Non-Constant Demand Elasticities, Firm Dynamics and Monetary Non-Neutrality: Role of Demand Shocks*

S. Borağan Aruoba† Eugene Oue‡ Felipe Saffie§
Jonathan L. Willis¶

February 21, 2024

Abstract

We develop a simple menu-cost model with non-constant elasticity of demand that features idiosyncratic productivity and demand shocks. The model is calibrated to match firm-level productivity and demand processes estimated from U.S. data. Despite its simplicity, the calibrated model delivers untargeted pricing dynamics and a markup distribution that are consistent with U.S. micro data. Moreover, it also generates sizable monetary non-neutrality that rivals more complicated alternative menu cost models that explicitly target pricing dynamics. The key in reconciling firm and pricing dynamics comes from the interaction between non-constant elasticity of demand and idiosyncratic demand shocks. Thus, this framework effortlessly unifies pricing, markup, and firm dynamics.

Keywords: Menu costs, strategic complementarities, demand shocks, sticky prices, monetary non-neutrality.

JEL Classifications: E30, E52, L11

*The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. The authors would like to thank Mark Bils, Laura Castillo-Martínez (discussant), John Haltiwanger, Pete Klenow, Virgiliu Midrigan, Luminita Stevens, Javier Turén, and seminar participants at the University of Maryland, the Federal Reserve Board of Governors, the University of Iowa, Georgetown University, PUC-Chile, Michigan State University, the European Central Bank, the Deutsche Bundesbank, the European University Institute, and North Carolina State University for useful comments and suggestions. Any remaining errors are the authors' responsibility.

†University of Maryland
‡Hong Kong Polytechnic University
§University of Virginia, Darden School of Business
¶Federal Reserve Bank of Atlanta

1 Introduction

Modeling the response of output and prices to monetary policy shocks has been a long-term quest in the macroeconomics literature. Quantitative estimates consistently indicate that monetary policy shocks have persistent impacts on real output (Christiano et al., 1999; Ramey, 2016). When it comes to modeling, early studies (Mankiw, 1985; Akerlof and Yellen, 1985) propose that nominal rigidities, like a menu cost, can explain real effects of monetary policy shocks. However, subsequent studies such as Golosov and Lucas (2007) show that a menu cost alone is insufficient to generate the observed non-neutrality from quantitative estimates. To this end, Ball and Romer (1990) demonstrate that real rigidities in combination with nominal rigidities can lead to sizable real effects of monetary policy.

Demand systems with non-constant elasticities, a general class of demand systems that deviate from the commonly used constant elasticity of substitution (CES) assumption, represent one promising form of real rigidity used in the money non-neutrality literature.[1] Kimball (1995) introduces such a demand system where price elasticity of demand varies with relative price and quantity, which creates a strategic complementarity in the price-setter's optimization problem. Eichenbaum and Fisher (2007) and Smets and Wouters (2007) incorporate this Kimball demand system into dynamic stochastic general equilibrium models in order to match the estimated responses of real output to a nominal shock, while also being consistent with micro evidence on the frequency of price adjustment. However, Klenow and Willis (2016) find that a menu-cost model with a Kimball demand system that matches firm-level pricing moments requires idiosyncratic productivity shocks that are much larger than observed in the data.[2]

In this paper, we propose a simple model featuring a demand system with non-constant elasiticities that can generate monetary non-neutrality, while also remaining consistent with the micro and macro evidence. The baseline specification is a menu-cost model with a real rigidity in the form of a Kimball demand system, in which the elasticity of substitution between a given variety and other varieties is decreasing in the relative quantity consumed.

[1]There is a broad literature featuring a range of alternative ways to generate real rigidities in models, including production networks (Carvalho, 2006; Nakamura and Steinsson, 2010; Midrigan, 2011; Blanco et al., 2022), segmented labor markets (Gertler and Leahy, 2008), decreasing returns to scale (Burstein and Hellwig, 2007), duopoloy (Mongey, 2021), and financial frictions (Gilchrist et al., 2017; Kim, 2021; Renkin and Züllig, 2024).

[2]Non-CES demand systems other than Kimball have also been considered in the literature. For example Bergin and Feenstra (2000) use a translog demand system along with staggered price setters to generate persistent real responses to monetary policy shocks.

The additional key element of this model is the specification of two forms of idiosyncratic shocks: productivity and demand. Despite its parsimony, the model generates sizable non-neutrality of monetary policy with realistic shocks, while remaining consistent with micro pricing facts.

Deviations from CES are worth considering, in fact, elsewhere in the economics literature, the use of demand systems with non-constant elasticities in models has been shown to be important for generating firm-level and aggregate dynamics needed to match key features of the data. Such demand systems have been used extensively to study exchange-rate pass-through in international macroeconomics (Gopinath and Itskhoki, 2010; Amiti et al., 2019; Berger and Vavra, 2019), variable markups in firm dynamics and international trade (Edmond et al., 2023; Arkolakis et al., 2019), and inflationary dynamics and optimal monetary policy (Harding et al., 2022, 2023; Fujiwara and Matsuyama, 2022). Given the success of these demand systems, reconciling pricing dynamics and non-constant demand elasticities, opens the door to consistently modeling real (investment, labor, and sales), nominal (pricing and non-neutrality), and competition (markup) dynamics in a parsimonious framework amenable to policy analysis.

The intuition for how our framework overcomes previous challenges in the literature is straightforward. Unlike a model with CES demand, where demand shocks fail to generate price adjustment, the Kimball demand system creates a channel where firms want to pass through portions of both productivity shocks and demand shocks to prices. In fact, a Kimball demand system generates a trade-off between the strength of strategic complementary and the pass-though of productivity shocks. Previous studies, including Klenow and Willis (2016), that exclusively focus on the role of idiosyncratic productivity shocks found that matching price-adjustment moments would require productivity shocks that seem unrealistically large. That is because in a typical calibration of the degree of strategic complementary, firms would only pass through 20 to 40 percent of productivity shocks into prices, therefore requiring very large productivity shocks to generate realistic sizes of price adjustments. But with the inclusion of idiosyncratic demand shocks, we are able to match price-adjustment moments using idiosyncratic productivity and demand processes that are calibrated based on firm-level data. Our calibrated model implies that firms pass through approximately 62 percent of idiosyncratic demand shocks and 38 percent of productivity shocks into prices. With demand shocks also playing a role in the pricing decisions of firms, our model can generate a distribution of price changes that is similar to the data while also remaining consistent with micro evidence on the persistence and volatility of idiosyncratic shocks.

Unlike typical menu-cost models which predominantly rely on pricing moments to discipline the size of idiosyncratic shocks, we calibrate the model using direct estimates of both idiosyncratic productivity and demand processes from the firm dynamics literature. Moreover, we show that the correlations between prices and measures of firm-level productivity are informative about the strength of real rigidities and use them to discipline the degree of deviation from CES demand in the model. Specifically, the model is calibrated to match the 5-yearly serial correlation as well as cross-sectional standard deviations of firm-level productivity and demand reported by Foster et al. (2008), who point to important, and separate, roles for these shocks using data from the Census of Manufactures.[3] We use two additional moments from Foster et al. (2008), namely the firm-level correlation between revenue-based total factor productivity (TFPR) and quantity-based total factor productivity (TFPQ) and the firm-level correlation between price and TFPQ, to pin down two parameters governing the curvature of the demand function. Despite using the frequency of price adjustment as the only targeted moment related to firm pricing behavior, the calibrated model effortlessly matches three untargeted pricing moments: the average size of a price change conditional on making a change, the fraction of adjustments that are positive, and the dispersion of non-zero price changes. Furthermore, the model delivers a downward-sloping pricing hazard function consistent with the data.

The incorporation of a Kimball demand system also allows our model to generate an untargeted cross-sectional markup distribution that closely resembles its empirical counterpart, which is impossible to achieve with CES demand. This is because the desired markup under CES is constant. Although nominal pricing rigidities can generate a non-degenerate distribution around the desired markup, such a model cannot replicate the large cross-sectional variance in the empirical markup distribution without unreasonably large nominal rigidities. In contrast, a Kimball demand system induces more dispersion in firm markups due to variable elasticities, as well as the incorporation of demand shocks, which affect the desired markup. Additionally, the calibrated model produces incomplete cost pass-through (pass-through of 38 percent), in line with the empirical literature.

To study the degree of monetary non-neutrality generated by the model, we expand the framework to incorporate nominal expenditure shocks. On impact, 82 percent of an increase in nominal expenditures is reflected in an increase in real output. The real output response decays with a half-life of five months, before fully dissipating after 20 months.

[3] Recent papers that jointly study and model both productivity and supply shocks are Aruoba et al. (2022) and Carlsson et al. (2022)

The cumulative response of real output is four times as large as a simple CES menu cost model calibrated to pricing moments. It is in the upper range of richer models that explore alternative sources of real rigidities without using micro estimates for their shock processes.[4] The two key features generating this non-neutrality are the Kimball demand system and the presence of both idiosyncratic productivity and demand shocks. The former generates a strategic pricing complementarity under which firms will temper their price adjustments in response to a shock because of an endogenous change in desired markups, whereas the latter dampens the selection effect in price adjustments to an aggregate shock and results in larger real output responses.

The remainder of the paper is structured as follows. Section 2 introduces a quantitative menu-cost model with idiosyncratic productivity and demand augmented with a Kimball demand system and explores its theoretical properties. Section 3 presents the calibration of the model. Section 4 discusses the model's implications for untargeted pricing moments, markup distribution, and non-neutrality of monetary shocks. Section 5 provides a discussion of why the two ingredients, Kimball demand system and idiosyncratic demand shocks, are particularly important for the results we obtain. Finally, Section 6 concludes.

2 Menu Cost Model

We build a quantitative menu-cost model following Golosov and Lucas (2007). The model features a representative household, a representative final-good producer, and a continuum of monopolistically competitive intermediate-variety producers who face nominal pricing frictions.

2.1 Households

A representative household supplies labor to firms in exchange for wage payments, purchases a complete set of Arrow-Debreu securities, $\mathbf{B_{t+1}}$, and consumes a final good, C_t. It also owns all firms in the economy and receives all accrued profits. The representative household solves the following problem

$$\max_{C_t, h_t, \mathbf{B_{t+1}}} \quad \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\log\left(C_t\right) - \chi h_t\right] \tag{1}$$

[4] See Golosov and Lucas (2007), Gertler and Leahy (2008), Burstein and Hellwig (2007), Nakamura and Steinsson (2010), Vavra (2014), and Mongey (2021).

subject to the budget constraint

$$P_t C_t + \mathbf{Q_t} \cdot \mathbf{B_{t+1}} \leq B_t + W_t h_t + \Pi_t, \tag{2}$$

where $\mathbf{Q_t}$ is a vector that contains the prices of the state-contingent securities, $\mathbf{B_{t+1}}$. B_t represents the payoff of the state-contingent security purchased in period $t-1$ that had a non-zero payoff in period t. P_t and W_t are the price of the final good and nominal wage, respectively, both of which are taken as given by the households. Π_t denotes the net dividends the household receives from the producers.

Household optimality requires

$$\frac{W_t}{P_t} = \chi C_t, \tag{3}$$

and we can also define the household's stochastic discount factor as

$$\Xi_{t,t+1} \equiv \beta \mathbb{E}_t \left(\frac{C_t}{C_{t+1}} \right). \tag{4}$$

2.2 Producers

Production is carried out by a continuum of perfectly-competitive final-good producers, who purchase varieties of intermediate goods and sell a combined final good to the households. A continuum of intermediate-good producers each produce a differentiated variety and are monopolistically competitive due to imperfect substitution across varieties.

2.2.1 Final Good Producers

A representative final-good firm combines intermediate varieties, y_t^i, to produce the final good, Y_t, using the Kimball (1995) aggregator. This aggregator is defined implicitly as

$$\int_0^1 G\left(\frac{n_t^i y_t^i}{Y_t}\right) di = 1, \tag{5}$$

where n_t^i represents an idiosyncratic variety-specific demand shifter. Following Dotsey and King (2005), we use the following specification for $G(.)$

$$G\left(\frac{n_t^i y_t^i}{Y_t}\right) = \frac{\omega}{1+\omega\psi}\left[(1+\psi)\frac{n_t^i y_t^i}{Y_t} - \psi\right]^{\frac{1+\omega\psi}{\omega(1+\psi)}} + 1 - \frac{\omega}{1+\omega\psi}, \tag{6}$$

where (5) and (6) show the only two deviations from a textbook menu cost model: the introduction of the variety-specific demand shifters and the Kimball aggregator. This specification nests the familiar constant elasticity of substitution (CES) Dixit-Stiglitz aggregator when $\psi = 0$. When this is the case, the final good Y_t can be expressed explicitly as

$$Y_t = \left[\int_0^1 \left(n_t^i y_t^i \right)^{\frac{1}{\omega}} \right]^{\omega}, \tag{7}$$

where the price elasticity of demand is given by $\frac{\omega}{1-\omega}$, elasticity of substitution is given by $\frac{\omega}{\omega-1}$ and the gross markup is given by ω. When $\psi \neq 0$, price elasticity of demand, elasticity of substitution and desired markup are no longer constant. We discuss how Kimball aggregation affects firms' pricing decisions in Section 2.3.

Taking as given variety prices, p_t^i, as well as P_t, n_t^i and aggregate demand, Y_t, the representative final-good producer chooses y_t^i to maximize profits

$$\max_{y_t^i} \quad 1 - \int_0^1 \frac{p_t^i y_t^i}{P_t Y_t} di \quad \text{subject to} \quad \int_0^1 G\left(\frac{n_t^i y_t^i}{Y_t} \right) di = 1. \tag{8}$$

The optimality condition of the final-good producer's maximization problem implicitly defines the demand function for each variety i

$$\frac{n_t^i y_t^i}{Y_t} = \frac{1}{1+\psi} \left[\left(\frac{p_t^i}{\lambda_t n_t^i P_t} \right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi \right], \tag{9}$$

where λ_t is the Lagrangian multiplier on the constraint in the optimization problem, which can be obtained by substituting (9) into (5) as

$$\lambda_t = \left[\int_0^1 \left(\frac{p_t^i}{n_t^i P_t} \right)^{\frac{1+\omega\psi}{1-\omega}} di \right]^{\frac{1-\omega}{1+\omega\psi}}. \tag{10}$$

The aggregate price index is derived from the zero-profit condition for the final-good producer as

$$P_t = \frac{1}{1+\psi} \left[\int_0^1 \left(\frac{p_t^i}{n_t^i} \right)^{\frac{1+\omega\psi}{1-\omega}} di \right]^{\frac{1-\omega}{1+\omega\psi}} + \frac{\psi}{1+\psi} \int_0^1 \frac{p_t^i}{n_t^i} di \tag{11}$$

2.2.2 Intermediate Variety Producers

There is a continuum of intermediate-good producers indexed by i, each producing a differentiated variety y_t^i. Intermediate producers are heterogeneous in their physical productivity, z_t^i, and face demand shocks for their variety, n_t^i. The production technology is linear with labor as the only input

$$y_t^i = z_t^i l_t^i \tag{12}$$

Idiosyncratic productivity, z_t^i, and idiosyncratic demand, n_t^i, evolve according to a VAR(1) process

$$\begin{pmatrix} \log(z_t^i) \\ \log(n_t^i) \end{pmatrix} = \begin{bmatrix} \rho_z & 0 \\ 0 & \rho_n \end{bmatrix} \begin{pmatrix} \log(z_{t-1}^i) \\ \log(n_{t-1}^i) \end{pmatrix} + u_t^i \text{ where } u_t^i \sim N\left(0, \begin{bmatrix} \sigma_z^2 & \sigma_{zn} \\ \sigma_{zn} & \sigma_n^2 \end{bmatrix}\right) \tag{13}$$

At the beginning of each period, intermediate-good producers decide whether or not to adjust their nominal prices and if so, by how much. Nominal price adjustments are subject to a fixed cost, f, in terms of labor. Given the demand schedule for individual varieties, the intermediate producers' gross profit when they charge price p is

$$\pi(p_t^i, z_t^i, n_t^i, \mathcal{S}_t) = \left(\frac{p_t^i}{P_t} - \frac{W_t}{z_t^i P_t}\right) \frac{Y_t}{n_t^i} \frac{1}{1+\psi} \left[\left(\frac{p_t^i}{\lambda_t n_t^i P_t}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi\right], \tag{14}$$

where $\mathcal{S}_t \equiv (P_t, W_t, Y_t, \lambda_t)$ collects all aggregate objects the firms need to know, and we assume that the firms know the law of motion for $\mathcal{S}_t$.

