北京大学中国经济研究中心
China Center for Ecnomic Research
讨论稿系列
Working Paper Series

E2022020 2022-11-10

# The Markup Elasticity of Monetary Non-Neutrality

Christian Höynck Minghao Li Donghai Zhang

**Abstract**

Firms' market power, measured by markups, has risen substantially and unequally across sectors. We develop a quantitative menu cost model that features multiple sectors with heterogeneous degrees of market competition to evaluate the implications of these trends for monetary non-neutrality. Two quantitative results stand out. First, the average markup elasticity of monetary non-neutrality in the United States equals 1. Second, the markup elasticity of monetary non-neutrality would be equal to 1.4 had the markup increased equally across sectors. We provide evidence to support the model mechanisms and predictions, and to differentiate the proposed mechanisms from the existing ones.

**Keywords:** Rising market power, Monetary non-neutrality, Menu cost model, Heterogeneous market power, Multi-sector model

**JEL Classification:** E31, E32, E52, E58

# The Markup Elasticity of Monetary Non-Neutrality*

Christian Höynck †  Minghao Li‡  Donghai Zhang§

November 9, 2022

*Click Here for the Latest Version*

**Abstract**

Firms' market power, measured by markups, has risen substantially and unequally across sectors. We develop a quantitative menu cost model that features multiple sectors with heterogeneous degrees of market competition to evaluate the implications of these trends for monetary non-neutrality. Two quantitative results stand out. First, the average markup elasticity of monetary non-neutrality in the United States equals 1. Second, the markup elasticity of monetary non-neutrality would be equal to 1.4 had the markup increased equally across sectors. We provide evidence to support the model mechanisms and predictions, and to differentiate the proposed mechanisms from the existing ones.

Keywords: Rising market power, Monetary non-neutrality, Menu cost model, Heterogeneous market power, Multi-sector model

JEL Classification: E31, E32, E52, E58

*We thank Jinhui Bai, Alessandro Ferrari, John Grigsby, Francesco Lippi, Matthias Meier, Michael Weber, Pablo Ottonello, Shengliang Ou, Jingyi Zhang (Discussant), as well as conference participants at China International Conference in Macroeconomics 2022 for their suggestions and discussions. The views expressed in this paper are those of the authors and they do not necessarily coincide with the views of Bank of Italy. All remaining errors are our own.

†Bank of Italy, chhoeynck@gmail.com

‡China Center for Economic Research (CCER), NSD, Peking University, mhli@nsd.pku.edu.cn

§University of Bonn, donghai.d.zhang@gmail.com

# 1. Introduction

Firms' market power (competition) has increased (decreased) substantially in the past decades: the aggregate markup in the United States has risen by about 30% (De Loecker et al. 2020). Moreover, the changes in markups are heterogeneous across sectors. These trends affect firms' price-setting strategies. When making pricing decisions, firms balance the fixed adjustment cost—menu cost—with the benefits of adjusting prices. Market competition affects this tradeoff, and therefore, alters firms' price-setting rules. Consequently, market power affects the potency of monetary policy since the latter depends on the responses of prices to monetary actions. In this paper, we ask the following questions. How does monetary non-neutrality change with the recent trend in the aggregate markup? What is the aggregate markup elasticity of monetary non-neutrality? Does the heterogeneous sectoral evolution of markups matter for aggregate monetary non-neutrality?

This paper addresses those questions using a multi-sector menu cost model with heterogenous market powers across sectors. Our quantitative analysis suggests that in the United States, the average markup elasticity of monetary non-neutrality is equal to 1 in the past three decades: i.e., the thirty percent increase in markups in the data raises monetary non-neutrality by thirty percent.[1] Moreover, the unequal changes in markups at the sectoral level act as a counterforce: the markup elasticity of monetary non-neutrality would be equal to 1.4 had the markup increased equally across sectors. The key mechanism is the endogenous relationship between the frequency of price adjustment (FPA) and market power. In other words, market power affects monetary non-neutrality by altering the *frequency effect*. We provide evidence to support the model mechanisms and predictions. We also show that the mechanism we emphasize is quantitatively important for the positive relationship between markups and monetary non-neutrality estimated in the data. We formulate these findings based on the following steps.

First, we build a *multi-sector* menu cost model with monopolistic competition that features random menu costs and leptokurtic idiosyncratic productivity shocks to generate a realistic distribution of price changes. Notably, the model allows for heterogenous desired markups (market power) across sectors. Firms operate in different sectors are subject to heterogeneous degrees of market competition. The latter is proportional to the elasticity of substitution across goods within a sector.

Second, we calibrate the model to match pricing and markup moments in the U.S. data, and use it as a laboratory to address the raised questions. In doing so, we construct the pricing and markup moments at the industry level, which is an additional contribution

[1] Monetary non-neutrality is defined as the cumulative effects of monetary policy shocks on real GDP.

valuable for calibrating this class of models. Our model makes the following quantitative predictions. First, the average markup elasticity of monetary non-neutrality is equal to 1 in the past decades—**Result 1**. Since 1980, monetary non-neutrality has increased by nearly 30% due to the increase in aggregate markup of the same magnitude. Second, through the lens of the model, we show that the unequal changes in sectoral markups documented in the data matters for monetary non-neutrality. In a counterfactual analysis, we show that had markups increased equally across sectors, the increase in the aggregate markup would have raised monetary non-neutrality by 42%. The resulting aggregate markup elasticity of monetary non-neutrality would be 1.4—**Result 2**.

We inspect the mechanism behind our quantitative findings. We find that the previous results are explained by the *increasing* and *concave* relationship between monetary non-neutrality and steady-state markups. Keeping everything equal, higher markup in a given sector leads to greater but incrementally smaller real effects of monetary policy. Understanding what drives this increasing concave relationship between markups and monetary non-neutrality is crucial to understanding our quantitative results.

Following Alvarez et al. (2022), we decompose the aggregate real output's response to monetary policy shocks into margins summarized by two sufficient statistics. The first is the *frequency effect*, which is proportional to the inverse of the FPA. The second is the *selection effect*, which is related to the kurtosis of the distribution of price changes. Both margins are increasing concave functions of the steady-state markups. Quantitatively, the frequency margin contributes approximately 90% to the increase in monetary non-neutrality originating from the increased markups. Our results are, therefore, mainly driven by the frequency margin, i.e., through the effects of market power on the frequency of price adjustment.

We finally provide three pieces of empirical evidence to support the predictions and the underlying mechanisms and, importantly, to differentiate the proposed mechanisms from the ones in the literature.

The first set of evidence demonstrates that monetary non-neutrality increases in the aggregate smoothed markup in the data. This evidence is based on estimating real GDP's impulse response functions (IRFs) to monetary policy shocks. The IRFs are estimated using Jordà (2005)'s local projection and the extended exogenous monetary policy shocks à la Gertler and Karadi (2015). Importantly, we allow the IRFs to depend on the smoothed markup in the empirical model.

The second set of evidence shows that the FPA has declined over time in the data, consistent with the frequency effect emphasized in the model. Moreover, the magnitude of this decline is aligned with the model's predictions.

The third piece of evidence differentiates our mechanisms from the existing ones. Specifically, we estimate a horse-race regression model that allows the IRFs to monetary policy shocks depending on the smoothed markup and the measured FPA. Existing theories (see, e.g., Wang and Werning 2022 and Baqaee et al. 2021) suggest that markup-elasticity of monetary non-neutrality can be positive due to channels that are unrelated to the frequency effect, i.e., orthogonal to the changes in the FPA. In this case, one would expect the horse-race regression model to deliver the same estimates of markup-elasticity of monetary non-neutrality as in the baseline empirical model. In contrast, our theoretical model predicts that including FPA, which controls for the frequency effect, in the horse-race model reduces the relevance of smoothed markup for monetary non-neutrality. We find that changing FPA contributes to about half of the effects of smoothed markup on monetary non-neutrality estimated in the baseline. This finding confirms the relevance of the frequency effect that we highlight in the quantitative model.

From the perspective of policy making, our paper provides a toolbox that assists central bankers in keeping track of monetary non-neutrality. This is important for determining the correct amount of the nominal demand stimulus package in times of recession. Our analysis suggests that ignoring cross-sector heterogeneity in market power overstates the stimulus power of monetary policy. The evolution of the aggregate markup is informative, but it is a noisy signal about the changes in aggregate monetary non-neutrality. A more accurate assessment of changes in monetary non-neutrality requires monitoring the evolution of market power at the industry level. Moreover, the markup elasticity of monetary non-neutrality decreases in markup. This requires central banks to continuously reassess the monetary non-neutrality in a world with changing market power.

**Related Literature** This paper contributes to a growing literature that models the relationship between market competition and monetary non-neutrality. Mongey (2021) shows that aggregate monetary non-neutrality is higher in a model with oligopoly competition than in a model with monopolistic competition. We focus on the markup elasticity of monetary non-neutrality. In a more related paper, Wang and Werning (2022) finds that higher market concentration, hence a higher market power, significantly amplifies the real effects of monetary policy in a model where firms play a Bertrand dynamic game but with stylized *Calvo* nominal rigidity. With Calvo pricing, the frequency of firms' price adjustment is fixed and unrelated to market power by construction. Based on a model with endogenous markup à la Kimball, Baqaee et al. (2021) show that the supply-side effects of monetary policy arise. Moreover, increased industrial concentration increases monetary non-neutrality through the amplified supply-side effects. Compared to Wang and Werning (2022) and Baqaee et al. (2021), we address a similar question in a different framework where market power is *endogenously* related to the frequency of price adjustment—the frequency effect. We show that increased market power enhances the potency of monetary policy even if the standard monopolistic competition assumption is assumed. We present empirical evidence that supports our mechanism. Moreover, we highlight the relevance of considering cross-sector heterogeneities in desired markups for monetary non-neutrality and the concave relationship between the desired markup and monetary non-neutrality. Meier and Reinelt (2022) argue the causal effect of price rigidities on markups charged by firms during business cycles. Specifically, they argue that firms with more rigid prices optimally set higher markups due to the precautionary price-setting motive. The current paper emphasizes the causal effects of market power (steady-state markups) on the implied nominal rigidity.

This paper is closely related to the literature that employs the menu cost as a microfoundation for price rigidity: Dotsey et al. (1999), Golosov and Lucas (2007), Gertler and Leahy (2008), Midrigan (2011), Vavra (2014), Alvarez et al. (2016), Karadi and Reiff (2019) and Alvarez et al. (2022). The existing literature focuses on pricing moments and their relationships with monetary non-neutrality. The current paper contributes to this line of research by demonstrating that market power, especially the distribution of market power across sectors, is a critical determinant of the price-change distribution and by quantifying the markup elasticity of monetary non-neutrality.

Finally, this paper is related to the literature that emphasized the relevance of sectoral heterogeneity in price rigidity for the aggregate monetary non-neutrality: see e.g., Carvalho (2006), Nakamura and Steinsson (2010), Carvalho and Schwartzman (2015), Gautier and Bihan (2018), Carvalho et al. (2020), Pasten et al. (2020), and Alvarez et al. (2022).[2] The current paper shows that heterogeneity in sectoral market power is another important determinant of monetary non-neutrality. Specifically, we demonstrate that heterogeneous nominal rigidities can arise endogenously due to heterogeneous market competitions. And we highlight the non-linear relationship between market competition and nominal rigidity that is crucial for Result 2.

The remainder of this paper is organized as follows. Section 2 presents the empirical findings that motivate the quantitative analysis. Section 3 describes the multi-sector menu cost model with heterogeneous markups. Section 4 presents the predictions of the model. Section 5 provides empirical evidence that supports the model's predictions and the underlying mechanisms. Section 6 concludes the paper.

[2]The key takeaway from this literature is the aggregation issue: monetary non-neutrality associated with a multi-sector model with heterogeneous price rigidities differs from the money non-neutrality derived assuming a one-sector economy with average price rigidity

## 2. Motivating Facts

This section reproduces the empirical findings uncovered by De Loecker et al. (2020). We highlight the cross-sector heterogeneity in markup development.

### 2.1 Data

**Firm-Level Markups** We use quarterly firm-level balance sheet data from 1980 - 2016 of publicly traded firms in Compustat to calculate firm-level markups. The data covers sales, employment, capital, and input factors of firms (cost of goods sold) over a long sample for a wide range of sectors covering manufacturing and service sector firms. We estimate firm-level markups following the single-input approach (De Loecker and Warzynski 2012). According to this approach, the markup $\mu_{i,t}$ of a firm $i$ at time $t$ can be computed from one flexible input, $X_i$, as the ratio of the output elasticity of the input, $\epsilon_{Q,X_i,t}$, to the revenue share of that input, $s_{R,X_i,t}$

$$\mu_{i,t} = \frac{\epsilon_{Q,X_i,t}}{s_{R,X_i,t}}. \tag{1}$$

Compustat reports a composite input called Cost of Goods Sold (COGS), which consists of intermediate and labor input and that will be used as the (partially) flexible input, $X_i$. We use a variant of the technique introduced by Olley and Pakes (1996) and described in De Loecker and Warzynski (2012) to estimate a Cobb-Douglas function and obtain a time-varying estimate of output elasticity at the sector level. The markups are then derived by dividing the former (estimated at the industry-year level) by the share of COGS to revenue (estimated at the firm-year level). In terms of implementation, we follow the procedure described in De Loecker et al. (2020) with the adjustments described in Baqaee and Farhi (2020). In particular, we estimate time-varying output elasticities and deflate using gross output price indices from KLEMS sector-level data.[3]

When transforming the data, we drop all firms in the government sector or the sector of the economy composed of finance, insurance, and real estate. We consider only observations that are positive and linear interpolate observations that are missing for one period. Additionally, we perform outlier adjustments by trimming at 1% of calculated markups[4].

One concern with Compustat is that it covers only publicly traded firms and thus is not representative of the distribution of the universe of firms. We account for a representativeness bias by using the weights of each sector in the Compustat data from the PCE

[3]KLEMS stands for capital (K), labor (L), energy (E), materials (M), and service (S).
[4]Our findings are robust to trimming at 5%.

expenditure shares to account for sectoral composition (while we still calculate markups from publicly traded firms).

**Smoothed Markups** The markups at the sectoral level ($\mu_{k,t}$) are the weighted average of firm-level markups using firms' sale share as the weight. We compute the smoothed markup ($\mu^{ss}_{k,t}$) for an industry $k$ at time t as the seven-year moving average of $\mu_{k,t}$ centered at year $t$.[5]

## 2.2 Empirical Results: Unequal Changes in Markups

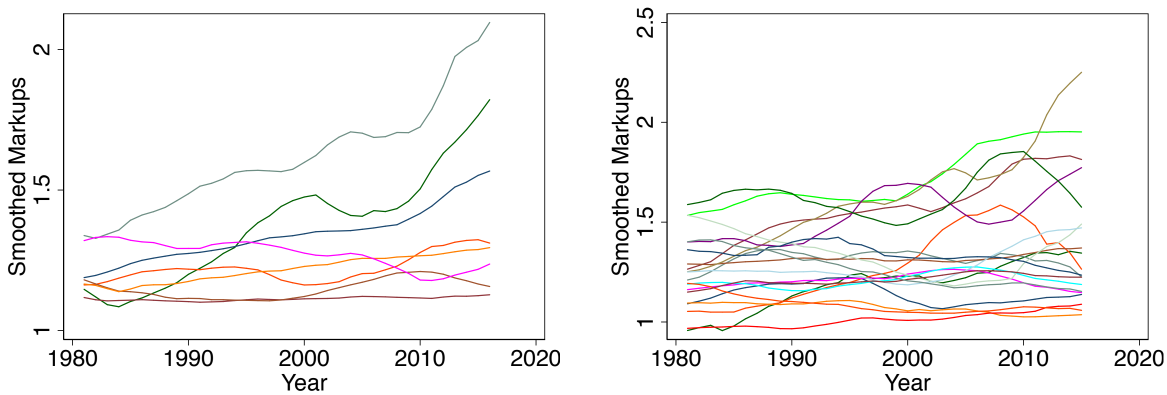
Figure A.1 confirms the overall increase in the aggregate markup documented by De Loecker et al. (2020). The smoothed aggregate markup raised from 1.2 in 1980 to 1.55 now—a nearly 30% increase. The current paper focuses on two observations related to cross-sector heterogeneity in markups.

**Observation 1: Markups are Heterogeneous** Figure 1 plots the smoothed markups at the industry level over time: See Panel (a) for eight one-digit NAICS sectors and Panel (b) for twenty two-digit NAICS sectors. The computed smoothed markups are highly heterogenous across sectors independent of the year of the observation. For example, in 2000, the smoothed markup ranges from nearly 1 to about 1.6 across industries, a difference of 60%.

**Observation 2: Unequal Changes in Markups across Sectors** The increased aggregate markup is not equally distributed cross sectors. Specifically, the changes in the right tail of the cross-sector markup distribution drive the increase in the aggregate markup. This can be seen in Panel (a) and (b) in Figure 1 (see also the red lines in Figure A.2 that plot the evolution of the smoothed markups by sector). In the eight-sector case, the increase in the aggregate markup documented in Figure A.1 is driven by three sectors that saw shape rises in markups. Markups in the other sectors remained relatively stable, with one sector witnessing a drop in markup.

