# Nominal rigidity and the inflation risk premium: identification from the cross section of equity returns

Hengjie Ai, Xinxin Hu, and Xuhui (Nick) Pan*

August 23, 2025

**Abstract:** Inflation risk premium is hard to identify in the data, because inflation induced by real shocks and that by nominal shocks carry risk premiums with opposite signs. We show that in the Calvo model of price rigidity, a firm's exposure to inflation risk—induced by monetary policy—is a monotonic function of its profit margin. Using profit margin sorted portfolios around pre-scheduled FOMC announcements, we identify an inflation risk premium from the cross-section of equity returns that supports the Calvo mechanism of price adjustment. We also develop a continuous-time Calvo model to guide our empirical analysis and provide an explanation for the inflation risk premium observed in the data.

**Keywords:** Inflation risk premium, New Keynesian model, FOMC announcements, cross section of expected returns.

*Hengjie Ai (hengjie.ai@wisc.edu) is from Wisconsin School of Business, University of Wisconsin-Madison. Xinxin Hu (xhu344@wisc.edu) is associated with Wisconsin School of Business, University of Wisconsin-Madison. Xuhui (Nick) Pan (xpan@ou.edu) is from the Price College of Business, The University of Oklahoma. We thank Yan Bai, Leyla Han, Jun Li, Jincheng Tong, and Lai Xu for their helpful comments.

# 1 Introduction

Firms that are subject to nominal rigidity are exposed to inflation risk. In equilibrium, such risk exposure should be compensated on financial markets and should receive an inflation risk premium. The study of this basic relationship between nominal rigidity and inflation risk premium, however, faces two major challenges. First, theoretically, the sign of inflation risk premium is ambiguous and depends on the driving force of inflation shocks. Because monetary policy responds to fundamental shocks, variations in inflation may be driven by expected responses to a productivity (supply) shock, or may be induced by an unexpected monetary policy decision. Depending on the nature of the driving force of the inflation shock, the sign of the inflation risk premium may be positive or negative. Second, because of the above complication, empirically, firms' risk exposure to inflation shocks can also be correlated with risk exposure to fundamental productivity shocks, making it hard to sort portfolios based on inflation beta while controlling for beta with respect to other sources of shocks.

In this paper, we study inflation beta and inflation risk premium by tackling both challenges. Theoretically, we show that, in the standard Calvo (1983) model of price rigidity, the exposure to monetary-policy induced inflation risk is a monotone function of firm-level profit margin. This is a robust conclusion of the Calvo model and does not depend on the details of the specification of monetary policy rules. Guided by this theoretical result, empirically, we sort portfolios right before pre-scheduled FOMC announcements and use high-frequency identification to measure portfolio risk exposure to monetary policy induced inflation shocks over short announcement windows. We verify that consistent with the theory, firms with high profit margins have a higher inflation beta. In addition, the high-minus-low portfolio that takes long positions in high profit margin firms and short positions in low profit margin firms earns a premium of 1.8% during the six-week period before the next FOMC announcement. Also, consistent with the implications of Calvo pricing, this premium decays exponentially over time after portfolio formation as firms adjust away from their previous price levels.

To understand the difficulty in identifying firm's response to monetary policy shocks, consider a New Keynesian model where monetary policy is specified using a simple Taylor rule of the form taken from Gali (2008):

$$i_t = i^* + \phi_\pi \pi_t + \phi_y \hat{y}_t + \nu_t, \tag{1}$$

where $\pi_t$ is inflation in period $t$, $\hat{y}_t$ is deviation of output from the full-employment level, and $\nu_t$ stands for monetary policy shock. Suppose the model has two shocks, a productivity shock $a_t$ and a monetary policy shock $\nu_t$. Suppose also, in equilibrium, $\pi_t = \pi(a_t, \nu_t)$ and $\hat{y}_t = \hat{y}(a_t, \nu_t)$ are both functions of the state variables $(a_t, \nu_t)$. In this case, an innovation in $i_t$ relative to expectation can happen for two reasons. First, a monetary policy shock affects $i_t$ both directly through Equation (1) and indirectly through policy functions $\pi(a_t, \nu_t)$ and $\hat{y}(a_t, \nu_t)$. In this case, shocks to $i_t$ and $\pi_t$ are induced by the unexpected monetary policy decision, $\nu_t$. Second, a surprise productivity $a_t$ can also affect $i_t$ and $\pi_t$ through the policy functions. In the latter case, shocks to inflation and the nominal rate are perfectly correlated with productivity shocks. Empirically, measuring portfolio beta with respect to monetary policy induced inflation shocks $(\nu_t)$, it will be different from beta with respect to monetary policy shocks and will be complicated and contaminated by productivity-shock-induced inflation changes.

We resolve the above problem in two steps. First, we start from a simple model with Calvo pricing. We demonstrate that firm-level profit margin is a sufficient statistic for a firm's risk exposure. In addition, we show that i) firms' risk exposure to monetary policy shock is monotonically increasing in profit margin; and ii) firms' risk exposure to productivity shock is monotonically decreasing in profit margin. We show that our results are robust to different specifications of the monetary policy rules. These results not only provide an explanation for the lack of robust relationship between inflation beta and expected returns in the data, but also generate sharp predictions that we can test empirically.

Our second step is to verify the above theoretical predictions by constructing markup sorted portfolios right before FOMC announcements and examine their beta with respect to inflation shocks over short FOMC announcement windows. Shocks to inflation over short FOMC announcement windows are more likely due to monetary policy surprises rather than productivity shocks. This allows us to measure monetary policy induced inflation beta accurately and test the above theory. We establish three results. First, monetary policy induced beta is monotonically increasing in markups. Second, also consistent with our theory model, CAPM beta is monotonically decreasing in markups. Third, consistent with Calvo pricing, markup sorted portfolios do receive an inflation risk premium after portfolio formation, but the inflation risk premium decays exponentially as firms adjust away from their previous price levels.

**Related literature** This paper contributes to the literature on inflation risk premia in the stock market. Most of the literature finds a time-varying or an ambiguous sign of the inflation risk premium. Boons, Duarte, de Roon, and Szymanowska (2020) uses monthly data to compute covariance between inflation and consumption with rolling regression and presents evidence of a time-varying sign. Fang, Liu, and Roussanov (2023) studies separately the pricing of core and energy inflation on asset prices, and finds a negative beta for core inflation and no significant beta for energy inflation. Cieslak and Pflueger (2023) argues that "good" inflation is linked to demand shocks while the "bad" inflation is to cost-push shocks via a textbook New-Keynesian model. The above evidence is consistent with our theoretical model where productivity shock induced inflation carries a negative risk premium and unexpected monetary conduct induced inflation carries a positive risk premium.

The ambiguity of the sign of inflation risk premium is also reflected in theoretical models of inflation. Structural models of price rigidity typically imply that unexpected inflation is associated with higher output and consumption and lower marginal utility. See, for example, Gali (2008). In contrast, many asset pricing models feature a positive inflation risk premium. For instance, Bansal and Shaliastovich (2013), Kung (2015), and Song (2017) all developed models with a positive inflation risk premium to account for a positive slope of the term structure of interest rates. Bhamra, Dorion, Jeanneret, and Weber (2023) find that inflation can lower the real equity prices because the cash flow is also sticky. Corhay and Tong (2024) demonstrates how inflation shocks can negatively affect the intermediary's equity via reduction in real debt burden.

The discussion of inflation risk premium is closely related to the literature on bond-stock return correlation. David and Veronesi (2013) develops a regime-switching model where inflation news may be positive or negative signals of future economic growth in order to explain stock and treasury bond co-movement. Gourio and Ngo (2020) resorts to the proximity of zero-lower-bound after the Great Recession to explain the changed correlation between stock returns and breakeven inflation rates derived from bond assets. Campbell, Pflueger, and Viceira (2020)connects the change of stock-bond correlation from positive into negative to the switch of sign of the correlation between inflation and the output gap, through a habit formation asset pricing model. Li, Zha, Zhang, and Zhou (2022) models negative consumption-inflation correlations in the monetary regime and positive consumption-inflation correlations in the fiscal regime to reproduce changes in stock-bond returns. We show that the time-varying sign of inflation risk premium is a robust implication of textbook New Keynesian models.

Also related is the literature that studies high frequency identification of monetary policy shocks over short FOMC announcement windows. Bernanke and Kuttner (2005) study stock market reaction to FOMC announcements. Nakamura and Steinsson (2018) document the Fed information effect and provide a high-frequency identification for the non-neutrality of monetary policy. Bauer and Swanson (2023b) estimate the effect on macroeconomy and asset prices from high-frequency changes in interest rate around FOMC announcements. Kekre and Lenel (2022) and Pflueger and Rinaldi (2022) developed structural models where monetary policy affects stock market valuation through the risk premium channel. Hanson and Stein (2015), Hanson, Lucca, and Wright (2021) and Kekre, Lenel, and Mainardi (2024) provide empirical evidence that the shocks to short-term policy rates can pass through the term-structure of treasury yields on FOMC days and attribute the sizable responses from long-term interest rates to heterogenous investors and changes in term premia.

This paper also connects to the literature that studies the connection between nominal rigidity and the cross-section of equity returns. Weber (2015) documents that it takes an annual premium as high as 4% for investors to hold firms which adjust prices more infrequently. Gorodnichenko and Weber (2016) discovers that conditional volatility of firm stock returns rise more after FOMC announcements when the firm's price is stickier. Augustin, Cong, Corhay, and Weber (2024) shows that the inflexible-price firms experience a larger increase in credit spreads upon monetary policy shocks.

The rest of the paper is organized as follows. Section 2 uses a simple static model to illustrate three robust asset pricing implications of the Calvo model. First, markup is a sufficient statistic for firm risk exposure. Second, beta with respect to inflation risk monotonically increases with markup, and third, beta with respect to productivity shock is monotonically decreasing with respect to markup. Section 3 presents empirical evidence by constructing markup sorted portfolios before FOMC announcements and examining its returns over short FOMC announcement windows. Section 4 concludes.

