FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Reviving Micro Real Rigidities: The Importance of Demand Shocks

S. Borağan Aruoba, Eugene Oue, Felipe Saffie, and Jonathan L. Willis

Working Paper 2023-3b
April 2023 (Revised January 2025)

Abstract: We revisit the role of micro real rigidities as a driver of monetary non-neutrality, using a simple menu-cost model featuring non-constant elasticity of demand with both idiosyncratic productivity and demand shocks. The model is calibrated to match firm-level productivity and demand processes estimated from US data. Despite its simplicity, the calibrated model overturns prior negative findings in the literature on micro real rigidities and generates sizeable monetary non-neutrality comparable to alternative frameworks. Additionally, the model reproduces untargeted pricing dynamics and a markup distribution consistent with US microdata. As a result, this framework effortlessly unifies pricing, markup behavior, and firm dynamics. The key to reconciling firm-level and pricing dynamics lies in the interaction between non-constant demand elasticity and idiosyncratic demand shocks.

JEL classification: E30, E52, L11

Key words: menu costs, strategic complementarities, demand shocks, sticky prices, monetary non-neutrality

https://doi.org/10.29338/wp2023-03

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. Any remaining errors are the authors' responsibility. The authors are grateful for the invaluable feedback they received on an earlier draft from Jón Steinsson. The authors would like to thank Isaac Baley, Mark Bils, Laura Castillo-Martínez (discussant), Basile Grassi, Susanto Basu, John Haltiwanger, Pete Klenow, Francesco Lippi, Virgiliu Midrigan, Luminita Stevens, Javier Turén, Rosen Valchev, and seminar participants at the NBER 2024 SI (ME), UC Berkeley, Boston College, the University of Maryland, the Federal Reserve Board of Governors, the Federal Reserve Banks of Atlanta and Richmond, ITAM, IMF, Banque de France, BIS, Bocconi, EIEF, CREI, 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.

S. Borağan Aruoba is with the University of Maryland and NBER. Eugene Oue is with Hong Kong Polytechnic University. Felipe Saffie is with the University of Virginia's Darden School of Business. Please address questions regarding content to Jonathan L. Willis, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA 30309, jon.willis@atl.frb.org.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at www.frbatlanta.org. Click "Publications" and then "Working Papers." To receive e-mail notifications about new papers, use frbatlanta.org/forms/subscribe.

1 Introduction

Modeling the response of output and prices to monetary policy shocks has long been central to macroeconomic research. Empirical studies consistently show that such shocks have persistent and substantial effects on real output (Christiano et al., 1999; Ramey, 2016). The non-neutrality of monetary policy, where real activity responds to nominal shocks, stems from incomplete price adjustments to nominal disturbances. When prices do not fully adjust, quantities must change, driving fluctuations in real activity. Traditionally, this sluggish price adjustment, or nominal rigidity, has been attributed to frictions such as menu costs or informational barriers that prevent firms from updating prices.

However, Ball and Romer (1990) argued that nominal frictions alone are insufficient to generate significant monetary non-neutrality. They emphasized that real rigidities, combined with nominal frictions, are necessary to generate large real effects from nominal shocks. Real rigidities arise from structural features of the economy, such as preferences, technology, and market competition, that limit firms' price adjustments even when nominal frictions are absent. Real rigidities are often manifested in strategic complementarity in pricing, where a firm's optimal pricing strategy depends on the pricing behavior of its competitors.

Following the classification of Nakamura and Steinsson (2010) and Gopinath and Itskhoki (2011), real rigidities can be grouped into two categories: micro (ω-type) and macro (Ω-type).[1] While micro-level real rigidities, often modeled using non-CES (constant elasticity of substitution) demand systems, are widely employed in fields such as international trade and industrial organization, their adoption in quantitative monetary models of pricing has been limited.[2] This limited use, as noted by Nakamura and Steinsson (2010), arises partly from the fact that micro real rigidities cannot match observed micro pricing facts under plausible parameter values. For example, Klenow and Willis (2016) argue that the idiosyncratic productivity shocks required to fit price-setting data under the non-constant elasticity demand system of Kimball (1995) are implausibly large and inconsistent with firm

[1]Examples of micro real rigidities include non-constant marginal costs (e.g., decreasing returns to scale, as in Burstein and Hellwig (2007)), strategic complementarities between large firms (Atkeson and Burstein, 2008; Mongey, 2021), and non-constant elasticities and markups (Kimball, 1995). Examples of macro real rigidities include production networks (Nakamura and Steinsson, 2010; Basu, 1995), segmented labor markets (Gertler and Leahy, 2008; Woodford, 2003) and real wage rigidity (Blanchard and Galí, 2007).. Section 2 develops a simple theoretical framework that clearly defines micro and macro real rigidities.