[5] All results presented in the paper are robust to the use of alternative length of moving average.

Figure 1: Heterogeneous Evolutions of Smoothed Markups cross Sectors




**(a)** Smoothed Markups by Sectors: One-digit

**(b)** Smoothed Markups by Sectors: Two-digit

*Note:* Authors' own calculation. This figures plots steady-state markups (measured as cost share) in Compustat from 1980-2017 for the one-digit and the two-digit NAICS sectors.

In the remainder of the paper, we investigate the implications of those observations for monetary non-neutrality based on a multi-sector menu cost model with heterogenous cross-sector market powers.

# 3. A Multi-sector Model with Heterogeneous Markups

In this section, we construct a multi-sector menu cost model. The model contains the standard ingredients of a second-generation menu cost model, such as random menu costs and leptokurtic idiosyncratic productivity shocks, to generate a realistic distribution of price changes, see, e.g., Vavra (2014), Karadi and Reiff (2019) and Alvarez et al. (2021). Motivated by the empirical observations, our main methodological innovation is to incorporate heterogeneous sectoral market powers into this class of model.

## 3.1 Household

There is a representative household in the economy and a continuum of monopolistically competitive firms in $K$ sectors. Firms are indexed by $(k, i)$, where $k = 1, 2, ..., K$ and $i \in [0, 1]$.

The preference of the representative household is given by:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[ \frac{C_t^{1-\gamma}}{1-\gamma} + \kappa \frac{N_t^{1-\lambda}}{1-\lambda} \right], \tag{2}$$

where $\beta$ is the discount factor, $\kappa$ controls the magnitude of disutility from working, $\gamma$ is the elasticity of inter-temporal substitution, and $\lambda$ is the inverse of the Frisch elasticity. $N_t$ represents aggregate labor supply and aggregate consumption bundle is denoted by $C_t$, defined as:

$$C_t = (\sum_{k=1}^{K} \omega_k^{\frac{1}{\eta}} C_{k,t}^{\frac{\eta-1}{\eta}})^{\frac{\eta}{\eta-1}}, \tag{3}$$

where $\eta$ is the elasticity of substitution over goods across sectors and $\{\omega_k\}$ are the sectoral weights. The final good in sector $k$ is given by:

$$C_{k,t} = \omega_k \left( \int_{i \in [0,1]} c_{k,i,t}^{\frac{\theta_k - 1}{\theta_k}} di \right)^{\frac{\theta_k}{\theta_k - 1}}, \tag{4}$$

where $\theta_k$ is the elasticity of substitution over goods within sector $k$. The inter-temporal budget constraint of the household at period $t$ is:

$$\sum_{k=1}^{K} \int p_{k,i,t} c_{k,i,t} di + Q_{t+1} B_{t+1} \leq B_t + W_t N_t + \Pi_t, \tag{5}$$

Here, $p_{k,i,t}$ is the price of goods produced by firm $i$ in sector $k$ at period $t$, $W_t$ is the wage rate, $Q_{t+1}$ is the price of state-contingent nominal bonds, and $\Pi_t$ are the profits from all firms.

**Household's Optimality Conditions** The representative household chooses consumption bundle $\{c_{k,i,t}\}$, labor supply $N_t$ and holdings of nominal bonds $B_{t+1}$ to maximize their sum of discounted expected utility expressed in (2), subject to the budget constraints (5).

Solving this problem gives the demand for differentiated goods, the inter-temporal Euler equation, and an intra-temporal equation for aggregate labor supply:

$$c_{k,i,t} = \left( \frac{p_{k,i,t}}{P_{k,t}} \right)^{-\theta_k} \left( \frac{P_{k,t}}{P_t} \right)^{-\eta} C_t, \tag{6}$$

$$Q_{t+1} = \beta \frac{P_t C_t}{P_{t+1} C_{t+1}} \tag{7}$$

$$\frac{W_t}{P_t} = \kappa C_t, \tag{8}$$

The sectoral price index ($P_{k,t}$) and the aggregate price index ($P_t$) are defined as:

$$P_{k,t} = \left( \int p_{k,i,t}^{1-\theta_k} di \right)^{\frac{1}{1-\theta_k}}, \tag{9}$$

$$P_t = \left( \sum_{k=1}^{K} \omega_k P_{k,t}^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{10}$$

It is important to note that allowing the elasticity of substitution to vary across sectors is equivalent to assuming a heterogeneous degree of competitiveness in the goods market across sectors. That is, firms that operate in different sectors have different market powers with different desired (steady-state) markups. Formally, the desired markup for firms in a sector $k$ is defined $\frac{\theta_k}{\theta_k - 1}$, $\forall \theta_k \in (1, \infty)$. The desired markup is a decreasing function of the degree of competitiveness measured by $\theta_k$. In the limiting case, as $\theta_k$ converges to $\infty$, the market converges to a perfectly competitive market. This paper corresponds the smoothed markups in the data to the steady-state markups in the model.

## 3.2 Firms

The differentiated good $c_{k,i,t}$ is produced by firm $i$ in sector $k$ by hiring $n_{k,i,t}$ units of labor and using the following linear technology:

$$y_{k,i,t} = a_{k,i,t} n_{k,i,t}.$$

where $a_{k,i,t}$ is the idiosyncratic productivity of firm $i$ in sector $k$, which evolves according to the following process:

$$\log a_{k,i,t} = \begin{cases} \rho_k^a \log a_{k,i,t-1} + \sigma_k^a \epsilon_{k,i,t}^a, & \text{with probability } \alpha_k \\ \log a_{k,i,t-1}, & \text{with probability } 1 - \alpha_k, \end{cases} \tag{11}$$

where $\epsilon_{k,i,t}^a \sim N(0, 1)$ is independent across firms. We denote the transition probability of this Markov chain as $\Pr_k \left( a_{k,i,t+1} | a_{k,i,t} \right)$.

We assume that firms adjust prices with random menu costs: In every period, with probability $\phi_k$, firms can adjust their prices freely. Otherwise, firms pay fixed costs $\overline{f}_k$ in labor units to change their nominal prices. To summarize, the menu cost $f_{k,i,t}$ is given by:

$$f_{k,i,t} = \begin{cases} 0, & \text{with probability } \phi_k \\ \bar{f}_k. & \text{with probability } 1-\phi_k \end{cases} \tag{12}$$

**Firms' Optimality Conditions** Firm $i$ in sector $k$ chooses its prices to maximize total real discounted profits:

$$\max_{\{p_{k,i,t}\}} E_0 \sum_{t=0}^{\infty} q_{0,t}\pi_{k,i,t},$$

where $\pi_{k,i,t}$ is the real profit at period $t$ and $q_{0,t} \equiv q_{0,1}q_{1,2}\dots q_{t-1,t}$ discounts future profits into present value. Note that $q_{t,t+1}$ is the real stochastic discount factor which satisfies $q_{t,t+1} = Q_{t+1}P_{t+1}/P_t$.

Let $\Gamma_{k,t}(p_{-1}, a, f)$ be the distribution over idiosyncratic states in sector $k$ at period $t$. We can formulate firms' decision problems recursively:

$$V_{k,t}(p_{k,i,t-1}, a_{k,i,t}, f_{k,i,t}) = \max_{p_{k,i,t}} \{u_{k,i,t} + \mathbb{E}q_t V_{k,t+1}(p_{k,i,t}, a_{k,i,t+1}, f_{k,i,t+1})\} \tag{13}$$

with

$$u_{k,i,t} \equiv \left(\frac{p_{k,i,t}}{P_t} - \frac{W_t}{a_{k,i,t}P_t}\right) c_{k,i,t} - \mathbb{1}_{\{p_{k,i,t-1} \neq p_{k,i,t}\}} f_{k,i,t} \frac{W_t}{P_t},$$

and subject to

$$P_{k,t} = \left(\int [\psi_{k,t}(p_{-1}, a, f)]^{1-\theta_k} \, \mathrm{d}\Gamma_{k,t}(p_{-1}, a, f)\right)^{\frac{1}{1-\theta_k}},$$

$$P_t = \left(\omega_k \sum_{k=1}^{K} P_{k,t}^{1-\eta}\right)^{\frac{1}{1-\eta}},$$

where $\psi_{k,t}(p_{-1}, a, f)$ is the policy function for the choice variable $p_{k,i,t}$ of firms in sector $k$ at period $t$, and demand $c_{k,i,t}$ is given by equation (6). The distribution $\Gamma_{k,t}(p_{-1}, a, f)$ evolves according to:

$$\Gamma'_{k,t+1}(\mathcal{B}, a', f') = \left[\phi_k \mathbb{1}_{\{f'=0\}} + (1-\phi_k)\mathbb{1}_{\{f'=\bar{f}\}}\right] \int_{\{(p_{-1},a):\ \psi_{k,t}(p,a,f)\in\mathcal{B}\}} \mathrm{Pr}_k(a'|a) d\Gamma_{k,t}(p_{-1}, a, f), \tag{14}$$

for all sets $\mathcal{B} \in \mathbb{R}$.

Note that the policy functions $\{\psi_{k,t}\}$ and the distributions $\{\Gamma_{k,t}\}$ are not stationary but vary over time since the aggregate shock is one-time and unexpected.[6] Let $\Gamma_t(p_{-1}, a, f)$ be the distribution over idiosyncratic states in the economy at period $t$. It is straightforward to show that:

$$\Gamma_t(p_{-1}, a, f) = \sum_{k=1}^{K} \omega_k \Gamma_{k,t}(p_{-1}, a, f) \tag{15}$$

## 3.3 The Central Bank

Following the tradition of the menu cost literature, we assume that the monetary authority controls the aggregate nominal spending $S_t = P_t C_t$, the log of which follows the following process:

$$\log S_t = \mu + \log S_{t-1} + \epsilon_t^S; \tag{16}$$

The current paper refers to $\epsilon_t^S$ as the monetary shock, which is a one-time unexpected shock. Throughout the paper, unless explicitly explained, monetary non-neutrality refers to the effects of $\epsilon_t^S$ on real GDP or $C_{t+h}$. Formally, as it is done in the literature, we define monetary non-neutrality as the total cumulative effects of $\epsilon_t^S$ on log real GDP: i.e., $\sum_{h=0}^{\infty} \frac{\partial c_{t+h}}{\partial \epsilon_t^s}$.

This formulation implies that monetary non-neutrality is negatively related to the effects of monetary policy shock on prices. To see this, take the log of the nominal spending:

$$s_t = c_t + p_t, \tag{17}$$

where a variable in a small letter denotes the log of the corresponding variable in a capital letter. According to this identity, a nominal stimulus policy ($\triangle s_t$) either passes to prices ($\triangle p_t$) or has a real effect ($\triangle c_t$). The more the aggregate price reacts to $\epsilon_t^S$, the less the real effects of a monetary policy shock.

## 3.4 The Equilibrium

Given the law of motions of exogenous shocks (11), (12) and the path of monetary policy shocks, an equilibrium of the economy consists of:

(i) the household's demand for differentiated goods $\{c_{k,i,t}\}$ given by condition (6) of households' decision problem,

[6] We consider one-time shocks instead of systematic monetary policy shocks (as in Nakamura and Steinsson 2010) because it is extremely difficult to reduce the infinite-dimensional state variable (distribution of last-period prices) to a computationally feasible finite number of states.

(ii) firms' value functions $\{\{V_{k,t}\}_{k=1}^{K}\}_{t=0}^{\infty}$ and policy functions $\{\{\psi_{k,t}\}_{k=1}^{K}\}_{t=0}^{\infty}$, that solve firms' price-setting problem (13),

(iii) aggregate output $C_t$ and sectoral outputs $\{C_{k,t}\}_{k=1}^{I}$ that satisfy equations (3) and (4),

(iv) aggregate price functions $W_t$, $Q_t$, $\{P_{k,t}\}_{k=1}^{K}$ and $P_t$ that are determined by equations (8), (9) and (10),

(v) the distributions on firms' individual states $\{\{\Gamma_{k,t}(p_{-1}, a, f)\}_{k=1}^{K}\}_{t=0}^{\infty}$ that evolve according to (14) and $\{\Gamma_t(p_{-1}, a, f)\}_{t=0}^{\infty}$ is given by (15),

(vi) and the aggregate nominal spending $S_t = P_t C_t$.

**Computing the Transition Path** In Section 4, we compute the transition path of the perfect foresight equilibrium in response to an unexpected monetary policy shock. We briefly describe the computational procedure. We first assume that the economy starts in the steady state and returns to it after 200 periods. We next conjecture the entire path of the sectoral prices over the transition path. Given these prices, we solve firms' pricing problems backward. We then simulate firms in the economy forward using a non-stochastic simulation algorithm similar to Young (2010) to get the distribution over individual state variables at every period. We compute the sectoral prices using this distribution and the firms' policy rules. If these prices differ from the conjectured ones, we update the guessed prices until the equilibrium converges. In doing so, we obtain the aggregate and sectoral impulse response functions to the monetary policy shock.

# 4. Results

This section begins with a detailed discussion of the construction of multi-sector pricing moments used to calibrate the model. Section 4.1 presents the evolution of monetary non-neutrality implied by the model. Section 4.2 demonstrates the relevance of the unequal changes of secotral markups for the evolution of monetary non-neutrality. Section 4.3 decomposes the evolution of monetary non-neutrality into three alternative sources. Section 4.4 presents the mechanism behind the model's predictions.

**Model Calibration and Targeted Moments** Throughout the paper, we consider three alternative models: (i) the one-sector model ($K = 1$), (ii) eight-sector model ($K = 8$), and (iii) twenty-sector model ($K = 20$). We begin with the discussions of model calibrations. Three sets of parameters are calibrated.

The first set of parameters are the elasticity of substitution in $K$ sectors $\{\{\theta_{k,t}\}_{t=1980}^{2016}\}_{k=1}^{K}$, where $\theta_{k,t}$ is set according to $\theta_{k,t} = \mu_{k,t}^{ss}/(\mu_{k,t}^{ss} - 1)$. Here, $\mu_{k,t}^{ss}$ is the estimate of smoothed markup in sector $k$ at year $t$ described in section 2.1.

The second set of parameters is set to conventional values used in the literature. The model is calibrated at a monthly frequency with $\beta = 0.96^{1/12}$. We choose $\gamma = 1$ and $\lambda = 0$ so that the utility is log in consumption and linear in labor. We set $\kappa$ such that labor supply in the flexible price steady state is 1/3. We calibrate $\mu = 0.0018$ to match the mean growth of nominal GDP minus real GDP during the period 1998-2005.[7] Following Nakamura and Steinsson (2010), We choose the persistence of idiosyncratic shock $\rho_z$ to be 0.7. These parameters are shown in Table 1

**Table 1:** Externally Calibrated Parameters

| Parameter | Description | Value |
|---|---|---|
| $\beta$ | Discount factor | $0.96^{1/12}$ |
| $\gamma$ | Elasticity of intertemporal substitution | 1 |
| $\lambda$ | Inverse of Frisch elasticity | 0 |
| $\kappa$ | Labor coefficient | 2.25 |
| $\mu$ | Mean growth rate of $S_t$ | 0.0018 |
| $\rho_z$ | Idiosyncratic productivity persistence | 0.7 |

For sector $k$, there are four remaining parameters: the standard deviation of idiosyncratic productivity shocks $\sigma^z$, the probability of idiosyncratic productivity shocks $\alpha$, the probability of zero menu cost $\phi$ and the magnitude of menu cost $\bar{f}$. These parameters govern the cross-sectional behavior of price changes. We must construct micro-pricing moments to calibrate these parameters at the sectoral level.

To this end, we build upon the micro pricing moments data that Nakamura and Steinsson (2008) provide, which covers the manufacturing and services sectors and focuses on a period with low inflation (1998-2005). [8] The data consists of various pricing moments for 270 Entry Level Items (ELIs) in the non-shelter component of the CPI over the period 1998-2005. The CPI is constructed at the BLS by collecting data on about 130,000 products per month from around 27,000 retail outlets across 87 geographical areas in the United States.

[7] The moments of price adjustment for calibration in Nakamura and Steinsson (2010) is calculated using data from 1998 to 2005.

[8] One of the primary sources for empirical studies on price-setting behavior at the microeconomic level is the CPI research database at the Bureau of Labor Statistics, which contains the product level data used to construct the consumer price index (CPI). It has been used by Bils and Klenow (2004), Nakamura and Steinsson (2008), Bils et al. (2012) and Nakamura et al. (2018).

The non-shelter components of the CPI represent about 70% of consumer expenditure.

To match pricing moments at the ELI level with markup data at the NAICS level, we construct a crosswalk between these two classifications by hand. In particular, we build many-to-many matches between 6-digit NAICS and ELI categories. The match is made by hand according to a comparison of product descriptions (as well as individual item names contained in the CPI Research Database). The micro pricing moments at the one-digit NAICS level are reported in Table A.2. Tables A.3 and A.4 report these moments at two-digit NAICS level. One valuable contribution of our paper is the construction of these moments because they are crucial for calibrating multi-sector menu cost models.