## 2 A textbook NK model

**Setup of the model** In this section, we formulate a standard textbook New Keynesian model with Calvo pricing in continuous time. The Calvo model has a simple form of heterogeneity, because firms that experience different histories of prices adjustments will have different price levels. We focus on the implications of this form of heterogeneity on the cross section of risk exposures and expected returns. The continuous time setup allows us to derive easily interpretable results on risk exposure and risk premium.

Time is infinite and continuous. The representative household maximizes life-time utility subject to a present value budget constraint:

$$\max E\left[\int_0^\infty e^{-\rho t} u\left(C_t, N_t\right) dt\right]$$
$$s.t. E\left[\int_0^\infty \Lambda_t \left(C_t - \frac{W_t}{P_t} N_t - \int \Psi_t\left(i\right) di\right)\right] \leq 0. \tag{2}$$

In the above formulation, $u\left(C_t, N_t\right)$ is the per-period utility flow as a function of consumption and labor supply and $\rho > 0$ is the discount rate. For simplicity, in this section, we focus on the case where $u\left(C, N\right) = \ln C - \frac{1}{1+\zeta} N^{1+\zeta}$. In maximize it life-time utility, the agent takes the stochastic discount factor, $\Lambda_t$, nominal wage, $W_t$, and the aggregate price level $P_t$ as given. The agent has two sources of income: labor income, when measured in real terms, equals $\frac{W_t}{P_t} N_t$, and net profit for owning a continuum of firms, where $\Psi_t\left(i\right)$ is the net profit of firm $i$ at time $t$.

As in standard New Keynesian models, the final goods producer produces the final output, $Y$, by using a continuum of intermediate inputs. At time $t$, the final goods producer maximizes nominal profit by solving the following problem.

$$\max_{Y,\{y_i\}_{i\in[0,1]}} P_t\left(s^t\right)Y-\int p_t\left(s^t,i\right)y_i di$$
$$s.t. Y\leq\left[\int y\left(i\right)^{1-\frac{1}{\eta}}di\right]^{\frac{1}{1-1/\eta}}.$$

The above profit maximization problem generates a demand function for all input varieties $i$: $y\left(\frac{p_i}{P}\middle|Y\right)=\left(\frac{p_i}{P}\right)^{-\eta}Y$.

The producer for variety $i$ is a monopoly in its own market and produces its output using a linear production function: $y_{i,t}=A_t n_{i,t}$, where $n_{i,t}$ is labor input and $A_t$ is aggregate productivity common across all firms. The real profit for a producer of $i$ with price level $p_i$ at time $t$ is

$$\Psi\left(\frac{p_i}{P_t}|w_t,Y_t,A_t\right)=\frac{p_i}{P_t}y\left(\frac{p_i}{P_t}\middle|Y_t\right)-\frac{w_t}{A_t}y\left(\frac{p_i}{P_t}\middle|Y_t\right). \tag{3}$$

Here, $w_t$ is real wage and $A_t$ is aggregate productivity at time $t$. The present value of the total profit of producer $i$ is given by:

$$E\left[\int_0^\infty\Lambda_t\Psi\left(\frac{p_{i,t}}{P_t}|w_t,Y_t,A_t\right)dt\right]. \tag{4}$$

Intermediate goods producers can adjust their price only infrequently. The opportunity of price adjustment arrives at Poisson rate $\lambda$ and the arrival of price adjustment opportunity is independent across firms. We assume the log productivity $a_t=\ln A_t$ follows a continuous-time AR(1) process (an Ornstein-Uhlenbeck process):

$$da_t=-\kappa_a a_t dt+\sigma_a dB_{a,t},$$

where $\kappa_a>0$ is the rate of mean reversion, and $dB_{a,t}$ represents productivity shocks modeled as increments of a Brownian motion.

Denote $\pi_t$ as aggregate inflation $\pi_t=\frac{dP_t}{P_t}dt$.[1] The monetary authority sets interest rates according to a Taylor rule that specifies nominal rate $i_t$ as a linear function of inflation and output gap:

$$i_t=\bar{i}+\phi_\pi\left(\pi_t-\bar{\pi}\right)+\phi_y\left(\ln Y_t-\ln\bar{Y}\right)-\nu_t, \tag{5}$$

[1] We focus on equilibrium in which the aggregate price level is a differentiable function of time.

where $\nu_t$ is monetary shock that is assumed to follow an Ornstein-Uhlenbeck process:

$$d\nu_t = -\kappa_\nu \nu_t dt + \sigma_\nu dB_{\nu,t}.$$

Here, we use a minus sign in front of $\nu_t$ in the specification of Taylor rule (5). This convention implies that a shock to $\nu_t$ can be interpreted as an expansionary inflation shock (as opposed to deflation shock). Formally, an equilibrium is a set of prices, including aggregate price level, real wage, and the stochastic discount factor, $\{P_t, w_t, \Lambda_t\}$, and quantities such that

1. given prices, households maximize life-time utility as in (2),
2. given prices, firms maximize profit (4) by setting optimal prices,
3. market clears, and,
4. interest rates satisfy the Taylor rule (5).

**Markov equilibrium** We look for Markov equilibria where prices and quantities are functions of the Markov state variables $(a, \nu)$. To save notation, we denote $z = (a, \nu)$ as the aggregate state variable whenever convenient. In the cross section, price rigidity leads to dispersion in prices across firms. Let $\chi_{i,t} = \frac{p_i}{P_t}$ denote the relative price level of firm $i$ at time $t$. Let $V(\chi, z)$ be the value function of a sticky price firm with relative price level $\chi$. Standard martingale method implies that $V$ has to satisfy the following HJB equation:[2]

$$\Lambda_t \Psi\left(\chi_t | w\left(z_t\right), Y\left(z_t\right), A_t\right) + \mathcal{L}\left[\Lambda_t V\left(\chi_t, z_t\right)\right] = 0. \tag{6}$$

From the definition of $\Psi$ in (3), we have

$$\Psi\left(\chi_t | w\left(z_t\right), Y\left(z_t\right), A_t\right) = \left[\chi_t^{1-\eta} - \frac{w\left(z_t\right)}{A_t}\chi_t^{-\eta}\right] Y\left(z_t\right).$$

Given the functional form of per-period profit $\Psi$, we guess (and later verify) that the value function takes the form

$$V(\chi, z) = \left[\delta_0(z) + \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}\right] Y(z). \tag{7}$$

[2]The $\mathcal{L}$ operator is defined as $\mathcal{L}[X_t] = \lim_{\Delta\to 0} \frac{1}{\Delta} E_t[X_{t+\Delta} - X_t]$.

Here, the policy functions $\{\delta_0(z), \delta_1(z), \delta_2(z), Y(z)\}$ are only functions of aggregate state variables and do not depend on $\chi$. Given these policy functions, the dependence of the value function $V(\chi, z)$ on firm-specific state variable $\chi$ is in closed form. We will obtain the value function by solving for the functions $\{\delta_0(z), \delta_1(z), \delta_2(z), Y(z)\}$.

Under log utility, $\Lambda_t = e^{-\rho t} Y_t^{-1}$. We can simplify (6) as

$$e^{-\rho t}\left[\chi_t^{1-\eta} - \frac{w(z_t)}{A_t}\chi_t^{-\eta}\right] + \mathcal{L}\left[e^{-\rho t}\left\{\delta_0(z_t) + \delta_1(z_t)\chi_t^{1-\eta} - \delta_2(z_t)\chi_t^{-\eta}\right\}\right] = 0,$$

or, equivalently,

$$\rho\left[\delta_0(z) + \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}\right] = \left[\chi^{1-\eta} - \frac{w(z)}{A}\chi^{-\eta}\right] + \mathcal{L}\left[\delta_0(z) + \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}\right]. \tag{8}$$

In our setup, $\chi_{i,t} = \frac{p_i}{P_t}$. Under the assumption of Calvo pricing $\chi_t$ adjusts at Poisson arrival times. Because the arrival of price adjustment opportunity is independent over time, the decision problem for all flexible price firms are identical. Let $\hat{\chi}(z_t)$ denote the optimal pricing decision as function of aggregate state variable $z_t$. The law of motion of $\chi_{i,t}$ can be described as follows. Before a price adjustment opportunity arrives, $p_i$ stays constant and $P_t$ changes at rate $\pi_t$. As a result, $d\chi_{i,t} = -\chi_{i,t}\pi_t dt$. At rate $\lambda$, $\chi_{i,t}$ adjusts discontinuously to its optimal level, $\hat{\chi}(z_t)$. This dynamics of $\chi_{i,t}$ allows us to write:

$$\mathcal{L}\left[\delta_1(z)\chi^{1-\eta}\right] = \mathcal{L}\left[\delta_1(z)\right]\chi^{1-\eta} - (1-\eta)\delta_1(z)\chi^{1-\eta}\pi + \lambda\delta_1(z)\left[\hat{\chi}^{1-\eta}(z) - \chi^{1-\eta}\right],$$

and

$$\mathcal{L}\left[\delta_2(z)\chi^{-\eta}\right] = \mathcal{L}\left[\delta_2(z)\right]\chi^{-\eta} + \eta\delta_2(z)\chi^{-\eta}\pi + \lambda\delta_2(z)\left[\hat{\chi}^{-\eta}(z) - \chi^{-\eta}\right].$$

The HJB equation (8) therefore takes the form of

$$\begin{aligned}\rho\left[\delta_0(z) + \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}\right] = \chi^{1-\eta} - \frac{w(z)}{A}\chi^{-\eta} + \mathcal{L}\left[\delta_0(z)\right] \\ +\mathcal{L}\left[\delta_1(z)\right]\chi^{1-\eta} - (1-\eta)\delta_1(z)\chi^{1-\eta}\pi + \lambda\delta_1(z)\left[\hat{\chi}^{1-\eta}(z) - \chi^{1-\eta}\right] \\ -\left\{\mathcal{L}\left[\delta_2(z)\right]\chi^{-\eta} + \eta\delta_2(z)\chi^{-\eta}\pi + \lambda\delta_2(z)\left[\hat{\chi}^{-\eta}(z) - \chi^{-\eta}\right]\right\}.\end{aligned}$$

Matching coefficients on both sides of the above equation, it is clear that $\delta_0$, $\delta_1$, and $\delta_2$ must satisfy the following HJB equations,

$$\rho\delta_0(z) = \mathcal{L}[\delta_0(z)] + \lambda\left[\delta_1(z)\hat{\chi}^{1-\eta}(z) - \delta_2(z)\hat{\chi}^{-\eta}(z)\right], \tag{9}$$

for $\delta_0$,

$$(\rho + \lambda + (1-\eta)\pi_t)\,\delta_1(z) = 1 + \mathcal{L}[\delta_1(z)], \tag{10}$$

for $\delta_1$ and

$$(\rho + \lambda - \eta\pi)\,\delta_2(z) = e^{-a_t}w(z) + \mathcal{L}[\delta_2(z)], \tag{11}$$

for $\delta_2$.