[2]Non-CES demand systems are crucial for explaining markup variability (Edmond et al., 2023; Arkolakis et al., 2019), exchange rate pass-through (Gopinath and Itskhoki, 2010; Amiti et al., 2019; Berger and Vavra, 2019), and inflation dynamics and optimal monetary policy (Harding et al., 2022, 2023; Fujiwara and Matsuyama, 2022).

dynamics.[3]

In this paper, we revisit the role of micro real rigidities in monetary economics using a menu cost model with a non-CES demand system à la Kimball (1995). In particular, we incorporate idiosyncratic demand shocks alongside standard idiosyncratic productivity shocks, which we argue may address the long-standing critique of micro real rigidities in the literature.[4] Our objective is to evaluate whether this quantitative model can reconcile micro-level evidence on both pricing and firm dynamics while also generating substantial monetary non-neutrality.

While quantitative menu cost models typically target pricing moments to discipline idiosyncratic shocks, we deviate from this approach by calibrating our model to firm-level evidence from Foster et al. (2008), ensuring consistency with the firm dynamics literature while leaving several pricing moments untargeted. Specifically, we calibrate the model to match key moments from Foster et al. (2008), including the autocorrelation and cross-sectional standard deviations of firm-level productivity and demand. Additionally, we target moments such as the correlation between revenue-based TFP (TFPR) and quantity-based TFP (TFPQ), as well as the correlation between price and TFPQ. Our calibrated model successfully replicates several untargeted micro-level pricing moments, namely the size and direction of price changes, and the dispersion of non-zero price adjustments. Furthermore, the model generates a downward-sloping pricing hazard function consistent with empirical data, and effortlessly delivers empirically-consistent distributions of markups and firm-level growth.

We demonstrate that our model generates substantial monetary non-neutrality, with cumulative output responses approximately four times larger than those produced by a comparable CES-based model—consistent with the findings of Golosov and Lucas (2007). Two key features drive these results: the Kimball demand system, which induces strategic pricing complementarities, and the inclusion of both idiosyncratic productivity and demand shocks, which mitigate the selection effect in price adjustments to aggregate shocks. The lion's share of the effect stems from pricing complementarities. Notably, our micro-calibrated demand system more than doubles monetary non-neutrality even in a Calvo (1983) model, where the selection effect is fully muted.

The key innovation of our framework lies in the inclusion of idiosyncratic demand shocks.

[3] Burstein and Hellwig (2007) also show that a calibrated model with decreasing returns to scale is unable to deliver significant non-neutrality.

[4] We focus on the Kimball (1995) demand system due to its widespread use across various fields of economics, its flexibility, and its central role in the critique by Klenow and Willis (2016). Additionally, we show that alternative sources of micro real rigidities, such as non-constant marginal costs, fail to align with evidence on firm dynamics.

Unlike models with CES demand systems, where demand shocks fail to generate price adjustments, the Kimball demand system allows firms to pass through portions of both productivity and demand shocks to prices. This demand system introduces a trade-off between strategic complementarities and the pass-through of productivity shocks. Previous studies that focused solely on productivity shocks found that matching price-adjustment moments required unrealistically large shocks (Klenow and Willis, 2016). By incorporating demand shocks, we achieve a realistic calibration that remains consistent with empirical evidence.

In conclusion, this paper demonstrates that a menu cost model with micro real rigidities can simultaneously align with micro-level pricing and firm dynamics data, and produce significant monetary non-neutrality. Our approach paves the way for future research that integrates real, nominal, and other firm-level decisions into a unified framework.[5]

The paper is organized as follows: Section 2 develops a simple theoretical framework that highlights our key contributions. Section 3 introduces the quantitative menu cost model with idiosyncratic shocks and a Kimball demand system. Section 4 discusses the calibration strategy and untargeted moments. Section 5 examines the implications for monetary non-neutrality. Finally, Section 6 concludes.

2 Real Rigidities and Demand Shocks

This section describes the main elements of our model and their interactions. The first key feature of our model is a real rigidity, which causes a reduction in the responsiveness of a firm's price to an nominal aggregate shock. This is the primary mechanism for delivering monetary non-neutrality. For the real rigidity specification in our model, we employ the demand system proposed by Kimball (1995). The second key feature of our model is the introduction of an idiosyncratic demand process in addition to the more standard idiosyncratic productivity process. We then demonstrate the interaction between the Kimball demand system and idiosyncratic demand shocks, which is critical for reconciling firm and pricing dynamics.