**Table 2:** Internally Calibrated Parameters

| Sector | Markup | $\sigma^z$ | $\alpha$ | $\phi$ | $\overline{f}$ |
|---|---|---|---|---|---|
| *One-sector model* | 1.36 | 0.081 | 0.064 | 0.035 | 0.067 |
| *Eight-sector model* | | | | | |
| Agriculture | 1.47 | 0.270 | 0.136 | 0.035 | 0.022 |
| Mining and Utilities | 1.17 | 0.144 | 0.104 | 0.069 | 0.007 |
| Manufacturing | 1.35 | 0.122 | 0.084 | 0.033 | 0.035 |
| Retail and Wholesale | 1.11 | 0.028 | 0.184 | 0.021 | 0.064 |
| Services (information, finance) | 1.64 | 0.056 | 0.086 | 0.055 | 0.037 |
| Education and Health Care | 1.14 | 0.042 | 0.213 | 0.015 | 0.082 |
| Services (entertainment etc.) | 1.24 | 0.048 | 0.142 | 0.027 | 0.057 |
| Other Services | 1.27 | 0.039 | 0.229 | 0.011 | 0.102 |

We calibrate the four sets of parameters ( $\sigma^z$, $\alpha$, $\phi$, and $\overline{f}$ for each sector $k$) to match the following price adjustment moments: the median frequency, the 25 percentile of the absolute size distribution, the median absolute size and the 75 percentile of the absolute size distribution. Table 2 presents the calibrated parameters across sectors for the one-sector and eight-sector model. When calibrating these parameters, we use the average sectoral markups from 1998 to 2005. Table A.1 shows these parameters of the twenty-sector model.

The model fits well both the pricing moments targeted and pricing moments not targeted. Table 3 shows the model fit for the one-sector model. Table A.2 in the Appendix shows the model fit of the eight-sector model. Tables A.3 and A.4 in the Appendix shows model fit of the twenty-sector model.

**Table 3:** Model Fit (one-sector model)

| Moment | Data | Model |
|---|---|---|
| **Moments targeted** | | |
| Frequency | 0.06 | 0.05 |
| Absolute size (median) | 0.06 | 0.06 |
| 25th percentile size | 0.03 | 0.03 |
| 75 percentile size | 0.12 | 0.11 |
| **Moments not targeted** | | |
| Size of price increase | 0.07 | 0.07 |
| Size of price decrease | 0.09 | 0.09 |

To compute monetary non-neutrality, we shock the economy with a one-time monetary shock of the size $\bar{\epsilon}$. Formally, we feed in a one-time unexpected monetary shock at $t = 1$ that dies out from $t = 2$ onwards. That is, $\epsilon_1^S = \bar{\epsilon}$ and $\epsilon_t^S = 0 \; \forall t > 1$. The impulse response functions (IRFs) are then computed using the perfect foresight assumption. Following the tradition of the menu cost literature, we define monetary non-neutrality as the cumulative changes in the log real GDP: i.e., $\sum_{h=0}^{\infty} \frac{\partial y_{t+h}}{\partial \epsilon_t^s} \bar{\epsilon}$. The size of the monetary shock is normalized to increase the nominal spending by 1% on impact.

Appendix C plots the IRFs to monetary policy shocks and provides a short discussion of monetary non-neutrality under alternative models. In the following, we focus on the paper's main findings: quantifying the markup elasticity of monetary non-neutrality and how it evolves.

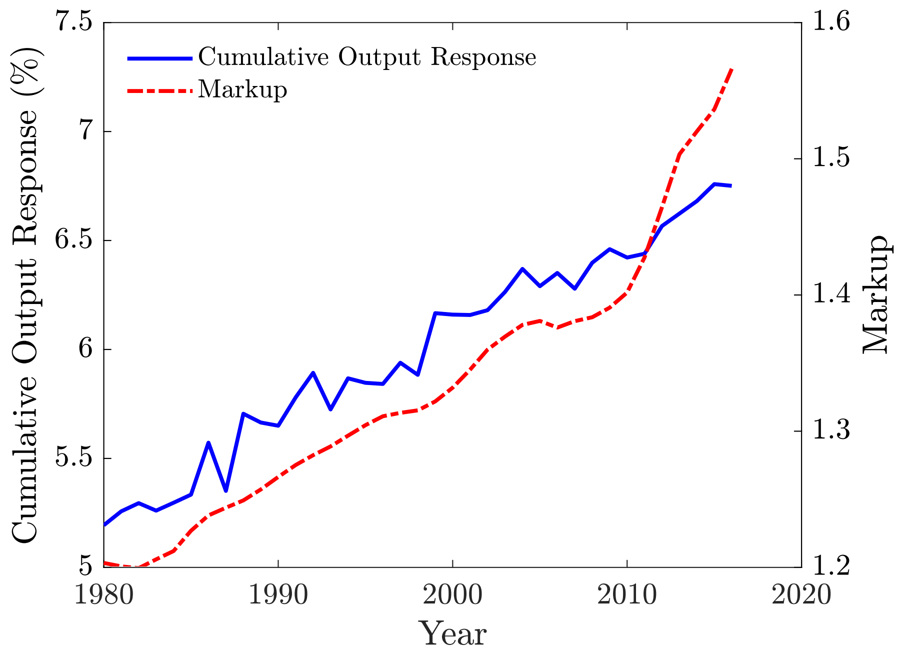
## 4.1 Result 1: the Rise of Markups and Monetary Non-neutrality Over Time

Markups have changed substantially over time, as it is documented in Figure 1. The economy-wide moving average markup, the sector-size weighted average of moving average markups at the industry level, has increased substantially: see the red dashed line in Figure 2. We now assess how the rise of markups in the data affects the monetary non-neutrality in the model. To this end, we calibrate the model for each year from 1980 to 2016. In each calibration, we change one set of parameters: the elasticities of substitution to match the changing moving average markups. We then compute measures of monetary non-neutrality for each year.

Figure 2 plots the evolution of the aggregate makeup in the U.S. (dashed red line) and the implied change of the aggregate monetary non-neutrality (solid blue line). Interestingly, aggregate monetary non-neutrality tracks the evolution of aggregate markups closely. Zooming to the sectoral level shows that similar patterns hold in each of the eight sectors considered in our calibration; see Figure A.2. The discussions about the economic mechanism is postponed to Section 4.4.

An important message for central bankers is that monitoring the evolution of aggregate markup is important for determining the right amount of nominal demand stimulus policy. According to our model, the average elasticity of monetary non-neutrality with respect to markup in the sample (1980-2016), defined as $\frac{\triangle log(\text{monetary non-neutrality})}{\triangle log(\text{markup})}$, is 1. A thirty percent increase in aggregate markup leads to a thirty percent increase in monetary non-neutrality.

**Figure 2:** Aggregate Monetary Non-neutrality Over Time (eight sectors)

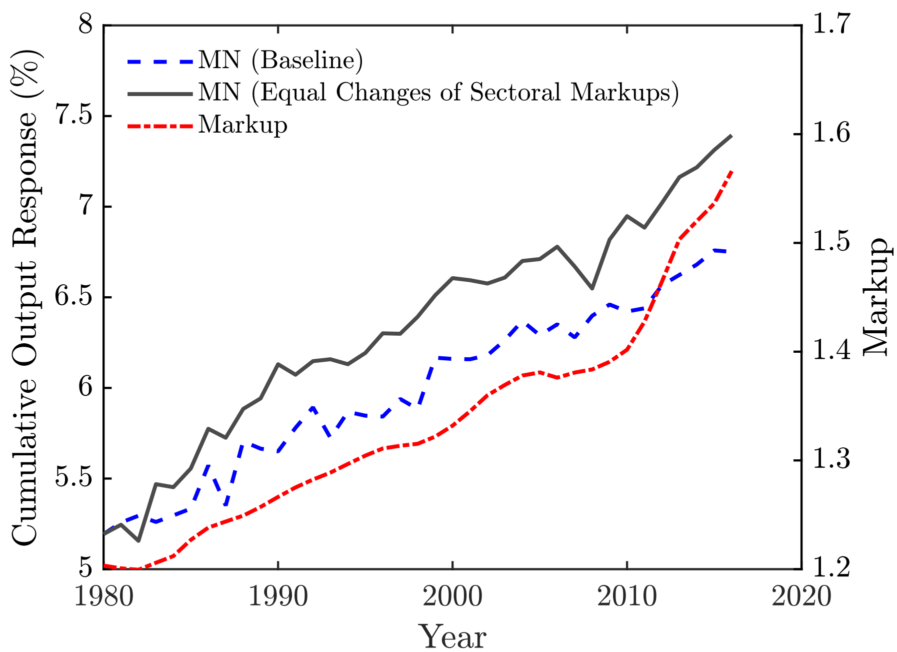



*Note:* This figure plots the evolution of the aggregate makeup in the U.S. (dashed red line) and the implied evolution of the aggregate monetary non-neutrality (solid blue line). Monetary non-neutrality is defined as the cumulative changes in the log real GDP following a monetary policy shock of the size $\bar{\epsilon}$: i.e., $\sum_{h=0}^{\infty} \frac{\partial y_{t+h}}{\partial \epsilon_t^s} \bar{\epsilon}$. The size of the monetary shock is normalized to increase the nominal spending by 1% on impact.

## 4.2 Result 2: the Role of Unequal Changes in Sectoral Markups

The second observation documented in Figure 1 is that changes in the moving average markups are heterogeneous across sectors. More specifically, the changes of sectoral markups have been unequal. How does the trend in the unequal changes of the smoothed markups affect aggregate monetary non-neutrality? The following experiment addresses this question.

**Figure 3:** Monetary Non-neutrality Over Time: An Equal-Change Counterfactual




*Note:* The blue dashed line plots the implied evolution of the aggregate monetary non-neutrality in our baseline calibration. The black solid line plots the implied evolution of the aggregate monetary non-neutrality for the counterfactual where the the aggregate markup is the same as the baseline but the sectoral markups have been increased equally. We achieve this by equalizing the annual markup increments across eight sectors. The red dash-dotted line shows the evolution of the aggregate makeup in the U.S.

We consider a counterfactual scenario, where markup changes since 1980 are equalized across sectors. The aggregate markup in this counterfactual analysis remains unchanged; the dashed red line in Figure 3. The model is re-calibrated for each year based on the counterfactual evolution of sectoral markups. The solid black line in Figure 3 plots the implied monetary non-neutrality in this counterfactual scenario. For comparison, the monetary non-neutrality according to the actual evolution of markups is plotted in the same figure using the blue dashed line.

The monetary non-neutrality measured by the cumulative output response would be 42% higher in 2016 than that of 1980. The associated markup elasticity of monetary non-neutrality is 1.4 in this counterfactual analysis.

### 4.3 Decomposing the Evolution of Monetary Non-Neutrality

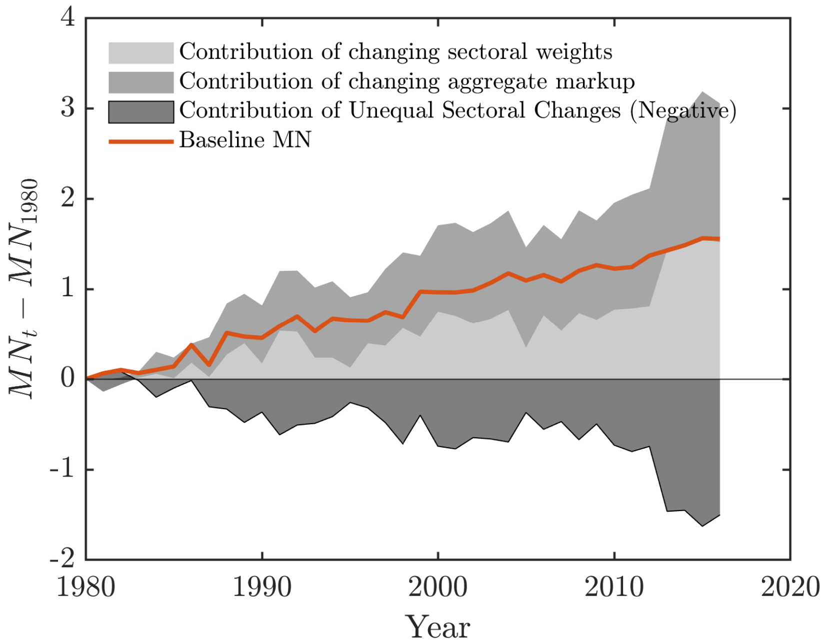
Three features account for the evolution of markups over time and, therefore, are important for aggregate monetary non-neutrality. First, the aggregate markup increases over time. Second, sector sizes, $\omega_{k,t}$, change over time. More specifically, the aggregate production is re-allocating to sectors with higher markups (De Loecker et al. 2020). Third, changes in markups are heterogeneous across sectors. Next, we decompose the evolution of aggregate monetary non-neutrality into these three components.

The foundation of the proposed decomposition exercise is the following. In the first step, we compute the evolution of aggregate monetary non-neutrality over time by using sectoral smoothed markups but fixing sector sizes at their values in 1980. The resulting evolution of aggregate monetary non-neutrality is labeled as $MN_t$(Constant Weight), plotted in dotted line in Figure A.5a. Interestingly, had the sectoral weights remained constant, the aggregate monetary non-neutrality would have remained stable, despite the increased aggregate markup (dashed black line in Figure A.5b). This result is due to unequal changes in markup.

To see this, in the second step, we construct counterfactual evolution of sectoral smoothed markups that feature equal changes in markups over time. And for the comparison to be meaningful, the resulting counterfactual aggregate markup is the same as in the first step. Sector sizes are fixed at their values in the year 1980. We then compute the implied evolution of aggregate monetary non-neutrality, labeled as $MN_t$(Constant Weight and Equal Changes): see the dashed line in Figure A.5a. The comparison between $MN_t$(Constant Weight and Equal Changes) and $MN_t$(Constant Weight) shows the importance of unequal changes in sectoral markups for the evolution of aggregate monetary non-neutrality.

Figure 4 plots the decomposition results based on the counterfactual analysis conducted above. The heights of the lightest colored area plots $MN_t$(Baseline) $-MN_t$(Constant weight): the difference between the aggregate monetary non-neutrality and $MN_t$(Constant Weight). It indicates the contribution of changes in sectoral sizes to aggregate monetary non-neutrality. The darkest colored area plots $MN_t$(Constant Weight) $-MN_t$(Constant Weight and Equal Changes), which indicates the contribution of markup dispersions to the aggregate monetary non-neutrality. This component is negative and quantitatively

**Figure 4:** Monetary Non-neutrality Over Time: A Decomposition




*Note:* This figure decomposes the evolution of aggregate monetary non-neutrality into three components reflecting the contribution of 1) changing sectoral weights, 2) unequal changes of sectoral markups, 3) equal increase in markups at the sectoral level.

significant. Moreover, the increase in absolute value over time of this component reflects the increased discrepancy across sectoral markups observed in the data. The intermediate colored area plots the remaining component: $MN_t$(Constant Weight and Equal Changes). It attributes the changes in the aggregate monetary non-neutrality to equal changes in markups at the sectoral level, holding sector sizes and dispersion constant.

Overall, the decomposition exercise illustrates the equal importance of (i) evolutions of sector sizes and (ii) unequal changes in sectoral markups for the changes in aggregate monetary non-neutrality.

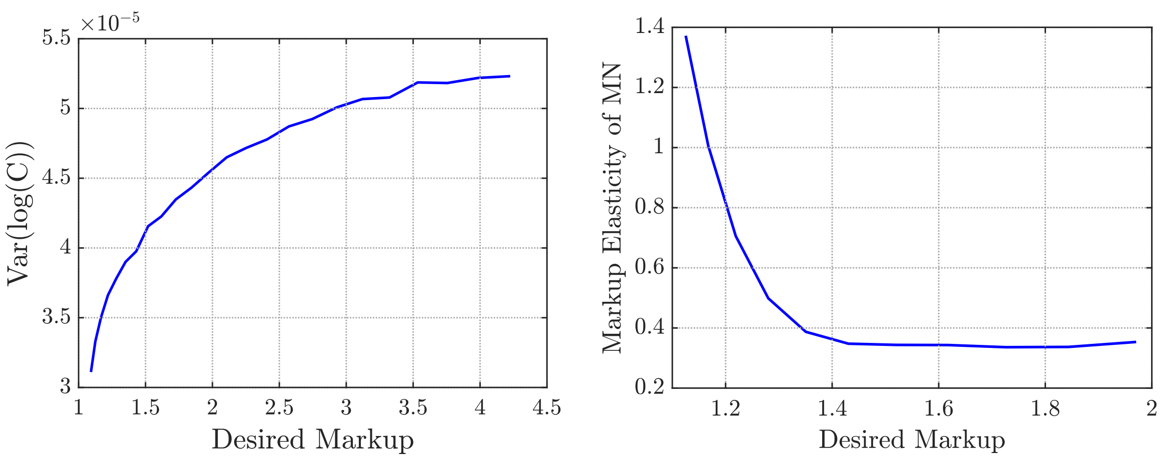
### 4.4 Inspecting the Mechanisms

This section shows that monetary non-neutrality is an increasing and concave function of firms' optimal markups, keeping everything else equal. Moreover, we assess the effects of market power on the frequency and the selection effects of monetary policy. Both margins are increasing and concave with firms' desired markups but quantitatively the frequency margin dominates.

