect upstream manufacturers' capacity to produce and deliver goods.

Each instrument takes the form of a shift-share variable (Bartik, 1991). Identification comes from plausibly exogenous shifts at the manufacturing level that reflect supply-side disruptions. Conditional on the specified controls, we assume (i) these shifts are orthogonal to unobserved shocks in downstream retail outcomes, and (ii) the mapped shifts affect retail price margins only through inventories. Fixing the exposure shares *ex ante* makes them predetermined but does not ensure exogeneity, since shares can co-move with persistent demand or pricing dynamics. Following Borusyak et al. (2022), we therefore place identification on the exogeneity of the shifts: under their quasi-experimental shift-share framework, a share-weighted average of shifts is as good as random when the shifts are exogenous, so instrument validity hinges on the shifts rather than on the exogeneity of the exposure shares. We next define the two shift-share instruments used in our analysis.

3.3.1 Reported material shortages from the QSPC

Production delays in manufacturing typically show up in capacity utilization data. Indeed, capacity utilization have taken center stage in recent work on inflationary pressures arising from production bottlenecks (e.g., Boehm and Pandalai-Nayar, 2022; Comin et al., 2023). Yet utilization is endogenous, reflecting both demand and supply forces. Unlike Boehm and Pandalai-Nayar (2022), who emphasize demand-driven movements in utilization and their interaction with supply conditions, our goal is to isolate upstream supply disturbances that shift utilization orthogonally to contemporaneous retail demand (conditional on controls).

Our first instrument exploits the establishment-level reason codes in the QSPC—the survey underlying official capacity utilization statistics. A key advantage of the QSPC is that it records *why* establishments operate below capacity, allowing us to isolate supply-driven variation in utilization. For each three-digit NAICS manufacturing industry, we collect the share of underutilizing establishments that cite "insufficient supply of materials."[5]

Firm-reported material shortages have been shown to track supply-driven disturbances

[5]See https://www.census.gov/programs-surveys/QSPC.html for details. The relevant ratio appears in Table 3 of the published survey results.

in production (Balleer and Noeller, 2024; Braun et al., 2024). Building on this evidence, we take the raw shortage shares and, industry-by-industry, residualize them on their own lag, measures of contemporaneous aggregate demand and industry-specific demand, and quarter fixed effects for seasonality. The aggregate demand proxy follows Shapiro (Forthcoming). The industry-specific demand proxy is the common component of gross output quantities and prices (extracted separately by industry) that loads positively on both. This procedure removes predictable demand movements and serial correlation, leaving an innovation that reflects unexpected, supply-driven disruptions.

Formally, our QSPC shift-share instrument is defined as:

$$Z_{i,t}^{QSPC} = \sum_{j} s_{ij,t_0} \times \Delta_k g_{j,t}^{QSPC}, \tag{4}$$

where the shift variable is,

$$\Delta_k g_{j,t}^{QSPC} = \sum_{\ell=0}^{k} \varepsilon_{j,t-\ell}^{QSPC}. \tag{5}$$

The shift variable is defined as the residualized share of plants in upstream industry j reporting material shortages (as described above). Meanwhile, the share variable, $s_{ij,t_0} \equiv \omega_{ij,t_0}/\omega_{i,t_0}$, is defined as the retailer i's exposure to a given manufacturing industry j in period t_0. The numerator, ω_{ij,t_0}, is the value of inputs purchased by retail industry i from manufacturing industry j in the base year t_0.[6] The denominator, $\omega_{i,t_0} = \sum_j \omega_{ij,t_0}$, represents the total value of inputs (manufacturing and non-manufacturing) used by retailer i. The term s_{ij,t_0} thus reflects the input share sourced from manufacturing industry j within retailer i's total input bundle.

3.3.2 Bai et al. (2024)'s supply chain disruption index

We next consider an instrument for global supply disruptions based on the ACR index of Bai et al. (2024), available from January 2017 to June 2024. The ACR index measures global port congestion and is regarded as an exogenous indicator of shipping disruptions.

[6] We calibrate ω_{ij,t_0} using data from the BEA Industry Input-Output Use Table, with $t_0 = 2007$, which is before the start of our sample period.

Our ACR shift-share instrument is defined as:

$$Z_{i,t}^{ACR} = \sum_{j} \left(s_{ij,t_0} \times \frac{m_{j,t_0}}{m_{t_0}} \right) \times \Delta_k g_t^{ACR}, \tag{6}$$

where the shift variable is,

$$\Delta_k g_t^{ACR} = \sum_{\ell=0}^{k} \varepsilon_{t-\ell}^{ACR}, \tag{7}$$

and ε_t^{ACR} is the residual from a regression of the ACR index on its own lag. Because the ACR shift variable is a time series, we define our share variable as a composite measure of exposure where s_{ij,t_0} denotes the retailers' exposure from the BEA 2007 Use Table, as before, and $m_{j,t_0}/m_{t_0}$ denotes the manufacturers' import share in 2007. Thus, the share variable is an exposure weight mapping global supply disruptions to manufacturers according to import exposure, and to retailers based on the use shares.

3.4 Regression results

We estimate equation (3) via OLS as well as with our instrumental variables, using the aforementioned industry-level data across retailers. The control variables specified in $X_{i,t}$ consist of the k-period log-differences of the following industry-level variables: real sales, average hourly earnings, total hours worked, and an exposure-weighted sum of upstream manufacturing suppliers' PPI. The upstream PPI for each retail industry is defined as:

$$PPI_{i,t}^{\text{upstream}} = \sum_{j} s_{ij,t_0} \times PPI_{j,t}, \tag{8}$$

where $PPI_{j,t}$ denotes the PPI of manufacturing industry j in period t. We also specify a range of fixed effects in $X_{i,t}$. First, we specify industry fixed effects to control for time-invariant unobserved retailer characteristics as well as the time-invariant component of the shifts. In addition, we specify industry-specific linear time trends to absorb gradual, deterministic retail industry-specific trends over time. Finally, we specify time fixed effects. For the QSPC instrumental variable regression, however, we specify an interaction of the

TABLE 1: Industry Regression Results

	Dependent variable: Gross price margin			
Estimation method	OLS	IV		
Instrument	–	QSPC	ACR	QSPC, ACR
Panel A: One-quarter Difference ($k = 1$)				
IS ratio	-0.143***	-0.579***	-0.583***	-0.648***
	(0.038)	(0.215)	(0.179)	(0.166)
First-stage F-stat		5.62	9.50	12.30
AR Wald F (p-value)		0.080	0.001	<0.001
Hansen J (p-value)				0.800
Observations	536	536	232	232
Sample	2008q1-2025q1	2008q1–2025q1	2017q2–2024q2	2017q2–2024q2
Panel B: Two-quarter Difference ($k = 2$)				
IS ratio	-0.121**	-0.786***	-0.446***	-0.570***
	(0.048)	(0.233)	(0.064)	(0.066)
First-stage F-stat		11.09	31.32	15.51
AR Wald F (p-value)		0.012	<0.001	<0.001
Hansen J (p-value)				0.180
Observations	536	528	224	224
Sample	2008q1–2025q1	2008q2-2025q1	2017q3–2024q2	2017q3–2024q2