Given the value function, the optimal choice of price satisfies $\hat{\chi}(z_t) = \arg\max\left\{\delta_1(z_t)\chi^{1-\eta} - \delta_2(z_t)\chi^{-\eta}\right\}$ which implies that

$$\hat{\chi}(z_t) = \hat{\mu}\frac{\delta_2(z_t)}{\delta_1(z_t)}, \tag{12}$$

where $\hat{\mu} = \frac{\eta}{\eta-1}$ is the optimal markup under flexible price. We summarize the above solution for firm's optimal price adjustment problem in the following lemma.

**Lemma.** *The firm value function takes the form of (7), where $\delta_1$ and $\delta_2$ are defined by (10) and (11), respectively. $\delta_0$ is given by:*

$$\rho\delta_0(z) = \mathcal{L}[\delta_0(z)] + \frac{\lambda}{\eta-1}\hat{\mu}^{-\eta}\delta_1(z)^{\eta}\,\delta_2(z)^{1-\eta}.$$

*Given the value function, the optimal pricing decision is given by (12).*

Under Calvo pricing, the consistency between individual price and aggregate price level implies that the firm-level pricing decision (12) and aggregate inflation must satisfy a consistency condition. We summarize this relationship in the following lemma.

**Lemma.** *Let $\hat{p}_t$ be the optimal pricing decision of a flexible-price firm at time $t$. Then, $\frac{\hat{p}_t}{P_t}$ must satisfy*

$$\frac{\hat{p}_t}{P_t} = \left[1 - \frac{\eta-1}{\lambda}\pi_t\right]^{\frac{1}{1-\eta}}, \tag{13}$$

*where $\pi_t$ is the rate of inflation.*

*Proof.* Under the CES production technology, aggregate price level $P_t$ and individual price $p_i$ are related by $P = \left[\int P(i)^{1-\eta}\, di\right]^{\frac{1}{1-\eta}}$. Under Calvo pricing, over a small time interval $(t, t+\Delta)$, a fraction $\lambda\Delta$ of firms adjust to the optimal price level $\hat{p}_{t+\Delta}$. This implies that the dynamics of price must satisfy:

$$P_{t+\Delta}^{1-\eta} = (1-\lambda\Delta)\, P_t^{1-\eta} + \lambda\Delta \hat{p}_{t+\Delta}^{1-\eta}.$$

Taking the continuous time limit, the above equation is written as $dP_t^{1-\eta} = \lambda \left(\hat{p}_t^{1-\eta} - P_t^{1-\eta}\right) dt$. Or, in $dP_t$ terms,

$$dP_t = \frac{1}{(1-\eta)\, P_t^{-\eta}} \lambda \left(\hat{p}_t^{1-\eta} - P_t^{1-\eta}\right) dt = \frac{\lambda}{1-\eta} \left(\hat{p}_t^{1-\eta} P_t^{\eta} - P_t\right) dt.$$

By definition, $\pi_t = \frac{dP_t}{P_t dt}$ must satisfy:

$$\pi_t = \frac{\lambda}{\eta - 1} \left(1 - \left(\frac{\hat{p}_t}{P_t}\right)^{1-\eta}\right), \tag{14}$$

which is equation (13). □

Motivated by the above lemma, we define

$$\chi^*(\pi) = \left[1 - \frac{\eta - 1}{\lambda}\pi\right]^{\frac{1}{1-\eta}}.$$

Intuitively, $\chi^*(\pi)$ is the level of relative price of a flexible price firm that is needed to implement an inflation level $\pi$. In equilibrium, $\chi^*(\pi)$ must be consistent with firms' optimal price setting decision (12). As a result, equilibrium inflation has to satisfy:

$$\hat{\mu} \frac{\delta_2(z_t)}{\delta_1(z_t)} = \chi^*(\pi(z_t)). \tag{15}$$

To completely characterize the equilibrium, we note that the households' optimal choice of labor implies $\frac{N^{\zeta}}{C^{-1}} = w$. Using the aggregate production function, $C = AN$, we can obtain aggregate labor supply $N(z_t) = \left[\frac{w(z_t)}{A_t}\right]^{\frac{1}{1+\zeta}}$ and aggregate consumption $C(z_t) = A_t^{\frac{\zeta}{1+\zeta}} w(z_t)^{\frac{1}{1+\zeta}}$ as functions of aggregate state variables. In addition, the household intertemporal optimality implies that the real interest rate must be consistent with the stochastic discount factor: $r_t = \rho - \frac{\mathcal{L}[C_t^{-1}]}{C_t^{-1}}$. Combining the above conditions with the Taylor rule, we have:

$$\bar{i} + (\phi_\pi - 1)\,\pi(z_t) + \phi_y \left[\frac{\zeta}{1+\zeta} a_t + \frac{1}{1+\zeta}\left(\ln w(z_t) - \bar{w}\right)\right] - \nu_t = \rho - \frac{\mathcal{L}\left[A_t^{-\frac{\zeta}{1+\zeta}} w(z_t)^{-\frac{1}{1+\zeta}}\right]}{A_t^{-\frac{\zeta}{1+\zeta}} w(z_t)^{-\frac{1}{1+\zeta}}}. \tag{16}$$

To summarize, a Markov equilibrium can be constructed by a pair of policy functions, $w(z_t)$ and $\pi(z_t)$ that jointly satisfy the functional equational equations (15) and (16), where $\delta_1(z)$ and $\delta_2(z)$ are constructed from recursions (10) and (11), respectively.

**Risk exposure** In the above model, aggregate fluctuations can be driven by two sources of shocks, the productivity shock $B_{a,t}$ and the monetary policy shock $B_{\nu,t}$. In our formulation, a positive shock to $\nu_t$ lowers the nominal rate and is expansionary. As a result, the signs of both shocks are positive in the sense that states with higher values of $a$ and $\nu$ are both high consumption and low marginal utility states. We summarize this observation in the following proposition.

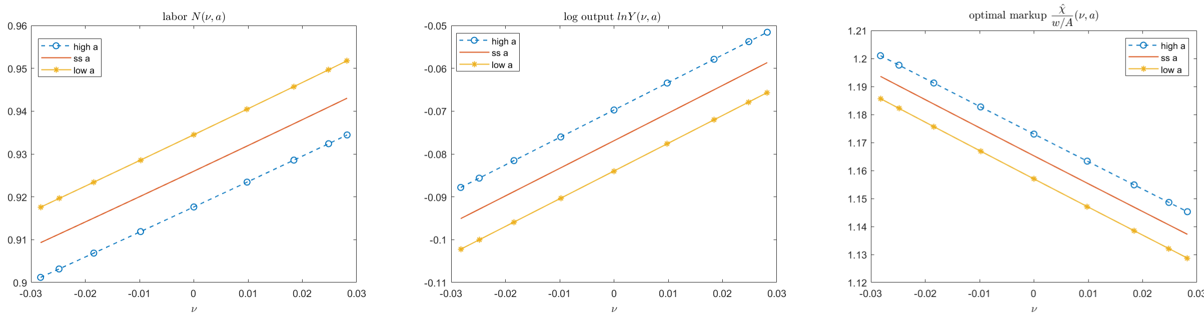
**Proposition 1.** *(Sign of shocks) Suppose the monetary policy is such that $\phi_\pi > 1$ and $\phi_y > 0$, then $\frac{\partial \ln Y(a,\nu)}{\partial a} > 0$ and $\frac{\partial \ln Y(a,\nu)}{\partial \nu} > 0$, where derivatives are evaluated at the deterministic steady state.*

*Proof.* See appendix. □

By the above proposition, aggregate output is increasing in both productivity shocks and monetary policy shocks. In this sense, the monetary policy, $\nu$ is expansionary. In Figure 1, we plot the policy functions of macroeconomic quantities, including labor supply ($N(a,\nu)$, left panel), log output ($\ln y(a,\nu)$, middle panel), and optimal markup ($\frac{A\hat{\chi}(a,\nu)}{w(a,\nu)}$, right panel) as functions of monetary policy shock, $\nu$, on the horizontal axis for three levels of aggregate productivity, high productivity (in circles), medium productivity (solid line), and how productivity (stars). These policy functions are obtained by using a global solution method which we detail in the appendix.

Focusing on the state variable $a$, equilibrium labor supply $N(a,\nu)$ is decreasing in productivity $a$, output is an increasing function of $a$, and the optimal markup for flexible price

Figure 1: Macroeconomic quantities




firms is also an increasing function of $a$. A higher productivity impacts labor supply through both the substitution effect and the income effect. In our model, the income effect dominates and labor supply drops upon a positive productivity shock. Despite the drop in labor supply, the overall impact of productivity on output is positive, and $Y(a, \nu)$ is an increasing function of $a$. The optimal markup is an increasing function of productivity because $a$ is a mean-reverting process. A high current productivity is often associated with a drop in productivity in the future. Because prices are sticky, lower productivity is associated with lower markups. Anticipating this future dynamics of markups, a flexible price firm optimally choose a higher current markup to offset the drops in future markups.