**Monotonicity and Concavity** To illustrate the monotonicity and concavity relationship, we introduce a systematic monetary policy shock into our calibrated one-sector model in the previous section. Specifically, we calibrate $\mu = 0.0018$ to match the mean growth of nominal GDP minus real GDP and $\sigma_S = 0.0032$ to match the standard deviation of nominal GDP growth rate during the period 1998-2005. We solve the model numerically using the method proposed by Krusell and Smith (1998).[9] We then vary the desired markups in the economy and solve for the equilibrium with different markups to compute monetary non-neutrality as a function of market power.

Figure 5a demonstrates that monetary non-neutrality, measured by the variance of log output, is an *increasing* and *concave* function of firms' optimal markups.[10] The concavity implies that the markup elasticity of monetary non-neutrality decreases in the desired markup: the higher the desired markup, the less the markup elasticity of monetary non-neutrality (see Figure 5b).

**Figure 5:** Markups and Monetary Non-neutrality




**(a)** Monetary Non-neutrality

**(b)** The Markup Elasticity of MN

*Note:* Panel (a) plots how monetary non-neutrality, measured by the variance of log output varies with the desired markup in our one-sector quantitative model. The y-axis in Panel (b) is the markup elasticity of monetary non-neutrality, which is defined as $\frac{\triangle log(\text{monetary non-neutrality})}{\triangle log(\text{markup})}$.

[9] Similar methods have been used in the class of menu cost models, for example, by Nakamura and Steinsson (2010), Midrigan (2011) and Vavra (2014). Previously, we considered one-time perfect foresight shocks in the multi-sector model instead of systematic monetary policy shocks. Because it is extremely difficult to reduce the infinite-dimensional state variable (distribution of last-period prices) to a computationally feasible finite number of states.

[10] Nakamura and Steinsson (2010) proposed to use of the variance of log output (conditional on nominal demand shocks) as a measure of monetary non-neutrality in a model solved with systematic monetary policy shocks.

We can now revisit our quantitative results through the lens of this monotonic and concave relationship. The monotonicity explains Result 1 discussed in Section 4.1, which shows that monetary non-neutrality increases with aggregate markups over time. The concavity, combined with the fact that the right tail mostly drives the unequal increase in sectoral markups, explains Result 2 discussed in Section 4.2. Result 2 highlighted that an equal increase in the sectoral markups would have led to a further increase in aggregate monetary non-neutrality.

Note that the concave relationship between the aggregate monetary non-neutrality and the aggregate desired markup is also present in the calibrated multi-sector model studied in Section 4.1. To illustrate this, we re-plot Figure 2 by displaying the aggregate monetary non-neutrality as a function of the aggregate desired markup. Figure A.6 plots the result with a fitted line using a polynomial function of order three. The concave relationship is apparent.

In the remaining part of this section, we inspect the fundamental mechanism behind this increasing and concave relationship.

**The Mechanisms** To understand the mechanisms behind the increasing concave relationship highlighted above, we follow the sufficient statistic approach described in Alvarez et al. (2022). Specifically, the cumulative output impulse response $\mathcal{M}(\Delta s; \mu_{ss})$ to a log nominal spending shock $\Delta s$ is characterized by two steady-state statistics: the frequency of price changes $Freq(\mu_{ss})$ and the kurtosis of the distribution of price changes $Kurt(\mu_{ss})$. More specifically,

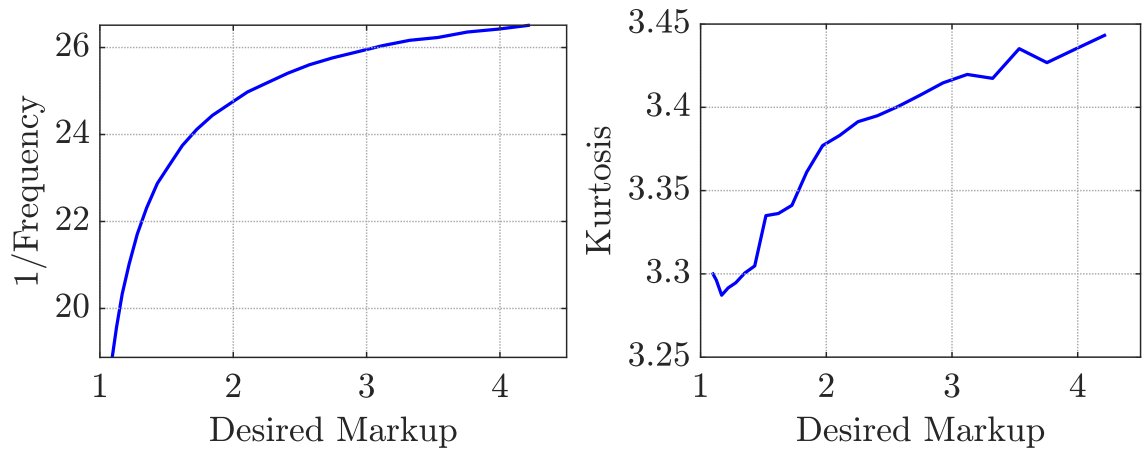
$$\mathcal{M}(\Delta s; \mu_{ss}) = \frac{1}{6} \frac{Kurt(\mu_{ss})}{Freq(\mu_{ss})} \times \Delta s + o(\Delta s^2). \tag{18}$$

Note that both the steady-state frequency of price changes and the kurtosis of the distribution of price changes depend on the steady-state desired markup $\mu_{ss}$. Two components, therefore, determine the monetary non-neutrality: the *frequency* effect, which is proportional to $1/Freq(\mu_{ss})$, and the *selection* effect, which is proportional to $Kurt(\mu_{ss})$.[11]

Figures 6a and 6b illustrate how the frequency and selection effects vary with steady-state markups, respectively. Similar to the aggregate monetary non-neutrality observed in Figure 5a, both effects are increasing concave functions of the desired markups. Quantitatively, the frequency margin contributes approximately 90% to the increase in monetary non-neutrality originating from the increased markups. The markup elasticity in our model is, therefore, mainly driven by the frequency margin, i.e., through the effects of market

[11] As shown in Alvarez et al. (2022), the kurtosis of the distribution of price changes is associated with the selection effect found in the previous literature (e.g., Golosov and Lucas (2007) and Midrigan (2011)).

**Figure 6:** Aggregate Monetary Non-neutrality: Frequency and Selection Effects




**(a)** The Frequency Effect

**(b)** The Selection Effect

*Note:* This figure plots how the frequency effect (panel a) and the selection effect (panel b) in response to a monetary policy shock vary with the desired markup. To plot this figure, we vary the desired markups in our calibrated one-sector model in the previous section and solve for the equilibrium of models with different markups.

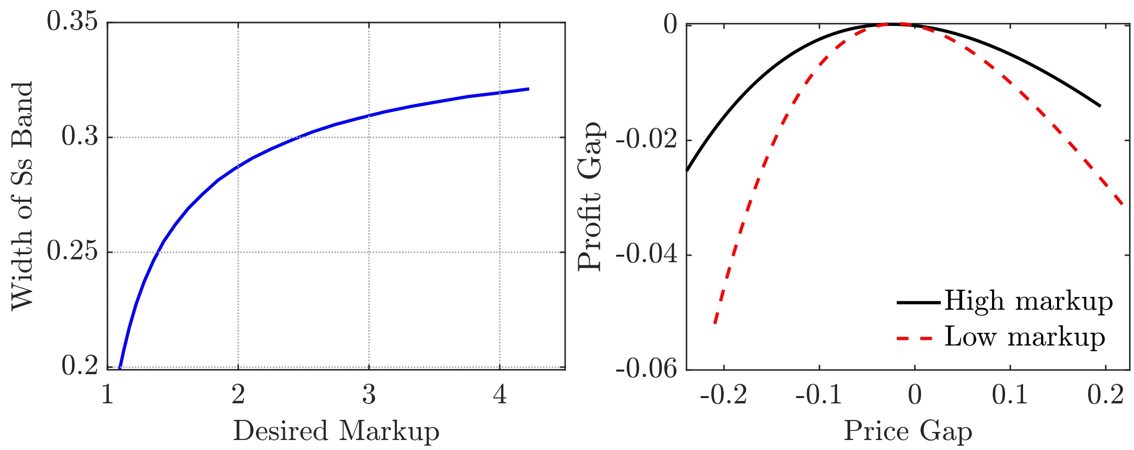
power on the frequency of price adjustment.

Given its quantitative significance, we focus on why the frequency effect is increasing and concave with steady-state markups. In our model, the steady-state frequency of price adjustment (FPA) is determined by 1) the distribution of price gaps and 2) the width of the Ss band that characterizes the non-adjusting region. Formally, the Ss band is defined as the interval between the upper and the lower bounds of price gap values within which a firm does not adjust its price. The price gap is defined as the distance between a firm's current price and its optimal price.

Our first observation is that the aggregate price gap distribution is almost invariant to variations in steady-state optimal markups. Figure A.7 shows that the price gap distribution is nearly identical when the desired markup varies. Therefore, we focus on how the width of the Ss band differs in models with different steady-state optimal markups.

Figure 7a shows that the Ss band is an increasing and concave function of the desired markups. The intuition behind this result is that as markups increase, firms' profits as functions of price gaps become less curved, as shown in Figure 7b. In other words, it is less costly to deviate from the optimal prices. Moreover, as markups increase, the marginal decrease of the curvature of the profit function is incrementally small. The width of the Ss band is, therefore, an increasing and concave function of firms' optimal markups. As

**Figure 7:** The Width of Ss Band and the Shape of Profit Function




**(a)** The Width of the Ss Band

**(b)** Profit Functions

*Note:* Panel (a) plots how the average width of the Ss band varies with the desired markup. Panel (b) plots the profit gap as a function of price gaps for different calibrations of markups: the high markup (1.6) and the low markup (1.2) cases. The profit gap is defined as the difference between a firm's profit given its price $p$ and the firm's profit under its optimal resetting price ($p^*$). Similarly, the price gap is defined as $p/P - p^*/P$.

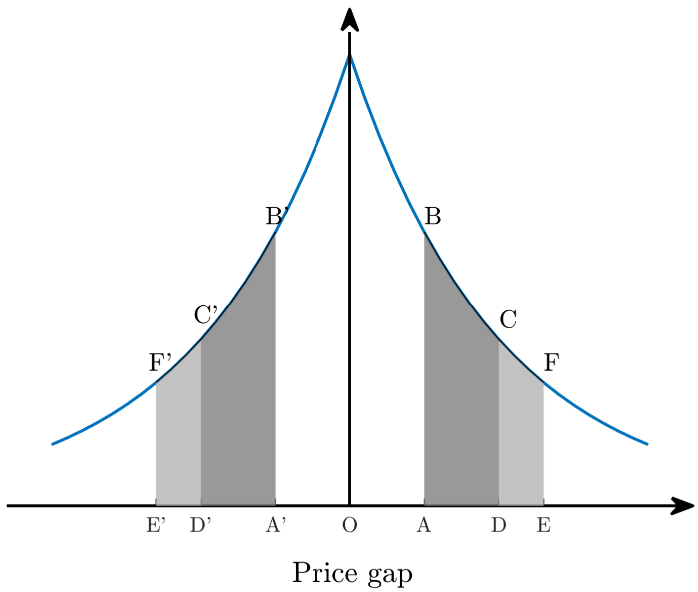
a result, the steady-state FPA is a decreasing and convex function of the desired markup. Correspondingly, its inverse, which reflects the frequency effect, is an *increasing* and *concave* function of the desired markup.

Figure 8 illustrates the change in the Ss band when the desired markup changes graphically. The Ss bands AB(A'B'), CD(C'D') and EF(E'F') correspond to models where markups are $\mu - \Delta$, $\mu$ and $\mu + \Delta$ respectively. The length of AD(A'D') is greater than that of DE(D'E') due to the increasing and concave relationship between the width of the Ss band and desired markup. The size of the area ABCD(A'B'C'D') is, therefore, larger than that of the area CDEF(C'D'E'F'). Consequently, as markup increases equally, the decrease in FPA is smaller. In other words, steady-state FPA is a *decreasing* and *convex* function of the desired markup.

To provide additional intuition, in Appendix D, we use a classic analytical menu cost model to show that the width of the Ss band is indeed an increasing and concave function of firms' optimal markups, and consequently, the frequency effect is an increasing and concave function of firms' optimal markups.[12]

[12]This type of models has been popularized by Barro (1972) and Dixit (1991). For recent progress, see Alvarez and Lippi (2014), and Alvarez et al. (2016) for examples.

**Figure 8:** Markups and Frequency of Price Adjustment




*Note*: This figure is for illustrative purpose. The Ss bands AB(A′B′), CD(C′D′) and EF(E′F′) correspond to models where markups are $\mu - \Delta$, $\mu$ and $\mu + \Delta$ respectively.

**Additional Discussions** The endogenous degree of nominal rigidity (FPA) arising from menu cost and the curvatures of firms' profit functions are essential for the findings of our paper. This is due to the decreased curvature of the profit function (as a function of price gaps) as steady markups increase, illustrated in Figure 7b.

Note that the curvatures of firms' profit functions that we emphasize are not specific to the monopolistic competition employed in this paper. Appendix E shows that the same features hold true in a model with the oligopolistic competition. This finding hints at the potential generalization of our results: our results hold as long as less market competition or greater market concentration makes firms' profits loss less sensitive to the price gaps.

## 5. Empirical Evidence

This section presents three pieces of empirical evidence that (i) support predictions of the model, (iii) support the model mechanisms, and (iii) differentiate the proposed mechanisms from the ones that exist in the literature.

## 5.1 Monetary Non-neutrality and Markup

First, we estimate monetary non-neutrality in the data and assess how it depends on the smoothed markup. Specifically, we estimate the following model:

$$y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + controls_t + \varepsilon_{t+h} \tag{19}$$

where, $y_t$ is the log of real GDP, $\epsilon_t^m$ denotes the monetary shock, $log(\mu_t^{ss})$ is the log of smoothed aggregate markup, and $\varepsilon_t$ is the residual. $controls_t \equiv \sum_{i=1}^p \beta_{i,x}^h X_{t-i} + \sum_{i=1}^p \gamma_{i,x}^h X_{t-i} \cdot log(\mu_t^{ss}) + \sum_{i=1}^p \gamma_{\mu,i}^h log(\mu_{t-i}^{ss}) + \gamma_3^h trend_t$, where, $X_t$ is a vector of control variables and $trend_t$ denotes the time trend. Importantly, the model includes a term that interacts with monetary shock with the smoothed aggregate markup ($\epsilon_t^m \cdot log(\mu_t^{ss})$) to allow for the interaction between the steady state markup and monetary non-neutrality.

The null hypothesis is $\beta_2^h = 0$, and $\beta_2^h < 0$ indicates that increased markup renders monetary policy more effective (provided that $\beta_1^h$ is negative). The vector of controls $X_t$ includes the log of real GDP and GDP deflator and the 2-year treasury yield. Four lags ($p = 4$) of control variables are included.

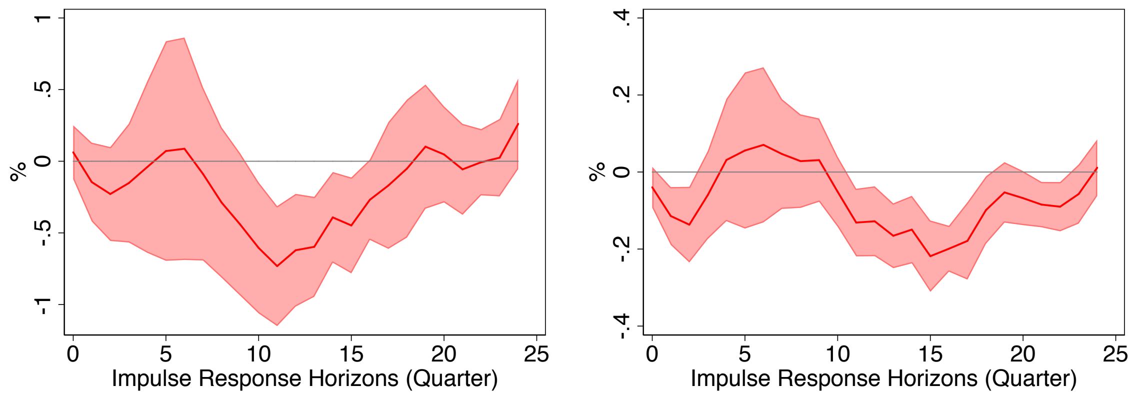
One needs an exogenous measure of monetary shocks to estimate the causal effects of monetary policy shocks on real GDP. We use the state-of-the-art monetary policy shock constructed based on a proxy VAR using high-frequency identified (HFI) monetary policy surprises as instruments. Specifically, we extend Gertler and Karadi (2015)'s monetary policy shock up to 2015Q4.