Notes: Driscoll–Kraay standard errors are reported in parentheses. Instruments for the IV columns are: reported material supply shortages from the Quarterly Survey of Plant Capacity (QSPC) and Average Congestion Rate (ACR). * p<0.10, ** p<0.05, *** p<0.01.

sum of exposure shares with time fixed effects.[7]

3.4.1 OLS estimation

We first estimate equation (3) using OLS. The results are reported in the OLS column of Table 1. Panels A and B present estimates for the one-quarter difference ($k = 1$) and the

[7]In our context, the exposure shares do not sum to one as we measure a retail industry's exposure based on its supply use from only manufacturers (ω_{ij,t_0}) even though its total supply use (ω_{i,t_0}) includes non-manufacturers. Without controlling for sum of share-by-time fixed effects, our QSPC instrument would be correlated with with sum of share which would render our QSPC instrument invalid.

two-quarter difference ($k = 2$), respectively. The OLS estimate in Panel A suggests that changes in the inventory-sales ratio are negatively related to changes in the retailer's gross price margin. We find a similar negative relation at the two-quarter horizon (Panel B). Specifically, a one percentage point decline in the inventory-sales ratio over two quarters is associated with a roughly 0.12–0.14 percentage point increase in the gross margin over the same period.

3.4.2 Two-stage least squares (2SLS) estimation

We next implement instrumental variables (IV) estimation using 2SLS, introducing the instruments first individually and then jointly. The corresponding results are reported in the final three columns of Table 1. At both horizons ($k = 1$ and $k = 2$), a one percentage point decline in the inventory-sales ratio leads to a statistically significant increase in retail gross margins. The estimated elasticities range from approximately -0.45 to -0.79. These estimates are economically significant, implying that a one standard deviation decrease in the inventory-sales ratios raises gross margins by about one standard deviation.

The first-stage F-statistics suggest that the instruments are generally strong, particularly at the two-quarter horizon and when using ACR or both instruments jointly. For the one-quarter specification with QSPC alone, the F-statistic is somewhat lower—possibly reflecting retail pricing frictions at higher frequencies—but inference remains supported by the Anderson-Rubin Wald test, which is robust to weak instruments.

In the final column of Table 1, we use both instruments jointly and continue to estimate a statistically significant and negative elasticity. The first stage F-statistics remain above conventional thresholds, and we can also implement a test of overidentifying restrictions, which provides additional support for the validity of our estimates.[8]

Relative to OLS, the IV estimates are more negative, indicating that the OLS estimates are biased upward. This pattern is consistent with the intuition described above and with

[8]In the final column of Table 1, we normalize the QSPC shares so that they sum to one. This normalization allows us to include time fixed effects rather than a full set of share-by-time fixed effects. Without normalization, the share-by-time fixed effects would subsume the ACR shift-share instrument, preventing us from using both instruments jointly. We view the normalization as a conservative adjustment, since it tends to weaken the strength of the QSPC instrument by not capturing the full variation in the non-normalized shares.

TABLE 2: Industry Regression Results (Including Other Input Costs)

	Dependent variable: Gross price margin			
Estimation method	OLS	IV		
Instrument	–	QSPC	ACR	QSPC, ACR
Panel A: One-quarter Difference ($k = 1$)				
IS ratio	-0.342*	-1.089***	-1.356	-0.986***
	(0.195)	(0.367)	(0.544)	(1.439)
First-stage F-stat		9.38	1.86	12.17
AR Wald F (p-value)		0.064	0.520	0.039
Hansen J (p-value)				0.81
Observations	96	96	86	86
Sample	2019q1–2025q1	2019q1–2025q1	2019q1–2024q2	2019q1–2024q2
Panel B: Two-quarter Difference ($k = 2$)				
IS ratio	-0.674***	-1.752***	-1.589**	-1.521***
	(0.221)	(0.426)	(0.764)	(318)
First-stage F-stat		13.59	3.69	17.49
AR Wald F (p-value)		0.0005	0.181	0.006
Hansen J (p-value)				0.91
Observations	91	91	81	81
Sample	2019q2-2025q1	2019q2–2025q1	2019q2–2024q2	2019q2–2024q2

Notes: Driscoll–Kraay standard errors are reported in parentheses. Instruments for the IV columns are: material shortages reported in the Quarterly Survey of Plant Capacity (QSPC) and Average Congestion Rate (ACR). * p<0.10, ** p<0.05, *** p<0.01.

the mechanism illustrated in the stylized model. In particular, an upward shift in the pricing curve results in higher margins and greater inventory stocks. Because OLS does not account for these unobserved shifts in the pricing curve, it induces a positive correlation between the inventory-sales ratio and gross margins, thereby biasing the OLS estimates upward.

3.4.3 Robustness

Retailers may face additional costs beyond the labor and acquisition costs that we explicitly control for. If these costs are omitted, we cannot rule out that the estimated increase in

gross margins reflects higher unobserved input costs. To address this concern, we estimate a version of our regressions that includes controls for other non-labor and non-capital costs, drawn from the BLS satellite series of *net inputs to industry price indexes*. These data are available only from December 2018 for NAICS 441, 445, and 452, and from December 2020 for the remaining retail and wholesale industries.

Because these series are available only over a shorter horizon, the sample size is reduced and the estimates are correspondingly noisier. Nonetheless, as reported in Table 2, the results remain broadly consistent with our baseline findings, indicating that our main conclusions are robust to the inclusion of these additional cost controls.

We report further robustness checks in the Appendix. These include first-stage regression results and a set of balance tests for the QSPC instrument at both at the manufacturing and retail levels. We also redefine the QSPC instrument to address potential measurement concerns.

3.5 Interpretation of the pricing response

Our findings suggest that retail prices tend to rise in response to exogenous supply disruptions, even after accounting for a broad set of controls, including measures of input costs. In the context of the NKPC literature, this evidence aligns with a strand of research emphasizing the role of time-varying retail price markups in shaping inflation dynamics. The next section builds on this perspective by further exploring its implications.

That said, we acknowledge that alternative interpretations of our results are possible and warrant further investigation. First, given our reliance on industry-level data, we are unable to observe the composition of retail goods. Retailers may respond to supply disruptions by shifting toward higher-margin products, which would mechanically raise measured price margins without requiring active price increases. Second, we cannot rule out unobserved cost pressures. In particular, absent data on retailers' expectations of future marginal costs, it remains possible that firms raise prices to maintain rather than expand markups. Third, we remain agnostic on manufacturer markups and do not study their interaction with retailer markups—a topic examined in detail by Alvarez-Blaser et al. (2025) using data from a large global manufacturer. These alternative mechanisms highlight the need for

richer data to fully disentangle pricing responses, which we leave for future work.