[5] Aruoba et al. (2022) underscore the flexibility of the Kimball framework using Chilean microdata, incorporating additional features such as leptokurtic shocks and news shocks. While distinct from our approach, their findings highlight the framework's value in explaining pricing behavior.

2.1 Micro and Macro Real Rigidities

Real rigidities are mechanisms that cause firms to refrain from fully adjusting their relative prices in response to changes in aggregate conditions, even when there are no nominal pricing frictions. As shown by Ball and Romer (1990), quantitative monetary models need to combine both nominal and real rigidities to ensure that monetary shocks have meaningful real aggregate effects. This section presents a simple price-setting problem to examine two key sources of real rigidities that lead to increased monetary non-neutrality.[6] Consider a static, frictionless price-setting problem for a firm with the profit function

$$\Pi\left(\frac{p_i}{P}, \frac{S}{P}, A_i\right), \tag{1}$$

where $\frac{p_i}{P}$ is the firm's relative price, $\frac{S}{P} \equiv Y$ is the real money supply—equal to real aggregate demand in equilibrium—and A_i represents a collection of additional idiosyncratic variables affecting the profit function. In the absence of nominal pricing frictions, the firm optimally chooses $\frac{p_i^*}{P}$ to maximize profit by satisfying the first-order condition

$$\Pi_1\left(\frac{p_i^*}{P}, \frac{S}{P}, A_i\right) = 0. \tag{2}$$

We define the degree of real rigidity by the responsiveness of the firm's desired relative price $\frac{p_i^*}{P}$ to a change in real aggregate demand $\frac{S}{P}$

$$\phi \equiv \frac{\partial\left(\frac{p_i^*}{P}\right)}{\partial\left(\frac{S}{P}\right)} = -\frac{\Pi_{12}\left(\frac{p_i^*}{P}, \frac{S}{P}, A_i\right)}{\Pi_{11}\left(\frac{p_i^*}{P}, \frac{S}{P}, A_i\right)}. \tag{3}$$

Assuming $\Pi_{12} > 0$, which ensures a stable equilibrium, and noting that profit maximization requires $\Pi_{11} < 0$, the firm's optimal relative price is increasing in real aggregate demand ($\phi > 0$).[7]

Real rigidities are stronger when ϕ is small. This occurs when Π_{12} is small or $|\Pi_{11}|$ is large. Intuitively, when Π_{12} is close to zero, changes in real aggregate demand have a small

[6]The expositions of Ball and Romer (1990) and Nakamura and Steinsson (2010) serve as the primary sources for this section.

[7]If $\Pi_{12} < 0$, an increase in S would trigger a chain of events where prices keep decreasing. If $\Pi_{12} = 0$, prices do not respond at all to changes in S so nominal shocks have permanent effects on real output. $\Pi_{12} > 0$ ensures that S/P is convergent following a nominal shock.

effect on Π_1. As a result, a firm's desired price, which is determined by $\Pi_1 = 0$, is less sensitive to aggregate demand shocks. Alternatively, ϕ is smaller when the profit function is more concave, or when $|\Pi_{11}|$ is large. This increases the slope of Π_1 so that firms' prices respond less to a given aggregate demand shock, as it is more costly for firms' relative prices to deviate from their competitors. This in turn also leads to a muted response of prices to a change in aggregate demand.

Real rigidities are closely linked to strategic interactions between firms' pricing decisions through Π_{11}. To illustrate this, we can define the strategic interaction between firms' price-setting decisions by

$$\zeta \equiv \frac{\partial p_i^*}{\partial P} = \frac{p_i^*}{P} + \frac{\Pi_{12}}{\Pi_{11}}\frac{S}{P}. \tag{4}$$

When $\zeta > 0$, firms' pricing decisions are said to be strategic complements, as their optimal prices comove positively with the aggregate price index. Conversely, firms' pricing decisions are strategic substitutes when $\zeta < 0$. In a symmetric equilibrium where all firms choose the same price, and by normalizing $Y \equiv S/P$ to 1, equations (3) and (4) yield $\phi = 1 - \zeta$, which shows that stronger real rigidities (ϕ close to zero) lead to stronger strategic complementarity in pricing (ζ close to unity).