At the beginning of the period, each intermediate-good producer inherits their price from the previous period p_{t-1}^i. At that point, they choose whether or not to change their prices by solving the problem

$$V\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right) = \max\left[V_N\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right), V_A\left(z_t^i, n_t^i, \mathcal{S}_t\right)\right], \tag{15}$$

where $V_N(.)$ and $V_A(.)$ are the values for the firm if they do not change and change their prices, respectively.

The value of not adjusting is

$$V_N\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right) = \pi\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right) + \mathbb{E}_t\left[\Xi_{t,t+1} V\left(p_{t-1}^i, n_{t+1}^i, z_{t+1}^i, \mathcal{S}_{t+1}\right)\right], \tag{16}$$

which is equal to the flow profit evaluated at last period's price plus a continuation value. If the firm chooses to adjust its price, it pays the fixed price adjustment cost and chooses p_t^i to maximize the sum of current flow profit and the present discounted value of future profits given by

$$V_A\left(z_t^i, n_t^i, \mathcal{S}_t\right) = -f\frac{W_t}{P_t} + \max_{p_t^i}\left\{\pi(p_t^i, z_t^i, n_t^i, \mathcal{S}_t) + \mathbb{E}_t\left[\Xi_{t,t+1}V\left(p_t^i, n_{t+1}^i, z_{t+1}^i, \mathcal{S}_{t+1}\right)\right]\right\}. \quad (17)$$

The intermediate-good producers solve this problem taking as given the laws of motion for the idiosyncratic state variables as in (13) and those for the aggregate variables in $\mathcal{S}_t$.

2.3 Kimball Aggregation and Pricing Decisions

The defining feature of the Kimball demand system is the variable price elasticity. In particular, when $\psi < 0$, the price elasticity of demand becomes an increasing function of the relative price of the variety p/P, and a decreasing function of idiosyncratic demand, n_t^i, and the effective market share of the variety, ny/Y.[5] The non-constant price elasticity implied by the Kimball demand system has two important consequences. First of all, unlike the CES case where desired markups are constant and independent of idiosyncratic demand, variable price elasticity leads to variable desired markups. Moreover, the desired markup depends on both the firm's idiosyncratic productivity as well as demand. For example, a firm with lower costs or higher demand would choose to have a higher markup relative to an average firm. Second, the Kimball demand system creates strategic complementarities in pricing among firms, because a deviation from the aggregate price index is costly.

To demonstrate how productivity and demand shocks affect pricing decisions of firms under a Kimball demand system, it is instructive to set $f = 0$ and focus on the problem of an individual intermediate-good producer. The first-order condition to the static profit-maximization problem of an intermediate-good producer is

[5] The price elasticity of demand is

$$\epsilon \equiv \frac{dy}{dp}\frac{p}{y} = \frac{\omega}{1-\omega}\frac{(1+\psi)\frac{ny}{Y} - \psi}{\frac{ny}{Y}} = \frac{\omega(1+\psi)}{1-\omega}\frac{\left(\frac{p}{\lambda nP}\right)^{\frac{\omega(1+\psi)}{1-\omega}}}{\left(\frac{p}{\lambda nP}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi}.$$

with $\frac{\partial\epsilon}{\partial(ny/Y)} = -\frac{\psi\omega}{\left(\frac{ny}{Y}\right)^2(\omega-1)}$ and $\frac{\partial\epsilon}{\partial n} = -\frac{\psi\omega Y}{n^2 y(\omega-1)}$. We can also compute the super-elasticity, defined as the elasticity of the demand elasticity with respect to price as

$$\gamma \equiv \frac{d\epsilon}{dp}\frac{p}{\epsilon} = \frac{\omega}{1-\omega}\cdot\frac{\psi(1+\psi)}{\left[\left(\frac{p}{\lambda nP}\right)^{\frac{\omega}{1-\omega}} + \psi\right]} = \frac{\omega}{1-\omega}\cdot\frac{\psi}{\frac{ny}{Y}}.$$

$$\left(\frac{p_i^*}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}} \left[1-\left(\frac{W}{z_i}-p_i^*\right)\left(\frac{\omega\,(1+\psi)}{1-\omega}\frac{1}{p_i^*}\right)\right]=-\psi \tag{18}$$

where p_i^* denotes the optimal price the firm chooses.

We show in Appendix A that log-linearizing (18) around a symmetric steady state and letting hatted variables denote log-deviations from the steady state yields the following expression for the optimal price

$$\hat{p}_i^* \;=\; \frac{\omega\psi}{\omega\psi-1}\left(\hat{\lambda}+\hat{P}+\hat{n}_i\right)+\frac{1}{\omega\psi-1}\hat{z}_i \tag{19}$$

Letting $\widehat{mc} \equiv -\hat{z}$ denote the log-deviation in marginal cost, which is inversely proportional to productivity, the price elasticities with respect to productivity and demand shocks, respectively, are given by

$$\frac{\partial \hat{p}_i^*}{\partial \widehat{mc}} \;=\; -\frac{1}{\omega\psi-1} \tag{20}$$

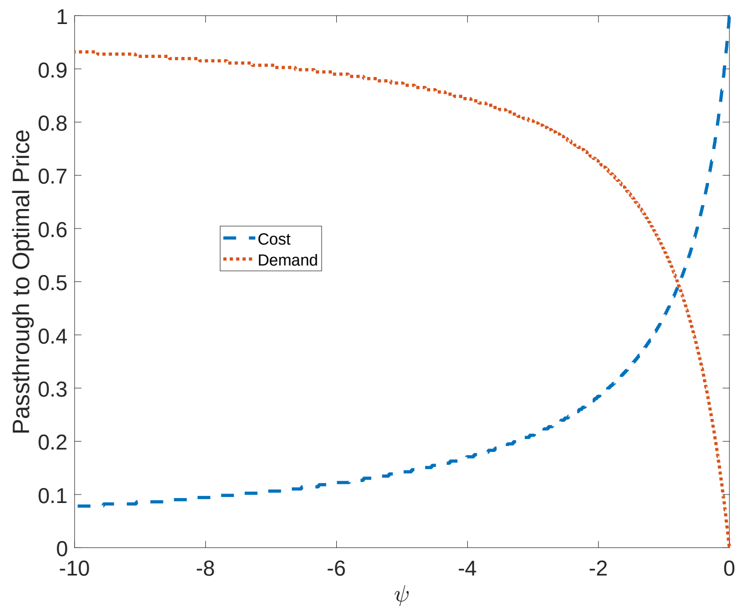
$$\frac{\partial \hat{p}_i^*}{\partial \hat{n}_i} \;=\; \frac{\omega\psi}{\omega\psi-1}, \tag{21}$$

In what follows we refer to these as cost and demand pass-through.

When $\psi=0$, cost pass-through is complete: price falls one to one with a positive z shock (negative mc shock). However, when $\psi<0$, the pass-through of a cost shock is incomplete. Variable demand elasticity is key to understanding this incomplete pass-through of cost. Consider a firm experiencing a positive productivity shock. As its marginal cost decreases, the firm finds it optimal to reduce its price, which yields greater sales. However, as it moves along the demand curve, with a larger effective market share, its price elasticity decreases, which dampens the increase in revenue from cutting prices. As such, the optimal price cut is smaller than in the CES case. Furthermore, this line of reasoning implies that in the nonlinear solution, the size of the cost pass-through is smaller (larger) for a larger reduction (increase) in cost. Because variable price elasticities attenuate gains from price deviations, the Kimball demand system (with $\psi<0$) generates strategic complementarities in that a firm wants to avoid moving its price too far away from its competitors. This is in line with the early literature such as Ball and Romer (1990) and Caplin and Leahy (1997), which emphasize the importance of strategic complementarities as a real rigidity.

Similarly, when $\psi=0$, demand pass-through is zero, which is the standard result under CES. When $\psi<0$, demand pass-through becomes positive: a firm receiving a demand shock

Figure 1: Pass-through of Demand and Cost Shocks to Price



Note: This plots pass-through of a small (1%) change in demand and productivity to the optimal frictionless price around a symmetric equilibrium with $\omega = 1.29$

chooses to increase its price or, equivalently, chooses a higher markup over its marginal cost. Under a Kimball demand system, the elasticity of demand decreases in the effective market share. Firms with stronger demand for their product can raise prices without losing as much sales, resulting in higher markups.

The relative importance of demand and cost shocks hinges on the degree of strategic complementarity. Figure 1 plots the pass-through of cost and demand to the optimal frictionless price as a function of ψ, holding ω constant at our calibrated value. As ψ decreases, the pass-through of cost (productivity) shocks decreases and the pass-through of demand shocks increases. This figure makes it clear that the value of ψ will be critical for determining the pass-through of idiosyncratic shocks. There is overwhelming evidence in the international-trade, international-finance and firm-dynamics literatures that the pass-through from cost shocks to prices is less than complete, some of which we turn to in Section 3.1.3. This suggests that a constant elasticity of substitution specification ($\psi = 0$) cannot produce a cost pass-through that matches empirical estimates and strongly points toward a role for strategic complementarities ($\psi < 0$).

2.4 Equilibrium

Money supply, S_t, which must be equal to nominal aggregate expenditures, P_tC_t, in equilibrium, follows the stochastic process

$$\log(S_t) = \mu + \log(S_{t-1}) + \sigma_S \epsilon_t \text{ where } \epsilon_t \sim N(0,1), \tag{22}$$

where money supply grows at rate of μ every period with stationary fluctuations around it given by ϵ_t. As standard in the literature, because a one-time change in ϵ_t creates a permanent change in money balances, we interpret it as a monetary policy shock. This shock is the only source of aggregate uncertainty in our model. In calibrating our model, we set $\sigma_S = 0$ as it has minimal influence on the model-implied moments used for calibration. We explain our computational strategy in more detail in Appendix B.

3 Calibration

Most quantitative papers in the literature that use menu cost models employ a single idiosyncratic shock in their setup. Sometimes this shock directly moves the desired price around (e.g. Caplin and Spulber (1987)), where the authors are agnostic about the fundamental source of this shock. In other instances, a productivity shock is used (e.g. Vavra (2014)), but the authors do not look to firm-level evidence on productivity to calibrate it. In both cases, the process that drives either the desired price or productivity is typically calibrated to match various moments related to the distribution of firm-level price changes.

In this study, we aim to have a model that respects a broader set of micro-level evidence, while also delivering significant monetary non-neutralities. To do so, we introduce two firm-level shocks in our model: productivity and demand. Furthermore, we calibrate the processes for these shocks to be consistent with direct firm-level evidence. This is a significant deviation from the common practice of directly targeting pricing dynamics and ignoring the empirical estimates of firm-level shocks. Before turning to the details of the calibration, we first review the moments we use from the data, either as calibration moments or as untargeted moments.

3.1 Calibration Targets and Untargeted Moments

In this section we report all of the data moments used throughout the paper.

3.1.1 Firm-Level Productivity and Demand Processes

Using the Census of Manufactures, in a seminal paper Foster et al. (2008) estimate firm-level productivity and demand for eleven product markets with minimal vertical differentiation.[6] Using data on sales, quantity sold, and input usage, they estimate the production function of firms assuming Cobb-Douglas technology and recover firm-level physical TFP (TFPQ) as the residual of the following estimation:

$$TFPQ_{it} = \ln q_{it} - \alpha_l \ln l_{it} - \alpha_k \ln k_{it} - \alpha_m \ln m_{it} - \alpha_e \ln e_{it}, \tag{23}$$

where $TFPQ_{it}$ is the firm-level physical TFP of firm i at time t, q_{it} is the quantity produced by the firm, l_{it} is the labor input, k_{it} is the capital input, m_{it} represents intermediate inputs used in production, and e_{it} is the energy used by the firm. Foster et al. (2008) also estimate revenue-based TFP, which can be obtained using the same method but replacing the quantity produced with the revenue of the firm,

$$TFPR_{it} = \ln p_{it}q_{it} - \alpha_l \ln l_{it} - \alpha_k \ln k_{it} - \alpha_m \ln m_{it} - \alpha_e \ln e_{it}. \tag{24}$$

To obtain firm-level estimates of idiosyncratic demand, Foster et al. (2008) estimate the demand function

$$\ln q_{it} = \alpha_0 + \alpha_1 \widehat{\ln p_{it}} + \sum_t \alpha_t \mathrm{YEAR}_t + \alpha_2 \ln(\mathrm{INCOME})_{mt} + n_t^i, \tag{25}$$

using an instrumental variable regression, where the log-price, $\ln p_{it}$, is instrumented by the estimate of TFPQ from (23), which serves as a supply shifter. The regression includes time fixed effects, and the average income in a plant's local market, m, is defined using the Bureau of Economic Analysis' Economic Areas. The residual of this equation is interpreted as a pure demand shifter for that firm.

For the eleven products in the analysis, Foster et al. (2008) report average five-yearly autocorrelations of 0.32 and 0.62 for idiosyncratic TFPQ and demand, respectively. The cross-sectional dispersion of TFPQ and demand are 0.26 and 1.16 respectively. This means that demand shocks are more persistent and more dispersed across firms. Furthermore, they report a correlation of –0.54 between firm-level prices and TFPQ and a correlation of 0.75 between firm-level TFPQ and TFPR.

[6] Examples include bread, block ice, and ready-mix concrete.