4 The NKPC with inventories

Our empirical investigation reveals a significantly positive response in the retail price margin to supply disruptions that adversely affect the stocking behavior of the retail industry. Building on this relationship, we examine whether the systematic link between the dynamics of the retail inventory-sales ratio and the retail price markup can help improve the empirical performance of the NKPC.

To address this question, we extend the stylized model introduced in Section 2. Specifically, we incorporate inventory management into the retail sector and introduce two standard sources of state-dependent price markups: the demand specification of Kimball (1995) and the deep habits framework of Ravn et al. (2006). These mechanisms generate time-varying price margins that align with our empirical findings and are widely utilized in the literature on inflation and business cycles.

Subsequently, we derive two representations of the linearized NKPC with inventories. We demonstrate that incorporating the retail inventory-sales ratio as an observable variable in the NKPC effectively captures time-varying price markups, which are otherwise difficult to measure directly.

4.1 Demand function

We first derive the retailers' demand function. To introduce time-varying price markups even without nominal rigidities, we specify the following aggregator where the aggregate habit-adjusted consumption, C_t, is implicitly defined as

$$\int_0^1 v_{j,t} \Upsilon \left(\frac{s_{j,t} - \theta s_{j,t-1}}{v_{j,t} C_t} \right) dj = 1.$$

The variables $s_{j,t}$ and $v_{j,t}$ denote the real sales and household preference of variety $j \in [0,1]$ in period t, respectively. The parameter $\theta \in [0,1]$ measures the degree of habit formation in consumption of each variety. This specification encompasses both the state-dependent

demand elasticity *a la* Kimball (1995) and deep-habit formation in customer markets analyzed in Ravn et al. (2006). Following Klenow and Willis (2016), we assume the aggregator $\Upsilon(x)$ such that

$$\Upsilon'(x) = \left(\frac{\eta - 1}{\eta}\right) \exp\left(\frac{1 - x^{\frac{\psi}{\eta}}}{\psi}\right),$$

This function is consistent with CES when $\psi > 0$ approaches zero and $\eta > 1$ is the price elasticity of demand under CES. More generally, the price elasticity of demand is expressed as

$$-\left(\frac{ds_{j,t}}{dp_{j,t}}\right)\left(\frac{p_{j,t}}{s_{j,t}}\right) = \frac{\Upsilon'(x_{j,t})}{x_{j,t}\Upsilon''(x_{j,t})}\frac{s_{j,t} - \theta s_{j,t-1}}{s_{j,t}} = \eta x_{j,t}^{-\frac{\psi}{\eta}}\frac{s_{j,t} - \theta s_{j,t-1}}{s_{j,t}},$$

where $x_{j,t} \equiv (s_{j,t} - \theta s_{j,t-1})/(v_{j,t}C_t)$. As discussed above, the price elasticity of demand is η in the special case of CES ($\theta = \psi = 0$). Otherwise when $\theta > 0$ or $\psi > 0$, the price elasticity of demand varies due to its state and history dependence, allowing more flexibility for the model to be consistent with our empirical findings of a negative inventory stock elasticity of the price margin.

As shown in the Appendix, the variety j demand function is derived as follows:

$$s_{j,t} = v_{j,t}C_t\Upsilon'^{-1}\left(\frac{p_{j,t}}{P_t}D_t\right) + \theta s_{j,t-1}, \tag{9}$$

where the demand index D_t is defined as

$$D_t = \int_0^1 \Upsilon'\left(\frac{s_{j,t} - \theta s_{j,t-1}}{v_{j,t}C_t}\right)\frac{s_{j,t} - \theta s_{j,t-1}}{C_t}dj,$$

and the price index P_t is implicitly defined in

$$1 = \int_0^1 v_{j,t}\Upsilon\left(\Upsilon'^{-1}\left(\frac{p_{j,t}}{P_t}D_t\right)\right)dj.$$

4.2 The retail firm problem

The monopolistically competitive retail firm j in period t purchases $y_{j,t}$ quantity of manufacturing goods at a unit price Q_t. It differentiates the good which is then stocked and sold to households at a price $p_{j,t}$. The stock of retail goods, $a_{j,t}$, is the sum of the undepreciated inventories from the previous period and the goods purchased in the current period:

$$a_{j,t} = (1-\delta)inv_{j,t-1} + y_{j,t}, \tag{10}$$

where $inv_{j,t-1}$ is the end-of-period inventories in period $t-1$ and δ is the inventory depreciation rate. In turn, the end-of-period inventories evolve as follows:

$$inv_{j,t} = a_{j,t} - s_{j,t}. \tag{11}$$

Dividing each side by $s_{j,t}$, we obtain a relationship between the inventory-sales ratio and the stock-sales ratio. Since they differ only by a constant, the two ratios move together. The inventory-sales ratio is defined as:

$$is_{j,t} = \frac{inv_{j,t}}{s_{j,t}}.$$

Demand for the retail good is given in (9). Following Bils and Kahn (2000) and Jung and Yun (2005), we assume that demand is also elastic to the stock of goods for sale:

$$v_{j,t} = \left(\frac{a_{j,t}}{A_t}\right)^{\zeta}, \tag{12}$$

where ζ is the sales elasticity to the stock of goods.[9] The aggregate stock is defined as

$$A_t = \left(\int_0^1 a_{j,t}^{\zeta} dj\right)^{\frac{1}{\zeta}}.$$

[9]We do not employ a stockout avoidance model, as the presence of deep habits introduces rich dynamics that make it difficult to track the relevant idiosyncratic state variables. As noted by Crouzet and Oh (2016) in the context of news shocks and by Kryvtsov and Midrigan (2013) for monetary policy shocks, the stock-elastic demand specification and the stockout avoidance model exhibit similar equilibrium implications for inventory dynamics, particularly with respect to the endogeneity of inventory-sales ratios.

Nominal price rigidity is introduced *a la* Rotemberg (1982). In detail, the retail firm faces a quadratic price adjustment cost in the following form:

$$\frac{\nu_p}{2}\left(\frac{p_{j,t}}{p_{j,t-1}}-1\right)^2 P_t S_t, \tag{13}$$

where ν_p represents the degree of nominal price rigidity and $S_t = \int_0^1 s_{j,t}dj$ is aggregate sales.

The retail firm maximizes the expected discounted sum of profits:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \Lambda_{0,t}\Phi_{j,t},$$

where the period-t profit of the retailer is

$$\Phi_{j,t} = p_{j,t}s_{j,t} - Q_t y_{j,t}.$$

The retailer is subject to the above conditions (9)-(13). The full optimality conditions are presented in the Appendix.