Turning to the impact of the monetary policy state variable $\nu$, we note that both labor supply and total output are increasing in $\nu$. This is a common implication in standard New Keynesian models. Nominal wage increases upon a positive shock to inflation. Because aggregate price is sticky, real wage rises and labor supply increases. The increase in labor supply simultaneously raises output. As shown in Proposition 1, output is an increasing function of $\nu$. Finally, we note that optimal markup is a decreasing function of $\nu$. Because $\nu$ is a mean-reverting process, a high current inflation is associated with anticipated drops in aggregate price and wage level in the future. Nominal rigidity implies a firm's markup will rise upon an drop in real wage. Because price setting decisions are forward looking, a flexible price firm optimally lowers current period markup in anticipation of this future dynamics of markups.

Innovations in inflation in our model is also driven by two sources of shocks, productivity shock $B_{a,t}$ and monetary policy shock, or unexpected monetary policy conduct, $B_{\nu,t}$. As we explain in the introduction of the paper, if the monetary authority follows a counter-cyclical monetary policy, that is, the monetary policy rule is to lower the nominal rate following a negative shock to output, $\phi_y > 0$, then productivity-induced inflation happens in states with low productivity. At the same time, because $\nu$ represents expansionary shocks, high monetary policy shock induced inflation is associated with positive surprises in total output. The following proposition formalizes the above intuition.

**Proposition 2.** *(Inflation beta) Let $\pi(a,\nu)$ denote the policy function in the Markov equilibrium. Suppose the monetary policy is such that $\phi_\pi > 1$ and $\phi_y > 0$, then $\frac{\partial}{\partial a}\pi(a,\nu) < 0$ and $\frac{\partial}{\partial \nu}\pi(a,\nu) > 0$, where derivatives are evaluated at the deterministic steady state.*

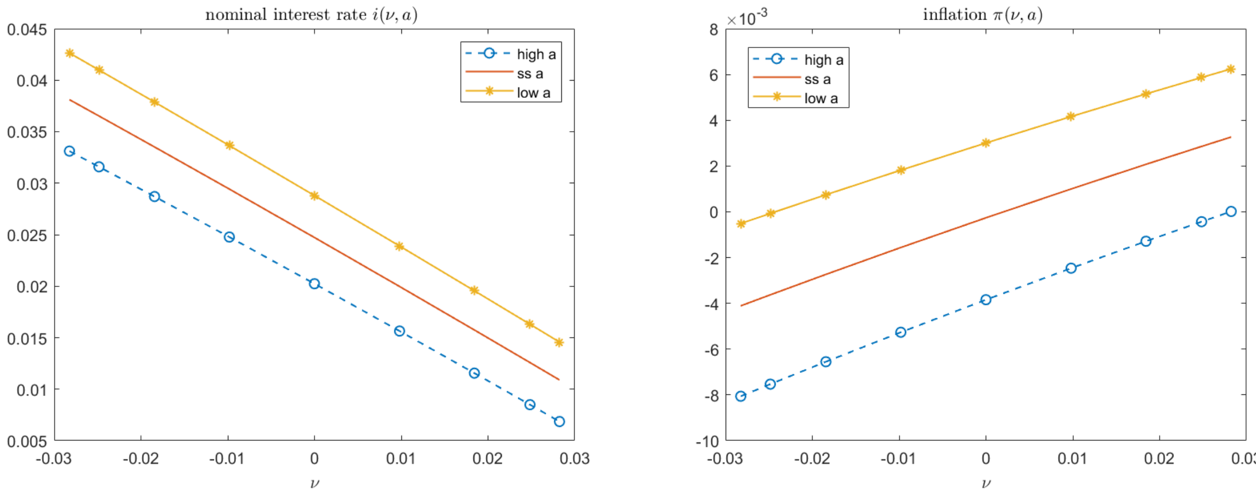
*Proof.* See appendix. □

Figure 2 plots inflation (left panel) and nominal interest rate (right panel) as functions of $\nu$ (horizontal axis) for three levels of productivity $a$. A higher productivity is associated with a lower inflation and a lower nominal rate. Because the level of productivity is mean reverting, a higher level of current productivity is associate a lower expected growth rate of output and consumption, and as a result, a lower real and nominal interest rate. Consistent with Proposition 2, inflation is a decreasing function productivity. Because the policy rule is to lower interest rate and raise inflation upon a negative output gap, productivity-induced innovations in inflation happens in low-productivity states (recessions).

The impact of monetary policy shock in our setup is also consistent with standard New Keynesian models. Positive shocks to $\nu$ is expansionary. As a result, higher values of $\nu$ are associated with lower nominal rates and higher inflations. It is clear that monetary policy shock induced inflation is associated with increases in total output and s in marginal utility.

Because both the productivity shock $a$ and the monetary policy shock $\nu$ are associated with high consumption and low marginal utility (Proposition 1), Proposition 2 implies that productivity induced inflation is counter-cyclical while monetary policy shock induced inflation is associated with positive surprises in consumption and output. This is a general feature of models with price rigidity and highlights the difficulty in identifying inflation risk premium in the data. Depending on the driving force of inflation, high inflation may be associated with high or low marginal utilities of the representative investor. If monetary policy is the main driving force of inflation, then inflation risk premium is positive. In contrast, if productivity shock is the most important driving force of inflation, then inflation risk premium is negative.

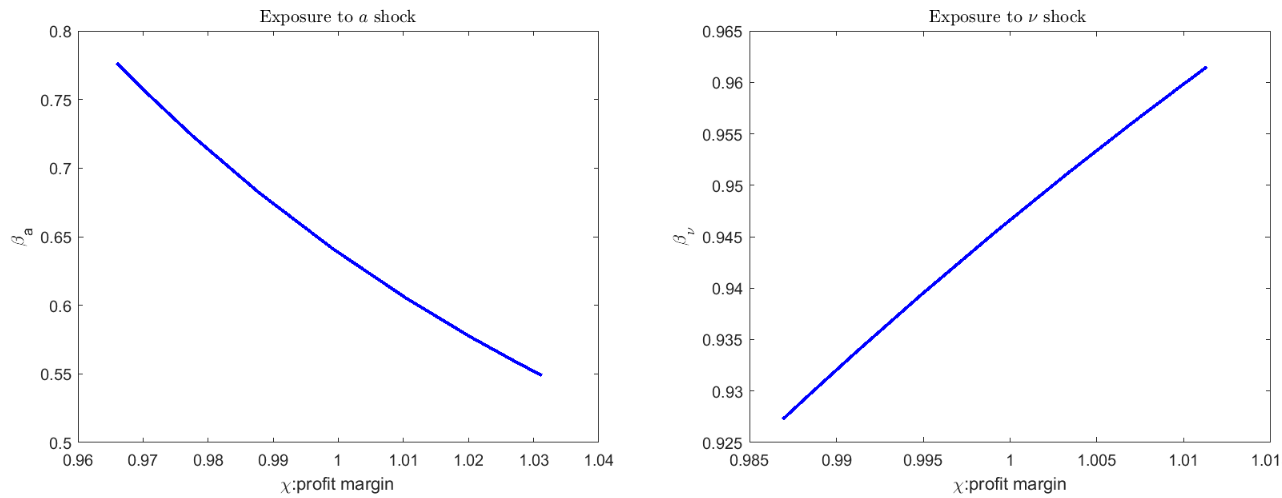
Figure 2: Inflation and nominal interest rate




In the data, inflation can be driven by both a productivity shock $a$ and a monetary policy shock $\nu$. Productivity induced inflation shock, strictly speaking, does not qualify as an inflation shock. It represents a surprise in inflation, but it is just an equilibrium response to the exogenous productivity shock $a$. The monetary policy shock $\nu$, however, represents a shock, because it is due to unexpected monetary policy conduct, rather than shocks that arise from the real side of the economy. To understand inflation risk premium, we are primarily interested in the second type of shocks, that is inflation risk driven by unexpected monetary policy conduct. The difficulty to identify inflation risk premium is that in the data, variations in inflation is not only driven by unexpected monetary policy conduct, $\nu$, but can also be confounded by endogenous responses of inflation to productivity shocks, which carries a risk premium with the opposite sign.

The key to our identification exercise is to use information from the cross section to identify the inflation risk premium. To understand the cross section of inflation risk premium, we now turn to firm value function and its implies inflation risk premium. Firm value function in our model as shown in Equation (7) has intuitive interpretations. The term $\delta_0(z)$ is the present value of cash flow obtained after the next price adjustment. This term is common across all firms and does not depend on $\chi$. Define the sum of the second and the third term in (7) as $\bar{V}(\chi, z) = [\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}]Y(z)$ as the present value of cash flow for the firm obtained before the next price adjustment. In our model, firm heterogeneity

Figure 3: Elasticities with respect to productivity and monetary policy shocks




is summarized by the state variable $\chi$ and the dependence of the value function on $\chi$ is completely summarized by $\bar{V}(\chi, x)$. To understand the heterogeneity in risk exposure, we focus on value function normalized by total output: $\bar{v}(\chi, z) = \frac{\bar{V}(\chi,z)}{Y(z)}$.

**Proposition 3.** *(Markup and inflation beta) Suppose the monetary policy is such that* $\phi_\pi > 1$ *and* $\phi_y > 0$*, then* $\frac{d \ln \bar{v}(\chi,a,\nu)}{d\nu}$ *is an increasing function of* $\chi$ *and* $\frac{d \ln \bar{v}(\chi,a,\nu)}{da}$ *is a decreasing function of* $\chi$*, where the derivatives are evaluated at the deterministic steady state.*

*Proof.* See appendix. □

We plot the above elasticities using the global solution from the model in Figure 3. The left panel is the elasticity of firm value with respect to productivity shocks and the right panel is the elasticity of firm value with respect to monetary policy shocks, both as functions of the state variable $\chi$. Clearly, high markups are associated with a lower risk exposure with respect to productivity shocks (left panel) and a higher exposure to shocks to unexpected monetary policy conduct (right panel).