3.1.2 Pricing Moments

For moments related to micro-level pricing behavior, we reference Vavra (2014) who reports pricing moments using CPI micro-data from the Bureau of Labor Statistics spanning the period from 1988 through 2012.[7] Price data are at the product-outlet level and temporary sales are discarded from the analysis. In his sample, Vavra (2014) reports a monthly frequency of a regular price change to be 11%, of which 65% are upward adjustments. The average size of a price change excluding non-adjustments is 7.7%, and the standard deviation of price changes is 0.075.

3.1.3 Markup and Pass-Through of Cost Shocks to Prices

Following the methods of De Loecker et al. (2020), we estimate the markup distribution of U.S. public firms using Standard and Poor's Compustat data. To be in line with the sample in Foster et al. (2008), we restrict the analysis to data between 1980 and 2000. We follow the production approach and compute firm-level markups as the ratio of sales to cost of goods sold, multiplied by the output elasticity of variable inputs estimated at the two-digit NAICS level.[8] In our sample, the average markup is 56% and the median markup is 33%.

A major theoretical implication of a Kimball demand system is the incompleteness of cost pass-through to prices. One of the ways of capturing empirically the magnitude of cost pass-through can be found in the international finance literature. This literature looks at the pass-through of exchange rate shocks to importer prices, with the understanding that the exchange rate movements are exogenous from the viewpoint of importers. The empirical evidence is overwhelmingly in support of an incomplete pass-through of costs even in the medium and long-run: Campa and Goldberg (2005) estimate the long-run pass-through in the US to be 42% whereas Bergin and Feenstra (2009) report 24%, Gopinath and Itskhoki (2010) find it to be between 20% to 40%, and Gopinath et al. (2010) find an aggregate pass-through of 30%. Estimation of cost pass-through is more challenging in a purely domestic setting, due to the scarcity of appropriate data and well-identified shocks. Using indirect estimates of marginal costs, De Loecker et al. (2016) report cost pass-through between 31% to 41% among manufacturing firms in India. Recent studies using merged data on both costs and prices recover cost pass-through estimates that are similar to the

[7] The same dataset is widely used in the literature, see Bils and Klenow (2004) and Nakamura and Steinsson (2008).

[8] Following the literature, we exclude the following two-digit industries: utilities, finance and insurance, real estate and rental and leasing, as well as public administration.

Table 1: Externally Calibrated Parameters

Parameter	Description	Value	Source
β	Discount Factor	0.9966	Annual discount rate of 4%
χ	Labor disutility	1	Normalization
μ	Growth rate of S	0.002	Annual inflation rate of 2.4%
σ_S	SD of shocks to nom. expenditure	0.0037	Vavra (2014)
σ_{zn}	Corr. b/w productivity and demand innov.	0	Foster et al. (2008)

Note: This table displays the externally calibrated parameters in the model.

international macro evidence. Using Chilean supermarket-supplier merged data, Aruoba et al. (2022) find that 29% of a supplier price change is passed onto the retail price conditional on a price change at the supermarket level. Carlsson et al. (2022) estimate that between 21% to 33% of innovations to firm productivity are passed through to prices using data on Swedish manufacturing firms. Overall, the evidence from both the open- and closed-economy literature points to incomplete cost pass-through to prices in the range of 20% to 40%.

3.2 Externally Calibrated Parameters

In this section we explain how we fix a subset of parameters. A period is a month, and we set the monthly discount rate β to 0.9966 such that the annual discount rate is 4%. Consistent with the usual choice in the literature (Golosov and Lucas, 2007), the disutility of labor χ is normalized to 1, so that the nominal wage, W_t, is equal to the money supply, S_t. The monthly growth rate of the money supply, μ, is 0.2%, which implies an annual inflation rate of approximately 2.4%. Following Foster et al. (2008), we assume that idiosyncratic demand and productivity innovations are uncorrelated ($\sigma_{zn} = 0$).[9] Finally, when we run monetary policy experiments, we set $\sigma_S = 0.0037$ following Vavra (2014) to match the volatility of nominal output growth in the U.S. Table 1 summarizes the five externally calibrated parameters.

3.3 Internally Calibrated Parameters

The former analysis leaves seven parameters to be internally calibrated. Four of these parameters govern the AR(1) processes for idiosyncratic productivity (ρ_z, σ_z) and idiosyncratic

[9] The estimation strategy of Foster et al. (2008) requires demand and productivity to be orthogonal. Using Colombian data and an alternative strategy that relaxes the assumptions about the covariance between demand and supply shocks, Eslava et al. (2023) report a correlation of -0.07 between demand and productivity.

demand (ρ_n, σ_n), two additional parameters govern the Kimball demand system (ω, ψ), and the final parameter is the fixed menu cost (f). In our baseline model, we jointly calibrate these seven parameters to match seven moments. We use a single pricing moment – the frequency of price changes – and six firm-dynamics moments from Foster et al. (2008): the five-yearly autocorrelation and the variance of demand and TFP, Corr(TFPQ,P) and Corr(TFPQ,TFPR). In Section 5 we turn to some alternative calibrations that fix a set of parameters in order to highlight how our baseline calibration works.

The model-based moments we need for calibration can only be computed via simulation. To that end, we simulate 20,000 firms for 700 periods and drop the first 100 periods before computing any statistics. Computing moments that are monthly is straightforward. In order to compute moments that have their data counterpart in Foster et al. (2008), we aggregate the simulated data to the corresponding frequency and replicate their methodology. In particular, we aggregate the simulated monthly data into annual frequency by taking simple sums of revenue, sales, and employment. We then construct a panel dataset with the same time structure as Foster et al. (2008), namely five waves of annual observations that are five years apart. Because labor is the only input and production technology is constant returns to scale in the model, we recover firm-level TFPQ and TFPR as,

$$
\begin{aligned}
TFPQ_{it} &= \ln q_{it} - \ln l_{it}, && (26)\\
TFPR_{it} &= \ln (p_{it} q_{it}) - \ln l_{it}. && (27)
\end{aligned}
$$

This is equivalent to mapping our unique inputs to their basket of inputs. We estimate the demand function using the same IV specification as Foster et al. (2008),

$$
\ln q_{it} = \alpha_0 + \alpha_1 \widehat{\ln p_{it}} + \text{Time FE} + \eta_{it}, \tag{28}
$$

where $\ln p_{it}$ is instrumented by $TFPQ_{it}$, and recover firm-level demand shifters as the residuals, η_{it}. At the end of this process, we obtain five-yearly measures that are direct counterparts of those computed by Foster et al. (2008). To be clear, in computing the model-implied moments we treat the model-generated data exactly the same way they treat actual data.

Before turning to the results, a discussion on the identification of parameters is in order. While all parameters influence the model's ability to match all calibration targets, some parameters are more responsible for matching specific target moments. Some of these are intuitive. The fixed cost, f, has a significant role in the model-implied frequency of price changes, and the shock-process parameters $(\rho_z, \sigma_z, \rho_n, \sigma_n)$ are mostly related to the corresponding five-yearly moments from Foster et al. (2008).

What may be less obvious or familiar is how ψ and ω are linked to the correlations of TFPQ with prices and TFPQ with TFPR. To understand this, it is instructive to start from a CES demand system (with $\psi = 0$). The profit-maximizing rule in that framework delivers a pricing strategy that sets price as a constant markup over marginal cost. As a result, a firm's optimal price is inversely proportional to its productivity, that is $\text{Corr}(P, TPFQ) = -1$. Moreover in a CES demand system, TFPR is equalized across firms and as such $\text{Corr}(TFPR, TPFQ) = 0$. This is because optimizing firms will operate at the point where the marginal product of labor ($p_{i,t}z_{i,t}$) is equal to the nominal market wage. Under a Kimball demand system, both productivity and demand factors affect the optimal price of a firm. In particular, deviations from CES, controlled by the parameter ψ, diminish the role of productivity in pricing relatively to demand. Because some price changes will be due to demand shocks, this, in turn, would reduce the perfect negative correlation between price and TFPQ. The parameter ω, on the other hand, governs the elasticity of substitution between varieties, with a higher value of ω indicating less substitutability across varieties. In a Kimball demand system, more productive firms can charge a higher price if the elasticity of substitution is lower (ω is higher) leading to a higher correlation coefficient between TFPQ and TFPR.

The results of the internal calibration are reported in Table 2. The top panel shows the targeted moments which were described in Sections 3.1.1 and 3.1.2. The second panel shows the seven parameters calibrated jointly. The results show that all seven moments are matched very closely. The calibrated value of $\psi = -1.27$ indicates substantial deviation from CES. The calibrated process for idiosyncratic demand is highly persistent with a monthly autocorrelation of 0.997, whereas the idiosyncratic productivity process exhibits less persistence with a monthly autocorrelation of 0.98. The standard deviation of the innovation to idiosyncratic productivity (0.06) is higher than that of idiosyncratic demand (0.02). Preferences appear to be very persistent and subject to relatively small shocks when compared to technology. However, the stationary distribution of idiosyncratic demand is highly dispersed due to the persistent nature of the process.

Having established that the model is successfully calibrated to match the *targeted* moments, in Section 4.1 we turn to the discussion of how the model performs in matching several *untargeted* moments. Before we do so, however, we discuss how the seven calibrated parameters are informed by the data and provide evidence of external validity.

Table 2: Internal Calibration

Moment	Data	Baseline Model
Frequency of price changes	0.11	0.12
5-yearly autocorrelation of z_t^i	0.32	0.32
Cross-sectional standard deviation of z_t^i	0.26	0.25
5-yearly autocorrelation of n_t^i	0.62	0.62
Cross-sectional standard deviation of n_t^i	1.16	1.05
Corr between TFPR and TFPQ	0.75	0.74
Corr between price and TFPQ	–0.54	–0.57
Parameter	**Description**	**Value**
ψ	Super-elasticity	–1.27
ω	Elasticity of Substitution	1.29
ρ_z	Persistence of z_t^i	0.98
σ_z	Standard deviation of z_t^i	0.06
ρ_n	Persistence of n_t^i	0.997
σ_n	Standard deviation of n_t^i	0.02
f	Menu cost	0.03

Note: The top panel of this table compares the targeted moments and model-implied moments. The bottom panel shows the parameter values for each calibration.

3.4 Identification of Model Parameters

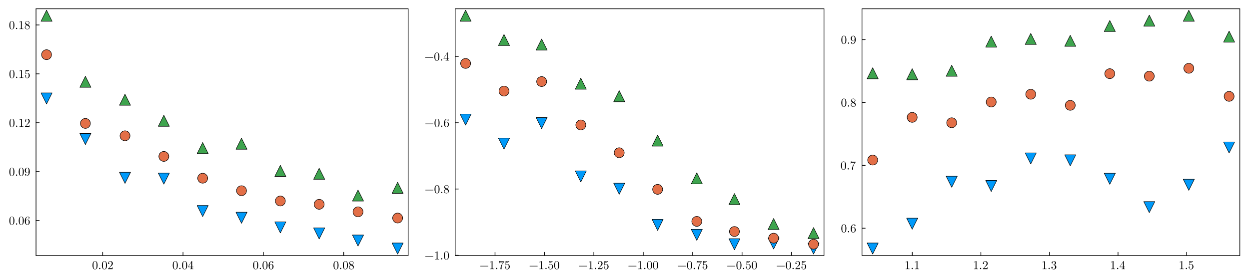
In order to demonstrate that the calibration targets are indeed very informative for the respective parameters we borrow an exercise from Daruich (2022), which mimics the first stage of the multistart global optimization proposed by Arnoud et al. (2022).

The main idea is to generate variation in the parameter space and investigate how the implied calibration targets are impacted – essentially taking a partial derivative. To do so, we first draw 500 parameter vectors from uniform Sobol points given a hypercube of the parameter space, which generates a quasi-random set of candidate parameter vectors.[10] Then, for each parameter vector, we solve and simulate the model to compute the relevant model-implied moments. This allows us to see how each of the seven parameters influences each of the seven calibration targets.

[10] A uniform Sobol sequence (Sobol, 1967) is a sequence of points that spans the n-dimensional hypercube in an even and quasi-random manner. For the purpose of the exercise, using quasi-random Sobol numbers are more efficient than drawing random numbers because Sobol numbers are designed to sample the space of possibilities evenly given the total number of draws, whereas a truly random sample is subject to sampling noise.

Figure 2: Identification of Internally-Calibrated Parameters

(a) Price Adjustment Freq. vs. f (b) Corr($TFPQ, P$) vs. ψ (c) Corr($TFPQ$,$TFPR$) vs. ω



Note: For each decile of a given parameter plotted on the horizontal axis, the red dot shows the median of the moment that is assigned to the parameter. The blue down-pointing triangles and green up-pointing triangles show the 25^{th} and 75^{th} percentiles respectively.

Figure 2 plots the values of three key model-implied target moment against the values of the parameter it is assigned to.[11] In particular, we group the values of each parameter in deciles, which we plot on the horizontal axis. Then, for each decile, we show the median value of the associated moment in red circled dots and the 25^{th} and 75^{th} percentiles in blue down-pointing triangles and green up-pointing triangles, respectively. The slope of the scatter plot is informative about the importance of that parameter, whereas the vertical dispersion reveals the influences of all other parameters on a particular moment.

The frequency of price adjustment exhibits a strong negative correlation with the menu cost f. Meanwhile, other parameters also play a role as is evident in the vertical dispersion. For example, for a fixed value of f, larger idiosyncratic shocks generate more frequent price changes. Consistent with our reasoning, we recover a strong negative relationship between Corr(TFPQ,P) and ψ. We also observe a weaker but visibly positive relationship between Corr(TFPQ,TFPR) and ω. The large variation in this correlation given a value of ω reveals that it is sensitive to the values of other parameters in addition to ω. In particular, we find that σ_n and σ_z, which determine the stationary distribution of idiosyncratic productivity and demand, have sizable effects on the level of this correlation. Given that all the parameters except for ω exhibit tight links with their associated targets, ω can be identified by Corr(TFPQ,TFPR) when all other parameters are fixed and matched to their respective targets.

The key takeaway from this exercise is that the links between the parameters and moments are quite tight. While it is too computationally intensive, if one were to consider a

[11] We delegate the figures for the stochastic properties of the demand and supply shocks to Appendix C.

formal generalized method of moments approach to estimating the parameters of interest, this analysis suggests that one would obtain fairly tight standard errors for the estimates.

3.5 External Validity

The work of Foster et al. (2008) is the only study for the U.S. with a systematic estimation of productivity and demand shocks at the firm level for different industries. This estimation requires price and quantity data at the product level along with other information such as inputs. By using a carefully selected set of firms in the Census of Manufactures that produce uniform products, they are able to separately estimate shock processes for productivity and demand. However, the external validity of the estimation of Foster et al. (2008) to the broader economy may be a concern.