4.3 Deriving the NKPC with inventories

In standard New Keynesian models without inventories, the NKPC is obtained by combining the firm's pricing decision with its sales condition, under the assumption that the firm satisfies all demand at its chosen price. Implicit in this setup is that the retail firm's stocking decision is equivalent to its production and sales decision.

With inventories, this equivalence breaks down: retailers can satisfy demand not only by purchasing new goods but also by drawing down existing stocks. This distinction introduces an additional margin of adjustment and implies that there are multiple routes to deriving the NKPC. Below, we present the key equilibrium conditions that give rise to alternative representations of the NKPC.

4.3.1 Equilibrium pricing condition

The log-linearized optimal pricing condition in symmetric equilibrium is

$$\hat{\pi}_t = \beta \mathbb{E}_t \hat{\pi}_{t+1} - \frac{1}{\nu_p}(\hat{\kappa}_t + \hat{c}_t - \hat{s}_t), \tag{14}$$

where hatted values denote log-deviations from the noninflationary steady state, and κ_t is the real marginal benefit of an additional unit of sales in period t—the Lagrange multiplier on the demand constraint (9), capturing the inframarginal revenue effect. Current inflation therefore depends on expected future inflation, the real marginal benefit of sales, and the gap between habit-adjusted consumption and sales. This condition also applies in the model without inventories. The key difference between models with and without inventories lies in how demand is met, as summarized by κ_t. In the model with inventories, κ_t is not directly linked to contemporaneous real marginal cost because retailers can satisfy demand by drawing down existing inventories rather than immediately purchasing additional goods to restock.

4.3.2 Sales-based NKPC

The log-linearized optimal sales condition is

$$\hat{\kappa}_t = \theta\beta \mathbb{E}_t(\hat{r}_{t,t+1} + \hat{\kappa}_{t+1}) - \left[\eta(1-\theta)\left(1 + \psi \ln\left(\frac{\eta-1}{\eta}\right)\right)^{-1} - 1 + \theta\beta\right] \mathbb{E}_t(\hat{r}_{t,t+1} + \widehat{mc}_{t+1}), \tag{15}$$

where $r_{t,t+1}$ is the real stochastic discount factor between period t and $t+1$ and mc_t the real marginal cost of producing a retail good in period t, equivalently the Lagrange multiplier on the law of motion for the inventory stock (10).

To interpret this condition, it is helpful to begin with the benchmark New Keynesian model without inventories and with CES preferences across varieties ($\theta = \psi = 0$). The sales condition then reduces to

$$\hat{\kappa}_t = -(\eta - 1)\widehat{mc}_t.$$

Because the markup without inventories is the inverse of real marginal cost, this relation shows that the real marginal benefit of sales is proportional to the markup. Substituting this expression into (14) yields the familiar Rotemberg-type NKPC:[10]

$$\hat{\pi}_t = \beta \mathbb{E}_t \hat{\pi}_{t+1} + \left(\frac{\eta - 1}{\nu_p} \right) \widehat{mc}_t.$$

In this benchmark, the sales and stocking conditions coincide, since all sales must be met by contemporaneous purchases of goods. This equivalence implies a one-for-one mapping between marginal cost and the marginal benefit of sales, which is why the standard NKPC can be written directly in terms of real marginal cost.

With inventories, however, this one-for-one mapping breaks down. Additional sales need not be met by contemporaneous purchases. Instead, the retailer recognizes that sales can be satisfied by drawing down inventories, so the optimal sales decision depends on the expected discounted future marginal cost rather than on the current marginal cost. When $\theta > 0$, habit formation adds an additional intertemporal channel: the expected discounted marginal benefit of future sales also matters. Combining (14) and (15) yields the following sales-based NKPC.

Proposition 1 (Sales-based NKPC)**.** *The sales-based NKPC is derived by combining* (14) *and* (15)*:*

$$\hat{\pi}_t = \beta \mathbb{E}_t \hat{\pi}_{t+1} + B_1 \mathbb{E}_t (\hat{r}_{t,t+1} + \widehat{mc}_{t+1}) - B_2 \mathbb{E}_t (\hat{r}_{t,t+1} + \hat{\kappa}_{t+1}) + B_3 (\hat{s}_t - \hat{c}_t),$$

where

$$B_1 = \frac{1}{\nu_p} \left[\eta(1-\theta) \left(1 + \psi \ln \left(\frac{\eta - 1}{\eta} \right) \right)^{-1} - 1 + \theta\beta \right], \quad B_2 = \frac{\theta\beta}{\nu_p}, \quad B_3 = \frac{1}{\nu_p}.$$

In the CES case ($\theta = \psi = 0$), the sales-based NKPC simplifies to

$$\hat{\pi}_t = \beta \mathbb{E}_t \hat{\pi}_{t+1} + \left(\frac{\eta - 1}{\nu_p} \right) \mathbb{E}_t (\hat{r}_{t,t+1} + \widehat{mc}_{t+1}).$$

[10] Without habits ($\theta = 0$), consumption equals sales so that $\hat{c}_t = \hat{s}_t$.

There are two main takeaways. First, without habits, expected future inflation and expected discounted real marginal cost are sufficient statistics for current inflation. Second, with habits, the sales-based NKPC is generally difficult to estimate empirically because it requires separate measures of the real marginal benefit of sales and the marginal cost.

4.3.3 Stock-based NKPC

Inventories also give rise to the equilibrium stocking condition, which characterizes how retailers choose inventory holdings. The log-linearized condition is

$$\hat{a}_t = \hat{c}_t + \frac{1}{1-\beta(1-\delta)}[\beta(1-\delta)\mathbb{E}_t(\hat{r}_{t,t+1} + \widehat{mc}_{t+1}) - \widehat{mc}_t] + \hat{\kappa}_t. \tag{16}$$

This condition highlights three key determinants of the retailer's stocking decision. First, *aggregate demand* ($\hat{c}_t$) matters, as retailers target a desired inventory-sales ratio; higher expected sales translate into higher desired stocks. Second, the *intertemporal substitution channel in stocking* captures the difference between the expected discounted future marginal cost and the current marginal cost: when costs are expected to rise, firms have an incentive to buy more today and carry inventories forward. This channel parallels the mechanism emphasized by Crouzet and Oh (2016). Third, the term $\hat{\kappa}_t$ reflects the *real marginal benefit of sales*: holding more stock increases sales capacity and boosts expected profits, reinforcing the incentive to accumulate inventories.

Substituting out $\hat{\kappa}_t$ and denoting $\overline{IS}$ as the steady-state inventory-sales ratio, we obtain the following stock-based NKPC.