Through the lens of the empirical model, monetary non-neutrality (effects of monetary shock on the log of real GDP) at the horizon h is defined as:

$$MN^h = \frac{\partial y_{t+h}}{\partial \epsilon_t^m} = \beta_1^h + \beta_2^h log(\mu_t^{ss}). \tag{20}$$

Figure 9 plots the results from estimating model (19) for each $h = 1, 2, \ldots, 20$. The left panel plots the average $MN^h$: $\hat{\beta}_1^h + \hat{\beta}_2^h \overline{log(\mu^{ss})}$ with $\overline{log(\mu^{ss})}$ denoting the average (log of) smoothed markup in the sample. The right panel plots $\hat{\beta}_2^h$: the additional response of real GDP due to a 1% increase in the average markup. The considered shock is a contractionary monetary shock. The shaded area indicates the 90% confidence interval constructed using Newey West standard errors.

**Figure 9:** The Markup-Dependent Effects of Contractionary Monetary Policy Shocks




**(a)** The average effects of monetary shocks

**(b)** Add. resp. due to 1% increase in $\mu^{ss}$

*Note:* This figure plots the markup-dependent IRF of real GDP to a contractionary monetary policy shock. Specifically, the panel (a) plots $\hat{\beta}_1^h + \hat{\beta}_2^h\overline{log(\mu^{ss})}$ and the panel (b) plots $\hat{\beta}_2^h$ from estimating the following model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + controls_t + \varepsilon_{t+h}$, where $\epsilon_t^m$ denotes Gertler and Karadi (2015) monetary policy shocks and $\overline{log(\mu^{ss})}$ is the average (log of) smoothed markup in the sample. The shaded area indicates the 90% confidence interval constructed using Newey West standard errors. Sample: 1986Q1-2015Q4.

The key takeaway from Figure 9 is that the null of monetary non-neutrality being orthogonal to the steady state level of markup is rejected at various horizons. More importantly, the estimation implies that the increase in smoothed markups observed in the past decades has led to an increase in monetary non-neutrality.

HFI monetary surprises might contain the Fed's private information as discussed in Nakamura and Steinsson (2018), Jarociński and Karadi (2020), Miranda-Agrippino and Ricco (2021), and Zhang (2022). To address this concern, we conduct a robustness check using Miranda-Agrippino and Ricco (2021) monetary policy shocks that are orthogonal to the Fed's internal forecasts. The main findings are unaffected: see Figure A.8.

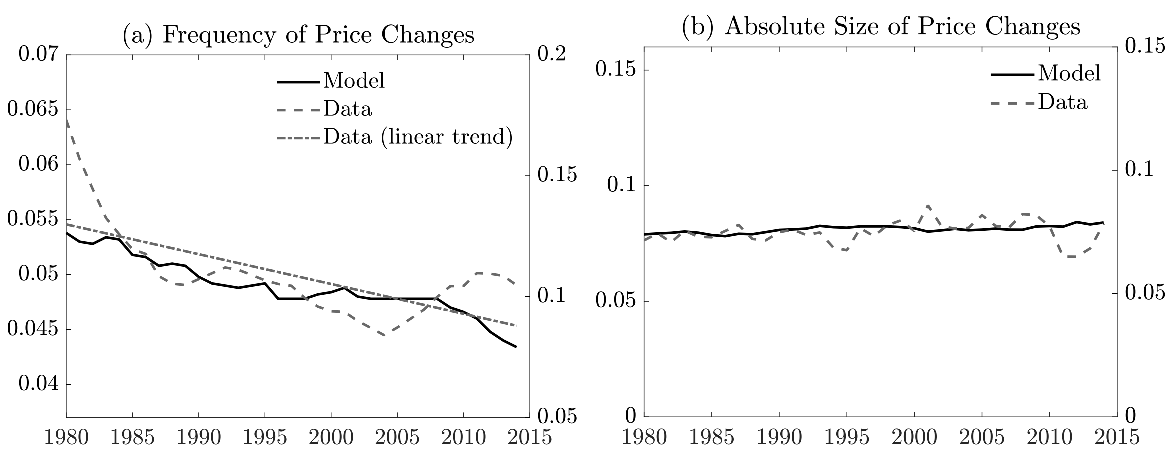
## 5.2 Differentiating from the Existing Mechanisms

This subsection provides evidence that supports our mechanism. More importantly, the evidence differentiates our mechanism from the existing ones (Wang and Werning 2022 and Baqaee et al. 2021) in explaining the increased monetary non-neutrality arising from reduced market competition.

**Pricing Moments Over Time** Our framework focuses on the endogenous relationship between market competition and the degree of sluggishness in adjusting prices. Specifically, our quantitative results rest on the mechanism that reduced market competition renders a decreased FPA, through which monetary non-neutrality increases (the frequency effect). The first relevant empirical question is whether the FPA has declined in the data and, more importantly, whether the *magnitude* of the decline is aligned with the predictions of the model.

Panel (a) in Figure 10 plots the frequency of price changes over time in the data together with the model-implied changes (solid line) resulting from the increase of the aggregate steady-state markup. The dashed (dash-dotted) line in panel (a) plots the moving average smoothed (linear trend) FPA in the data (source: Nakamura et al. 2018). The solid line in panel (a) plots the model implied FPA over time. The message is clear: The data support the model's mechanism both qualitatively and quantitatively. Moreover, Panel (b) in Figure 10 plots the evolution of the absolute size of price changes in the data and the model. Again, the two lines are well aligned, which serves as another out-of-the-sample validation of the model.

**Figure 10:** Frequency and Absolute Size of Price Changes




*Note*: The dashed (dash-dotted) line in panel (a) plots the moving average smoothed (linear trend) FPA in the data (source: Nakamura et al. 2018). The solid line in panel (a) plots the model implied FPA over time. Panel (b) depicts the evolutions of the absolute size of price changes in the data and the model.

**The Relevance of the Frequency Margin** To further differentiate our mechanism from the existing ones (Wang and Werning 2022 and Baqaee et al. 2021), we present additional evidence that suggests the relevance of our mechanism—the frequency margin—in explaining the rising monetary non-neutrality in the data.

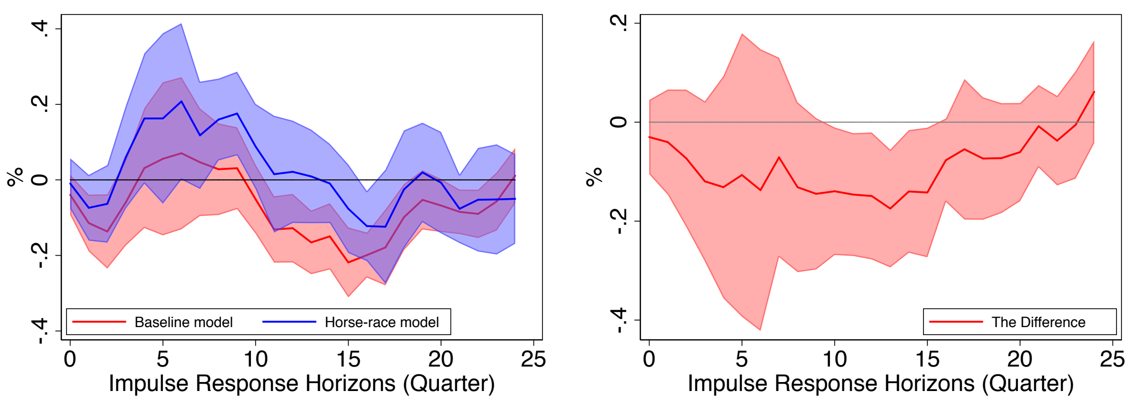
To this end, we estimate the following horse-race model:

$$y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + \beta_3^h \epsilon_t^m \cdot FPA_t + controls_t + \varepsilon_{t+h}. \tag{21}$$

The interaction between monetary policy shock and time-varying FPA ($\epsilon_t^m \cdot FPA_t$) is included in the model, in addition to the baseline regression specification (19). The set of control variables $controls_t$ is identical to those included in the model (19) such that any changes in the estimated $\beta_2^h$ can be attributed to the inclusion of $\epsilon_t^m \cdot FPA_t$.[13]

Two results can be interesting to readers. First, the mainstream monetary economics models predict that $\beta_3^h$ to be positive: *reduced* frequency of price adjustment renders monetary policy shocks more potent in affecting real GDP. We verify this prediction in the data: see Figure A.9.

**Figure 11:** The Relevance of the Frequency Effect




**(a)** $\widehat{\beta}_2^h$: Baseline v.s Horse-race

**(b)** Baseline $\widehat{\beta}_2^h$ − Horse-race $\widehat{\beta}_2^h$

*Note:* The red line in panel (a) plots $\hat{\beta}_2^h$ from estimating the baseline model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + controls_t + \varepsilon_{t+h}$. The blue line in panel (a) plots $\hat{\beta}_2^h$ from estimating the horse-race model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + \beta_3^h \epsilon_t^m \cdot FPA_t + controls_t + \varepsilon_{t+h}$, where $\epsilon_t^m$ and $FPA_t$ indicate Gertler and Karadi (2015) monetary policy shocks and the smoothed frequency of price adjustments, respectively. Panel (b) depicts the difference in the estimated $\hat{\beta}_2^h$s across the two models. The shaded area indicates the 90% confidence interval. Sample: 1986Q1-2015Q4.

A more important result emerges by comparing $\beta_2^h$ from estimating (21) with the corresponding parameters from estimating the baseline specification (19). If markups

[13]Note that while the model predicts a perfect correlation between $FPA_t$ and $\mu_t^{ss}$, this is not the case in data. Thus, there is no issue of perfect collinearity between $\epsilon_t^m \cdot log(\mu_t^{ss})$ and $\epsilon_t^m \cdot FPA_t$ in the data. This is not surprising, given that stylized models can never fully capture the richness of actual data,

affect monetary non-neutrality through channels that were orthogonal to the ones we stress in the paper (i.e., FPA), then one should expect $\beta_2^h$ to remain unchanged. On the other extreme, if the effects of markup on monetary non-neutrality were entirely driven by the frequency margin ($\epsilon_t^m \cdot FPA_t$), then one should expect $\beta_2^h$ to be zero in the horse-race model. In general, the magnitude of changes in $\beta_2^h$ across the two models indicates the relevance of the frequency margin highlighted in our quantitative model.

Figure 11 plots the findings. Panel (a) plots $\hat{\beta}_2^h$s from estimating the baseline model (in red) and the horse-race model (in blue). In contrast to the baseline model, the contribution of markup to monetary non-neutrality is less significant economically and statistically. Moreover, the differences in the estimated $\hat{\beta}_2^h$s across the two models are significant: see Panel (b). The shaded area indicates the 90% confidence interval. These results suggest that the mechanism (frequency margin) that we emphasize in the paper explains a big chunk of changes in monetary non-neutrality arising from changing markups.

Those findings are robust to the use of an alternative measure of monetary policy shocks (Miranda-Agrippino and Ricco 2021): see Figure A.10.

# 6. Conclusion

Firms' market power, measured by markups, has risen substantially in the past decades. Moreover, the changes in markups are heterogeneous across sectors. This paper assesses the implications of these trends for the real effects of monetary policy. We develop a quantitative menu cost model that features multiple sectors with heterogeneous degrees of market competition.

Two quantitative results stand out. First, the average markup elasticity of monetary non-neutrality in the United States is equal to 1 in the past three decades: i.e., the thirty percent increase in markups in the data raises monetary non-neutrality by thirty percent. Second, the unequal changes in markups at the sectoral level act as a counterforce: the markup elasticity of monetary non-neutrality would be equal to 1.4 had the markup increased equally across sectors.

These results are due to (i) a decrease in the frequency of firms' price adjustments resulting from increased market power and (ii) a concave relationship between markups and monetary non-neutrality.

We provide evidence supporting the model's predictions and mechanisms. Moreover, we present evidence suggesting that the proposed mechanism—frequency effect—accounts for a big chunk of markup elasticity of monetary non-neutrality observed in the data.

Our paper provides a toolbox that assists central bankers in keeping track of monetary non-neutrality. Our calculation of the markup elasticity of monetary non-neutrality can inform central banks to determine the correct amount of nominal demand stimulus package in the current and future economy with rising markups.

## References


Alvarez, F., H. Le Bihan, and F. Lippi (2016). The real effects of monetary shocks in sticky price models: a sufficient statistic approach. *American Economic Review 106*(10), 2817–51.

Alvarez, F. and F. Lippi (2014). Price setting with menu cost for multiproduct firms. *Econometrica 82*(1), 89–135.

Alvarez, F., F. Lippi, and A. Oskolkov (2022). The macroeconomics of sticky prices with generalized hazard functions. *The Quarterly Journal of Economics 137*(2), 989–1038.

Alvarez, F. E., A. Ferrara, E. Gautier, H. Le Bihan, and F. Lippi (2021). Empirical investigation of a sufficient statistic for monetary shocks. Technical report, National Bureau of Economic Research.

Atkeson, A. and A. Burstein (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review 98*(5), 1998–2031.

Baqaee, D., E. Farhi, and K. Sangani (2021). The supply-side effects of monetary policy. Technical report, National Bureau of Economic Research.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Barro, R. J. (1972). A theory of monopolistic price adjustment. *The Review of Economic Studies 39*(1), 17–26.

Bils, M. and P. J. Klenow (2004). Some evidence on the importance of sticky prices. *Journal of political economy 112*(5), 947–985.

Bils, M., P. J. Klenow, and B. A. Malin (2012). Reset price inflation and the impact of monetary policy shocks. *American Economic Review 102*(6), 2798–2825.

Carvalho, C. (2006). Heterogeneity in price stickiness and the real effects of monetary shocks. *Frontiers in Macroeconomics 2*(1).

Carvalho, C., N. A. Dam, and J. W. Lee (2020). The cross-sectional distribution of price stickiness implied by aggregate data. *Review of Economics and Statistics 102*(1), 162–179.

Carvalho, C. and F. Schwartzman (2015). Selection and monetary non-neutrality in time-dependent pricing models. *Journal of Monetary Economics 76*, 141–156.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics 135*(2), 561–644.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *American economic review 102*(6), 2437–71.

Dixit, A. (1991). Analytical approximations in models of hysteresis. *The Review of Economic Studies 58*(1), 141–151.

Dotsey, M., R. G. King, and A. L. Wolman (1999). State-dependent pricing and the general equilibrium dynamics of money and output. *The Quarterly Journal of Economics 114*(2), 655–690.

Gautier, E. and H. L. Bihan (2018). Shocks vs menu costs: Patterns of price rigidity in an estimated multi-sector menu-cost model. *The Review of Economics and Statistics*, 1–45.

Gertler, M. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Gertler, M. and J. Leahy (2008). A phillips curve with an ss foundation. *Journal of political Economy 116*(3), 533–572.

Golosov, M. and R. E. Lucas (2007). Menu costs and phillips curves. *Journal of Political Economy 115*(2), 171–199.

Jarociński, M. and P. Karadi (2020). Deconstructing monetary policy surprises - the role of information shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review 95*(1), 161–182.

Karadi, P. and A. Reiff (2019). Menu costs, aggregate fluctuations, and large shocks. *American Economic Journal: Macroeconomics 11*(3), 111–46.

Krusell, P. and A. A. Smith, Jr (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy 106*(5), 867–896.

Li, M. (2022). Countercyclical dispersion of price changes: Causes and implications. *Working paper*.

Meier, M. and T. Reinelt (Forthcoming 2022). Monetary policy, markup dispersion, and aggregate tfp. *The Review of Economics and Statistics*.

Midrigan, V. (2011). Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica 79*(4), 1139–1180.

Miranda-Agrippino, S. and G. Ricco (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics 13*(3), 74–107.

Mongey, S. (2021). Market structure and monetary non-neutrality. Technical report, National Bureau of Economic Research.

Nakamura, E. and J. Steinsson (2008). Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics 123*(4), 1415–1464.

Nakamura, E. and J. Steinsson (2010). Monetary non-neutrality in a multisector menu cost model. *The Quarterly journal of economics 125*(3), 961–1013.

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics 133*(3), 1283–1330.

Nakamura, E., J. Steinsson, P. Sun, and D. Villar (2018). The elusive costs of inflation: Price dispersion during the us great inflation. *The Quarterly Journal of Economics 133*(4), 1933–1980.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Pasten, E., R. Schoenle, and M. Weber (2020). The propagation of monetary policy shocks in a heterogeneous production economy. *Journal of Monetary Economics 116*, 1–22.

Vavra, J. (2014). Inflation dynamics and time-varying volatility: New evidence and an ss interpretation. *The Quarterly Journal of Economics 129*(1), 215–258.

Wang, O. and I. Werning (2022). Dynamic oligopoly and price stickiness. *The American Economic Review (forthcoming)*.

Young, E. R. (2010). Solving the incomplete markets model with aggregate uncertainty using the krusell–smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control 34*(1), 36–41.

Zhang, D. (2022). Term structure, forecast revision, and the signaling channel of monetary policy. *Journal of the European Economic Association 20*(4), 1522–1553.