Going back to real rigidities, the literature typically refers to mechanisms that operate through Π_{12} as macro real rigidities (termed Ω-type strategic complementarity by Nakamura and Steinsson (2010)) and those that operate through Π_{11} as micro real rigidities (termed ω-type strategic complementarity). Macro real rigidities, which work through smaller Π_{12}, make firms less responsive to aggregate shocks. This can occur, for example, if firms' costs are sticky and do not respond to changes in aggregate conditions. Examples of macro real rigidities include real wage rigidities (Blanchard and Galí, 2007) and sticky input prices in production networks (Basu, 1995).

In contrast, due to the concavity of the profit function as captured by $|\Pi_{11}|$, micro real rigidities make adjusting one's own relative price costly. This can dampen firms' pricing responses to aggregate shocks when paired with nominal pricing frictions. This type of real rigidity can have two sources: the demand function and the cost function. To see this, consider the real profit function

$$\Pi\left(\frac{p_i}{P};\cdot\right) = \frac{p_i}{P} D\left(\frac{p_i}{P};\cdot\right) - C\left[D\left(\frac{p_i}{P};\cdot\right)\right], \tag{5}$$

where $D\left(\frac{p_i}{P};\cdot\right)$ is the demand schedule, which depends on the firm's relative price and other factors, and $C\left[D\left(\frac{p_i}{P};\cdot\right)\right]$ is the total real cost of production. The second derivative of the profit function with respect to the firm's relative price is given by

$$\Pi_{11}\left(\frac{p_i}{P};\cdot\right) = 2D_1\left(\frac{p_i}{P};\cdot\right) + \left[\frac{p_i}{P} - C_1\left[D\left(\frac{p_i}{P};\cdot\right)\right]\right] D_{11}\left(\frac{p_i}{P};\cdot\right) - C_{11}\left[D\left(\frac{p_i}{P};\cdot\right)\right]\left[D_1\left(\frac{p_i}{P};\cdot\right)\right]^2.$$

A more concave demand function ($D_{11} << 0$) can increase $|\Pi_{11}|$, as firms incur larger losses when their relative prices deviate from the optimal level. Examples of demand-side micro real rigidities include CES demand systems that allow for more curvature, such as those in Kimball (1995) and translog demands (Bergin and Feenstra, 2000). Alternatively, firms may also find relative price adjustments costly if they face an increasing marginal cost curve ($C_{11} >> 0$). When the marginal cost curve is increasing, the average cost rises whenever firms cut prices to increase quantity sold, reducing profit per unit sold and offsetting the gains from lowering prices. This effect is more pronounced when the marginal cost curve is more convex. Examples of cost-side mechanisms include decreasing returns to scale production technology Burstein and Hellwig (2007) and segmented input markets leading to upward-sloping cost curves at the firm level (Woodford, 2003; Gertler and Leahy, 2008).[8]

Despite its theoretical soundness, the literature has largely regarded micro real rigidities (or ω-type strategic complementarity) of limited relevance in quantitative pricing models. As noted by Klenow and Willis (2016), the scale of idiosyncratic productivity shocks or price adjustment costs needed to replicate the empirically observed price change sizes is implausibly high within menu cost models that incorporate micro real rigidities.

To clarify this critique, observe the expression

$$\frac{\partial p_i^*}{\partial A_i} = -\frac{\Pi_{13}}{\Pi_{11}} P. \tag{6}$$

Assuming Π_{13} is invariant to the presence of real rigidities, this demonstrates that when micro real rigidities are present—that is, when $|\Pi_{11}|$ is large —the sensitivity of firms' prices to changes in idiosyncratic variables, such as productivity shocks, is reduced. As a result, to achieve a certain magnitude of price changes under stronger micro real rigidities, larger disturbances to idiosyncratic variables would be required.

[8]Alternative mechanisms that do not fall into the demand or cost side mechanisms discussed here are typically concerned with market structure and competition, such as the oligopolistic competition framework in Mongey (2021).

2.2 Revisiting Micro Real Rigidities

Our focus in this paper is on demand-side micro real real rigidities that increase the concavity of profit function through concave demand functions, though we also examine cost-side mechanisms for robustness.[9] To capture these demand-side rigidities, we employ the demand system proposed by Kimball (1995), which produces variable price elasticity and can accommodate any curvature of the demand function.