Several other countries have similar data for a wider set of firms, and researchers have estimated some of the moments that we use for our identification strategy. For instance, Eslava et al. (2013) use Colombian firm-level data covering the entire manufacturing industry to separately identify productivity and demand processes at the firm level. They report similar values for $\text{Corr}(TFPQ, TFPR)$ and $\text{Corr}(TFPQ, P)$, which are crucial for pinning down the Kimball demand system parameters. Specifically, they report $\text{Corr}(TFPQ, TFPR) = 0.69$ and $\text{Corr}(TFPQ, P) = –0.65$ versus the Foster et al. (2008) values of 0.75 and –0.54, respectively.[12]

Given that idiosyncratic demand and productivity processes determine the ergodic properties of firm growth, we can also compare the cross-sectional dispersion of output growth rates computed from model-simulated data with external evidence to gauge if the estimates from Foster et al. (2008) can be generalized beyond the eleven industries. We benchmark our estimates to Davis et al. (2006), a study using the Longitudinal Business Database and thus a good measure of the U.S. business dynamics. They estimate the cross-sectional standard deviation of firm revenue growth rate to be 0.39 over the period 1982-1997, while in our simulated data this untargeted moment is 0.41. Therefore, we conclude that the estimation of Foster et al. (2008) can be used to shed some light about the general behavior firms.

[12] Moreover, they estimate a correlation between productivity and demand innovations in the neighborhood of zero, consistent with Foster et al. (2008) and our assumption.

Table 3: Untargeted Pricing Moments

Moments	Data	Baseline
Average Size	0.08	0.07
Fraction Up	0.65	0.58
SD(Δp)	0.08	0.07

Note: This table shows the three untargeted moments: average size of adjustment conditional on a price change, the fraction of adjustments that are positive, and the standard deviation of price changes excluding zeros from the data and from model simulated data.

4 Results

We calibrated our model using firm-dynamics moments from Foster et al. (2008) and one pricing moment, the average frequency of price adjustments. In this section, we investigate the implications of the model in three other dimensions: additional pricing moments, the distribution of markups and the pass-through of shocks, and monetary non-neutrality, none of which was targeted in our calibration.

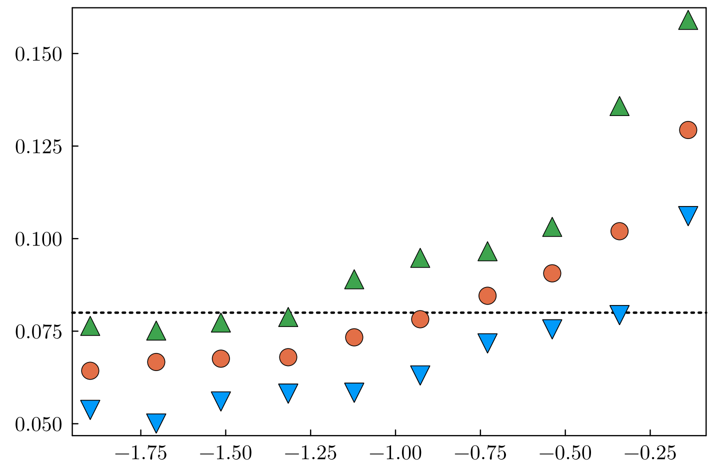
4.1 Untargeted Pricing Moments

Table 3 shows three important pricing moments we do not target in the calibration of the baseline model, along with their data counterparts. These are the average size of a price change conditional on a change, the fraction of adjustments that are positive, and the dispersion of non-zero price changes. Our model matches all three moments very well.

One may wonder if the success of the model has much to do with the precise calibration of ψ, or whether any deviation from CES by reducing ψ below zero would have been sufficient. In Figure 3 we plot the average size of non-zero price changes versus different values of ψ, similar to Figure 2, where the vertical variation for a given level of ψ is due to the differences in other parameters. This figure shows that there is a tight relationship between the average size of adjustments and ψ. As ψ falls, the increasing real rigidities (strategic complementarities) make the firms less and less willing to deviate from their competitors, thus avoiding large price changes. This leads to a decline in the average size of price changes. Therefore, it is even more remarkable that our model matches these moments as well as it does, given that a different value of ψ from the model calibrated to firm-dynamics moments may have led to a completely different value for the average size of price changes.

We also examine the hazard function of price change from the model. The hazard of a

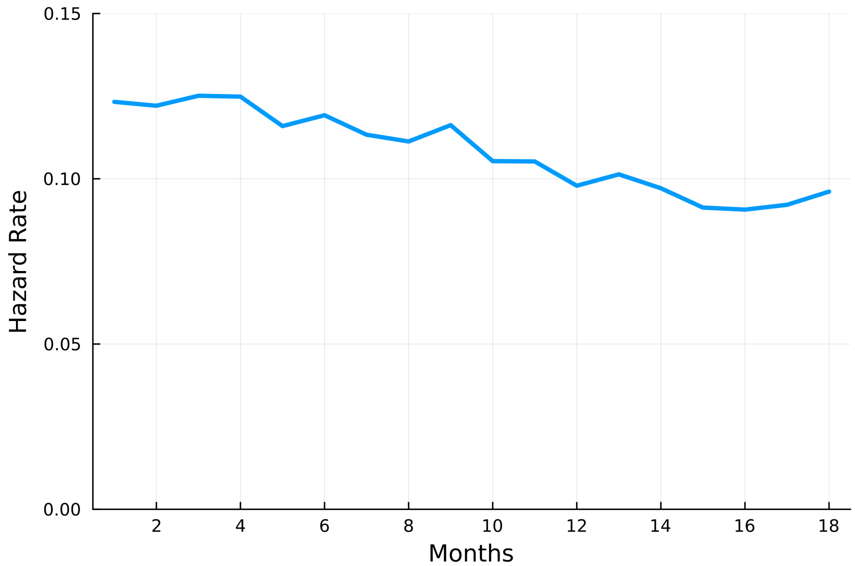
Figure 3: Value of ψ and Average Size of Non-zero Price Changes



Note: We group values of ψ into deciles. For each decile of ψ values plotted on the horizontal axis, the red dots, blue down-pointing triangles, and green up-pointing triangles show the median, 25^{th}, and 75^{th} percentiles of a given untargeted pricing moment respectively. The underlying data is the same as that used for Figure 2. Specifically, they are random draws from a hypercube of parameter space.

price change is the probability that a price will change t periods after the last adjustment, conditional on the price spell lasting t periods. Empirically, the hazard function is found to be either downward-sloping (Nakamura and Steinsson, 2008) or flat (Klenow and Kryvtsov, 2008). As pointed out by Nakamura and Steinsson (2008), simple menu cost models are typically not able to generate hazard functions that are consistent with the empirical evidence. This largely hinges on the calibration of the idiosyncratic processes. In a model with trend inflation, the hazard function is upward sloping when idiosyncratic shocks are small. Larger idiosyncratic shocks and more persistent idiosyncratic processes flatten the hazard function as they lead to temporary price changes that are often reversed quickly. When idiosyncratic shocks are sufficiently large, a simple menu cost is able to generate a downward-sloping hazard. Nakamura and Steinsson (2008) argue that such calibrations are unrealistic, due to the fact that they are inconsistent with micro pricing facts. This conclusion has spurred alternative pricing models that seek to rationalize a downward-sloping hazard, such as Baley and Blanco (2019), who introduce firm-level uncertainty and learning to generate frequent price changes shortly following an adjustment. Figure 4 shows the mildly downward-sloping hazard curve generated by our model, which is consistent with the empirical results in the literature we cite above: the hazard rate falls from 12% at one month to 10% at one and a half years.

Figure 4: Hazard Function of Price Change



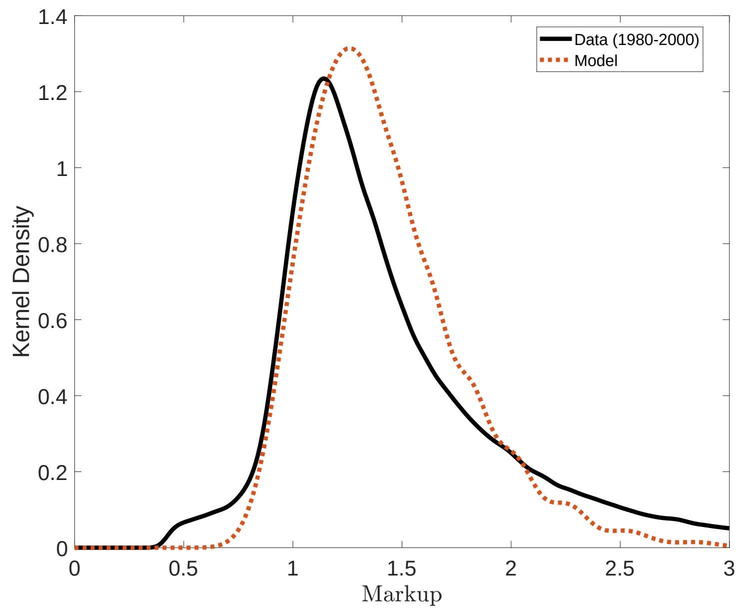
Note: This figure plots the pricing hazard for the four model calibrations over the first 18 months.

4.2 Markups and Cost Pass-through

In a flexible-price model with CES demand and symmetric firms, all prices are set to a fixed markup above marginal cost and thus all firms have the same markup, leading to a degenerate markup distribution. With pricing rigidities, firms that cannot change their price in a period deviate from their desired markup. As a firm's marginal cost changes via changes in productivity, so would its price, reflecting complete pass-through of cost to price. Because demand shocks do not change prices, they have no impact on markups. Thus it is clear that a model with CES demand would produce a counterfactual markup distribution, which is near-degenerate around the desired markup.

The desired markup of a firm facing a Kimball demand system depends on both its idiosyncratic productivity and demand. Because the pass-through of marginal cost to price is incomplete, more productive firms do not fully pass on their cost advantage to price, resulting in higher markups. Also, firms with larger idiosyncratic demand optimally choose higher prices and hence higher markups. The cross-sectional distribution of productivity and demand, alongside pricing frictions, result in a non-degenerate markup distribution in the model. Figure 5 plots the kernel density of the cross-sectional distribution of gross markups both from the data and the model. The markup distribution from the model mimics the untargeted empirical distribution reasonably well. The median gross markup in the model

Figure 5: Cross-Sectional Distribution of Gross Markup: Model vs. Data



The figure plots the kernel density of the empirical markup distribution from publicly traded firms in the U.S. as well as the kernel density of the markup distribution in the ergodic distribution of the model. Both kernel densities are computed using the optimal bandwidth for normal densities.

is 1.35, compared to 1.33 in the data. The model-implied markup distribution, however, exhibits lower variance relative to the empirical distribution, much of which stems from the tails. Our model does not generate as many firms that have markups larger than 200%, nor does it deliver markups deep into the negative territory. Matching these extreme tails of the distribution would require additional features such as non-Gaussian shocks to demand, monopolies (large positive markups), customer capital, or firm exit (negative markups).

The success of our model in replicating the empirical markup distribution is perhaps not very surprising given the work of Edmond et al. (2023), among others, who use Kimball demand systems for modeling firm markups, although we do not target any empirical moments related to the markup distribution. Surprisingly, while our model is calibrated to six demand, supply and productivity moments from Foster et al. (2008) that rely on selected manufacturing industries, the model provides a remarkable fit to the untargeted distribution of estimated markups for a broad set of industries.

In our model, as we explain in Section 2.3, there is a direct link between ψ and the strength of the pass-through of cost shocks. This parameter, in turn, is linked to the strength of the strategic complementarity and is calibrated primarily using the correlation between TFPQ and prices from Foster et al. (2008). The calibrated model has a cost pass-through of 38%, which is in line with the empirical estimates in the range of 20 to 40 percent reported in the literature, as reviewed in Section 3.1.3.

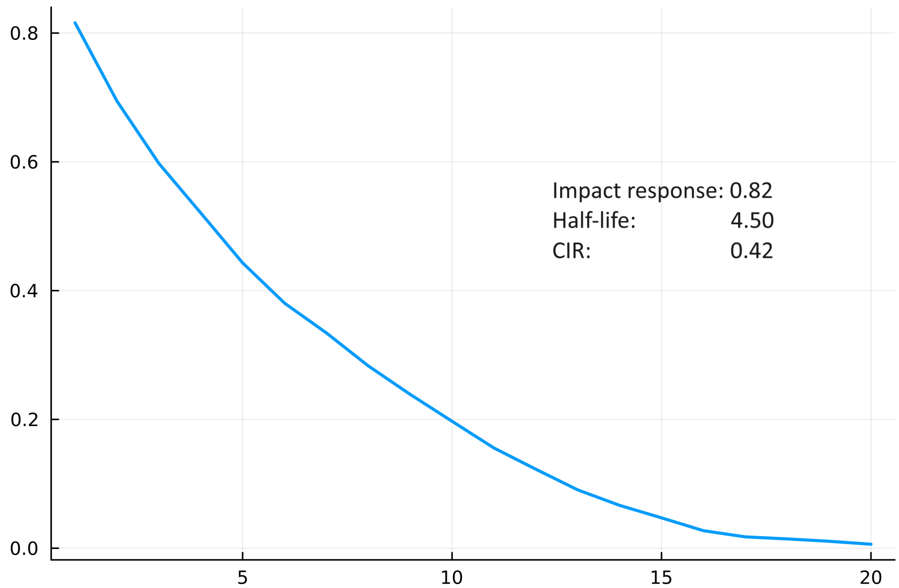
4.3 Monetary Policy and Non-Neutrality

Nominal rigidities alone are typically insufficient at generating sizable real effects to nominal shocks (Caplin and Spulber, 1987; Golosov and Lucas, 2007). A solution put forth by Ball and Romer (1990) is to introduce real rigidities in conjunction with nominal pricing frictions. Since then, the macroeconomics literature has explored the plausibility of various potential sources of real rigidities. Strategic complementarities in pricing induced by a Kimball demand system is one such mechanism. When a monetary policy shock hits the economy, some firms choose not to respond due to the presence of *nominal* rigidities, which makes monetary policy effective in stimulating real activity, or non-neutral. With strategic complementarities, the firms that adjust their prices choose to adjust less than under CES demand in order to remain closer to their competitors who choose not to adjust their prices, and this *real* rigidity adds to the degree of non-neutrality. This intuition is formalized in Alvarez et al. (2023), who derive analytic results in a menu cost model featuring strategic complementarity casted as a Mean Field Game and show that complementarity makes the impulse response of output to a nominal shock larger at each horizon.