In the cross section, $\chi_{i,t} = \frac{p_{i,t}}{P_t}$ is a monotone function of markup or profit margin, $\frac{p_{i,t}}{P_t(w_t/A_t)}$. As a result, Proposition 3 implies that among all sticky price firms, firms with a high markup has a high $\beta$ with respect to unexpected monetary policy conduct-induced inflation shocks but a lower sensitivity with respect to productivity shocks. This is the key implication of the model that forms the basis of our identification exercise. Sorting portfolios by their inflation $\beta$

does not identify risk exposure with respect to unexpected monetary policy conduct, because in the data, variations in inflation are driven by both productivity shocks and monetary policy shocks. For the same reason, naturally inflation risk prone assets, such as treasury notes, do not necessarily produce an inflation risk premium compared to their inflation-protected counterparts, such as treasury inflation protected securities. Our model provides a theory-based identification from the cross section of stock returns: because $\beta$ with respect to monetary policy shock induced inflation is increasing markups, markup sorted portfolios should differ in their inflation risk exposure and earn an inflation risk premium.

The intuition for the above proposition can be explained from the form of the value function: $\bar{V}(\chi, z) = [\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}]Y(z)$. Under the CES production function, final goods producer's demand function $y(\chi_i|Y) = \chi_i^{-\eta}Y$ is proportional to $\chi_i^{-\eta}$. As a result, revenue $\frac{p}{P}y$ is proportional to $\chi^{1-\eta}$ and cost $\frac{wy}{A}$ is proportional to $\chi^{1-\eta}$ for a stick price firm. The first term $\delta_1(z)\chi^{1-\eta}Y(z)$ represents the present value of revenue and the second term, $\delta_2(z)\chi^{-\eta}Y(z)$ represents the present value cost. Interpreted this way, the difference $\bar{v}(\chi, z) = \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}$ is the present value of markups. Upon a monetary policy shock, the value of all firms increases, and $Y(z)$ represents the common impact of aggregate shocks on all firms. Firm heterogeneity is completely captured by the term $\bar{v}(\chi, z) = \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}$. For a sticky price firm, a monetary policy shock that triggers inflation and raises aggregate price level increases economy-wide real wage. Due to price rigidity, the markup for stick price firms, $\bar{v}(\chi, z) = \delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}$, drops. The drop in markup partially offsets the initial increase in output and acts as an insurance against inflations shocks.

The above insurance mechanism differs across firms. Because risk exposure (beta of return with respect to shocks) depends on the percentage change of firm value in response of shocks, the magnitude of the insurance mechanism is high for low markup firms and relatively small for high markup firms. For low-markup firms, the term $\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}$ is close to zero. As a result, a small change in inflation represents a large percentage in markup. In fact, this insurance mechanism can completely offsets the initial positive impact of expansionary monetary shock, resulting in a negative risk exposure for low markup firms with respect to monetary policy shocks. For high markup firms, because the present value of markup, $\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}$ is high, a small inflation results in a small drop in markup in percentage terms and the insurance mechanism is almost negligible. As a result, high markup firms have a high overall risk exposure with respect to inflation shocks.

In our model, markup (price divided by marginal cost) and profit margin (total profit divided by total cost) are the same. Thanks to the linear production function and the absence of fixed cost, marginal cost in our model is the same as the average cost. It is clear from the above discussion, under more general assumptions on the production function, the magnitude of the insurance mechanism discussed above depends on profit margin. For robustness, our empirical portfolio sorting exercises uses both measures of markup and measures of profit margin.

Proposition 3 ranks the risk exposure of all sticky price firms and does not apply to flexible price firms that have an opportunity to adjust their prices. The following proposition demonstrates that in terms of the ranking of risk exposure, price stickiness is irrelevant: markup is a sufficient statistic for inflation beta regardless of price stickiness.

**Proposition 4.** *(Irrelevance of sticky price) Let $V^F(a,v)$ be the value function of flexible price firms that adjust price optimally. Then* $\frac{\partial \ln V^F(a,\nu)}{\partial \nu} = \left.\frac{\partial \ln V(\chi,a,\nu)}{\partial \nu}\right|_{\chi=\hat{\chi}(a,\nu)}$.

*Proof.* The above Proposition is a consequence of the envelop theorem. The proof is simple and intuitive, so we provide it here. Because flexible price firms choose their price optimally,

$$V^F(a,v) = \max_{\chi} V(\chi,a,\nu) = V(\hat{\chi}(a,\nu),a,\nu),$$

where $\hat{\chi}(a,\nu)$ is the optimal pricing function. As a result, $\frac{\partial V^F(a,\nu)}{\partial \nu} = \frac{\partial V(\hat{\chi}(a,\nu),a,\nu)}{\partial \nu} + \left.\frac{\partial V(\chi,a,\nu)}{\partial \chi}\right|_{\chi=\hat{\chi}(a,\nu)} \times \frac{\partial \hat{\chi}(a,\nu)}{\partial \nu}$, but $\left.\frac{\partial V(\chi,a,\nu)}{\partial \chi}\right|_{\chi=\hat{\chi}(a,\nu)} = 0$ due to the first-order optimality condition. □

Thanks to the above proposition, markup is a sufficient statistic for risk exposure to monetary policy shocks among all firms, including sticky price firms and flexible price firms. This implication of the model allows us to sort portfolios by markups and identify risk exposure to monetary policy shocks.

In this section, we have used to the term profit margin and markup interchangeably. In the model, there is no distinction between profit margin (price divided by average cost) and markup (price divided by marginal cost), because of the constant return to scale technology and because of the lack of overhead cost. In fact, to be consistent with the language of the New Keynesian literature, we adopted the term markup in most of this section. In the data, however, profit margin and markup can differ. In our empirical exercise, we will primarily use profit margin and only markup as a robustness check for two reasons. First, an inspection of the proof of Proposition 3 implies that the main mechanism for the monotonicity of risk exposure with respect to $\chi$ is due to a leverage effect that is directly related to profit margin, not markup. A small inflation can have a large impact on firms' profit if profit margin is low. A high markup firm with a low profit margin will still be severely impacted by an inflation shock due the rise of average cost. Second, profit margin depends on average cost is much easier to measure in the data than markups, which is a function of marginal cost. Despite all of these complications, one should expect that profit margin and markup are highly correlated in cost section. As we show below, our empirical evidence is robust to both measures.

# 3 Empirical evidence

To operationalize the above identification strategy, we use the cross section of equity returns to identify inflation risk exposure and inflation risk premium. As we explain earlier, in the data, typically, monetary policy induced variations in inflation can be confounded by monetary authority's endogenous responses to productivity shocks. However, over a short FOMC announcement window, changes in inflation and inflation expectations can only be due to monetary policy shock and not to productivity shocks under the assumption of rational expectations. Building on this intuition, our empirical exercises consist of two main components. First, we sort portfolios based on profit margin (and markup) and examine their inflation beta on FOMC announcement days. Second, we hold the portfolios until the next FOMC announcement to study their inflations risk premium.

**Construction of profit margin sorted portfolios.** The first step of our exercise is to construct profit margin sorted portfolios. Our data sample for stock return consists of all NYSE, AMEX, and NASDAQ listed common stocks that are included in both CRSP and Compustat. Daily stock returns are from CRSP and we exclude those stocks with closing prices lower than $5 per share. Our main measure of profit margin is defined as operating income divided by sales following Fairfield and Yohn (2001) and Jansen, Ramnath, and Yohn (2012). We construct each individual firm's profit margin using the Compustat database and we exclude financial and utility firms in all analyses. As a robustness analysis, we also construct each individual firm's markup by following De Loecker, Eeckhout, and Unger (2020). We report empirical findings based on markup sorted portfolios in the appendix and our results are consistent with the empirical evidence when using profit margin.

In our model, the only source of variation of profit margin is due to price rigidity. In reality, heterogeneity in profit margin may to due to many different reasons, such as difference in technology and markup competitiveness. To avoid these complications, we sort portfolios based on profit margin within narrowly defined industries, which presumably have similar technology and market competitiveness. In portfolio formation, we rank each firm's profit margin relative to other firms in the same industry, defined as the four-digit SIC code. For example, a firm is assigned to the bottom (top) portfolio if its profit margin is lower (higher) than eighty percent (80%) of firms in the same industry.

To examine $\beta$ with respect to identified monetary policy induced inflation shocks, we form portfolios right before pre-scheduled FOMC announcement days. We focus on pre-scheduled FOMC announcement days spanning from 1994 to 2023, with 239 FOMC announcements in total. Before each pre-scheduled FOMC announcement day, we sort stocks into quintile portfolios based on their intra-industry profit margin. We then hold these portfolios until two-days before the next FOMC announcement and record value-weighted portfolio returns. We repeat this procedure and resort stocks into portfolios before each FOMC announcement days for our entire data period.

**Inflation beta.** We use changes in inflation swaps on FOMC announcement days to measure innovations of inflation induced by unexpected monetary policy conduct. Inflation swaps are derivative contracts where one party agrees to exchange fixed payments for floating payments tied to the inflation rate, for a given notional amount and period of time. Market participants use inflation swaps to hedge inflation risk and (or) to speculate on the course of inflation in the future. Therefore, inflation swaps provide timely real market-based proxy for inflation expectations. An increase in an inflation swap index indicates a positive innovation in investors' expectation of inflation for inflation by the maturity of the swap contract. Our identification assumption is that under rational expectations, the response of monetary policy with respect to macroeconomic productivity shock must already be priced in all inflation swap contracts right before FOMC announcements. Innovations in the inflation swap indices during a short FOMC announcement window must therefore reflect unexpected monetary policy conduct by the monetary authority. We obtain daily inflation swap data from Bloomberg. The data start from July 21, 2004 and end in 2023.

To measure the impact of inflation on asset prices, we focus on inflation swaps with the shortest maturity, which is one year. In our model, because pricing decisions are forward looking, realized inflation and inflation expectation can have drastically different implications on firm policy and firm valuation. As we explain earlier, an unexpected inflation raises real wage and labor supply and represents an expansionary shock. This is about realized inflation and not expected future expectation. In fact, holding current inflation constant, a rise in expected long-term inflation in the future lowers the market clearing level of real wage and has the opposite effect on labor supply and total output. We therefore carefully choose the shortest maturity inflation swaps to isolate the impact of inflation, which is the key economic mechanism of the paper, from the impact of long-run inflation expectations. In the data, the standard deviation of changes in one-year tenor inflation swaps is at least two times higher than that of longer tenor inflation swaps (e.g., 20- or 30-yr tenor) on FOMC days. This indicates that realized short-term inflation can be quite a bit more volatile than long-run inflation expectations. The higher volatility of the shortest term inflation swap is consistent with the fact that it is good measure of realized short term inflation rather than expected long-run inflation.