The Kimball demand framework defines an aggregator G that combines varieties y_i into a composite good Y through the implicit equation

$$1 = \int G\left(\frac{y_i}{Y}\right) di, \tag{7}$$

where G satisfies $G(1) = 1$, $G'(x) > 0$, and $G''(x) < 0$ for all $x > 0$. Under this demand structure, a producer of variety y_i faces an inverse demand function[10]

$$p_i = \frac{\lambda}{Y} G'\left(\frac{y_i}{Y}\right). \tag{8}$$

A key feature of the Kimball demand system is that the price elasticity of demand, ϵ, depends on the relative output $\frac{y_i}{Y}$ of the firm's product

$$\frac{1}{\epsilon\left(\frac{y_i}{Y}\right)} = -\frac{y_i}{Y}\frac{G''}{G'}. \tag{9}$$

With an appropriate choice of the $G(.)$ function, any desired relationship between demand elasticity and relative output—and, therefore, any desired degree of curvature in the demand function—can be generated. To produce micro real rigidity and strategic complementarity, we focus on the case where demand elasticity decreases with relative output. In this setting, the incremental output from a price reduction decreases as firms move down the demand curve, resulting in greater concavity of the demand function.[11] This greater concavity in the

[9] Appendix D.3 examines the quantitative performance of a cost-side model, comparing our baseline model to Burstein and Hellwig (2007), which uses a CES framework with decreasing returns to scale and demand shocks.

[10] This inverse demand function is derived from the cost minimization problem $\min_{y_i} \int p_i y_i di$, subject to $1 = \int G\left(\frac{y_i}{Y}\right) di$. The associated Lagrangian is $\mathcal{L} = \int p_i y_i di - \lambda\left[1 - \int G\left(\frac{y_i}{Y}\right) di\right]$.

[11] The super-elasticity of demand, S, defined as the elasticity of demand elasticity with respect to price, is insightful here. For a generic demand function $D(p)$, it is given by $S \equiv \frac{d\epsilon}{dp}\frac{p}{\epsilon} = \frac{D''(p)\cdot p\cdot D(p)+D'(p)\cdot D(p)-D'(p)^2\cdot p}{D'(p)\cdot D(p)}$. As concavity in $D(p)$ increases (i.e., more negative $D''(p)$), the super-elasticity becomes more positive, indicating that demand elasticity declines faster as price decreases and relative output rises.

profit function penalizes deviations from the optimal relative price $\frac{p_i}{P}$, leading to strategic complementarity in pricing.

A significant implication of variable demand elasticity is its effect on desired markups and cost pass-through. For a firm experiencing a positive productivity shock, its marginal cost decreases, encouraging a price reduction to boost sales. Under CES demand, prices adjust proportionally to changes in marginal cost and markup remains constant. However, with Kimball demand, demand elasticity decreases as prices fall and relative output increases, which lessens the revenue gains from further price reductions. As a result, the firm adjusts prices by less than one-for-one relative to marginal cost changes, leading to higher markups. This incomplete pass-through of costs forms the basis of the critique by Klenow and Willis (2016), who argue that because micro real rigidities dampen price responses to idiosyncratic productivity shocks, implausibly high productivity shock volatility would be needed to match observed price changes.

We propose that this critique can be addressed by incorporating idiosyncratic demand n_i into the Kimball aggregator, modifying it as $1 = \int G\left(\frac{n_i y_i}{Y}\right) di$. Here, n_i alters each firm's effective market share, impacting its demand elasticity and, thus, its desired markups and prices. Intuitively, varieties with higher n_i attract greater consumer preference, leading to smaller demand elasticities and higher permissible markups. Consequently, demand shock volatility contributes directly to price volatility, offsetting the reduced impact of idiosyncratic productivity shocks on price adjustments. Put differently, for a given degree of price volatility, the presence of demand shocks reduces the need for large productivity shocks.

In our quantitative analysis, following Dotsey and King (2005), we specify $G(.)$ as:

$$G\left(\frac{n_i y_i}{Y}\right) = \frac{\omega}{1+\omega\psi}\left[(1+\psi)\frac{n_i y_i}{Y} - \psi\right]^{\frac{1+\omega\psi}{\omega(1+\psi)}} + 1 - \frac{\omega}{1+\omega\psi}, \tag{10}$$

where n_i is an idiosyncratic demand shifter. Here, ψ plays a crucial role, influencing the relationship between demand elasticity and effective market share, the curvature of the demand function, and the pass-through of idiosyncratic productivity and cost changes to prices.