Proposition 2 (Stock-based NKPC)**.** *The stock-based NKPC is derived by combining* (14) *and* (16)*:*

$$\hat{\pi}_t = \beta\mathbb{E}_t\hat{\pi}_{t+1} + \tilde{B}_1[\beta(1-\delta)\mathbb{E}_t(\hat{r}_{t,t+1} + \widehat{mc}_{t+1}) - \widehat{mc}_t] - \tilde{B}_2\widehat{is}_t,$$

where

$$\tilde{B}_1 = [\nu_p(1-\beta(1-\delta))]^{-1}, \quad \tilde{B}_2 = \overline{IS}(\nu_p(\overline{IS}+1))^{-1}.$$

Two insights stand out. First, inflation is driven by three observable statistics: expected future inflation, the expected change in discounted real marginal cost, and the inventory-sales ratio. Second, this representation is agnostic to the sources of endogenous markup variation arising from the Kimball aggregator or deep habits, since these demand effects are absorbed by the inventory-sales ratio. The dynamics of the ratio thus serve as a sufficient statistic for markup behavior in a wide class of models, including CES, Kimball demand, and deep habits.

4.4 Discussion

We find that the stock-based NKPC offers greater flexibility by remaining agnostic about the sources of time-varying markups. Its key observable is the inventory-sales ratio, which—after purging the intertemporal substitution channel—primarily captures the inflationary role of the marginal benefit of sales, i.e., markups. This provides an empirically tractable proxy for time-varying markups in inflation dynamics, which we analyze in the next section.

It is also worth noting that, while our industry-level IV results exploit exogenous variation from short-lived supply disruptions, the stock-based NKPC does not take a stand on the origin of shocks. Its relative improvement in fit should therefore not be interpreted as evidence of the importance of supply disruptions *per se*, since any inflationary disturbance can propagate through inventories. Put differently, while our IV strategy isolates shifts in the stocking curve, the fluctuations in the inventory-sales ratio that enter the stock-based NKPC reflect both shifts of the curve and endogenous movements along it.

5 Empirical Phillips curve with inventories

In this section, we estimate the stock-based NKPC using macroeconomic data and show that it outperforms both the canonical specification and the sales-based NKPC under CES demand. The improvement comes from including the inventory-sales ratio as an observable, which captures the time-varying markup channel missing from standard NKPCs. We further show that this specification accounts for two prominent episodes—the 2009–2011 missing disinflation and the COVID-era surge—underscoring the role of inventories as an

indicator of inflationary pressures. Note that this section provides a reduced-form empirical assessment of aggregate inflation dynamics; causal effects of the inventory channel are established earlier in the industry-level regressions.

5.1 Empirical specification and data

We estimate three reduced-form (semi-structural) Phillips curve specifications with theory-motivated exclusion and equality restrictions:

$$\hat{\pi}_t = \alpha_0 + \alpha_1 \hat{\pi}_{t-1} + \alpha_2 \mathbb{E}_t \hat{\pi}_{t+1} + \alpha_3 \widehat{mc}_t + \alpha_4 (\mathbb{E}_t \hat{r}_{t,t+1} + \mathbb{E}_t \widehat{mc}_{t+1}) + \alpha_5 \widehat{is}_t + \varepsilon_t, \tag{17}$$

where ε_t denotes the regression error. The first sets $\alpha_4 = \alpha_5 = 0$, yielding the canonical NKPC without inventories, as in Galí and Gertler (1999), with both backward- and forward-looking inflation expectations. The second sets $\alpha_3 = \alpha_5 = 0$, corresponding to the sales-based NKPC under CES demand (Proposition 1). The third imposes the equality restriction $\alpha_3 = -\alpha_4$, delivering the stock-based NKPC (Proposition 2) with expected marginal cost growth expressed under zero inventory depreciation and no average time discounting.[11]

To estimate these regressions, we compile macroeconomic variables at quarterly frequency. Since retail inventory management is primarily relevant within the goods sector, our analysis focuses on accounting for goods inflation, though we also report results using headline inflation for comparison. Accordingly, our baseline regression defines $\hat{\pi}_t$ as the annualized quarterly inflation of the goods component of Personal Consumption Expenditures (PCE). For expected inflation $\mathbb{E}_t \hat{\pi}_{t+1}$, we use both the mean professional inflation forecast from the Survey of Professional Forecasters (SPF) and the mean household inflation forecast from the Michigan Survey, each reflecting year-ahead expectations. To our knowledge, no survey provides separate forecasts specifically for goods inflation. By estimating an unrestricted coefficient on inflation expectations, we allow for proportionality between expectations of headline and goods inflation.

To proxy for the real marginal cost, $\widehat{mc}_t$, we use the unemployment gap (scaled by -1), a commonly used measure of economic slack (e.g., Stock and Watson, 2020). Specifically,

[11] Although results are robust to relaxing the restriction, estimating α_3 and α_4 separately induces multicollinearity due to the high correlation between current and expected marginal cost.

we define the unemployment gap as the difference between the quarterly unemployment rate and the quarterly estimate of the Non-Accelerating Inflation Rate of Unemployment (NAIRU) from the Congressional Budget Office (CBO). Because our Phillips curve specification includes expected future real marginal cost, $\mathbb{E}_t\widehat{mc}_{t+1}$, using the unemployment gap is particularly useful, as we can construct a comparable expectation-based proxy. We define the expected future unemployment gap as the difference between the SPF's year-ahead unemployment forecast and the year-ahead NAIRU estimate from the CBO.

Alternative proxies for the real marginal cost used in the literature include the output gap, the labor share, the vacancy-to-unemployment ratio, and unit labor cost. Because the NKPC model with inventories requires a measure of expected future real marginal cost, however, constructing such a measure using these alternatives is more challenging.[12]

We define the expected real interest rate, $\mathbb{E}_t\hat{r}_{t,t+1}$, as the effective Federal Funds rate minus the SPF's year-ahead inflation forecast.

Finally, for the inventory-sales ratio, $\widehat{is}_t$, we aggregate inventory and sales data from the retail and wholesale sectors. We detrend this variable using the Hamilton (2018) filter.

The sample spans 1983q3–2025q1. The inventory-sales ratio is constructed from 1980q4 onward—when wholesale inventories become available—and this earlier start is used to implement the Hamilton filter, which regresses the ratio on its 8th–11th lags.

5.2 NKPC regression results

We now present regression estimates for the three Phillips curve specifications.

5.2.1 Evidence for the stock-based NKPC

Our derivation of multiple NKPC representations is useful because, under CES demand, the sales-based and stock-based specifications should be empirically equivalent. If their empirical fit differs in favor of the stock-based specification, this provides evidence against CES and, through the lens of our model, implies that aggregate inflation dynamics require channels of time-varying markups beyond those captured by nominal rigidity.

[12]Nevertheless, in the Appendix we show that our main results are robust to the choice of the marginal cost proxy in a canonical NKPC specification augmented with the inventory-sales ratio, suggesting that our findings are unlikely to be driven by the specific proxy used.