In this section we turn to assessing whether our model is able to generate sizable real responses to nominal shocks. To that end, we consider four measures of non-neutrality. The first measure is the unconditional standard deviation of consumption in a long simulation. Our model features no aggregate shocks other than the nominal expenditure shock ϵ_t in (22), which can also be interpreted as a monetary policy shock. As such, if monetary policy was perfectly neutral, then aggregate consumption would be constant. Thus, the standard deviation of aggregate consumption serves as a measure of deviation from neutrality. We find a standard deviation of 0.52%, which is very sizable. The other three measures are obtained from a response of the economy to a one-time change in ϵ_t. In particular, we compute the response of real output to a positive nominal expenditure shock of size 0.2%, which doubles the monthly growth rate of aggregate expenditure. The first object we compute from this response is the peak response of output as a fraction of the size of the shock, which in a model like ours without any internal propagation, will happen on impact.[13] The second

[13] If the aggregate price does not respond at all, then the entirety of the increase in nominal expenditure is reflected in an increase in real output, leading to complete transmission to real output. On the other

Figure 6: Impulse Response of Real Output to a Nominal Expenditure Shock



Note: This figure plots the impulse response of real output expressed as a fraction of the nominal expenditure shock on the vertical axis and periods elapsed since the shock on the horizontal axis.

object we can compute from this response is the half-life of the shock, which tells us how persistent the effect of the shock is. Finally, these two measures can be summarized by the cumulative impulse response (CIR), which adds up the response of output as a fraction of the shock over the period when it is non-zero and divides by the number of periods in a year – 12 in our case.

Before we turn to the results, it is worth considering a benchmark. As we show in Appendix A.2 in detail, in a CES model with Calvo pricing, where firms receive an i.i.d. shock that determines when they can adjust their prices, the output response as a fraction of the shock is exactly $(1-\alpha)^h$ where α is the probability that firms can change their price and h is the horizon where $h = 1$ is the period of the shock. Based on the evidence we presented in Section 3.1.2, α would be set to 0.11 in such a model, indicating that the initial (and peak) response of output would be 89% of the shock. The cumulative response according to the Calvo model would be 0.76.[14] This coincides with the result of Alvarez et al. (2016) who show that the CIR can be expressed in terms of the kurtosis and frequency of price changes. According to their formula, the Calvo model has a CIR of 0.76 whereas a

extreme, if the aggregate price is perfectly flexible, the real effect of the nominal shock would be zero.

[14]The CIR in the Calvo model is given by $\sum_{t=0}^{\infty} \frac{0.89^t}{12} = 0.76$

menu cost model à la Golosov and Lucas (2007) has a CIR of 0.13.

Figure 6 plots the impulse response of real output expressed as a fraction of the size of the shock for the baseline model. In our model, on impact, approximately 82% of the increase in nominal expenditure is reflected in an increase in the real output. The real effects of the shock dampen as time goes by and eventually die out about 20 months after the shock. The half-life of the shock is 4.5 months. The larger response on impact coupled with slower decay is reflected in the relatively large CIR of 0.42. The degree of monetary non-neutrality that the model generates is larger than a simple menu cost model. The real effect is close to the Calvo version, which generates a cumulative response that is six times as large as the Golosov-Lucas model (Alvarez et al., 2016). It is fair to say that a response of this magnitude is considered appropriately large in the literature as summarized in Mongey (2021).[15]

The model exhibits substantial monetary non-neutrality due to two key features. First, as discussed, a Kimball demand system introduces strategic complementarity in pricing among firms as deviations from the prices of competitors are costly. Therefore, firms are hesitant to respond aggressively, if at all, to the nominal shock if a sizeable share of other firms are not responding. Here, we show that a degree of strategic complementarity, as controlled by ψ and summarized by the cost pass-through to prices, that is consistent with empirical evidence contributes significantly to non-neutrality. Second, the presence of both productivity and demand shocks plays a role in weakening the selection effect in price adjustments to an aggregate shock, which further contributes to a large output response. As shown in Golosov and Lucas (2007), the real response to nominal shocks hinges not on how many prices adjust but which prices adjust. In the absence of idiosyncratic shocks, prices respond only to aggregate shocks and the only prices that adjust are those that are most out of line with the aggregate shock. Adding idiosyncratic shocks weakens this so-called selection effect as firms respond not only to aggregate shocks but also to disturbances to their idiosyncratic states. Having two idiosyncratic shocks further weakens the selection as firms in our model respond to two orthogonal shocks in productivity and demand on top of shocks to aggregate nominal

[15] As Table A.1 of Mongey (2021) summarizes, menu-cost models without real rigidities typically generate a peak output response in the range of 0.35 to 0.50 (Golosov and Lucas, 2007). Some of the other studies exploring alternative sources of real rigidities find higher numbers. Nakamura and Steinsson (2010) consider a multi-sector model with a round-about production structure and obtain a peak output response of 0.80. Gertler and Leahy (2008) model segmented labor markets and report a real response of 0.75. Burstein and Hellwig (2007) incorporate decreasing returns to scale and wage rigidity and obtain 0.70, whereas Blanco et al. (2022) builds a similar single-product model with DRS and find a peak output response of 0.80. Lastly, Mongey (2021) studies a pricing model with duopoly competition and finds a peak response of 0.74. Relative to the standard setup with monopolistic competition, the duopoly model generates a CIR that is 2.3 times as large.

Table 4: Internal Calibration and Alternative Calibration: Target Size of Price Changes

Moment	Data	Baseline	Target Size	Leptokurtic
Frequency of price changes	0.11	**0.12**	0.09	**0.12**
Fraction of price increases	0.65	0.58	0.63	0.61
Size of price changes	0.08	0.07	**0.08**	0.08
Raw Kurtosis of Price Changes	4.50	1.75	1.38	**4.13**
5-yearly autocorr of z_t^i	0.32	**0.32**	**0.31**	**0.32**
Cross-sectional SD of z_t^i	0.26	**0.25**	**0.27**	**0.24**
5-yearly autocorr of n_t^i	0.62	**0.62**	**0.68**	**0.58**
Cross-sectional SD of n_t^i	1.16	**1.05**	**1.07**	**1.07**
Corr b/w TFPR and TFPQ	0.75	**0.74**	**0.73**	**0.58**
Corr b/w price and TFPQ	–0.54	**–0.57**	**–0.59**	**–0.55**
Parameter	**Description**			
ψ	Super-elasticity	**–1.27**	**–1.32**	**–1.10**
ω	Elasticity of Substitution	**1.29**	**1.25**	**1.32**
ρ_z	Persistence of z_t^i	**0.98**	**0.98**	**0.98**
σ_z	Standard deviation of z_t^i	**0.06**	**0.05**	**0.05**
ρ_n	Persistence of n_t^i	**0.997**	**0.996**	**0.82**
σ_n	Standard deviation of n_t^i	**0.02**	**0.02**	**0.18**
p_n	Poisson prob. for n_i^t shock	–	–	**0.025**
f	Menu cost	**0.03**	**0.06**	**0.01**
Impact Response of Monetary Policy		0.82	0.77	0.77

Note: The top panel of this table compares the targeted moments and model-implied moments for the two model specifications, where the bolded numbers highlight moments that are targeted in the calibration. The bottom panel shows the parameter values for each calibration.

expenditure.

4.4 Robustness

Before we turn to discussing our results and putting them in context of similar studies in the literature, we explore the robustness of our results with respect to two calibration choices we made. First, in the baseline calibration we targeted the frequency of price changes and left the size of price changes as an untargeted moment. Here we consider the reverse. Second, it is well known that the distribution of price changes implied by a standard menu-cost model like our baseline model exhibits negative excess kurtosis, in contrast to the positive excess kurtosis found in the empirical distribution. Our baseline model delivers a raw kurtosis of 1.75, which falls short of the U.S. estimates of 4.5, which is the middle of the range reported in Alvarez et al. (2016). We enhance the model to have leptokurtic demand shocks, and target the kurtosis of price changes.

The results are presented in Table 4. The first two columns replicate the results in Table 2, where we continue to use boldface to emphasize the moments being targeted and parameters used to do so. In results not shown, the two versions considered here perform as good as the baseline model in how they deliver the untargeted moments. The third column reports the results from the calibration where we target the size of non-zero price changes. The results are very close to the baseline results with the frequency of price changes and kurtosis being somewhat smaller. The impact response is down to 0.77 but still very much near the high end of the results from the literature.

The last column in Table 4 uses a leptokurtic shock, following Midrigan (2011), applied to the idiosyncratic demand shock, in order to target the kurtosis of non-zero price changes. In particular, we assume that the demand shock n_t^i follows the persistent AR(1) process in (13) with probability p_n and remains unchanged with probability $(1 - p_n)$ from one month to the next. We add p_n to the list of parameters being calibrated, and add the kurtosis of non-zero price changes at 4.5 as a target to match. The calibration delivers a kurtosis of 4.13 and is able to match rest of moments. p_n is calibrated to be 0.025, which means 97.5% of the time a firm inherits the demand from the previous month.[16] This calibration delivers the large degree of persistence found in the results of Foster et al. (2008) using a low p_n while the ρ_n is substantially reduced and σ_n is substantially increased. This delivers a highly leptokurtic path for demand shocks for a firm, which is flat for long periods and experiences large jumps when it changes, and this translates in to prices that display high kurtosis in price changes.

5 Discussion

In the preceding sections, we presented our main results. We constructed a model which is simultaneously consistent with: (a) firm dynamics moments, (b) pricing moments, (c) pass through of cost shocks, (d) firm-level markup distribution, and (e) large monetary non-neutralities. We argued that the two key ingredients of the success of the model in matching all these moments was a non-CES demand system and the introduction of idiosyncratic demand shocks. In this section we demonstrate that the presence of both of these elements is necessary for the success of the model. In order to do so, we review several alternative models, some of which are existing models in the literature, and some of which are special

[16]In comparison, Vavra (2014) obtains $p_z = 0.13$ where he applies the leptokurtic process on firm-level TFP. We choose to apply the leptokurtic process on demand rather than TFP since the former is much more persistent according to Foster et al. (2008) and this process is a convenient way of delivering this persistence.

cases of our baseline model, in terms of their implications for (a) through (e) above. We show that *none* of the alternatives are successful in simultaneously matching all of these elements. It is also worth noting that we discipline the two key features in our baseline model using firm dynamics moments without targeting (b) through (e), with the exception of the frequency of price adjustment. Therefore, the relative success of the baseline model compared to alternative setups is not simply due to having additional degrees of freedom.

The results are summarized in Table 5 for models that feature CES demand and Table 6 for models with non-CES demand. We turn to the details of each model below. In both tables, the first column shows data moments which were introduced in Section 3.1.[17]

5.1 Alternative Models with CES Demand

We consider three models that feature a CES demand system. The first two are the simple model of Nakamura and Steinsson (2008) and the model of Vavra (2014) that includes leptokurtic supply shocks and random menu costs in the form of random possibility of changing prices at no cost (commonly referred to as Calvo-plus), neither of which features an idiosyncratic demand shock. We also include a version of our baseline model that features CES demand and idiosyncratic demand shocks. In Appendix D.1, we show two additional versions of our model with CES demand but without idiosyncratic demand shocks.

All of these models are calibrated to match the frequency of price changes, and Nakamura and Steinsson (2008) and Vavra (2014) also target the average size of a price change. Looking at the first block of results in the first panel of Table 5, the two versions without idiosyncratic demand shocks, Nakamura and Steinsson (2008) and Vavra (2014), fail to match any of the firm dynamics moments: firm-level TFP is almost non-persistent and very small, and the two correlations are either too weak or too strong relative to the data, though less so for Nakamura and Steinsson (2008).[18] In column 4 we introduce the idiosyncratic demand shock to the model with CES demand. We match the first four moments by picking appropriate parameters for the two shocks and set $\omega = 1.33$ to match the median markup in our data. This leads to price changes that are too large (0.14 versus 0.08 in the data) and the model

[17] Some of the papers we review target slightly different numbers than we show in the tables (and thus may look like they do not match the targets as well as they should) but this does not make a material difference in the results.

[18] In both the Nakamura and Steinsson (2008) and Vavra (2014) columns in Table 5 as well as the No Demand and Klenow and Willis (2016) versions in Table 6, the standard deviation of n_t^i is positive despite the models not featuring a demand shock. This is because in computing these statistics we continue to follow the procedure we explain in 3.1.1 where the demand shocks are identified as the residual. This is the same reason why the two correlations are different from 0 and -1, respectively.

Table 5: Alternative Models: Models with CES Demand

Moment	Data	NS (2008)	Vavra (2014)	CES with Demand
5-yearly autocorr of z_t^i	0.32	0.00	0.06	**0.32**
Cross-sectional SD of z_t^i	0.26	0.03	0.09	**0.26**
5-yearly autocorr of n_t^i	0.62	0.01	0.00	**0.62**
Cross-sectional SD of n_t^i	1.16	0.01	0.06	**1.18**
Corr b/w TFPR and TFPQ	0.75	0.77	0.25	0.28
Corr b/w price and TFPQ	–0.54	–0.64	–0.90	–0.97
Frequency of Price Changes	0.11	**0.08**	**0.10**	**0.11**
Average Size of Price Changes	0.08	**0.08**	**0.06**	0.14
Pass-through of Supply Shocks	20%-40%	100%	100%	100%
Average Markup	1.56	1.34	1.17	1.33
Cross-sectional SD of Markup	0.79	0.11	0.06	0.09
Impact Response	NA	0.59	0.65	0.69

Note: Boldface denotes calibration targets for each model. NS (2008) refers to Nakamura and Steinsson (2008). They target a frequency of 0.08 and they match that. All results for Nakamura and Steinsson (2008) use the codes they supply except for the impact response, which we compute using our codes.

fails to match the two firm-dynamics correlations given the CES structure. It is instructive to compare the CES model with demand shocks to our baseline model in terms of matching the average size of price changes. They both match the same shock moments and have very similar parameters for the laws of motion of z and n. The z shocks are passed through to prices 100% (and n shocks are not passed through), yielding an average price change of 0.14. A back-of-the-envelope calculation using the 38% cost pass-through for our baseline model would suggest that for the same size productivity shocks, the size of price changes will be about 0.053. In the baseline model firms also change prices in response to demand fluctuations. The introduction of demand shocks increases this number to around 0.07, which is very close to the moment in the data at 0.08.

Turning to pass-through of supply shocks and the firm-level markup distribution, the results are as expected. A model with CES demand system will feature a 100% pass-through of supply shocks, failing to match the 20%-40% range we reported for the data. Moreover, in the absence of price rigidities, all firms would have an identical markup. When some firms are unable to change their prices, this leads to a non-degenerate distribution of markups but one that is symmetric, unlike the data. Moreover, the cross-sectional standard deviation of this distribution is about an order of magnitude smaller than what is in the data.

Finally, turning to non-neutrality, the CES models feature an impact response that ranges from 0.59 for the basic version to 0.69 for the one with idiosyncratic demand. This range is substantially lower than the Calvo response, and in line with the results reported in Alvarez et al. (2016) for a Golosov-Lucas-type menu cost model. Monetary policy in the version with demand shocks exhibits slightly greater non-neutrality due to the fact that these shocks act as random menu costs.[19] In the presence of random menu costs, firms are responding not only to the aggregate shock but also to idiosyncratic realizations of the adjustment cost, thereby weakening the selection effect of responding to idiosyncratic productivity shocks and raising monetary non-neutrality.

To summarize, the CES models we consider, even the one with idiosyncratic demand, are unable to match all of the moments that are matched in our baseline model, and these alternative models deliver a smaller amount of monetary non-neutrality.