Given the aggregator $G(.)$ in (10), solving the final producer's cost-minimization problem yields the demand function for each variety i[12]

$$\frac{n_i y_i}{Y} = \frac{1}{1+\psi}\left[\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi\right], \tag{11}$$

where λ is the Lagrange multiplier on the aggregator G in the cost-minimization problem and P is the aggregate price index.[13] Following (9), the price elasticity of the demand system for firm i is given by

$$\epsilon_i \equiv \frac{\partial y_i}{\partial p_i}\frac{p_i}{y_i} = \frac{\omega}{1-\omega}\frac{(1+\psi)\frac{n_i y_i}{Y} - \psi}{\frac{n_i y_i}{Y}} = \frac{\omega(1+\psi)}{1-\omega}\frac{\left(\frac{p}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}}}{\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi}.$$

When $\psi < 0$, the price elasticity of demand increases with the relative price of the variety, $\frac{p_i}{P}$, and decreases with the effective market share of the variety, $\frac{n_i y_i}{Y}$.[14] The opposite effects hold when $\psi > 0$. When $\psi = 0$, this specification simplifies to the familiar constant elasticity of substitution (CES) Dixit-Stiglitz aggregator, which results in iso-elastic demand. The parameter ψ also controls the super-elasticity of demand, given by

$$\gamma_i \equiv \frac{\partial \epsilon_i}{\partial p_i}\frac{p_i}{\epsilon_i} = \frac{\omega}{1-\omega}\frac{\psi}{\frac{n_i y_i}{Y}}. \tag{12}$$

Assuming $\omega > 1$, the super-elasticity becomes more negative as $\psi < 0$ decreases, which creates greater curvature in the demand function. Thus, more negative values of ψ lead to stronger real rigidity measured by ϕ in (3), as illustrated in Figure A-1.

The parameter ψ is also critical for addressing the Klenow and Willis (2016) critique, as it determines the pass-through of idiosyncratic demand and cost to firms' desired prices in conjunction with ω. To see this, we log-linearize the optimal pricing problem of a variety producer who does not face any pricing frictions and react to idiosyncratic shocks z and n.

[12] The full final producer's problem is presented in Section 3.2.1.

[13] The Lagrange multiplier λ can be obtained by substituting final producer's first-order condition into the Kimball aggregator: $\lambda = \left[\int_0^1 \left(\frac{p^i}{n^i P}\right)^{\frac{1+\omega\psi}{1-\omega}} di\right]^{\frac{1-\omega}{1+\omega\psi}}$. The aggregate price index is derived from the zero-profit condition of the final producer: $P = \frac{1}{1+\psi}\left[\int_0^1 \left(\frac{p^i}{n^i}\right)^{\frac{1+\omega\psi}{1-\omega}} di\right]^{\frac{1-\omega}{1+\omega\psi}} + \frac{\psi}{1+\psi}\int_0^1 \frac{p^i}{n^i} di$.

[14] Appendix A.1 show the partial derivatives of the demand elasticity with respect to relative price, demand, and market share.

The optimal price is

$$\hat{p}_i^* \quad = \quad \frac{\omega\psi}{\omega\psi - 1}\left(\hat{\lambda} + \hat{P} + \hat{n}_i\right) + \frac{1}{\omega\psi - 1}\hat{z}_i, \tag{13}$$

where hatted variables denote log-deviations from the steady state. Note that $\widehat{mc}_i \equiv -\hat{z}_i$ denotes the log-deviation in marginal cost, which is inversely related to productivity under constant returns to scale. The price elasticities with respect to productivity and demand shocks, respectively, are given by

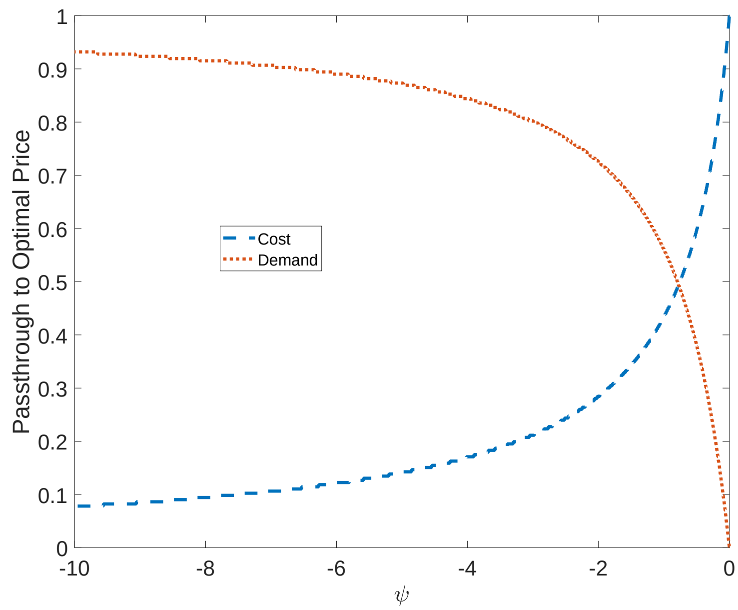
$$\frac{\partial \hat{p}_i^*}{\partial \widehat{mc}_i} = -\frac{1}{\omega\psi - 1} \quad \text{and} \quad \frac{\partial \hat{p}_i^*}{\partial \hat{n}_i} = \frac{\omega\psi}{\omega\psi - 1}.$$

We refer to these as cost and demand pass-through, respectively.