Table 1 reports the results when we regress the excess portfolio returns on daily changes in one-year tenor inflation swaps and market excess returns on FOMC days. In Panel A, we use observed inflation swap change on FOMC announcement days as the measure of monetary policy induced inflation shocks. Profit margin sorted portfolio returns monotonically react to inflation expectation changes induced by monetary policy announcements. High profit margin firms are more sensitive to inflation risk as evidenced by the regression coefficients. When investors expect a higher inflation rate after the FOMC announcements, high profit margin firms strongly react with a positive coefficient while low profit margin firms experience an decrease in returns. The inflation $\beta$ of the long-short portfolio is statistically significant. In addition, there is a monotonic decreasing pattern of the market $\beta$ for profit margin sorted portfolios, consistent with our model prediction on the relationship between profit margin and $\beta$ with respect to productivity shocks (Proposition 3).

Recent literature, such as Bauer and Swanson (2023a) and Bauer and Swanson (2023b)

Table 1: Inflation Beta of Profit Margin-Sorted Portfolios

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| Panel A: Inflation Swap | | | | | | |
| Inflation β | -2.1815 | -0.3466 | -0.2305 | -0.1807 | 0.1167 | 2.3029 |
| | | | | | | (2.24) |
| Market β | 1.1134 | 1.0771 | 1.0071 | 0.9744 | 0.9049 | -0.2080 |
| | | | | | | (-4.45) |
| Panel B: Orthogonalized Inflation Swap | | | | | | |
| Inflation β | -2.1983 | -0.4491 | -0.3605 | -0.2079 | 0.1997 | 2.4041 |
| | | | | | | (2.37) |
| Market β | 1.1161 | 1.0785 | 1.0087 | 0.9749 | 0.9039 | -0.2116 |
| | | | | | | (-4.50) |

This table reports the coefficients when we regress profit margin sorted portfolio excess returns on changes in one-year tenor inflation swap and market excess returns on FOMC days. In Panel A, we use the daily changes in one-year tenor inflation swaps. In Panel B, we use the daily changes in one-year tenor inflation swaps orthogonalized by log changes of core CPI, unemployment rate, and GDP during the last three months before each FOMC announcements. Our regression sample covers from July 2004 to December 2023 with 154 pre-scheduled FOMC announcements. Newey-West $t$-statistics with 6 lags are reported in parentheses.

has shown that measured "monetary policy surprises" over short FOMC windows can be correlated with public macroeconomic information that is available before the FOMC announcements. To address the concern that part of the measured innovations in inflation swaps on FOMC announcement days may be expected by the market, we follow Bauer and Swanson (2023b) and compute the component of inflation swap changes on FOMC days that is orthogonal to macroeconomic news that precedes the announcements. We first regress daily changes in inflation swap on macroeconomic news, including the latest changes in core CPI, unemployment rate, and GDP during the last three months before each FOMC announcements. We use the regression residual that is orthogonal macroeconomic news as an alternative measure of inflation surprises induced by monetary policy announcements. Panel B of Table 1 reports the results when we regress profit margin sorted portfolio returns on the orthogonalized inflation swap changes and market excess returns. The result that inflation $\beta$ monotonically increases with profit margin and market $\beta$ monotonically decreases with profit margin are robust to this alternative construction of inflation surprises.

**Inflation risk premium** Having identified portfolios that are monotone with respect to exposure to inflation risk (induced by unexpected monetary policy conduct), the next step of our analysis is to examine the existence and the magnitude of inflation risk premium. We hold the profit margin sorted portfolio until the next FOMC announcement and report in Table 2 the time series average of expected portfolio returns, their alpha from the CAPM, that from the Fama-French three-factor model and the same from the five-factor model.[3]

Table 2: Expected Returns of Profit Margin-Sorted Portfolios

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| Return (%) | 0.3933 | 1.1714 | 1.3218 | 1.2763 | 1.5912 | 1.1979 |
| | (0.64) | (2.50) | (3.11) | (3.79) | (4.71) | (3.14) |
| CAPM $\alpha$ (%) | -1.3682 | -0.4594 | -0.0844 | 0.0029 | 0.3922 | 1.4829 |
| | (-3.40) | (-2.74) | (-0.60) | (0.03) | (3.33) | (3.96) |
| FF3 $\alpha$ (%) | -1.1697 | -0.3566 | -0.0462 | 0.0443 | 0.4391 | 1.3328 |
| | (-3.74) | (-2.67) | (-0.34) | (0.41) | (5.02) | (4.11) |
| FF5 $\alpha$ (%) | -0.6637 | -0.2097 | 0.0286 | -0.0234 | 0.3818 | 0.7693 |
| | (-3.14) | (-1.46) | (0.26) | (-0.20) | (4.36) | (3.39) |

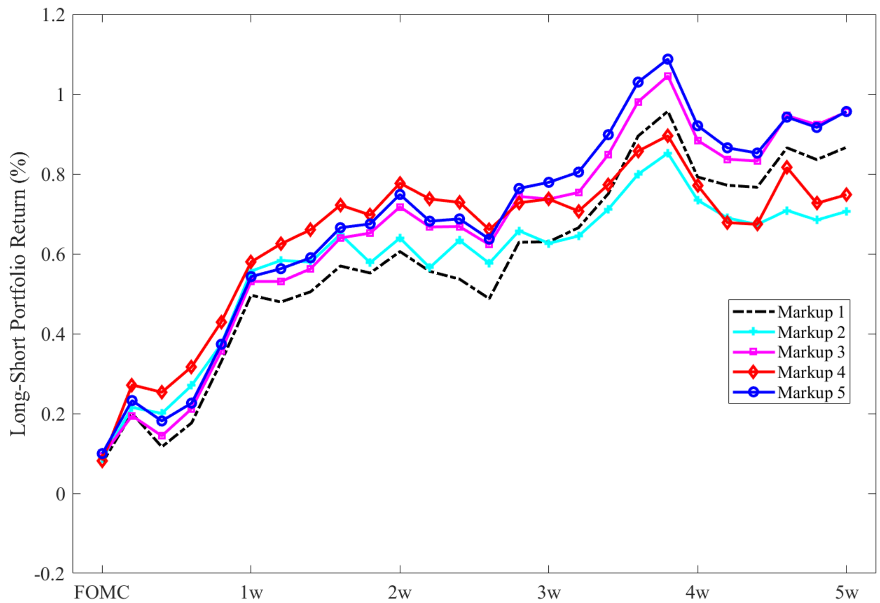
This table reports expected returns when we sort stocks based on intra-industry profit margin before each pre-scheduled FOMC announcement days. We hold portfolios till two-days before the next pre-scheduled FOMC announcements. We also report alphas from the CAPM, Fama-French three-factor and five-factor models. Newey-West $t$-statistics with 6 lags are reported in parentheses. The data sample is from 1994 to 2023.

We observe that high profit margin firms earn higher average returns than low profit margin firms do after the FOMC announcements. The long-short portfolio generates a highly significant expected return of 1.20% on average during the holding period, which is about 9.52% when accumulated for a year since there are eight pre-scheduled FOMC announcements each year. The alpha from the CAPM is 1.48% per FOMC cycle and 11.8% (1.48%×8) per year. The alpha computed from the Fama-French three-factor model is 10.64% (1.33%×8) per year and is also monotonic in profit margins. The long-short portfolio delivers a Fama-French five-factor alpha of 6.15% per year and remains highly significant

[3] We report the intra-industry profit margin sorted portfolio returns when using the four-digit SIC codes to define industries. Our portfolio sorting results hold if we use adjusted SIC code following Dou, Ji, and Wei (2021). Our results also remain robust when we assign firms into the Fama-French 49 industries either using SIC or adjusted SIC codes.

even after controlling for the profitability and investment factors.

Figure 4: Average Cumulative Expected Return of the Long-Short Portfolio

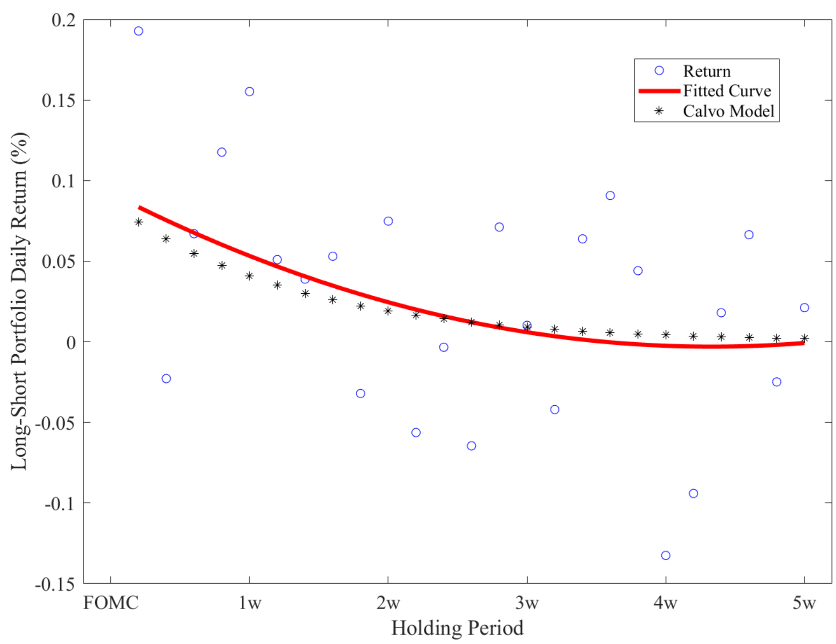



This figure plots the average cumulative expected return of the profit margin-sorted long-short portfolio. Since FOMC announcement days are irregularly scheduled and the lowest gap between two FOMC announcements is 34 calendar days, we plot the portfolio returns five (5) weeks after the FOMC announcement days. We consider various measures of profit margin. Profit margin 1-5 are defined as operating income after depreciation divided by sales, following Soliman (2008); operating income divided by revenues, following Fairfield and Yohn (2001); operating income before depreciation divided by sales, following Grullon, Larkin, and Michaely (2019); net income divided by revenues, following Berrone, Fosfuri, Gelabert, Gomez-Mejia (2013); and operating income divided by sales, following Jansen, Ramnath, Yohn (2011).