5.2 Alternative Models with Non-CES Demand

In this section we turn to three alternative models that feature non-CES demand systems: a version of our model without an idiosyncratic demand shock, the Klenow and Willis (2016) model, and a version of our model that features a translog demand system instead of a Kimball demand system.

Results are reported in Table 6, where for comparison the second column replicates the results for the baseline model. In the third column, we remove the idiosyncratic demand shocks from our baseline model and recalibrate. In doing so, naturally, we need to give up on matching the demand moments. Moreover, the model can no longer match the two firm-dynamics correlations and as such we fix the values for ψ and ω at their baseline values. The failure of the model to match these four moments is the main downside of this version of the model. Moreover, the average price change is also somewhat smaller and the markup distribution is less dispersed. The latter results arise from the absence of demand shocks that cause the firm the change its price. This version of the model displays similar results to the baseline model in other dimensions.

The next column reports the results from Klenow and Willis (2016), where we use the calibration in which they target a frequency of price change of 0.09.[20] This model has three

[19]Random menu costs were first introduced by Dotsey et al. (1999). Recent work by Nakamura and Steinsson (2010), Midrigan (2011), and Alvarez et al. (2016) explicitly explore the implications of random menu costs on monetary non-neutrality.

[20]The baseline calibration in Klenow and Willis (2016) ($\theta = 5, \epsilon = 10$ using their specification) translates

Table 6: Alternative Models: Models with Non-CES Demand

Moment	Data	Baseline	No Demand	KW (2016)	Translog
5-yearly autocorr of z_t^i	0.32	**0.32**	**0.32**	0.00	**0.29**
Cross-sectional SD of z_t^i	0.26	**0.25**	**0.26**	0.27	**0.27**
5-yearly autocorr of n_t^i	0.62	**0.62**	0.04	0.00	**0.68**
Cross-sectional SD of n_t^i	1.16	**1.05**	0.42	0.23	**1.11**
Corr b/w TFPR and TFPQ	0.75	**0.75**	0.99	0.98	**0.55**
Corr b/w price and TFPQ	–0.54	**–0.57**	–0.98	–0.77	**–0.76**
Frequency of Price Changes	0.11	**0.12**	**0.11**	**0.09**	**0.10**
Average Size of Price Changes	0.08	0.07	0.06	0.14	0.12
Pass-through of Supply Shocks	20%-40%	38%	38%	28%	60%
Average Markup	1.56	1.42	1.38	1.64	1.46
Cross-sectional SD of Markup	0.79	0.39	0.28	0.57	0.31
Impact Response	NA	0.82	0.83	0.85	0.73

Note: Boldface denotes calibration targets for each model. KW (2016) refers the calibration in the third row of Table 6 in Klenow and Willis (2016) where they target a frequency of price change of 0.09.

main problems. First, because the model does not feature an idiosyncratic demand shock, it fails to match moments related to demand, as well as the correlation between TFPR and TFPQ. The correlation of price and TFPQ is stronger than what is in the data, but still considerably less than 1 in absolute value due to the incomplete and non-linear pass-through of cost (productivity) shocks, which is 28% with their calibration. Second, the average size of price changes is too high at 0.14. Third, in order to match other price moments they use (standard deviation and some sectoral price moments) they need large firm-level TFP innovations. Their calibration for TFP has $\rho_z = 0.89$ and $\sigma_z = 0.18$, where the former yields virtually no autocorrelation at the 5-yearly frequency and the latter is about three times as large as our calibration. Comparing our results with that of Klenow and Willis (2016), we see that the key is the inclusion of demand shocks. In order to match the same pricing moments, one needs much smaller TFP shocks in our model because the demand shocks, with their 72% pass-through, create more reasons for the firm the change its price. What is a key result, however, is the outcome that once TFP and demand shocks are disciplined by the evidence in Foster et al. (2008), the size of price changes turns out just right.

We also consider the translog demand system, which is frequently used in the trade and

roughly to parameter values $\omega = 1.25$ and $\psi = -2$ under our specification of the Kimball aggregator, neither of which are too far from our calibrated values. ψ is more negative indicating stronger pricing complementarities and thus a smaller pass-through of cost shocks.

industrial organization literatures as an alternative deviation from CES demand. As Bergin and Feenstra (2000) explains, this system achieves the same goals as the original Kimball (1995) paper, but does so in a more explicitly parametric way. They start with a sub-utility function defined by the dual expenditure function that has the form

$$\ln P_t = \sum_{i=1}^{n} \alpha_i \ln P_{it} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \ln P_{it} \ln P_{jt} \tag{29}$$

with restrictions $\gamma_{ij} = \gamma_{ji}$, $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{i=1}^{n} \gamma_{ij} = 0$, where N is the number of distinct intermediate goods and P_{it} is the price of good i.

In order to solve our model using this demand system, we take advantage of results provided in Mrázová and Neary (2017) who provide formulas to nest a translog demand system inside the Kimball demand system, under the assumptions of symmetric firms and at a steady state. This makes computation much easier and also enables us to easily compare the results to our baseline model. The detailed derivations in Appendix D.2 show that a restriction of $\psi = -\frac{1}{\omega^2}$ in a Kimball demand system would lead to a translog demand system with an elasticity of demand given by $\frac{\omega}{1-\omega}$.

The last column in Table 6 shows the results from a calibrated version of our model with a translog demand system. To make it as comparable to our baseline as possible, we calibrate its parameters to the same moments as our baseline model. Given that a translog demand system is equivalent to a Kimball demand system with a restriction on the parameters ψ and ω, this implies that we have one more moment than parameters in the calibration. Our results are robust to dropping one of the correlation moments and calibrating a balanced system. The model featuring a translog demand system matches the four moments that come from the shock processes. However the correlation between TFPR and TFPQ is too weak and the correlation between price and TFP is too strong relative to the data (and the baseline model). This is because the pass-through of supply shocks turns out to be 60%, well outside the relevant range. In fact, we show in Appendix D.2 that cost pass-through in the translog model is constrained to be between 50% and 100%.[21] Relatedly, given the size

[21] We compute this pass-through in two ways. First, we use the formula we derived for Kimball in Appendix A as we did for the baseline model. Second, Mrázová and Neary (2017) provide a formula for pass-through for all demand systems they consider, one of which is translog. Plugging in the parameter values in to that formula yields the identical result. Bergin and Feenstra (2000) apply an approximation, which is valid for small shares of expenditures or equivalently small markups, that shows the pass-through to cost shocks is 50%. Rodriguez-Lopez (2011) derive an exact formula for the cost pass-through and show that it is in general different from 50%.

of the underlying shocks, the size of price changes turn out to be 12%, which is much larger than the data. Turning to the effects of monetary policy, the impact response is somewhat smaller at 0.73. Considering all results together, the baseline model with a Kimball demand system is more consistent with the firm-level evidence regarding prices, pass-through and productivity than a model with a translog demand system. Due to having one additional degree of freedom and no restrictions on the cost pass-through, a Kimball demand system is more flexible over a translog system and thus more desirable to work with. Nonetheless, as a model with non-CES demand, the translog demand system certainly performs better than models using CES demand.

6 Conclusion

The advent of rich micro datasets has spurred the advancement of models of firm dynamics over the past twenty years to study a range of macroeconomic and international topics, including monetary non-neutrality, exchange-rate pass-through, and firm markups. From each of these literatures, key modeling ingredients have emerged as important for explaining aggregate and firm-level dynamics. From the exchange-rate pass-through and markup-dynamics literatures, models of strategic complementarity, or demand system with non-constant elasticities more generally, have been shown to play an important role. From the price-setting literature, nominal rigidities play a central role, but additional features in the form of real rigidities have also been necessary to approximate the degree of monetary non-neutrality from quantitative estimates.

However, each of these literatures has faced limitations that have prevented the emergence of a single modeling framework able to deliver the main results across all of these areas of study. From the price-setting literature, Klenow and Willis (2016) show that in order to generate realistic price-adjustment moments, a model with strategic complementarities requires a large magnitude of idiosyncratic productivity shocks that is inconsistent with micro evidence. A related limitation across these literatures has been the absence of micro data sets containing both prices and quantities at the firm level. Thus, most of these studies estimate or calibrate the parameters of idiosyncratic productivity processes to match pricing moments, because these pricing datasets lack data on quantities.

We propose a parsimonious framework that brings together the modeling elements from these literatures and also resolves prior limitations on the selection of modeling ingredients. First, we calibrate separate idiosyncratic shock processes for demand and productivity using direct empirical estimates from the firm-dynamics literature that employ both prices and quantities, along with revenue and inputs. This eliminates the need to calibrate these shock process parameters to produce observed pricing moments without any connection to quantities. Second, this approach allows us to re-investigate the role for strategic complementarities in a richer structure than was used in prior studies.

Our calibrated model is able to generate real and nominal dynamics that match key features of interest across these literatures. The combination of menu costs, a Kimball demand system, and idiosyncratic shocks for demand and productivity in the model produces cost pass-through in line with previous studies along with a distribution of markups that is similar to that estimated from the data. The model also produces simulated moments consistent with untargeted moments on the size, direction, and dispersion of price changes. And when the model is extended to feature nominal expenditure shocks, it produces real effects from a nominal shock that are within the range of results from other studies of monetary non-neutrality. The two features generating this non-neutrality are the Kimball demand system and the presence of both idiosyncratic productivity and demand shocks, the latter of which dampens the selection effect in price adjustments to an aggregate shock and results in a larger real output response. By calibrating the model to directly match firm-level empirical estimates of the underlying shock processes for productivity and demand, we are able to avoid the critique of Klenow and Willis (2016), while also highlighting the importance of the joint inclusion of idiosyncratic productivity and demand shock processes in the study of models of strategic complementarities. Thus, we view our work as opening the door to future research that jointly models real (e.g. investment, employment, entry / exit), nominal (price setting) and other (e.g. markup, pass-through) decisions of firms using one unified framework.

References

Akerlof, George A and Janet L Yellen, "Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?," *The American Economic Review*, 1985, *75* (4), 708–720.

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis, "Price Setting With Strategic Complementarities as a Mean Field Game," *Econometrica*, 2023, *91* (6), 2005–2039.

— , Hervé Le Bihan, and Francesco Lippi, "The real effects of monetary shocks in sticky price models: a sufficient statistic approach," *American Economic Review*, 2016, *106* (10), 2817–2851.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings, "International Shocks, Variable Markups, and Domestic Prices," *The Review of Economic Studies*, 2019, *86* (6), 2356–2402.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare, "The Elusive Pro-competitive Effects of Trade," *The Review of Economic Studies*, 2019, *86* (1), 46–80.

Arnoud, Antoine, Fatih Guvenen, and Tatjana Kleineberg, "Benchmarking Global Optimizers," Technical Report 2022.

Aruoba, S Boragan, Andrés Fernández, Daniel Guzmán, Ernesto Pastén, and Felipe Saffie, "Pricing under Distress," 2022.

Baley, Isaac and Andrés Blanco, "Firm uncertainty cycles and the propagation of nominal shocks," *American Economic Journal: Macroeconomics*, 2019, *11* (1), 276–337.

Ball, Laurence and David Romer, "Real Rigidities and the Non-neutrality of Money," *The Review of Economic Studies*, 1990, *57* (2), 183–203.

Berger, David and Joseph Vavra, "Shocks versus Responsiveness: What Drives Time-Varying Dispersion?," *Journal of Political Economy*, 2019, *127* (5), 2104–2142.

Bergin, Paul R and Robert C Feenstra, "Staggered Price Setting, Translog Preferences, and Endogenous Persistence," *Journal of Monetary Economics*, 2000, *45* (3), 657–680.

— and — , "Pass-through of exchange rates and competition between floaters and fixers," *Journal of Money, credit and Banking*, 2009, *41*, 35–70.

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

Blanco, Andres, Corina Boar, Callum Jones, and Virgiliu Midrigan, "Nonlinear Inflation Dynamics in Menu Cost Economies," 2022.

Burstein, Ariel and Christian Hellwig, "Prices and Market Shares in a Menu Cost Model," Working Paper 13455, National Bureau of Economic Research September 2007.

Campa, Jose Manuel and Linda S Goldberg, "Exchange rate pass-through into import prices," *Review of Economics and Statistics*, 2005, *87* (4), 679–690.

Caplin, Andrew and John Leahy, "Aggregation and optimization with state-dependent pricing," *Econometrica*, 1997, pp. 601–625.

Caplin, Andrew S and Daniel F Spulber, "Menu Costs and the Neutrality of Money," *The Quarterly Journal of Economics*, 1987, *102* (4), 703–725.

Carlsson, Mikael, Alex Clymo, and Knut-Eric Joslin, "Dispersion Over The Business Cycle: Passthrough, Productivity, And Demand," Technical Report 2022.

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

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans, "Monetary Policy Shocks: What Have We Learned and to What End?," *Handbook of Macroeconomics*, 1999, *1*, 65–148.

Daruich, Diego, "The Macroeconomic Consequences of Early Childhood Development Policies," 2022.

Davis, Steven J, John Haltiwanger, Ron Jarmin, Javier Miranda, Christopher Foote, and Eva Nagypal, "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," *NBER Macroeconomics Annual*, 2006, *21*, 107–179.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

_ , Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik, "Prices, markups, and trade reform," *Econometrica*, 2016, *84* (2), 445–510.

Dotsey, Michael and Robert G King, "Implications of State-dependent Pricing for Dynamic Macroeconomic Models," *Journal of Monetary Economics*, 2005, *52* (1), 213–242.

_ , _ , and Alexander L Wolman, "State-dependent pricing and the general equilibrium dynamics of money and output," *The Quarterly Journal of Economics*, 1999, *114* (2), 655–690.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How Costly Are Markups?," *Journal of Political Economy*, 2023, *131* (7), 1619–1675.

Eichenbaum, Martin and Jonas DM Fisher, "Estimating the Frequency of Price Re-optimization in Calvo-style Models," *Journal of Monetary Economics*, 2007, *54* (7), 2032–2047.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler, "Trade and Market Selection: Evidence from Manufacturing Plants in Colombia," *Review of Economic Dynamics*, 2013, *16* (1), 135–158.

_ , _ , and Nicolas Urdaneta, "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand, and Wedges," *The Review of Economic Studies*, 02 2023, *91* (1), 259–300.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Fujiwara, Ippei and Kiminori Matsuyama, "Competition and the Phillips Curve," Technical Report 2022.

Gertler, Mark and John Leahy, "A Phillips Curve with an Ss Foundation," *Journal of Political Economy*, 2008, *116* (3), 533–572.

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajšek, "Inflation dynamics during the financial crisis," *American Economic Review*, 2017, *107* (3), 785–823.

Golosov, Mikhail and Robert E. Jr. Lucas, "Menu Costs and Phillips Curves," *Journal of Political Economy*, 2007, *115* (2), 171–199.

Gopinath, Gita and Oleg Itskhoki, "Frequency of Price Adjustment and Pass-through," *The Quarterly Journal of Economics*, 2010, *125* (2), 675–727.