# Appendix

## A. Tables and Figures

**Figure A.1:** The evolution of the Aggregate Markup in the U.S

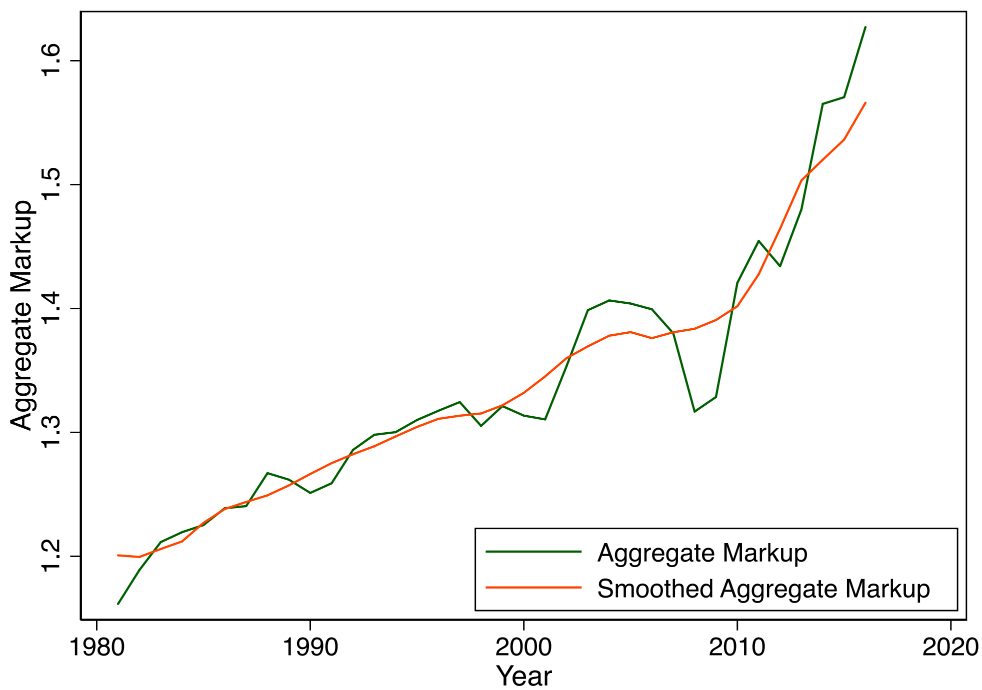
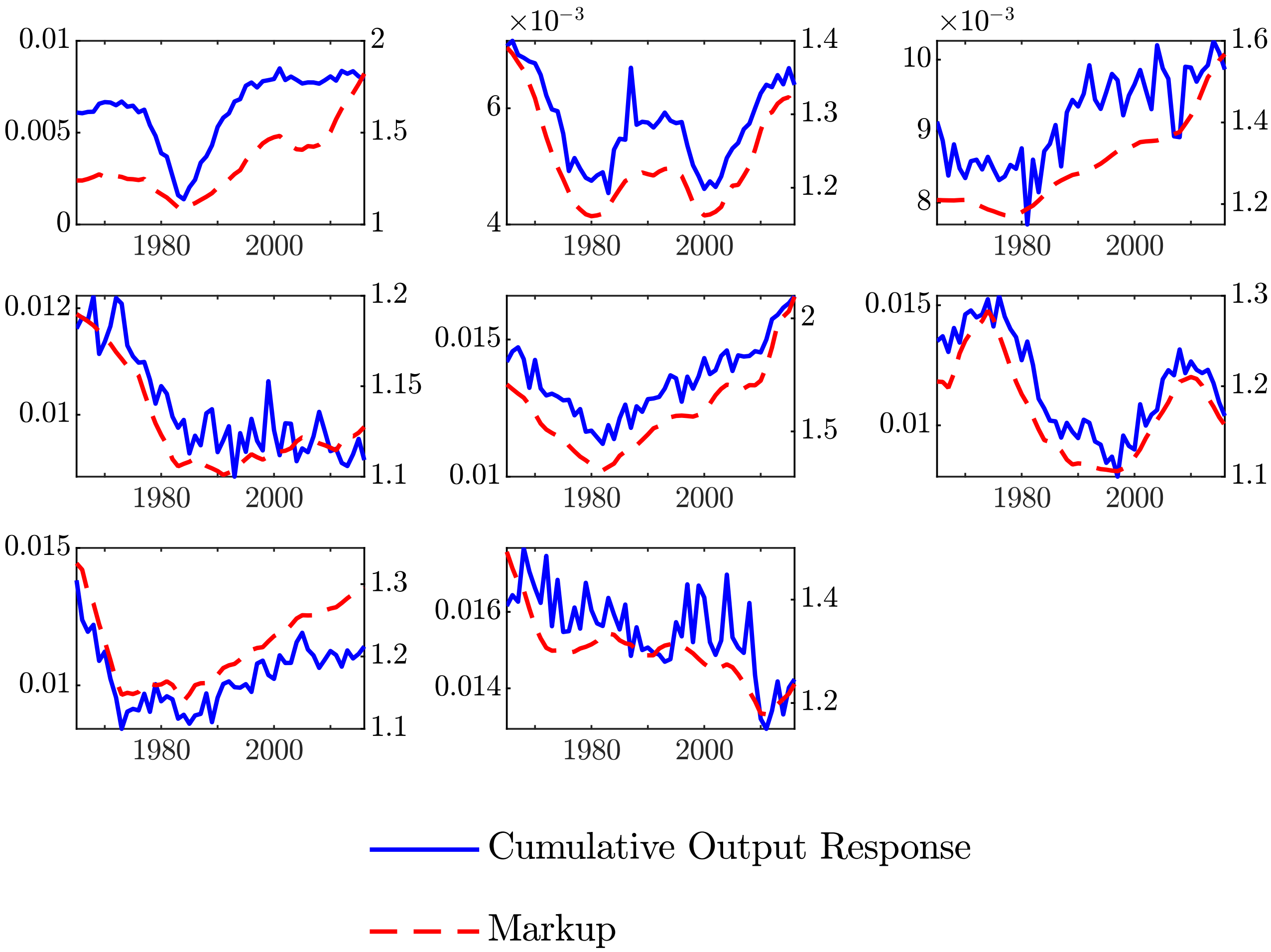
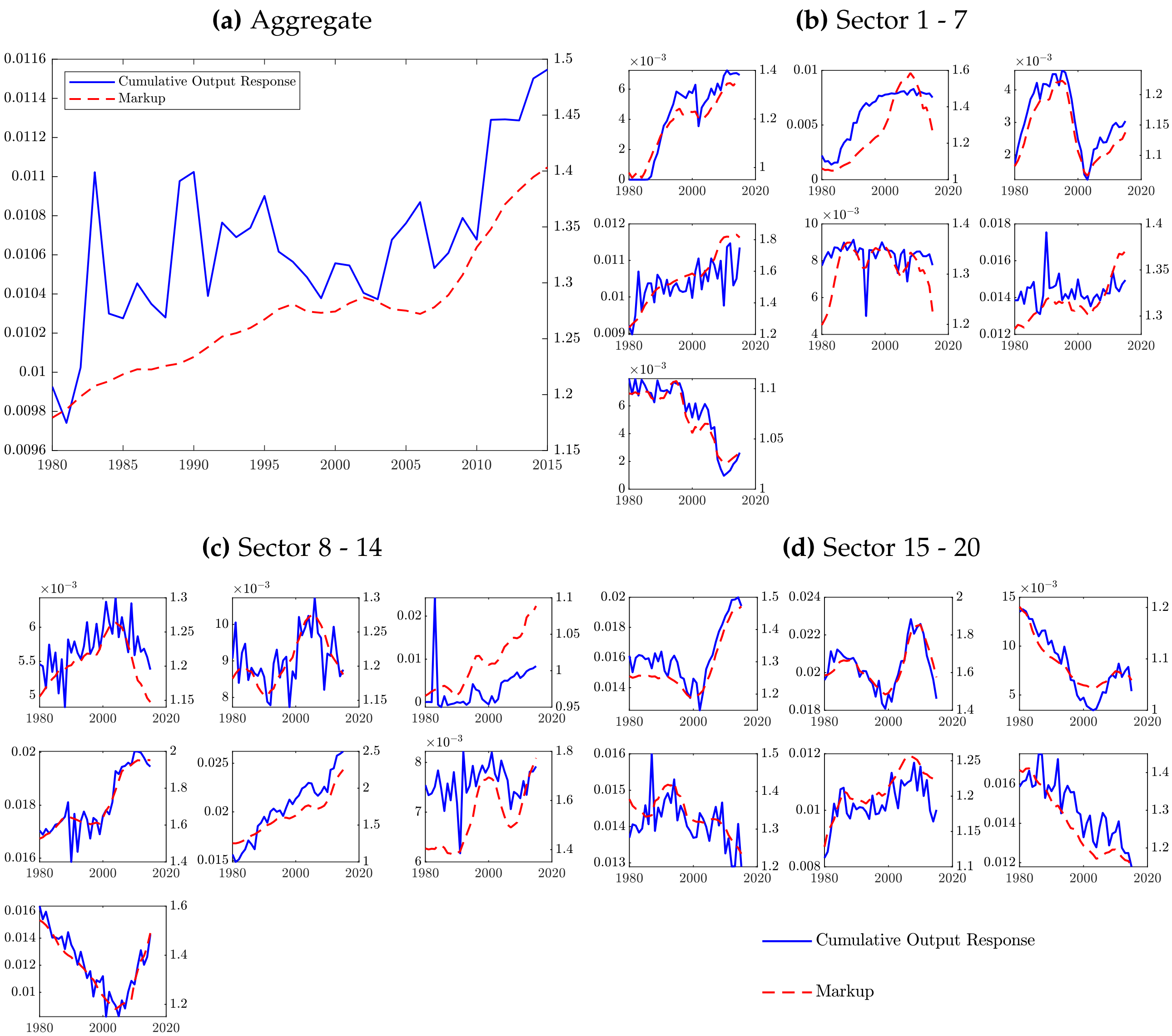


**Figure A.2:** Sectoral Cumulative Output Response Over Time (eight sectors)



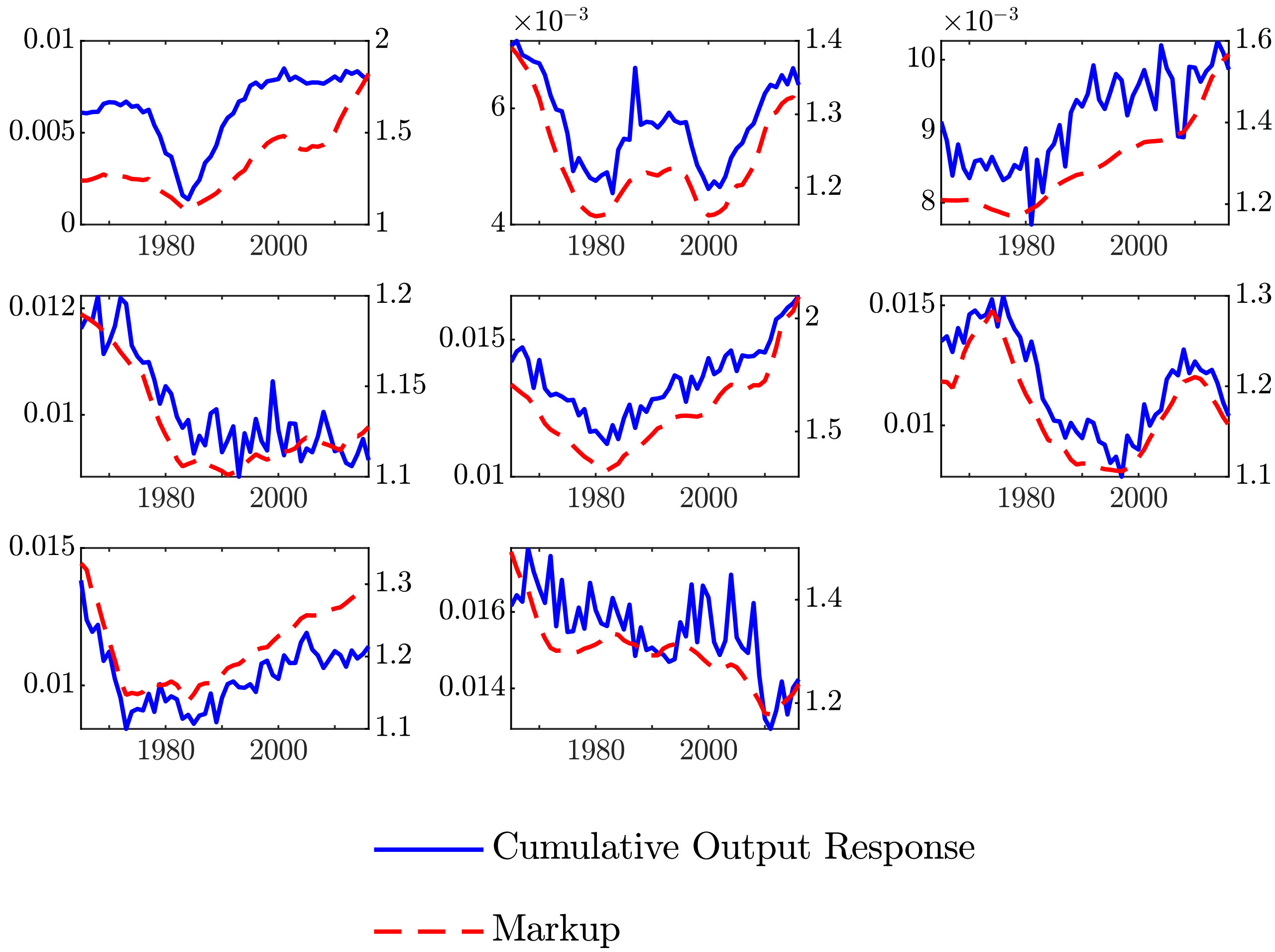

*Note*: This figure plots the evolution of the sectoral makeup in the U.S. (dashed red line) and the implied evolution of the sectoral monetary non-neutrality (solid blue line), measured by the cumulative output response, in our eight-sector model.

**Figure A.3:** Sectoral Cumulative Output Response Over Time (twenty sectors)



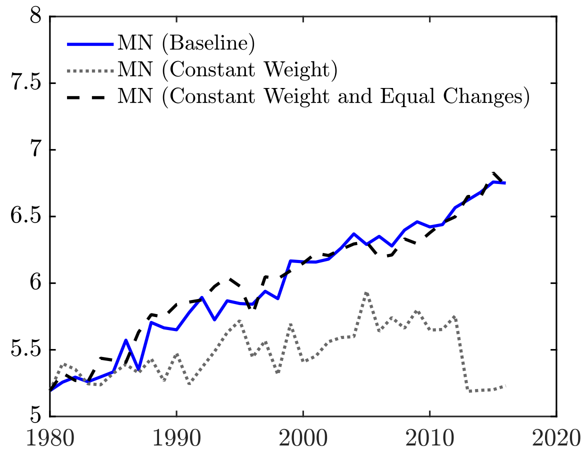

*Note*: This figure plots the evolution of the sectoral makeup in the U.S. (dashed red line) and the implied evolution of the sectoral monetary non-neutrality (solid blue line), measured by the cumulative output response, in our twenty-sector model.

**Figure A.4:** Sectoral Cumulative Output Response Over Time (Equal Changes)

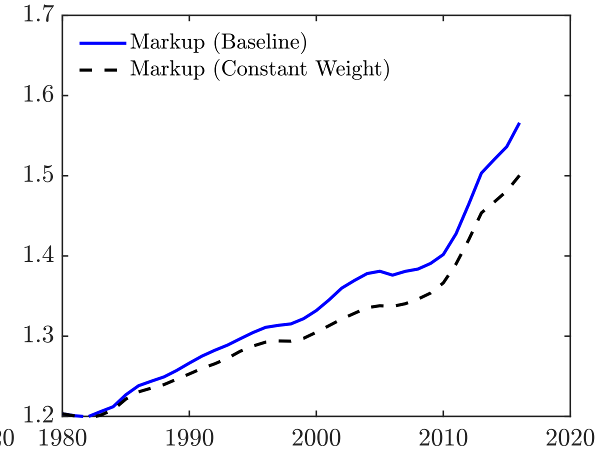



*Note*: this figure plots the evolution of the sectoral makeup in the U.S. had the markups increased equally across sectors (dashed red line) and the implied evolution of the sectoral monetary non-neutrality (solid blue line).