Figure 4 plots the average cumulative expected returns of the long-short portfolio based on various measures of profit margin proposed in the literature. We form quintile portfolios before each FOMC announcement days and hold portfolios till two days before another FOMC announcements. FOMC announcement days are irregularly scheduled. The lowest gap between two FOMC announcements in our sample is 34 calendar days. So we plot the average cumulative expected returns five (5) weeks after the FOMC announcement without being interfered by another announcement. The long-short portfolio has a strong path of returns during the first two weeks after the FOMC days and the return gradually decay afterwards. In summary, we show high profit margin firms earn higher inflation risk premium than low profit margin firms do after the FOMC announcements.

In Figure 5, we plot the average daily return of the profit margin sorted portfolios across an FOMC announcement cycle (circles), a fitted curve for the expected return (solid line), and the expected return implied by the Calvo model (stars). The horizontal axis is weeks after a pre-scheduled FOMC announcement, and the vertical axis is the average daily return of a long-short portfolio during our sample period. The fitted curve of daily inflation risk premium in our sample period is very close to that one implied by the Calvo model of periodic price adjustment, where inflation risk premium decays exponentially over time as firms adjusts away from their previous price levels over time.

Figure 5: Daily returns over an FOMC cycle




This figure plots the average daily return of the profit margin sorted portfolios across an FOMC announcement cycle (circles), a fitted curve for the expected return (solid line), and the expected return implied by the Calvo model (stars).

# 4 Conclusion

In this paper, we present a theory-guided identification for the existence and the magnitude of inflation risk premium. We show theoretically that when monetary policy is countercyclical, variations inflation can either be due to active monetary policy responses to fundamental productivity shocks or due to unexpected monetary policy conduct. The first type of inflation shock commands a negative risk premium while the second one commands a positive risk premium. We show that a robust implication of models with price rigidity is that exposure to unexpected monetary policy conduct is increasing in firms' profit margin. Using profit margin sorted portfolios, we identify a significant inflation risk premium over FOMC announcement cycles of about 11% per year. Our evidence for inflation risk premium is robust to controlling for convention risk factors such as the Fama-French five factors. We also show that the evolution of inflation risk premium is consistent with Calvo-type of models of slow price adjustment.

# References


Augustin, Patrick, Linxiao Francis Cong, Alexandre Corhay, and Michael Weber, 2024, Price Rigidities and Credit Risk, *Working paper.*

Bansal, Ravi, and Ivan Shaliastovich, 2013, A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets, *The Review of Financial Studies* 26, 1–33.

Bauer, Michael, and Eric Swanson, 2023a, An Alternative Explanation for the "Fed Information Effect", *American Economic Review* 113, 664–700.

Bauer, Michael, and Eric Swanson, 2023b, A Reassessment of Monetary Policy Surprises and High-Frequency Identification, *NBER Macroeconomics Annual* 37, 87–155.

Bernanke, Ben S., and Kenneth N. Kuttner, 2005, What Explains the Stock Market's Reaction to Federal Reserve Policy?, *The Journal of Finance* 60, 1221–1257.

Bhamra, Harjoat S., Christian Dorion, Alexandre Jeanneret, and Michael Weber, 2023, High Inflation: Low Default Risk and Low Equity Valuations, *The Review of Financial Studies* 36, 1192–1252.

Boons, Martijn, Fernando Duarte, Fransand de Roon, and Marta Szymanowska, 2020, Time-varying inflation risk and stock returns, *Journal of Financial Economics* 136, 444–470.

Calvo, Guillermo, 1983, Staggered Prices in a Utility Maximizing Framework, *Journal of Monetary Economics* 3, 383–398.

Campbell, John, Carolin Pflueger, and Luis Viceira, 2020, Macroeconomic Drivers of Bond and Equity Risks, *Journal of Political Economics* 128, 3148–3185.

Cieslak, Anna, and Carolin Pflueger, 2023, Inflation and Asset Returns, *Annual Review of Financial Economics* 15, 433–448.

Corhay, Alexandre, and Jincheng Tong, 2024, Inflation risk and the finance-growth nexus, *Working paper.*

David, Alexander, and Pietro Veronesi, 2013, What Ties Return Volatilities to Price Valuations and Fundamentals?, *Journal of Political Economics* 121, 682–746.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, 2020, The Rise of Market Power and the Macroeconomic Implications, *The Quarterly Journal of Economics* 135, 561–644.

Dou, Winston, Yan Ji, and Wu Wei, 2021, Competition, Profitability, and Discount Rates, *Journal of Financial Economics* 140, 582–620.

Fairfield, Patricia, and Teri Yohn, 2001, Using Asset Turnover and Profit Margin to Forecast Changes in Profitability, *Review of Accounting Studies* 6, 371–385.

Fang, Xiang, Yang Liu, and Nikolai Roussanov, 2023, Getting to the Core: Inflation Risks Within and Across Asset Classes, *Working paper*.

Gali, Jordi, 2008, *Introduction to Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. (Princeton University Press).

Gorodnichenko, Yuriy, and Michael Weber, 2016, Are Sticky Prices Costly? Evidence from the Stock Market, *American Economic Association* 106, 165–199.

Gourio, Francois, and Phuong Ngo, 2020, Risk Premia at the ZLB: A Macroeconomic Interpretation, *Working paper*.

Hanson, Samuel G, David O Lucca, and Jonathan H Wright, 2021, Rate-Amplifying Demand and the Excess Sensitivity of Long-Term Rates, *The Quarterly Journal of Economics* 136, 1719–1781.

Hanson, Samuel G., and Jeremy C Stein, 2015, Monetary policy and long-term real rates, *Journal of Financial Economics* 115, 429–448.

Jansen, Ivo, Sundaresh Ramnath, and Teri Yohn, 2012, Diagnostic for Earnings Management Using Changesin Asset Turnover and Profit Margin, *Contemporary Accounting Research* 29, 221–251.

Kekre, Rohan, and Moritz Lenel, 2022, Monetary Policy, Redistribution, and Risk Premia, *Econometrica* 90, 2249–2282.

Kekre, Rohan, Moritz Lenel, and Federico Mainardi, 2024, Monetary Policy, Segmentation, and the Term Structure, *Working paper*.

Kung, Howard, 2015, Macroeconomic linkages between monetary policy and the term structure of interest rates, *Journal of Financial Economics* 115, 42–57.

Li, Erica X.N., Tao Zha, Ji Zhang, and Hao Zhou, 2022, Does fiscal policy matter for stock-bond return correlation, *Journal of Monetary Economics* 128, 20–34.

Nakamura, Emi, and Jon Steinsson, 2018, High Frequency Identification of Monetary Non-Neutrality: The Information Effect, *Quarterly Journal of Economics* 133, 1283–1330.

Pflueger, Carolin, and Gianluca Rinaldi, 2022, Why does the Fed move markets so much? A model of monetary policy and time-varying risk aversion, *Journal of Financial Economics* 146, 71–89.

Song, Dongho, 2017, Bond Market Exposures to Macroeconomic and Monetary Policy Risks, *The Review of Financial Studies* 30, 2761–2817.

Weber, Michael, 2015, Nominal Rigidities and Asset Pricing, *Working paper*.

# 5 Appendix

## 5.1 Proof for Proposition 3

Given the form of the $\bar{v}(\chi, a, \nu)$ function,

$$\begin{aligned}\frac{\partial \ln \bar{v}(\chi, a, \nu)}{\partial \nu} =& \frac{\delta_1(z)\chi^{1-\eta}}{\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}} \frac{\partial \ln \delta_1(z)}{\partial \nu} - \frac{\delta_2(z)\chi^{-\eta}}{\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}} \frac{\partial \ln \delta_2(z)}{\partial \nu} \\ =& \frac{\partial \ln \delta_1(z)}{\partial \nu} + \omega(\chi, z)\left[\frac{\partial \ln \delta_1(z)}{\partial \nu} - \frac{\partial \ln \delta_2(z)}{\partial \nu}\right],\end{aligned}$$

where

$$\omega(\chi, z) = \frac{\delta_2(z)\chi^{-\eta}}{\delta_1(z)\chi^{1-\eta} - \delta_2(z)\chi^{-\eta}} = \left[\frac{\delta_1(z)}{\delta_2(z)}\chi - 1\right]^{-1}$$

is a strictly decreasing function of $\chi$. A similar expression holds for $\frac{\partial \ln \bar{v}(\chi,a,\nu)}{\partial a}$. As a result, $\frac{\partial \ln \bar{v}(\chi,a,\nu)}{\partial \nu}$ being increasing in $\chi$ is equivalent to $\frac{\partial \ln \delta_1(a,\nu)}{\partial \nu} - \frac{\partial \ln \delta_2(a,\nu)}{\partial \nu} < 0$ and $\frac{\partial \ln \bar{V}(\chi,a,\nu)}{\partial a}$ being decreasing in $\chi$ is equivalent to $\frac{\partial \ln \delta_1(a,\nu)}{\partial a} - \frac{\partial \ln \delta_2(a,\nu)}{\partial a} > 0$.