_ , _ , and Roberto Rigobon, "Currency choice and exchange rate pass-through," *American Economic Review*, 2010, *100* (1), 304–336.

Harding, Martín, Jesper Lindé, and Mathias Trabandt, "Resolving the Missing Deflation Puzzle," *Journal of Monetary Economics*, 2022, *126*, 15–34.

Harding, Martín, Jesper Lindé, and Mathias Trabandt, "Understanding post-COVID inflation dynamics," *Journal of Monetary Economics*, 2023, *140*, S101–S118. Inflation: Drivers and Dynamics 2022.

Kim, Ryan, "The effect of the credit crunch on output price dynamics: The corporate inventory and liquidity management channel," *The Quarterly Journal of Economics*, 2021, *136* (1), 563–619.

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

Klenow, Peter J and Jonathan L Willis, "Real Rigidities and Nominal Price Changes," *Economica*, 2016, *83* (331), 443–472.

_ and Oleksiy Kryvtsov, "State-dependent or time-dependent pricing: Does it matter for recent US inflation?," *The Quarterly Journal of Economics*, 2008, *123* (3), 863–904.

Krusell, Per and Anthony A Smith Jr, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Mankiw, N Gregory, "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *The Quarterly Journal of Economics*, 1985, *100* (2), 529–537.

Midrigan, Virgiliu, "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *Econometrica*, 2011, *79* (4), 1139–1180.

Mongey, Simon, "Market Structure and Monetary Non-neutrality," Technical Report, National Bureau of Economic Research 2021.

Mrázová, Monika and J Peter Neary, "Not so Demanding: Demand Structure and Firm Behavior," *American Economic Review*, 2017, *107* (12), 3835–3874.

Nakamura, Emi and Jón Steinsson, "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

_ and _ , "Monetary Non-neutrality in a Multisector Menu Cost Model," *The Quarterly Journal of Economics*, 2010, *125* (3), 961–1013.

Ramey, Valerie A, "Macroeconomic Shocks and their Propagation," *Handbook of Macroeconomics*, 2016, *2*, 71–162.

Renkin, Tobias and Gabriel Züllig, "Credit Supply Shocks and Prices: Evidence from Danish Firms," *American Economic Journal: Macroeconomics*, 2024.

Rodriguez-Lopez, Jose Antonio, "Prices and Exchange Rates: A Theory of Disconnect," *The Review of Economic Studies*, 2011, *78* (3), 1135–1177.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 2007, *97* (3), 586–606.

Sobol, Ilya Meerovich, "On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals," *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 1967, *7* (4), 784–802.

Vavra, Joseph, "Inflation Dynamics and Time-varying Volatility: New Evidence and an Ss Interpretation," *The Quarterly Journal of Economics*, 2014, *129* (1), 215–258.

Internet Appendix (For Online Publication)

A Derivations

A.1 Pass-through

Consider the static optimization problem of an intermediate firm without any pricing frictions. The nominal profit of an intermediate firm is

$$\pi_i = \left(\frac{p_i}{P} - \frac{W}{P z_i}\right) \frac{Y}{n_i} \frac{1}{1+\psi} \left[\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi\right] \tag{A-1}$$

The first-order condition with respect to p_i is given by

$$\begin{aligned}
\frac{1}{p_i - \frac{W}{z_i}} + \frac{\frac{\omega(1+\psi)}{1-\omega}\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}-1} \frac{1}{\lambda n_i P}}{\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi} &= 0 && \text{(A-2)} \\
\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}} \left[1 - \left(\frac{W}{z_i} - p_i\right)\left(\frac{\omega\left(1+\psi\right)}{1-\omega}\frac{1}{p_i}\right)\right] &= -\psi && \text{(A-3)}
\end{aligned}$$

Log-linearizing the first-order condition around a symmetric steady state yield

$$\left(\hat{p}_i - \hat{\lambda} - \hat{P} - \hat{n}_i\right) + \frac{1 + \left(\frac{W}{\bar{z}_i} - \bar{p}_i\right)\left(\frac{1}{\bar{p}_i}\right)}{1 - \left(\frac{W}{\bar{z}_i} - \bar{p}_i\right)\left(\frac{\omega(1+\psi)}{1-\omega}\frac{1}{\bar{p}_i}\right)}\hat{p}_i + \frac{\frac{W}{\bar{z}_i}\left(\frac{1}{\bar{p}_i}\right)}{1 - \left(\frac{W}{\bar{z}_i} - \bar{p}_i\right)\left(\frac{\omega(1+\psi)}{1-\omega}\frac{1}{\bar{p}_i}\right)}\hat{z}_i = 0 \tag{A-4}$$

where hat variables denote log-deviations from the steady state and bar variables denote the steady state values.

Note that in a symmetric steady state, all firms are identical, have the same market share, and set the same price. Specifically, the optimal price is a fixed markup over cost $\bar{p}_i = \omega\frac{W}{\bar{z}_i}$. Substituting this into the log-linearized first-order condition gives

$$\hat{p}_i = \frac{\omega\psi}{\omega\psi - 1}\left(\hat{\lambda} + \hat{P} + \hat{n}_i\right) + \frac{1}{\omega\psi - 1}\hat{z}_i \tag{A-5}$$

A.2 Transmission of Nominal Shock under Calvo Pricing and CES

Consider a simple model where CES demand and Calvo pricing. Each period, a random fraction α of firms can adjust their prices freely while the other fraction $1-\alpha$ cannot change their prices. For simplicity, assume that there are no aggregate risks and the economy is initially in a symmetric equilibrium where all firms set nominal prices to $\bar{p}$ and nominal expenditure $S = PC$ is equal to $\bar{S}$.

Under the Calvo setup, the aggregate price index P_t can be written as a combination of the optimal reset price X_t, which is the price chosen by firms that are adjusting, and the lagged price index, which summarizes the prices of non-adjusters, as follows

$$P_t = \left[\alpha X_t^{1-\theta} + (1-\alpha) P_{t-1}^{1-\theta}\right]^{\frac{1}{1-\theta}} \tag{A-6}$$

Log-linearizing around the initial steady state where $P_t = X_t = \bar{p}$ and using hatted variables to denote log-deviations from the steady state yields

$$\hat{P}_t^{1-\theta} = \alpha \hat{X}_t^{1-\theta} + (1-\alpha) \hat{P}_{t-1}^{1-\theta} \tag{A-7}$$

Suppose that in period $t = 1$, there is an unanticipated permanent shock $\mu > 0$ to nominal expenditure shock S, such that $\hat{S}_1 = \mu$. Because the nominal wage is proportional to nominal expenditure, firms with the opportunity to adjust will respond to the shock by increasing their prices by μ as the optimal markup is a constant $\frac{\theta}{\theta-1}$ over the marginal cost. In other words, the optimal reset price is $\hat{X} = \mu$.

As such, the aggregate price in period one is given by

$$\hat{P}_1 = \alpha\mu \tag{A-8}$$

Iterating forward, the aggregate price in period h can be written as

$$\hat{P}_h = \left(\sum_{i=1}^{h} (1-\alpha)^{i-1}\right) \alpha\mu \tag{A-9}$$

The response of real output is therefore given by

$$\begin{aligned}\hat{C}_h &= \hat{S} - \hat{P}_h && \text{(A-10)}\\ &= \mu - \left(\sum_{i=1}^{h} (1-\alpha)^{i-1}\right)\alpha\mu && \text{(A-11)}\\ &= (1-\alpha)^h \mu && \text{(A-12)}\end{aligned}$$

which implies that the output response as a fraction of the shock is $(1-\alpha)^h$ at period h.

B Model Solution

B.1 Rewriting the Problem

Note that firms need to observe $\mathcal{S}$ and know its law of motion in order to solve their problem. These relevant aggregate variables in $\mathcal{S}$ can be summarized by a single aggregate state variable P_{t-1}/S_t.

Because money supply $S_t = P_t C_t$ exhibits positive growth on average, nominal prices are also ever-increasing. We normalize all nominal variables by S_t to ensure that the state variables are stationary. As such, we can rewrite the firm's profit function

$$\pi\left(\frac{p_t^i}{S_t}; n_t^i, z_t^i, \frac{P_t}{S_t}, \lambda_t\right) = \left(\frac{p_t^i/S_t}{P_t/S_t} - \frac{1}{z_t^i P_t/S_t}\right)\frac{(P_t/S_t)^{-1}}{n_t^i(1+\psi)}\left[\left(\frac{p_t^i/S_t}{\lambda_t n_t^i (P_t/S_t)}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi\right] \tag{A-13}$$

where we also use $Y_t = C_t = \frac{P_t}{S_t}$ from goods market clearing and $W_t = S_t$ from the household's intratemporal optimality condition.

P_t/S_t and λ_t are the collective results of the pricing decisions of all firms. To know these, firms must know the entire firm distribution over the idiosyncratic states which is an infinite-dimensional object. Following the application of the Krusell and Smith (1998) algorithm in menu-cost models (Nakamura and Steinsson, 2010; Midrigan, 2011; Vavra, 2014), we conjecture the following forecasting rules for P_t/S_t and λ_t

$$\log\left(\frac{P_t}{S_t}\right) = F\left(\frac{P_{t-1}}{S_t}\right) = \alpha_0 + \alpha_1 \log\left(\frac{P_{t-1}}{S_t}\right) \tag{A-14}$$

$$\log(\lambda_t) = G\left(\frac{P_{t-1}}{S_t}\right) = \beta_0 + \beta_1 \log\left(\frac{P_{t-1}}{S_t}\right) \tag{A-15}$$

Using these, the law of motion of the aggregate variable P_{t-1}/S_t is also given by

$$\log\left(\frac{P_t}{S_{t+1}}\right) = \log\left(\frac{P_t}{S_t}\right) + \log\left(\frac{S_t}{S_{t+1}}\right) \tag{A-16}$$

$$= \alpha_0 + \alpha_1 \log\left(\frac{P_{t-1}}{S_t}\right) - (\mu + \sigma_S \epsilon_{t+1}) \tag{A-17}$$

Now, we rewrite the intermediate producers' problem using these state variables. At the beginning of a period, each intermediate producer starts off with a price p^i_{t-1}/S_t, idiosyncratic demand n^i_t, and idiosyncratic productivity z^i_t. They also observe P_{t-1}/S_t and forecast P_t/S_t and λ_t using the aforementioned laws of motion. The value of not adjusting is

$$V_N\left(\frac{p^i_{t-1}}{S_t}; n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right) = \pi\left(\frac{p^i_{t-1}}{S_t}; n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right) + \mathbb{E}_t\left[\Xi_{t,t+1} \cdot V\left(\frac{p^i_t}{S_{t+1}}, n^i_{t+1}, z^i_{t+1}, \frac{P_t}{S_{t+1}}\right)\right] \tag{A-18}$$

which is equal to the flow profit evaluated at last period's price adjusted for inflation plus a continuation value.

If the firm chooses to adjust its price, it pays the fixed price adjustment cost and chooses p^i_t to maximize the sum of the current flow profit and the present discounted value of future profit

$$V_A\left(n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right) = -f^i_t \frac{P_t}{S_t} + \max_{p^i_t}\left\{\pi\left(\frac{p^i_t}{S_t}; n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right) + \mathbb{E}_t\left[\Xi_{t,t+1} \cdot V\left(\frac{p^i_t}{S_{t+1}}, n^i_{t+1}, z^i_{t+1}, \frac{P_t}{S_{t+1}}\right)\right]\right\} \tag{A-19}$$

A firm chooses to adjust its price if and only if the value of doing so exceeds the value of inaction. Therefore, the value function of the firm is

$$V\left(\frac{p^i_{t-1}}{S_t}; n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right) = \max\left[V_N\left(\frac{p^i_{t-1}}{S_t}; n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right), V_A\left(n^i_t, z^i_t, \frac{P_{t-1}}{S_t}\right)\right] \tag{A-20}$$

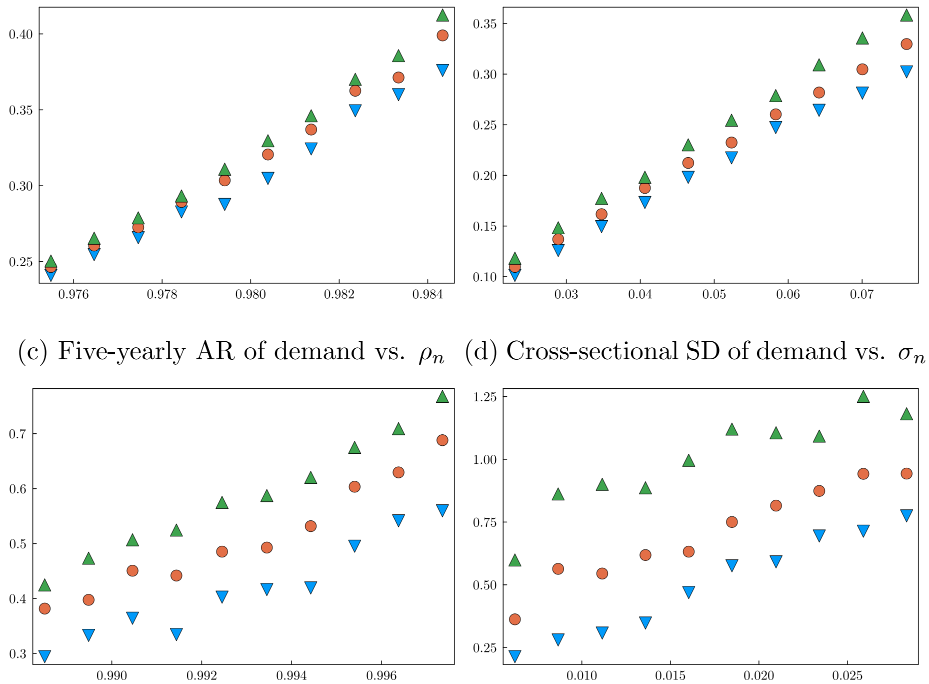
B.2 Computational Strategy

A sketch of the computation algorithm is as follows. We first make guesses of the coefficients $(\alpha^0_0, \alpha^0_1, \beta^0_0, \beta^0_1)$ in the forecasting equations F and G. Given the guesses, use value function iteration to solve for the intermediate-good producers' value functions as well as the optimal pricing rules. Using pricing rules, simulate the model for a large number of periods and obtain simulated sequences of $\frac{P_t}{S_t}$, λ_t, and $\frac{P_{t-1}}{S_t}$. Estimate the regressions F and G with model simulated data and obtain estimated coefficients $(\alpha^1_0, \alpha^1_1, \beta^1_0, \beta^1_1)$ which is then used to

Figure A-1: Identification of Internally-Calibrated Parameters

(a) Five-yearly AR of TFPQ vs. ρ_z (b) Cross-sectional SD of TFPQ vs. σ_z

(c) Five-yearly AR of demand vs. ρ_n (d) Cross-sectional SD of demand vs. σ_n



Note: For each decile of a given parameter plotted on the horizontal axis, the red dot shows the median of the moment that is assigned to the parameter. The blue down-pointing triangles and green up-pointing triangles show the 25^{th} and 75^{th} percentiles respectively.

update the initial guesses. Repeat this process until the coefficient guesses and sufficiently close to the estimated coefficients from the linear regressions. In doing so, we find that the conjectured law of motion approximates the true law of motion from the model simulation well, as the regression yields an R^2 larger than 0.99.