**Figure A.5:** Monetary Non-neutrality Over Time: Some Counterfactuals




**(a)** Cumulative Output Response Over Time




**(b)** Markups Over Time

**Figure A.6:** Markups and Monetary Non-neutrality (Multi-sector model)

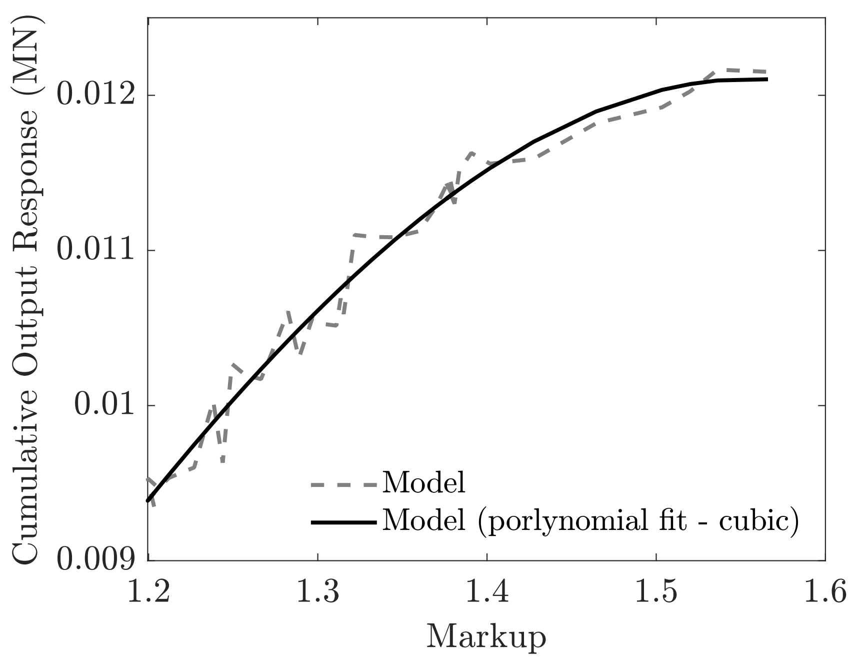
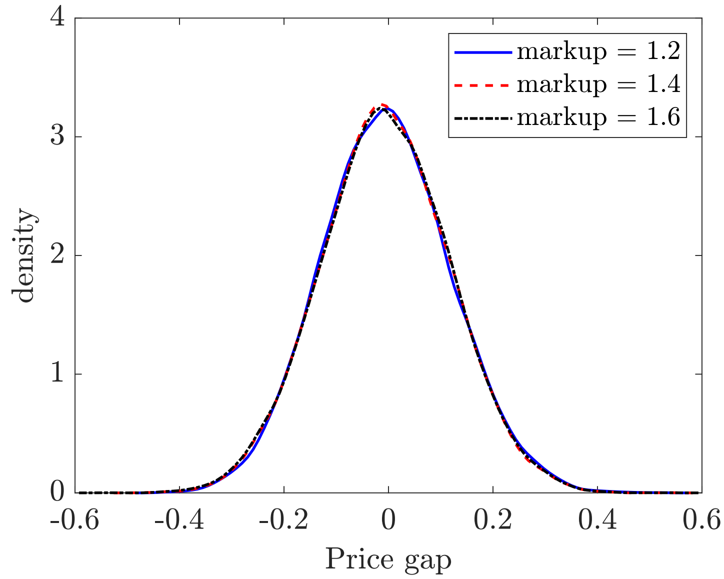
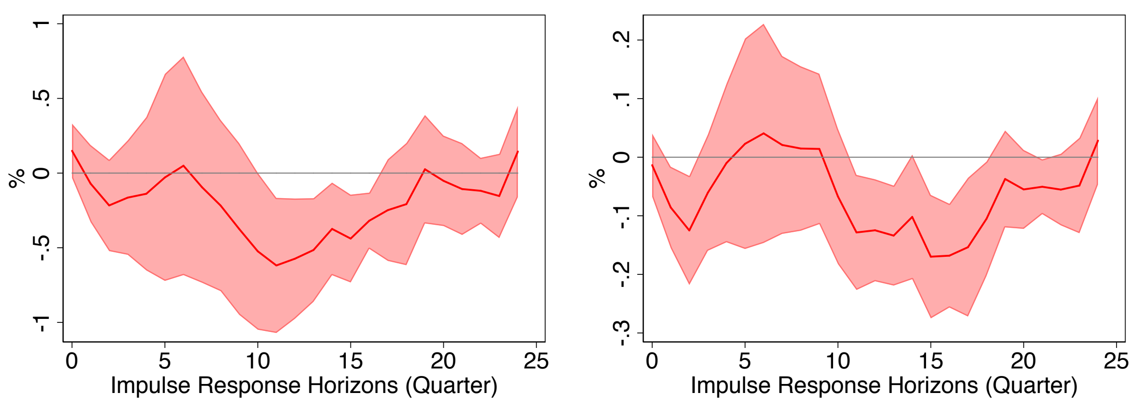


**Figure A.7:** Markup and the Distribution of Price Gaps




*Note*: This figure plots the price gap distribution when aggregate markup is equal to 1.2, 1.4 and 1.6 respectively.

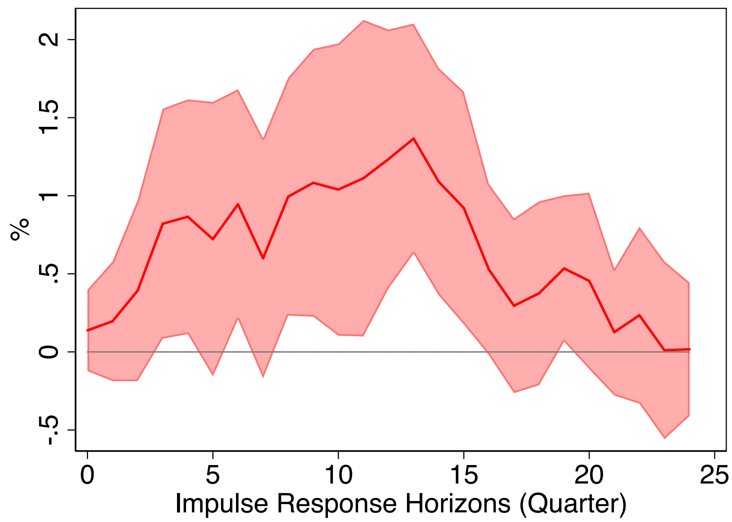
**Figure A.8:** The Markup-Dependent Effects of Monetary Policy Shocks: MR shock




**(a)** The average effects of monetary shocks **(b)** Add. Resp.due to 1% Increase in $\mu^{ss}$

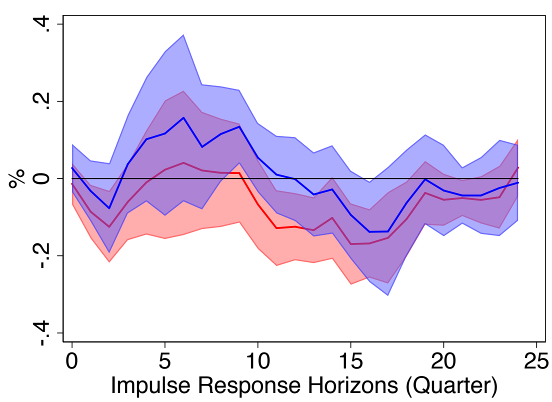
*Note*: This figure plots the markup-dependent IRF of real GDP to a contractionary monetary policy shock. Specifically, the panel (a) plots $\hat{\beta}_1^h$ and the panel (b) plots $\hat{\beta}_2^h$ from estimating the following model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + controls_t + \varepsilon_{t+h}$, where $\epsilon_t^m$ denotes Miranda-Agrippino and Ricco (2021) monetary policy shocks. The shaded area indicates the 90% confidence interval constructed using Newey West standard errors. Sample: 1986Q1-2015Q4.

**Figure A.9:** The Effects of FPA on Monetary Non-neutrality: $\widehat{\beta}_3^h$

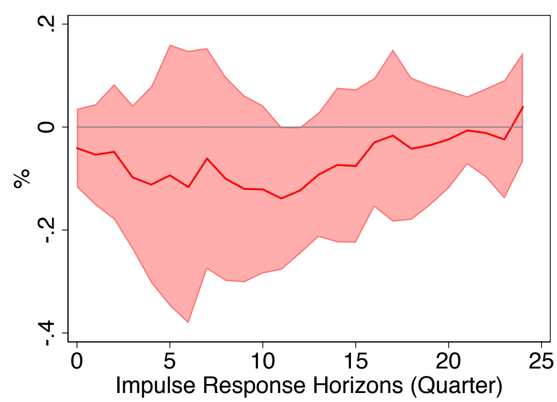



*Note:* This figure plots $\hat{\beta}_3^h$ from estimating the horse-race model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + \beta_3^h \epsilon_t^m \cdot FPA_t + controls_t + \varepsilon_{t+h}$, where $FPA_t$ indicates the smoothed frequency of price adjustments. The shaded area indicates the 90% confidence interval constructed using Newey West standard errors. Sample: 1986Q1-2015Q4.

**Figure A.10:** The Relevance of the Frequency Effect: Alternative Monetary Shock




**(a)** $\widehat{\beta}_2^h$: Baseline v.s Horse-race




**(b)** Baseline $\widehat{\beta}_2^h$ − Horse-race $\widehat{\beta}_2^h$

*Note:* The red line in panel (a) plots $\hat{\beta}_2^h$ from estimating the baseline model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + controls_t + \varepsilon_{t+h}$. The blue line in panel (a) plots $\hat{\beta}_2^h$ from estimating the horse-race model: $y_{t+h} = \beta_0^h + \beta_1^h \epsilon_t^m + \beta_2^h \epsilon_t^m \cdot log(\mu_t^{ss}) + \beta_3^h \epsilon_t^m \cdot FPA_t + controls_t + \varepsilon_{t+h}$, where $\epsilon_t^m$ and $FPA_t$ indicate Miranda-Agrippino and Ricco (2021) monetary policy shocks and the smoothed frequency of price adjustments, respectively. Panel (b) depicts the difference in the estimated $\hat{\beta}_2^h$ across the two models. The shaded area indicates the 90% confidence interval. Sample: 1986Q1-2015Q4.

**Table A.1:** Internally Calibrated Parameters (Twenty Sectors)

| Sector | Markup | $\sigma^z$ | $\alpha$ | $\phi$ | $f$ |
|---|---|---|---|---|---|
| Agriculture | 1.22 | 0.270 | 0.136 | 0.035 | 0.022 |
| Mining, Oil and Gas | 1.41 | 0.108 | 0.114 | 0.054 | 0.020 |
| Utilities | 1.08 | 0.121 | 0.123 | 0.101 | 0.005 |
| Manufacturing (1) | 1.56 | 0.107 | 0.099 | 0.043 | 0.041 |
| Manufacturing (2) | 1.34 | 0.111 | 0.111 | 0.051 | 0.035 |
| Manufacturing (3) | 1.31 | 0.061 | 0.070 | 0.020 | 0.076 |
| Wholesale Trade | 1.06 | 0.067 | 0.163 | 0.010 | 0.085 |
| Retail Trade (1) | 1.26 | 0.090 | 0.178 | 0.056 | 0.039 |
| Retail Trade (2) | 1.25 | 0.078 | 0.137 | 0.024 | 0.076 |
| Transportation | 1.01 | 0.026 | 0.178 | 0.033 | 0.061 |
| Information | 1.71 | 0.052 | 0.071 | 0.054 | 0.060 |
| Finance and Insurance | 1.70 | 0.041 | 0.073 | 0.036 | 0.081 |
| Real Estate | 1.65 | 0.105 | 0.137 | 0.050 | 0.029 |
| Scientific and Tech Services | 1.21 | 0.077 | 0.135 | 0.037 | 0.095 |
| Administrative Services | 1.20 | 0.016 | 0.200 | 0.040 | 0.095 |
| Education | 1.56 | 0.037 | 0.083 | 0.020 | 0.093 |
| Health Care | 1.04 | 0.074 | 0.070 | 0.024 | 0.111 |
| Arts and Entertainment | 1.32 | 0.028 | 0.170 | 0.019 | 0.084 |
| Accommodation and Food Services | 1.22 | 0.048 | 0.141 | 0.027 | 0.058 |
| Other Services | 1.19 | 0.039 | 0.229 | 0.011 | 0.102 |

**Table A.2:** Data and Model Moments (Eight Sectors)

| Moment | Agriculture | Mining and Utilities | Manufacturing | Retail and Wholesale | Services (information, finance) | Education and Health Care | Services (entertainment etc.) | Other services |
|---|---|---|---|---|---|---|---|---|
| **Data Moments** | | | | | | | | |
| *Moments Targeted* | | | | | | | | |
| Frequency | 0.099 | 0.132 | 0.079 | 0.035 | 0.063 | 0.027 | 0.049 | 0.025 |
| Absolute size (median) | 0.154 | 0.068 | 0.081 | 0.046 | 0.041 | 0.076 | 0.056 | 0.069 |
| 25th percentile size | 0.095 | 0.014 | 0.039 | 0.030 | 0.010 | 0.035 | 0.029 | 0.054 |
| 75th percentile size | 0.382 | 0.125 | 0.157 | 0.098 | 0.091 | 0.116 | 0.096 | 0.128 |
| *Moments Not Targeted* | | | | | | | | |
| Size of price increase | 0.236 | 0.062 | 0.076 | 0.062 | 0.052 | 0.104 | 0.069 | 0.105 |
| Size of price decrease | 0.272 | 0.064 | 0.097 | 0.075 | 0.062 | 0.133 | 0.107 | 0.142 |
| **Model Moments** | | | | | | | | |
| *Moments Targeted* | | | | | | | | |
| Frequency | 0.119 | 0.132 | 0.082 | 0.041 | 0.062 | 0.035 | 0.048 | 0.022 |
| Absolute size (median) | 0.201 | 0.066 | 0.081 | 0.050 | 0.042 | 0.038 | 0.059 | 0.082 |
| 25th percentile size | 0.083 | 0.014 | 0.026 | 0.024 | 0.012 | 0.083 | 0.023 | 0.049 |
| 75th percentile size | 0.383 | 0.151 | 0.155 | 0.076 | 0.091 | 0.104 | 0.099 | 0.125 |
| *Moments Not Targeted* | | | | | | | | |
| Size of price increase | 0.214 | 0.076 | 0.087 | 0.053 | 0.052 | 0.076 | 0.065 | 0.090 |
| Size of price decrease | 0.296 | 0.156 | 0.133 | 0.027 | 0.068 | 0.075 | 0.062 | 0.021 |

**Table A.3:** Data and Model Moments (Twenty Sectors), Part 1

| Moment | Agriculture | Mining, Oil and Gas | Utilities | Manufacturing (1) | Manufacturing (2) | Manufacturing (3) | Wholesale Trade | Retail Trade (1) | Retail Trade (2) | Transportation |
|---|---|---|---|---|---|---|---|---|---|---|
| **Data Moments** | | | | | | | | | | |
| *Moments targeted* | | | | | | | | | | |
| Frequency | 0.114 | 0.106 | 0.150 | 0.089 | 0.079 | 0.021 | 0.066 | 0.106 | 0.033 | 0.035 |
| Absolute size (median) | 0.154 | 0.082 | 0.043 | 0.078 | 0.082 | 0.059 | 0.082 | 0.086 | 0.095 | 0.031 |
| 25th percentile size | 0.095 | 0.038 | 0.009 | 0.041 | 0.039 | 0.046 | 0.037 | 0.038 | 0.034 | 0.030 |
| 75th percentile size | 0.382 | 0.150 | 0.125 | 0.167 | 0.172 | 0.148 | 0.167 | 0.154 | 0.154 | 0.085 |
| *Moments not targeted* | | | | | | | | | | |
| Size of price increase | 0.236 | 0.068 | 0.062 | 0.089 | 0.072 | 0.066 | 0.084 | 0.068 | 0.089 | 0.059 |
| Size of price decrease | 0.272 | 0.059 | 0.064 | 0.127 | 0.078 | 0.122 | 0.102 | 0.056 | 0.120 | 0.039 |
| **Model Moments** | | | | | | | | | | |
| **Moments targeted** | | | | | | | | | | |
| Frequency | 0.119 | 0.102 | 0.150 | 0.076 | 0.088 | 0.036 | 0.047 | 0.107 | 0.045 | 0.055 |
| Absolute size (median) | 0.201 | 0.082 | 0.043 | 0.081 | 0.084 | 0.073 | 0.104 | 0.090 | 0.100 | 0.040 |
| 25th percentile size | 0.083 | 0.022 | 0.009 | 0.026 | 0.023 | 0.042 | 0.053 | 0.034 | 0.033 | 0.023 |
| 75th percentile size | 0.383 | 0.152 | 0.125 | 0.165 | 0.166 | 0.130 | 0.148 | 0.152 | 0.154 | 0.081 |
| **Moments not targeted** | | | | | | | | | | |
| Size of price increase | 0.214 | 0.080 | 0.062 | 0.089 | 0.089 | 0.078 | 0.092 | 0.092 | 0.098 | 0.049 |
| Size of price decrease | 0.296 | 0.139 | 0.129 | 0.135 | 0.144 | 0.112 | 0.127 | 0.112 | 0.119 | 0.088 |

**Table A.4:** Data and Model Moments (Twenty Sectors), Part 2

| Moment | Agriculture | Mining, Oil and Gas | Utilities | Manufacturing (1) | Manufacturing (2) | Manufacturing (3) | Wholesale Trade | Retail Trade (1) | Retail Trade (2) | Transportation |
|---|---|---|---|---|---|---|---|---|---|---|
| **Data Moments** | | | | | | | | | | |
| *Moments targeted* | | | | | | | | | | |
| Frequency | 0.053 | 0.024 | 0.079 | 0.041 | 0.043 | 0.012 | 0.027 | 0.031 | 0.042 | 0.025 |
| Absolute size (median) | 0.028 | 0.034 | 0.095 | 0.050 | 0.035 | 0.063 | 0.076 | 0.057 | 0.056 | 0.069 |
| 25th percentile size | 0.010 | 0.015 | 0.044 | 0.030 | 0.014 | 0.009 | 0.035 | 0.041 | 0.029 | 0.054 |
| 75th percentile size | 0.088 | 0.083 | 0.188 | 0.145 | 0.067 | 0.118 | 0.147 | 0.095 | 0.096 | 0.128 |
| **Moments not targeted** | | | | | | | | | | |
| Size of price increase | 0.049 | 0.055 | 0.073 | 0.069 | 0.050 | 0.054 | 0.104 | 0.090 | 0.065 | 0.105 |
| Size of price decrease | 0.046 | 0.062 | 0.102 | 0.089 | 0.041 | 0.105 | 0.133 | 0.142 | 0.107 | 0.101 |
| **Model Moments** | | | | | | | | | | |
| **Moments targeted** | | | | | | | | | | |
| Frequency | 0.055 | 0.038 | 0.091 | 0.061 | 0.045 | 0.026 | 0.051 | 0.029 | 0.049 | 0.027 |
| Absolute size (median) | 0.030 | 0.036 | 0.100 | 0.063 | 0.035 | 0.060 | 0.084 | 0.059 | 0.059 | 0.070 |
| 25th percentile size | 0.013 | 0.016 | 0.031 | 0.021 | 0.014 | 0.016 | 0.044 | 0.035 | 0.023 | 0.059 |
| 75th percentile size | 0.088 | 0.084 | 0.175 | 0.137 | 0.067 | 0.121 | 0.138 | 0.098 | 0.097 | 0.130 |
| **Moments not targeted** | | | | | | | | | | |
| Size of price increase | 0.049 | 0.055 | 0.095 | 0.080 | 0.040 | 0.076 | 0.086 | 0.067 | 0.062 | 0.112 |
| Size of price decrease | 0.051 | 0.041 | 0.136 | 0.091 | 0.005 | 0.046 | 0.114 | 0.024 | 0.068 | 0.092 |

# B. Data Description

## B.1 Markup Estimation

To construct firm-level markup data, we use firm-level data from Compustat for all public firms in the U.S. covering all years since 1950. We exclude firm-year observations with negative or missing observations for sales, COGS or capital and investment measures, and all firms that did not report SIC or NAICS indicators or have their headquarter outside the U.S.