When $\psi = 0$, cost pass-through is complete: price decreases one-to-one with a positive z shock (negative mc shock). However, when $\psi < 0$, cost pass-through becomes incomplete. 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 experiencing a demand shock chooses to increase its price or, equivalently, applies a higher markup over its marginal cost. Thus the relative importance of demand and cost shocks in determining prices depends on the degree of micro real rigidity, as controlled by ψ. Figure 1 shows the pass-through of cost and demand to the optimal frictionless price as a function of ψ for a given value of ω. As ψ decreases, the pass-through of productivity shocks declines, while the pass-through of demand shocks increases.

Interestingly, although the analysis around (6) shows that stronger micro real rigidities (larger $|\Pi_{11}|$) dampen the response to productivity shocks, this conclusion is reversed for demand shocks. The key to this difference lies not in the behavior of $|\Pi_{11}|$ but in the relationship between ψ and Π_{13} for different types of idiosyncratic shocks. As shown in Appendix A.4, under a symmetric equilibrium, the value of Π_{13} is independent of ψ for productivity shocks, implying that (6) is dominated by the increase in $|\Pi_{11}|$ triggered by more negative values of ψ. In contrast, Π_{13} increases with ψ when considering idiosyncratic demand shocks. This increase dominates the dynamics of (6), producing the behavior illustrated in Figure 1. This is why demand shocks within a Kimball demand system have the potential to overturn the volatility-dampening effect documented by Klenow and Willis (2016).

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 $\omega = 1.29$

3 Quantitative Menu Cost Model

This section builds a quantitative menu-cost model similar to Golosov and Lucas (2007) and Nakamura and Steinsson (2008), consisting of a representative household, a representative final-good producer, and a continuum of monopolistically competitive intermediate-variety producers facing nominal pricing frictions. The key difference is the inclusion of the elements described in Section 2, i.e., a Kimball (1995) demand system with idiosyncratic demand shocks.

3.1 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

$$\max_{C_t, h_t, \mathbf{B_{t+1}}} \quad \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\log\left(C_t\right) - \chi h_t\right] \tag{14}$$

subject to the budget constraint

$$P_tC_t + \mathbf{Q_t} \cdot \mathbf{B_{t+1}} \leq B_t + W_th_t + \Pi_t, \tag{15}$$

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

$$\frac{W_t}{P_t} = \chi C_t, \tag{16}$$

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

$$\Xi_{t,t+1} \equiv \beta \mathbb{E}_t \left(\frac{C_t}{C_{t+1}} \right). \tag{17}$$

3.2 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.

3.2.1 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 defined in (10). The representative final-good producer chooses y_t^i to maximize profits taking as given variety prices, p_t^i, as well as P_t, n_t^i and the real aggregate demand, Y_t, solving the problem

$$\max_{y_t^i} \quad 1 - \int_0^1 \frac{p_t^i y_t^i}{P_t Y_t} di \quad \text{subject to} \quad \int_0^1 G\left(\frac{n_t^i y_t^i}{Y_t}\right) di = 1. \tag{18}$$

The optimality condition of the final-good producer's maximization problem implicitly defines the demand function for each variety in (11) and the aggregate price index follows from the zero-profit condition.

3.2.2 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 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

$$y_t^i = z_t^i l_t^i. \tag{19}$$

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

$$\begin{pmatrix} \log(z_t^i) \\ \log(n_t^i) \end{pmatrix} = \begin{bmatrix} \rho_z & 0 \\ 0 & \rho_n \end{bmatrix} \begin{pmatrix} \log(z_{t-1}^i) \\ \log(n_{t-1}^i) \end{pmatrix} + u_t^i \text{ where } u_t^i \sim N\left(0, \begin{bmatrix} \sigma_z^2 & \sigma_{zn} \\ \sigma_{zn} & \sigma_n^2 \end{bmatrix}\right) \tag{20}$$