C Calibration Details

Figure A-1 exhibit the link between the parameters governing the idiosyncratic productivity and demand processes and the corresponding empirical moments. The parameters (ρ_z, σ_z) are strongly correlated with the five-yearly autocorrelation and cross-sectional distribution of firm productivity, whereas other parameters play a minimal role as can be seen in the tight vertical variation in the scatter plots. For (ρ_n, σ_n), we observe a similar relationship, but there is noticeably more noise in the cross-sectional standard deviation of demand. This is mainly because at a given decile of σ_n, the remaining parameters, including ρ_n are randomly drawn. Because the value of ρ_n is generally very close to one, the resulting cross-sectional dispersion of demand is very sensitive to the value of ρ_n in addition to σ_n.

D Alternative Models and Calibrations

D.1 Models with CES Demand

We present two alternative calibration of the model with CES demand, neither of which feature a demand shock. In CES I, we use the standard, or agnostic, approach in the literature and calibrate (f, ρ_z, σ_z) to match three pricing moments: frequency of price changes, fraction of positive changes and the average size of price changes. In CES II, we take the first step in trying to be consistent with the firm-dynamics facts from Foster et al. (2008) and calibrate (ρ_z, σ_z) to match two moments: the five-yearly autocorrelation and the variance of TFP . We still calibrate f to match the frequency of price changes. In both versions we set $\omega = 1.33$ to obtain a desired markup of 33%, which is the median markup in our data, and, naturally $\psi = 0$ so that CES aggregation is obtained.

Results are presented in Table A-1. Starting with CES I, which uses (ρ_z, σ_z, f) to match the first three pricing moments in the first panel, we see that it matches those moments very well. However, it completely misses the firm-dynamics moments. The size and persistence of the z process, which is used to match pricing moments, generates very small and nearly transitory movements in idiosyncratic productivity, as can be seen in the first two rows of the firm-dynamics moments. This version is, by definition, unable to say anything about the remaining firm-dynamics moments due to the absence of demand shocks. Turning to CES II, where we now give up on matching pricing moments, except for the frequency of price changes, and use (ρ_z, σ_z) to match the first two firm-dynamics moments, the calibration is successful in the sense that the targeted moments are matched, including the dynamics of idiosyncratic productivity. However, this version misses the two untargeted pricing moments, especially the last one completely. In an effort to match the more volatile and persistent idiosyncratic productivity process, the model generates price changes that are about twice as large on average than the data. Moreover, due to the absence of demand shocks, the four remaining firm-dynamics moments are also not matched.

Figure A-2 plots the hazard function generated by our model across the four specifications. In CES I, which is the approach commonly taken by the literature, the hazard function is increasing over the first few months and flattens out afterwards. This is consistent with the baseline calibration of Nakamura and Steinsson (2008). The other three

Table A-1: Two Alternative Versions with CES Demand

Moment		Data	Baseline	CES I	CES II
Frequency of price changes		0.11	**0.12**	**0.11**	**0.11**
Fraction of price increases		0.65	0.58	**0.64**	0.61
Size of price changes		0.08	0.07	**0.08**	0.15
5-yearly autocorr of z_t^i		0.32	**0.32**	0.00	**0.32**
Cross-sectional SD of z_t^i		0.26	**0.25**	0.03	**0.26**
5-yearly autocorr of n_t^i		0.62	**0.62**	0.01	0.00
Cross-sectional SD of n_t^i		1.16	**1.05**	0.01	0.04
Corr b/w TFPR and TFPQ		0.75	**0.74**	0.63	0.28
Corr b/w price and TFPQ		–0.54	**–0.57**	–0.85	–0.99
Parameter	**Description**				
ψ	Super-elasticity		**–1.27**	0	0
ω	Elasticity of Substitution		**1.29**	1.33	1.33
ρ_z	Persistence of z_t^i		**0.66**	**0.98**	
σ_z	Standard deviation of z_t^i		**0.98**	**0.04**	**0.05**
ρ_n	Persistence of n_t^i		**0.997**	–	–
σ_n	Standard deviation of n_t^i		**0.02**	–	–
f	Menu cost		**0.03**	**0.01**	**0.06**

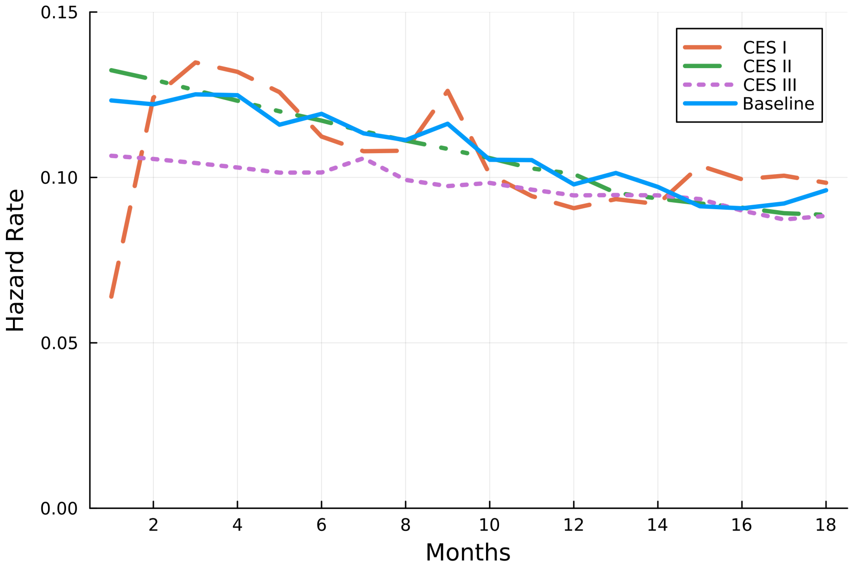
Note: The top panel of this table compares the targeted moments and model-implied moments for the four model specifications, where the bolded numbers highlight moments that are targeted in the calibration. The bottom panel shows the parameter values for each calibration.

model specifications, where idiosyncratic processes are calibrated to firm dynamics evidence, instead exhibit a downward-sloping hazard. (Here CES III is the version with CES and idiosyncratic demand shocks presented in the main text.) This shows that calibrations that imply downward-sloping hazards are not necessarily unrealistic. In fact, a simple menu cost model with Kimball aggregation, calibrated to firm dynamics estimates generates a price change hazard that is much more consistent with the empirical counterpart, while remaining consistent with micro pricing facts.

Figure A-3 plots the kernel density of markup distribution for our CES versions along with its data counterpart. For all CES versions the markup varies very little and does so symmetrically around $\omega = 1.33$, or a net markup of 33%. This is completely at odds with the distribution we obtain from the data, which has a mode just above 0% with a very wide right tail reaching a level of 200%, though markups of as low as –50% are also observed.

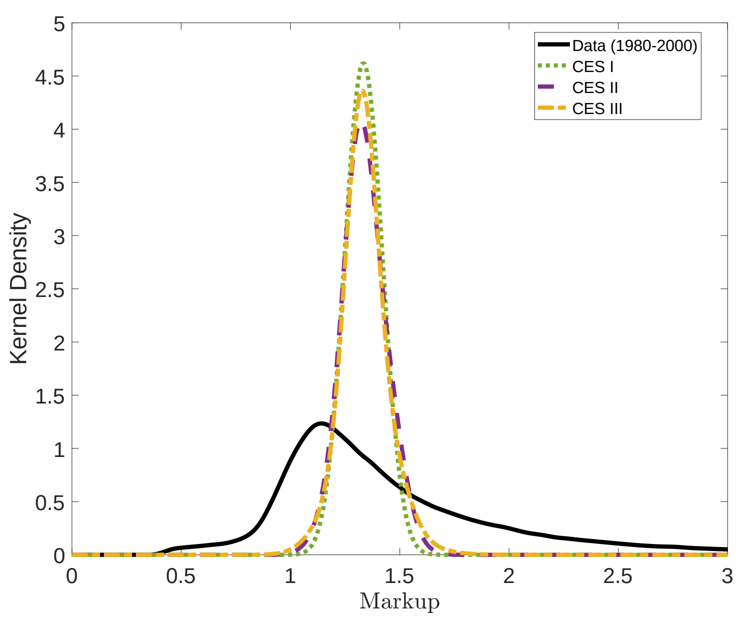
Figure A-4 plots the impulse response of real output expressed as a fraction of the size of the shock for the baseline model as well as the three CES versions we introduced earlier.

Figure A-2: Hazard Function of Price Change



Note: This figure plots the pricing hazard for the four model calibrations over the first 18 months.

Figure A-3: Gross Markup Distribution: Model vs. Data (CES Versions)



This figure plot the kernel density of the empirical markup distribution from publicly traded firms in the U.S. as well as the kernel density of the markup distribution in the ergodic distribution of the three versions of the model. Both kernel densities are computed using the optimal bandwidth for normal densities.

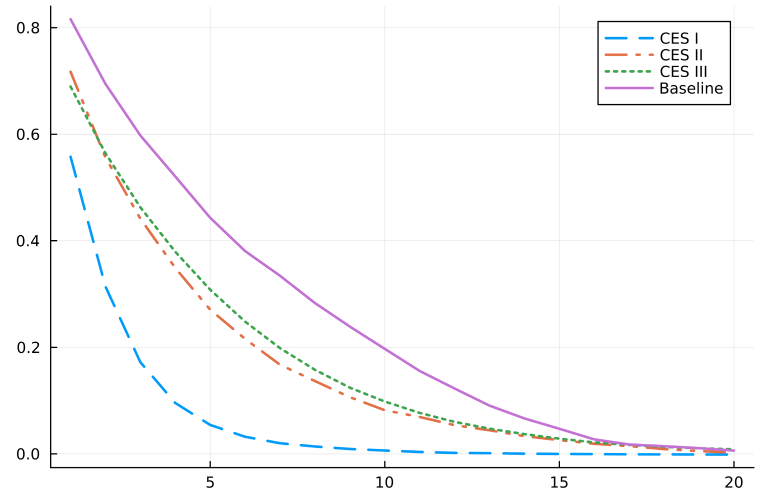


Figure A-4: Impulse Response of Real Output to a Nominal Expenditure Shock

Note: This figure plots the impulse response of real output expressed as a fraction of the nominal expenditure shock on the vertical axis and periods elapsed since the shock on the horizontal axis.

Table A-2: Measures of Monetary Non-neutrality

Moment	CES I	CES II	CES III	Baseline
SD(C)	0.22%	0.39%	0.44%	0.52%
Impact	0.56	0.69	0.69	0.82
Half-life	1.25	2.90	3.46	4.50
CIR	0.11	0.28	0.30	0.42

Note: This table displays four measures of monetary non-neutrality for the four model calibrations.

Table A-2 contains the four statistics that summarize the degree of monetary non-neutrality we discussed above. CES I (blue dashed line), which was calibrated to pricing moments, deliver a peak response of 0.56, which is substantially lower than the Calvo response. The response is also fairly short-lived with a half-life of 1.25 months. The resulting cumulative impulse response is 0.11, which is close to the number reported in Alvarez et al. (2016) for a Golosov-Lucas type menu cost model. CES II (orange dashed dotted line) which calibrates the model to the productivity process of Foster et al. (2008) produces more non-neutrality with a peak response of 0.69. It is also somewhat more long-lived with a half-life of 2.9 months and has a larger CIR of 0.28 compared to CES I. Neither of these versions delivers a reasonable level of non-neutrality, which is the key result of Golosov and Lucas (2007).

D.2 Translog Demand System

D.2.1 Link Between Translog and Kimball Demand Systems

This appendix establishes the link between a translog demand system and the Kimball demand system. The derivations follow Mrázová and Neary (2017) closely, who define a demand system as a locus of (ϵ, ρ) where

$$\epsilon(x) \equiv -\frac{p(x)}{xp'(x)} > 0 \text{ and } \rho(x) \equiv -\frac{xp''(x)}{p'(x)} \tag{A-21}$$

where $p(x)$ is the inverse demand function satisfying $p'(x) < 0$. In this notation ϵ is the minus of the elasticity of demand and ρ is a measure of demand convexity. In their paper they show how many popular demand systems can be expressed as a mapping between ρ and ϵ, under the assumptions of a steady state and symmetric firms.

For our purposes, they show that for translog demand, the mapping is given by

$$\rho^T(\epsilon) = \frac{3\epsilon - 1}{\epsilon^2} \tag{A-22}$$

For Kimball demand, they define the super-elasticity as $S = b\epsilon$ for some $b > 0$ and the mapping is given by

$$\rho^K(\epsilon, b) = \frac{(1-b)\epsilon + 1}{\epsilon} \tag{A-23}$$

Our goal is to represent the translog mapping from ϵ to ρ as one that holds in Kimball, where we pick a particular b for every ϵ by setting $\rho^T(\epsilon) = \rho^K(\epsilon, b)$ and solving for the b. Doing so yields

$$b(\epsilon) = \left(\frac{\epsilon - 1}{\epsilon}\right)^2 \tag{A-24}$$

So far we used the notation of Mrázová and Neary (2017). To convert the restriction in (A-24) to our notation, based on our derivations of the elasticity of demand and super-elasticity in footnote 5 we get

$$\epsilon = \frac{\omega}{\omega - 1} \text{ and } \psi = -b \tag{A-25}$$

Using the relationships in (A-25) in (A-24) we obtain

$$\psi = -\frac{1}{\omega^2} \tag{A-26}$$

which shows that given an ω we can a ψ such that the Kimball system with (ψ, ω) corresponds to a translog system with an elasticity of demand of $\frac{\omega}{1-\omega}$ under the assumptions in Mrázová and Neary (2017) listed above. (Note that the ϵ in the notation of Mrázová and Neary (2017) was minus the elasticity of demand.)

D.2.2 Cost Pass-through with Translog and Kimball Demands

Mrázová and Neary (2017) show that for a general demand system, the cost pass-through to demand can be obtained as

$$\frac{d\log p}{d\log mc} = \frac{\epsilon - 1}{\epsilon}\frac{1}{2-\rho} \tag{A-27}$$

where mc denotes the marginal cost of the firm and ϵ and ρ are specific to the particular demand system used. Given the representation for translog demand in (A-22), this simplifies to

$$\frac{d\log p}{d\log mc} = \frac{\epsilon}{2\epsilon - 1} \tag{A-28}$$

which is bounded between 0.5 and 1 for $\epsilon \in (1, \infty)$.

As we derived in Appendix A.1, our Kimball specification yields a cost pass-through of $\frac{1}{1-\psi\omega}$ in the symmetric steady state. Using (A-23) and (A-25), the formula in (A-27) simplifies exactly to the same formula.