To compare the partial derivatives of $\delta_1$ and $\delta_2$, we look for a log-linear approximation of the solution around the deterministic steady state. We look for linear policy functions of the form $\pi(a, \nu) = \pi_a a + \pi_\nu \nu$, and $w(a, \nu) = \bar{w}e^{w_a a + w_\nu \nu}$, where $\bar{w}$ is the level of real wage at the deterministic steady state. Here we use the fact that steady-state inflation is zero. We write Equation (10) as

$$(\rho + \lambda + (1 - \eta)\pi(a, \nu)) = \delta_1^{-1}(a, \nu) + \frac{\mathcal{L}[\delta_1(a, \nu)]}{\delta_1(a, \nu)}.$$

The above Equation implies the steady state $\delta_1$ is $\bar{\delta}_1 = \frac{1}{\rho+\lambda}$. Assume a log linear form for $\delta_1(a, \nu)$, that is $\delta_1(a, \nu) = \frac{1}{\rho+\lambda}e^{\delta_{1,a}a + \delta_{1,\nu}\nu}$. With a first order approximation for $\delta_1^{-1}(a, \nu)$ , we write the above equation as

$$\rho + \lambda + (1 - \eta)[\pi_a a + \pi_\nu \nu] = (\rho + \lambda)(1 - \delta_{1,a}a - \delta_{1,\nu}\nu) - \kappa_a \delta_{1,a}a - \kappa_\nu \delta_{1,\nu}\nu.$$

Matching coefficients, we have:

$$\delta_{1,a} = \frac{(\eta - 1)\pi_a}{\rho + \lambda + \kappa_a};\ \delta_{1,\nu} = \frac{(\eta - 1)\pi_\nu}{\rho + \lambda + \kappa_\nu}.$$

Similarly, Equation (11) implies $\bar{\delta}_2 = \frac{\bar{w}}{\rho+\lambda}$. The log linear version of (11) is:

$$-\eta\left[\pi_a a + \pi_\nu \nu\right] = (\rho+\lambda)\left[(w_a - 1)\, a + w_v \nu\right] - (\rho+\lambda)\left[\delta_{2,a} a + \delta_{2,\nu}\nu\right] - \kappa_a \delta_{2,a} a - \kappa_\nu \delta_{2,\nu}\nu,$$

which allows us to solve for $\delta_{2,a}$ and $\delta_{2,\nu}$ as functions of the policy functions of $\pi$ and $w$:

$$\delta_{2,a} = \frac{\eta\pi_a + (\rho+\lambda)(w_a - 1)}{\rho+\lambda+\kappa_a};\ \ \delta_{2,\nu} = \frac{\eta\pi_\nu + (\rho+\lambda)\, w_\nu}{\rho+\lambda+\kappa_\nu}.$$

Using the above log linear solution, we can compare the elasticities of $\delta_1$ and $\delta_2$:

$$\begin{aligned} \frac{d\ln\delta_1(z)}{d\nu} - \frac{d\ln\delta_2(z)}{d\nu} &= \delta_{1,\nu} - \delta_{2,\nu} = -\frac{\pi_\nu + (\rho+\lambda)\, w_\nu}{\rho+\lambda+\kappa_\nu}; \\ \frac{d\ln\delta_1(z)}{da} - \frac{d\ln\delta_2(z)}{da} &= \delta_{1,a} - \delta_{2,a} = -\frac{\pi_a + (\rho+\lambda)(w_a - 1)}{\rho+\lambda+\kappa_a}. \end{aligned}$$

To sign the above expressions, we need to solve for the policy functions of $\pi$ and $w$. Equation (15) implies that at the zero inflation steady state, $\bar{w} = \frac{1}{\bar{\mu}}$. The log linear version of this equation is

$$\frac{1}{\lambda}\left(\pi_a a + \pi_\nu \nu\right) = \left(\delta_{2,a} - \delta_{1,a}\right) a + \left(\delta_{2,\nu} - \delta_{1,\nu}\right)\nu. \tag{17}$$

This equation implies that $\delta_{1,a} - \delta_{2,a}$ has the same sign as $-\pi_a$ and $\delta_{1,\nu} - \delta_{2,\nu}$ has the same sign as $-\pi_\nu$. As a result, to prove the proposition, we only need to show $\pi_\nu > 0$ and $\pi_a < 0$. We combine (17) and the following log linear version of the Taylor rule,

$$(\phi_\pi - 1)\left(\pi_a a + \pi_\nu \nu\right) + \phi_y\left(\frac{\zeta + w_a}{1+\zeta} a + \frac{w_\nu}{1+\zeta}\nu\right) - \nu = -\kappa_a \frac{\zeta + w_a}{1+\zeta} a - \kappa_\nu \frac{w_\nu}{1+\zeta}\nu. \tag{18}$$

to solve for the policy functions. Matching coefficients for $a$ in Equations (17) and (18), we have

$$\frac{1}{\lambda}\frac{\rho+\kappa_a}{\rho+\lambda}\pi_a - w_a = -1;\ \ (\phi_\pi - 1)\,\pi_a + \frac{\phi_y + \kappa_a}{1+\zeta} w_a = -\frac{\phi_y + \kappa_a}{1+\zeta}\zeta.$$

Matching coefficients coefficients for $\nu$ in Equations (17) and (18), we have,

$$\frac{\rho+\kappa_\nu}{\lambda}\pi_\nu = (\rho+\lambda)\, w_\nu;\ \ (\phi_\pi - 1)\,\pi_\nu + \frac{\phi_y + \kappa_\nu}{1+\zeta} w_\nu = 1.$$

Combining the above two equations, we can solve for $\pi_a$ and $\pi_\nu$:

$$\pi_a = -\left[\frac{\phi_\pi - 1}{\phi_y + \kappa_a} + \frac{1}{\lambda}\frac{\rho + \kappa_a}{(\rho + \lambda)(1 + \zeta)}\right]^{-1} < 0,$$

and

$$\pi_\nu = \left[(\phi_\pi - 1) + \frac{(\phi_y + \kappa_\nu)(\rho + \kappa_\nu)}{\lambda(\rho + \lambda)(1 + \zeta)}\right]^{-1} > 0,$$

as needed.

## 5.2 Empirical evidence from markup-sorted portfolios

We construct each individual firm's markup which is defined as *price-marginal cost ratio* by following De Loecker, Eeckhout, and Unger (2020). More specifically, we estimate markup as the output elasticity of the variable input multiplied by the ratio of sales and operating expenses, which is calculated as cost of goods sold + 0.7 * selling, general and administrative expense. The elasticity parameter is estimated using the same procedure of De Loecker, Eeckhout, and Unger (2020) and is found persistent. We use the quarterly Compustat database and we exclude financial and utility firms. We use adjusted SIC code following Dou, Ji, and Wei (2021). Before each pre-scheduled FOMC announcement days, we sort firms into quintile portfolios based on their markup relative to other firms in the same industry and hold the portfolios till two days before another FOMC days.

Table 3 reports the results when we regress markup-sorted portfolio returns on daily changes in inflation swap and market excess returns on FOMC days. High markup firms have a higher inflation β than low markup firms, no matter we use observed inflation swap or the one orthogonalized by macro variables. Meanwhile, high markup firms have a lower market β comparing to low markup firms in general.

Table 4 reports the holding period returns and alphas of markup-sorted portfolio. High markup firms earn higher expected returns and alphas than low markup firms. The long-short portfolio generates an annualized return of 6.91% (0.86%×8). Alphas of the long-short portfolio in the CAPM and Fama-French three- and five-factor models are positive and highly significant. For instance, the annualized Fama-French three factor alpha is 6.73% (0.84%×8) with a t-statistic of 3.04.

Table 3: Inflation Beta of Markup-Sorted Portfolios

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| Panel A: Inflation Swap | | | | | | |
| Inflation $\beta$ | -1.3497 | -0.5675 | -0.3547 | -0.3577 | 0.0144 | 1.3657 |
| | | | | | | (1.89) |
| Market $\beta$ | 1.1727 | 0.9849 | 1.0204 | 0.9798 | 0.9221 | -0.2500 |
| | | | | | | (-6.65) |
| Panel B: Orthogonalized Inflation Swap | | | | | | |
| Inflation $\beta$ | -1.4469 | -0.5964 | -0.4756 | -0.3978 | 0.1126 | 1.5656 |
| | | | | | | (2.23) |
| Market $\beta$ | 1.1752 | 0.9858 | 1.0220 | 0.9806 | 0.9210 | -0.2535 |
| | | | | | | (-6.59) |

This table reports the coefficients when we regress markup-sorted portfolio excess returns on changes in one-year tenor inflation swap and market excess returns on FOMC days. In Panel A, we use the daily changes in one-year tenor inflation swaps. In Panel B, we use the daily changes in one-year tenor inflation swaps orthogonalized by log changes of core CPI, unemployment rate, and GDP during the last three months before each FOMC announcements. Our regression sample covers from July 2004 to December 2023 with 154 pre-scheduled FOMC announcements. Newey-West $t$-statistics with 6 lags are reported in parentheses.

Table 4: Expected Returns of Markup-Sorted Portfolios

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| Return (%) | 0.7085 | 1.3147 | 1.4092 | 1.4211 | 1.5723 | 0.8639 |
| | (1.39) | (3.06) | (3.55) | (4.05) | (4.30) | (2.93) |
| CAPM $\alpha$ (%) | -0.9315 | -0.1723 | -0.0048 | 0.0749 | 0.2823 | 0.9363 |
| | (-2.91) | (-1.03) | (-0.03) | (0.58) | (1.78) | (3.08) |
| FF3 $\alpha$ (%) | -0.7535 | -0.1031 | 0.0321 | 0.1372 | 0.3641 | 0.8417 |
| | (-3.06) | (-0.73) | (0.22) | (1.11) | (3.09) | (3.04) |
| FF5 $\alpha$ (%) | -0.4762 | 0.0019 | 0.0206 | 0.0148 | 0.3354 | 0.5353 |
| | (-2.10) | (0.01) | (0.14) | (0.12) | (2.84) | (1.96) |

This table reports expected returns when we sort stocks based on intra-industry markup before each pre-scheduled FOMC announcement days. We hold portfolios till two-days before the next pre-scheduled FOMC announcements. We also report alphas from the CAPM, Fama-French three-factor and five-factor models. Newey-West $t$-statistics with 6 lags are reported in parentheses. The data sample is from 1994 to 2023.