TABLE 3: NKPC Estimates for Goods and Headline Inflation (SPF Expectations)

	Goods PCE Inflation			Headline PCE Inflation		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged inflation	0.342**	0.343***	0.047	0.455***	0.463***	0.197
	(0.131)	(0.130)	(0.100)	(0.135)	(0.140)	(0.120)
Exp. inflation (SPF)	0.369	0.373	1.173**	0.321**	0.263**	0.665***
	(0.284)	(0.250)	(0.459)	(0.160)	(0.128)	(0.225)
Marginal cost	0.030			0.074		
	(0.145)			(0.063)		
Exp. marginal cost		0.001			0.038	
		(0.092)			(0.047)	
Exp. marginal cost growth			0.048			0.089
			(0.166)			(0.094)
IS ratio			-0.430***			-0.189***
			(0.092)			(0.048)
Constant	-0.282	-0.319	-2.299**	0.461	0.548*	-0.028
	(0.868)	(0.761)	(1.158)	(0.313)	(0.299)	(0.530)
Observations	167	167	167	167	167	167
Adjusted R-squared	0.130	0.130	0.375	0.315	0.313	0.500

Notes: Variable definitions are provided in the main text. Expected inflation is from the SPF; results using the Michigan Survey are reported in the Appendix. Newey–West standard errors are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table 3 reports OLS estimates of the three specifications in equation (17), using goods and headline PCE inflation as dependent variables and SPF inflation expectations. For goods inflation, the stock-based NKPC in column (3) attains a noticeably higher adjusted R-squared than both the canonical specification (column 1) and the sales-based NKPC under CES demand (column 2). The same pattern holds for headline inflation: column (6) dominates columns (4) and (5). The better fit of the stock-based NKPC relative to the CES-demand specification indicates that the data favor our non-CES specification with time-varying markups. Taken together with the industry-level evidence, these results

provide plausible evidence that time-varying markups matter for inflation dynamics.[13]

5.2.2 Why the stock-based NKPC fits better: The role of inventories

Why does the stock-based NKPC fit better? Inspecting columns (3) and (6), the inventory-sales ratio enters with a negative and statistically significant coefficient, as predicted by equation (16), while variables related to marginal cost dynamics—including both current and expected marginal cost—are not significant. The magnitudes are economically meaningful: a one-standard-deviation decline in the inventory-sales ratio is associated with roughly a 2 percentage-point increase in goods inflation and a 0.9 percentage-point increase in headline inflation. By comparison, a one-standard-deviation rise in SPF expected inflation is associated with increases of about 1.1 and 0.6 percentage points, respectively. These results indicate that the inventory-sales ratio plays an important role in accounting for inflation dynamics, likely operating through goods prices.

In column (3), once the inventory-sales ratio is included, the coefficient on SPF expectations strengthens substantially, while the lagged goods inflation term becomes negligible. This suggests that lagged goods inflation had been proxying for omitted inventory dynamics, attenuating the role of expectations. With inventories included, the SPF term becomes more informative, consistent with goods-price dynamics being shaped jointly by professional forecasts and retailers' inventory management decisions.

Using the Michigan Survey for inflation expectations yields a complementary message, as shown in the Appendix. The coefficients of the goods NKPC are stable, and lagged goods inflation is insignificant even in specifications without inventories, consistent with evidence that household expectations—rather than backward-looking indexation—play a central role in the Phillips curve (Coibion et al., 2018). Adding the inventory-sales ratio in this setting leaves coefficients largely unchanged but increases explanatory power, indicating that inventories provide incremental information largely orthogonal to the channel captured by Michigan expectations. In this sense, inventories complement household expectations by improving fit without undermining the structural stability of the NKPC.

[13] In the Appendix, we show that other metrics such as the Akaike information criterion or the one-quarter-ahead root mean squared errors, which also generally favor the stock-based NKPC.

TABLE 4: Semipartial R-squared (Unique Contribution of Regressors)

	Goods PCE Inflation	Headline PCE Inflation
Panel A: SPF Expectations		
Lagged inflation	0.16	2.39
Exp. inflation	4.24	5.03
Exp. marginal cost growth	0.04	0.56
IS ratio	24.43	18.86
Total R-squared	39.03	51.18
Panel B: Michigan Expectations		
Lagged inflation	0.02	0.08
Exp. inflation	9.08	12.23
Exp. marginal cost growth	1.54	5.04
IS ratio	15.07	10.62
Total R-squared	43.87	58.38

Notes: Semipartial R-squared values report each regressor's marginal contribution conditional on all others. All values are percentages.

Our findings are robust across a wide range of specifications and data choices.[14] Re-estimating the model with alternative marginal cost proxies commonly used in the literature confirms that the inventory-sales ratio remains a strong and significant predictor of inflation.[15] We further estimate the NKPC using core goods inflation—adjusting the inventory-sales ratio accordingly—and again obtain results consistent with our main conclusion. Finally, testing a nonlinear specification of the inventory term shows that the linear mechanism continues to dominate.

5.2.3 Orthogonal contribution of inventories

To isolate the contribution of the inventory–sales ratio in the stock-based NKPC and compare it with other regressors, we report the semipartial R-squared. Following the Frisch-

[14] See Appendix for details.

[15] As discussed earlier, we estimated the canonical NKPC augmented with the inventory-sales ratio, rather than the stock-based NKPC, because expected marginal cost proxies are not available for this robustness check.

Waugh-Lovell logic, we orthogonalize each regressor with respect to the other controls and then computing its explanatory power. This statistic captures the regressor's unique (orthogonal) contribution and, in general, does not sum to total R-squared across regressors. Recall that, in our framework, the component of the inventory-sales ratio orthogonal to expected marginal cost growth primarily reflects movements in time-varying markups; accordingly, its semipartial R-squared provides a conservative measure of the markup channel.

Table 4 reports results for both goods and headline inflation using expectations from either the SPF or the Michigan Survey. For SPF expectations (Panel A), the inventory-sales ratio accounts for a substantial 24.4 percent, far exceeding the marginal cost and inflation expectations channels in terms of orthogonal contributions. This indicates that variation in the inventory-sales ratio orthogonal to marginal cost and inflation expectations is quantitatively important for goods inflation. A similar, though smaller, role appears for headline inflation, where the inventory-sales ratio's semipartial R-squared is about 18.9 percent.

With Michigan expectations (Panel B), expected inflation absorbs a larger share of variation in goods inflation (about 9.1 percent). Even so, the inventory-sales ratio remains the single largest semipartial R-squared component (roughly 15 percent). For headline inflation, expected inflation contributes more (about 12.2 percent) than the inventory-sales ratio (about 10.6 percent), but the latter remains economically meaningful. In both cases, the inventory-sales ratio adds distinct, orthogonal information not contained in Michigan expectations and therefore complements them in accounting for inflation dynamics.