For the estimation of markups, we use the production function estimation approach PF1 (traditional production function) described by De Loecker et al. (2020) and follow the variations described in Baqaee and Farhi (2020): namely (i) using Olley and Pakes (1996) rather than Levinsohn and Petrin (2003) and (ii) deflating variables at the sectoral level instead of using aggregate deflators. According to the production function estimation approach, the markup $\mu_{i,t}$ of a firm $i$ at time $t$ can be computed from one variable input, $X_i$, as the ratio of the output elasticity of the input, $\epsilon_{Q,X_i}$, to the revenue share of that input, $s_{R,X_i}$

$$\mu_{i,t} = \frac{\epsilon_{Q,X_i}}{s_{R,X_i}}. \tag{B.1}$$

There are therefore two ingredients to calculate markups: the revenue share and the output elasticity of the flexible input. Compustat reports a composite input called Cost of Goods Sold (COGS), which consists of intermediate and labor input and that we will use as the (partially) flexible input, $X_i$. We use a variant of the technique introduced by Olley and Pakes (1996) and described in De Loecker and Warzynski (2012) to estimate a Cobb-Douglas production function and obtain a time-varying estimate of output elasticity at the sector level. The markups are then derived by dividing the former (estimated at the industry-year level) by the share of COGS to revenue (estimated at the firm-year level). In terms of implementation, we follow the procedure described in De Loecker et al. (2020) with the adjustments described in Baqaee and Farhi (2020). In particular, we estimate time-varying output elasticities and deflate variables at the sectoral level using gross output price indices from KLEMS sector-level data. Specifically, we apply the correspondence between 3-digit NAICS codes and BEA industry segments outlined in Baqaee and Farhi (2020). We exclude the Financial sector (SIC codes 6000-6999 or NAICS3 codes 520-525) and firms with BEA code 999 because there is no BEA depreciation available for them. We use CAPX as the instrument and COGS as a variable input. We trim the dataset by excluding all firms with COGS-to-sales and XSGA-to-sales ratios in the top and bottom 1%

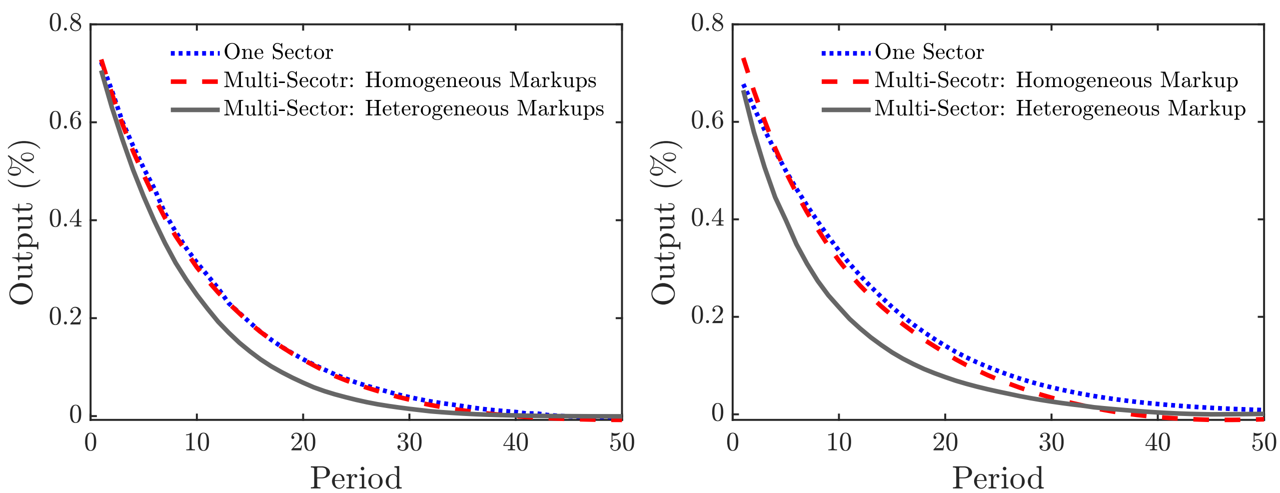
of the corresponding year-specific distributions.

The output-elasticities estimation is implemented using the prodest Stata package. This estimation requires to define a set of variables where we follow Baqaee and Farhi (2020) and choose (i) log sales as outcome variable, (ii) log COGS as free variable (variable inputs), (iii) the log capital stock (measured as log PPEGT in the Compustat data) as state variable, (iv) log investment (CAPX) as an instrument for productivity, and finally (v) SIC 3-digit and SIC 4-digit firm sales shares to control for markups.

We estimate markups based on time-varying output-elasticities at the 2-digit NAICS level. To do so, we run the estimation procedure for every sector and every year using a 3-year rolling windows.

## C. Heterogeneous Markups and Monetary Non-neutrality

**Figure C.1:** Heterogeneous Markups and Monetary Non-neutrality




**(a)** Eight-sector Model

**(b)** Twenty-sector Model

*Note:* This figure plots the IRFs of real output to a monetary shock in various models. The solid black lines (dashed red lines) depict the results in a multi-sector model with heterogeneous (homogeneous) steady-state markups. The dotted blue lines represent the results in a one-sector model. We consider an Eight-sector version (panel a) and a Twenty-sector version (panel b) of multi-sector models.

Figure C.1 plots the IRFs of real output to a monetary shock ($\frac{\partial y_{t+h}}{\partial \epsilon_t^s}$) in various models. The size of the monetary shock is normalized to increase the nominal spending by 1% on impact. The solid black lines (dashed red lines) depict the results in a multi-sector model with heterogeneous (homogeneous) steady-state markups. The dotted blue lines represent the results in a one-sector model. We consider an Eight-sector version (panel a) and a Twenty-sector version (panel b) of multi-sector models.

Monetary non-neutrality is reduced in a multi-sector model with heterogenous steady-state markups (solid black lines) as compared to a one sector model (blue lines). To provide a quantitative number, the cumulative output response in the Eight-sector model with heterogenous steady-state markups is 30.6 percent lower compared to the one-sector model. Note that the multi-sector model with heterogenous steady-state markups differ from the benchmark one-sector model in many dimensions. Particularly, the multi-sector model features other sources of heterogeneities: productivity distribution and frequency of price adjustments.

We conduct a semi-decomposition exercise to decompose the total reduction in monetary non-neutrality into the component that arises from the introduction of heterogenous markups and the component due to other multi-sector features. To do so, we recalibrate a multi-sector model that share the same features and target to the same moments except that steady-state markups are imposed to be homogenous across sectors. These IRFs are plotted in dashed red lines. Comparing the two versions Eight-sector models, the cumulative output response in the baseline model with heterogenous market power is reduced by 23.3 percent. In contrast, the cumulative output response in the Eight-sector model with homogeneous market powers is merely 9.6 percent lower than that of the one-sector model.

Similarly, the cumulative output response in the Twenty-sector model with heterogenous steady-state markups is 31% lower compared to the one-sector model and 23% lower compared to the homogeneous steady-state markups model.

In summary, our quantitative analyses show that heterogenous steady-state markup reduces monetary non-neutrality.

## D. An Analytical Menu Cost Model

In an analytical menu cost model with monopolistic competition, this section shows that 1) the width of the Ss band is an increasing concave function of the desired markup, and 2) the steady-state FPA is a decreasing convex function of the desired markup, and consequently the frequency effect is an increasing and concave function of steady-state markups.

Time is continuous. Firms set optimal prices in a monopolistically competitive environment, facing standard CES demand. The elasticity of substitution across the differentiated goods is $\theta$. The frictionless profit-maximizing price $e^{p^*}$ is given by a constant markup $\mu = \theta/(\theta - 1)$ over the marginal cost $W$. The log optimal price $p^*$ follows a Browning motion, where $dp^* = \sigma dw$ and $dw$ is the increment of a Wiener process. Firms can adjust prices by paying a fixed cost $\psi$, in units of flow profits at the profit-maximizing price. The price gap is denoted by $z = p - p^*$. Firms discount payoffs at rate $r$ and the aggregate output and price are denoted by $Y_I$ and $P_I$. The firm's flow profit can be approximated, up to second order, with respect to the log price $p$ around the log frictionless profit-maximizing price $p^*$. We therefore obtain a quadratic loss function in lemma 1.

Let $\pi(p_i)$ denote the firm $i$'s profit as a function of $p_i$, and $\pi(p_i^*)$ denote the the desired level of profit without nominal rigidity (zero price adjustment cost). A firm's profit gap is defined as $|\pi(p_i) - \pi(p_i^*)|$.

**Lemma 1.** *A firm's profit gap can be approximated by a quadratic loss function*

$$|\pi(p_i) - \pi(p_i^*)| \approx \frac{1}{2}\frac{\mu}{(\mu-1)^2}z^2.$$

*Proof.* With CES demand, firm $i$'s profit function can be written as

$$\Pi(P_i) = (P_i - \mu P_i^*) Y_I \left(\frac{P_i}{P_I}\right)^{-\theta},$$

Taking second-order Taylor approximation around the steady state we have

$$\begin{aligned}\Pi(P_i) - \Pi(P) = \Pi'(P)P(p_i - p) + \frac{1}{2}\Pi''_{PP}(P)P^2(p_i - p)^2 \\ + \Pi''_{PP^*}(P)P^2(p_i - p)(p_i^* - p) + \text{terms to be cancelled} + o(2), \quad \text{(D.1)}\end{aligned}$$

and

$$\begin{aligned}\Pi(P_i^*) - \Pi(P) = \Pi'(P)P(p_i^* - p) + \frac{1}{2}\Pi''_{PP}(P)P^2(p_i^* - p)^2 \\ + \Pi''_{PP^*}(P)P^2(p_i^* - p)^2 + \text{terms to be cancelled} + o(2), \quad \text{(D.2)}\end{aligned}$$

where $P$ is the steady state price. Substituting the following results into the difference between equation (D.1) and (D.2)

$$\begin{aligned}\Pi'(P) = 0, \quad \Pi(P) = \frac{PY}{\theta}, \\ \Pi''_{PP} = -(\theta - 1)\frac{Y}{P}, \\ \Pi''_{PP^*} = (\theta - 1)\frac{Y}{P}.\end{aligned}$$

gives firms' quadratic loss function

$$|\pi(p_i) - \pi(p_i^*)| \approx \frac{1}{2}\theta(\theta-1)(p_i - p_i^*)^2$$
$$= \frac{1}{2}\frac{\mu}{(\mu-1)^2}z^2.$$

■

Given this lemma, we can write down the Hamilton-Jacobi-Bellman (HJB) equation of firms' optimization problem

$$rv(z) = \frac{1}{2}\frac{\mu}{(\mu-1)^2}z^2 + \frac{\sigma^2}{2}v''(z), \quad \text{for } z \in (-\bar{z}, \bar{z}).$$

Applying the value matching and smooth pasting condition $v(\bar{z}) = v(0) + \phi$ and $v'(\bar{z}) = 0$, we obtain the following result.

**Proposition 1.** *Firms' do not adjustment prices when $|z| < \bar{z}^*$ where $\bar{z}^* = \left(\frac{12\psi\sigma^2(\mu-1)^2)}{\mu}\right)^{1/4}$. The width of the Ss band is given by $S = 2\bar{z}^*$. The frequency of price adjustment is given by $f = \left(\frac{\mu}{12\psi(\mu-1)^2}\right)^{1/2}\sigma$.*

Taking first and second order derivatives, it is straightforward to show that

**Corollary 1.**

$$\frac{\partial S}{\partial \mu} > 0, \quad \frac{\partial f}{\partial \mu} < 0$$
$$\frac{\partial^2 S}{\partial \mu \partial \mu} < 0, \quad \frac{\partial^2 f}{\partial \mu \partial \mu} > 0$$

We therefore prove that 1) the width of the Ss band is an increasing concave function of the desired markup, and 2) the FPA is a decreasing convex function of the desired markup. The frequency effect is therefore an increasing concave function of the desired markup.

# E. A Model with Oligopolistic Competition

We show that under a simple model with oligopolistic competition, firms' profit function becomes more curved when there are less number of firms in a given sector, or firms have more market power.

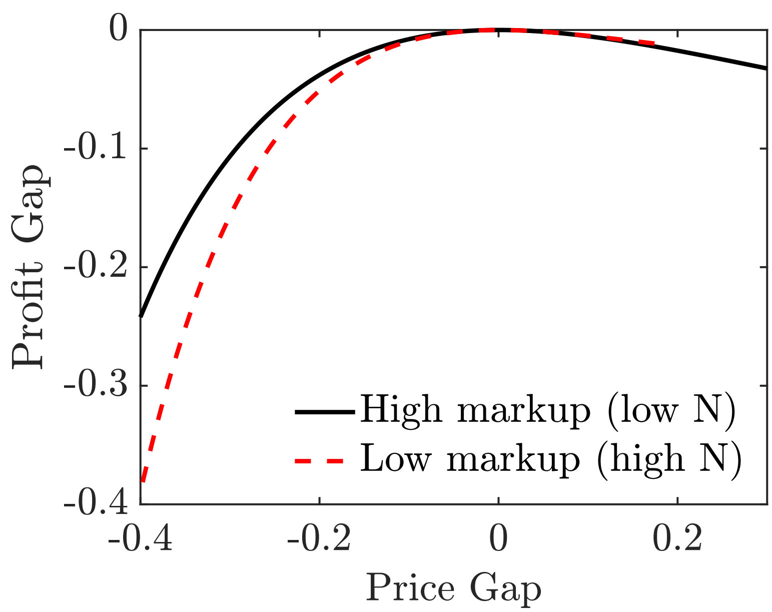
We closely follow the oligopolistic setup in Atkeson and Burstein (2008). There is a continuum of sectors and within each sector there is a finite number of firms. Since we focus on firms' pricing strategies, it is natural to assume that firms play a game of Bertrand price competition. Households' demand features a nested CES structure. The elasticity of substitution between sectoral goods is $\eta$ and the elasticity substitution between goods within sectors is $\theta$. Following the tradition of literature, we assume $\theta > \eta > 1$. For simplicity, we also assume that within each sector there are $N$ symmetric firms, such that each firm's market share is $1/N$.

The standard prediction of this type of model is that firms' *steady-state* markup $\mu$ is *increasing* in its market share (or *decreasing* in number of firms $N$) in its sector:

$$\mu(N) = 1 + \frac{1}{\theta - (\theta - \eta)/N - 1},$$

so that firms' frictionless profit-maximizing price $e^{p^*}$ is a constant markup $\mu N$ over the marginal cost $W$. We now plot the profit of an individual firm as a function of its price gap for two levels of competition: $N = 3$ and $N = 6$. We choose conventional values for parameters in the literature. Figure E.1 shows that as market becomes more competitive, the profit function becomes more curved.

**Figure E.1:** Market Power and Profit Functions




*Note:* This graph plots the profit gap as a function of price gaps for different calibrations of markups in a model with oligopolistic competition: the high markup (N = 3) and the low markup (N = 6) cases. The profit gap is defined as the difference between a firm's profit given its price $p$ and the firm's profit under its optimal resetting price ($p^*$). Similarly, the price gap is defined as $p/P - p^*/P$.