5.3 The stock-based NKPC and recent inflation episodes

Having established the superior fit of the stock-based NKPC in the full sample, we next examine whether the framework can also account for two specific inflation episodes that have attracted considerable attention in the literature.

5.3.1 The 2009–2011 missing disinflation puzzle

The 2009–2011 missing disinflation puzzle has been widely discussed in the literature, with a prominent resolution proposed by Coibion and Gorodnichenko (2015). Using the SPF,

they show that a version of the canonical NKPC substantially underpredicts inflation over this period. However, when substituting inflation expectations from the Michigan Survey—where household expectations rose more sharply—the puzzle largely disappears. This implies that the puzzle is conditional on the choice of the inflation expectation measure, particularly the SPF.

We offer an alternative explanation using the stock-based NKPC while still conditioning on SPF expectations. Panel C of Figure 2 plots actual headline PCE inflation alongside inflation predicted by both the canonical and stock-based NKPCs. The canonical specification predicts four-quarter inflation to remain below 2 percent between 2009 and 2011 (the first shaded region), while observed inflation reached 3 percent. In contrast, the stock-based NKPC successfully tracks the actual inflation path. This improvement coincides with a period of below-trend inventory-sales ratios, consistent with widespread inventory liquidation by financially constrained firms following the 2008 credit shock, as analyzed by Kim (2021).

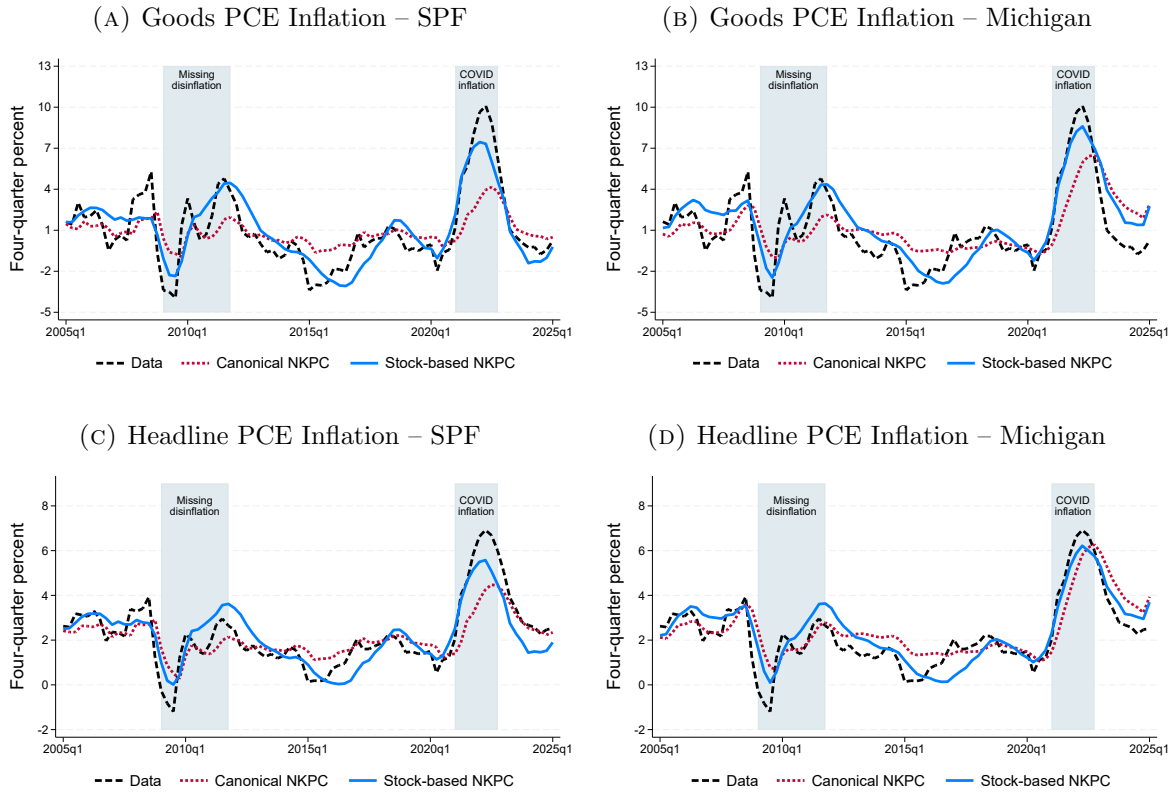
For comparison, Panel D uses a specification similar to Coibion and Gorodnichenko (2015) with Michigan inflation expectations. In this case, the canonical NKPC also aligns more closely with observed inflation, and the stock-based NKPC remains consistent with the data.

5.3.2 The COVID-era inflation surge

The stock-based NKPC also captures the surge in inflation—particularly for goods—during the COVID period. As shown in Panels A and B of Figure 2, goods inflation rose to 10 percent between 2021 and 2022 (the second shaded region), a surge that the stock-based NKPC better tracks in both panels. In contrast, the canonical NKPC explains only a smaller portion of the rise, predicting a peak inflation of roughly 4 percent when using SPF expectations (Panel A) and 6 percent with Michigan expectations (Panel B). Moreover, the stock-based NKPC better captures the sharp reversal in goods inflation following the peak, underscoring the complementary role of the inventory-sales ratio in accounting for goods inflation dynamics, especially during periods of swift adjustment.

Turning to headline inflation, the tight inventory-sales ratio, which played a central role in the surge in goods inflation, also contributed significantly to the rise in headline

FIGURE 2: Model Fit



Notes: Panel A shows the model fit to goods PCE inflation using year-ahead SPF inflation forecasts. Panel B presents the model fit to goods PCE inflation using year-ahead Michigan inflation forecasts. Panel C shows the model fit to headline PCE inflation with year-ahead SPF forecasts, while Panel D presents the fit using year-ahead Michigan forecasts. In all panels, the first shaded region corresponds to 2009q1–2011q4 and the second to 2021q1-2022q4.

inflation. The stock-based NKPC not only better matches the magnitude of the headline surge than the canonical specification, but also more accurately predicts its earlier timing—showing a peak that arrives sooner and aligns more closely with the data. This improvement is pronounced when using SPF expectations (Panel C), and while the magnitude remains similar with Michigan expectations (Panel D), the earlier predicted peak remains consistent with the observed timing.[16]

[16]In the Appendix, COVID-era pseudo-out-of-sample inflation forecasts show a similar pattern, though all models exhibit a delayed pickup in inflation relative to the data.

6 Conclusion

This paper investigates whether the retail price margin channel—whereby inventory scarcity amplifies retail price margins—can help explain observed inflation dynamics. By modeling the retailer's joint decision over pricing and inventory management, and constructing industry-level instruments from supply chain disruptions that plausibly affect only the retailer's stocking behavior, we identify the elasticity of price margins with respect to inventory stocks. We estimate this elasticity to be negative, implying that tighter retail inventories are systematically associated with higher price margins. This evidence supports the view that retail prices are shaped not only by cost-push factors but also by the strategic pricing behavior of firms facing inventory constraints.

Integrating this mechanism into a NKPC framework, we show that the inventory-sales ratio serves as a powerful and observable proxy for retail markups. The resulting stock-based NKPC delivers a substantially improved empirical fit, successfully capturing key inflationary episodes such as the 2009–2011 missing disinflation and the COVID-era surge that originated from the goods sector. These findings suggest that incorporating retail inventory dynamics enhances our understanding of inflation—particularly during periods when supply chains are constrained—by capturing a critical source of time-varying markups often overlooked in standard NKPC models.

Looking ahead, several avenues for further research remain. One is to examine how inventory dynamics interact with nonlinear NKPC specifications, for example by allowing the slope of the Phillips curve to vary across regimes of inventory tightness or demand pressure. Another is to investigate whether the inventory-based mechanism can account for the time variation in the slope of the NKPC documented in the empirical literature. More broadly, extending the framework to incorporate sectoral heterogeneity or to model the interaction between retail inventories and upstream supply chain frictions could yield new insights into how microeconomic bottlenecks shape aggregate inflation dynamics.

References

Alessandria, George, Shafaat Yar Khan, Armen Khederlarian, Carter Mix, and Kim J. Ruhl, "The Aggregate Effects of Global and Local Supply Chain Disruptions: 2020–2022," *Journal of International Economics*, 2023, *146*, 103788.

Alvarez-Blaser, Santiago, Alberto Cavallo, Alexander MacKay, and Paolo Mengano, "Markups and Cost Pass-through Along the Supply Chain," Working Paper 34110, National Bureau of Economic Research August 2025.

Bai, Xiwen, Jesús Fernández-Villaverde, Yiliang Li, and Francesco Zanetti, "The Causal Effects of Global Supply Chain Disruptions on Macroeconomic Outcomes: Evidence and Theory," Working Paper 32098, National Bureau of Economic Research February 2024.

Balleer, Almut and Marvin Noeller, "Monetary Policy in the Presence of Supply Constraints: Evidence from German Firm-level Data," Discussion Paper 18732, CEPR 2024.

Bartik, Timothy J., "Who Benefits from State and Local Economic Development Policies?," W. E. Upjohn Institute for Employment Research 1991.

Bils, Mark and James A. Kahn, "What Inventory Behavior Tells Us about Business Cycles," *American Economic Review*, June 2000, *90* (3), 458–481.

Boehm, Christoph E. and Nitya Pandalai-Nayar, "Convex Supply Curves," *American Economic Review*, December 2022, *112* (12), 3941–69.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *Review of Economic Studies*, 2022, *89* (1), 181–213.

Braun, Robin, Aaron Flaaen, and Sinem Hacioglu-Hoke, "Supply vs Demand Factors Influencing Prices of Manufactured Goods," FEDS Notes, Board of Governors of The Federal Reserve System February 2024.

Cavallo, Alberto and Oleksiy Kryvtsov, "What Can Stockouts Tell Us about Inflation? Evidence from Online Micro Data," *Journal of International Economics*, 2023, *146*, 103769.

Coibion, Olivier and Yuriy Gorodnichenko, "Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation," *American Economic Journal: Macroeconomics*, January 2015, *7* (1), 197–232.

_ and _ , "Inflation, Expectations and Monetary Policy: What Have We Learned and to What End?," Working Paper 33858, National Bureau of Economic Research May 2025.

_ , _ , and Rupal Kamdar, "The Formation of Expectations, Inflation, and the Phillips Curve," *Journal of Economic Literature*, 2018, *56* (4), 1447–1491.

Comin, Diego A, Robert C Johnson, and Callum J Jones, "Supply Chain Constraints and Inflation," Working Paper 31179, National Bureau of Economic Research April 2023.

Crouzet, Nicolas and Hyunseung Oh, "What Do Inventories Tell Us about News-driven Business Cycles?," *Journal of Monetary Economics*, May 2016, *79*, 49–66.

Ferrante, Francesco, Sebastian Graves, and Matteo Iacoviello, "The Inflationary Effects of Sectoral Reallocation," *Journal of Monetary Economics*, 2023, *140*, S64–S81.

Galí, Jordi and Mark Gertler, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 1999, *44* (2), 195–222.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 2020, *110* (8), 2586–2624.

Hamilton, James D., "Why You Should Never Use the Hodrick-Prescott Filter," *The Review of Economics and Statistics*, 12 2018, *100* (5), 831–843.

Jung, Yongseung and Tack Yun, "Monetary Policy Shocks, Inventory Dynamics, and Price-setting Behavior," Working Paper Series 2006-02, Federal Reserve Bank of San Francisco 2005.

Kim, Ryan, "The Effect of the Credit Crunch on Output Price Dynamics: The Corporate Inventory and Liquidity Management Channel," *The Quarterly Journal of Economics*, February 2021, *136* (1), 563–619.

Kimball, Miles S., "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, November 1995, *27* (4), 1241–1277.

Klenow, Peter J. and Jonathan L. Willis, "Real Rigidities and Nominal Price Changes," *Economica*, July 2016, *83*, 443–472.

Kryvtsov, Oleksiy and Virgiliu Midrigan, "Inventories and Real Rigidities in New Keynesian Business Cycle Models," *Journal of the Japanese and International Economies*, June 2010, *24* (2), 259–281.

_ and _, "Inventories, Markups, and Real Rigidities in Menu Cost Models," *The Review of Economic Studies*, January 2013, *80* (1), 249–276.

Lubik, Thomas A. and Wing Leong Teo, "Inventories, Inflation Dynamics and the New Keynesian Phillips Curve," *European Economic Review*, April 2012, *56* (3), 327–346.

Mavroeidis, Socrates, Mikkel Plagborg-Møller, and James H. Stock, "Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve," *Journal of Economic Literature*, 2014, *52* (1), 124–188.

Ortiz, Julio L., "Spread Too Thin: The Impact of Lean Inventories," International Finance Discussion Papers 1342, Federal Reserve Board 2022.

Ravn, Morten, Stephanie Schmitt-Grohé, and Martin Uribe, "Deep Habits," *The Review of Economic Studies*, 2006, *73*, 195–218.

Reis, Ricardo, "Four Mistakes in the Use of Measures of Expected Inflation," *AEA Papers and Proceedings*, May 2023, *113*, 47–51.

Rotemberg, Julio J., "Sticky Prices in the United States," *Journal of Political Economy*, December 1982, *90* (6), 1187–1211.

Shapiro, Adam H., "Decomposing Supply- and Demand-Driven Inflation," *Journal of Money, Credit and Banking*, Forthcoming.

Stock, James H. and Mark W. Watson, "Slack and Cyclically Sensitive Inflation," *Journal of Money, Credit and Banking*, 2020, *52* (S2), 393–428.