At the beginning of each period, intermediate-good producers inherit their prices from the previous period p_{t-1}^i and observe the realizations of z_t^i and n_t^i. They then 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_t^i is

$$\pi(p_t^i, z_t^i, n_t^i, \mathcal{S}_t) = \left(\frac{p_t^i}{P_t} - \frac{W_t}{z_t^i P_t}\right) \frac{Y_t}{n_t^i} \frac{1}{1+\psi} \left[\left(\frac{p_t^i}{\lambda_t n_t^i P_t}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi\right], \tag{21}$$

where $\mathcal{S}_t \equiv (P_t, W_t, Y_t, \lambda_t)$ collects all the relevant state variables, and the firms can approximate well the law of motion for $\mathcal{S}_t$.

The firms choose whether or not to change their prices by solving the problem

$$V\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right) = \max\left[V_N\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right), V_A\left(z_t^i, n_t^i, \mathcal{S}_t\right)\right], \tag{22}$$

where $V_N(.)$ and $V_A(.)$ are the values for the firm not adjusting and adjusting their prices, respectively.

The value of not adjusting prices is

$$V_N\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right) = \pi\left(p_{t-1}^i, z_t^i, n_t^i, \mathcal{S}_t\right) + \mathbb{E}_t\left[\Xi_{t,t+1} V\left(p_{t-1}^i, n_{t+1}^i, z_{t+1}^i, \mathcal{S}_{t+1}\right)\right], \tag{23}$$

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

$$V_A\left(z_t^i, n_t^i, \mathcal{S}_t\right) = -f\frac{W_t}{P_t} + \max_{p_t^i}\left\{\pi(p_t^i, z_t^i, n_t^i, \mathcal{S}_t) + \mathbb{E}_t\left[\Xi_{t,t+1}V\left(p_t^i, n_{t+1}^i, z_{t+1}^i, \mathcal{S}_{t+1}\right)\right]\right\}. \quad (24)$$

The intermediate-good producers solve this problem taking as given the laws of motion for the idiosyncratic state variables as in (20) and those for the aggregate variables in $\mathcal{S}_t$.

3.2.3 Aggregate Nominal Expenditure

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

$$\log\left(S_t\right) = \mu + \log\left(S_{t-1}\right) + \sigma_S\epsilon_t,\ \epsilon_t \sim N(0,1), \quad (25)$$

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.[15]

4 Calibration

This section outlines the calibration strategy and demonstrates that a simple menu cost model incorporating micro real rigidities aligns with firm-level estimated shocks and generates appropriate pricing dynamics. Additionally, the parsimonious model captures several non-targeted empirical regularities documented in the literature.

The calibration approach departs from standard menu cost models in one significant way. In some quantitative menu cost models there is a shock that directly affects the desired price (e.g., Caplin and Spulber (1987)). More commonly, a productivity shock is introduced (e.g., Golosov and Lucas (2007); Nakamura and Steinsson (2010); Midrigan (2011); Vavra (2014)) however its properties are calibrated to match various pricing moments without any reference to empirical evidence on firm-level productivity. We take a novel approach by disciplining idiosyncratic demand and productivity processes with direct firm-level evidence.

[15] Appendix B.2 describes our solution method.

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.

Our methodology also contributes to the macroeconomic literature using Kimball demand systems (Smets and Wouters, 2007; Klenow and Willis, 2016; Harding et al., 2022, 2023), demonstrating how micro moments can discipline key parameters characterizing the demand system.

The model involves calibrating 12 parameters. Five parameters $(\beta, \chi, \mu, \sigma_S, \sigma_{zn})$ are externally calibrated, while the remaining seven $(\rho_z, \sigma_z, \rho_n, \sigma_n, \omega, \psi, f)$ are determined internally.

4.1 Externally Calibrated Parameters

We begin by detailing the externally calibrated parameters. The model is calibrated to U.S. data, with each model period representing one month. The monthly discount factor β is set to 0.9966, corresponding to an annual discount rate of 4%. Following standard practice in the literature (Midrigan, 2011), the disutility of labor, χ, is normalized to 1, ensuring that the nominal wage, W_t, equals the money supply, S_t. The monthly growth rate of the money supply, μ, is set at 0.2%, implying an annual inflation rate of approximately 2.4%. For firm-level processes, we assume that idiosyncratic demand and productivity innovations are uncorrelated ($\sigma_{zn} = 0$).[16] Lastly, in monetary policy experiments, we calibrate $\sigma_S = 0.0037$, following Vavra (2014), to match the observed volatility of nominal output growth in the U.S. Table 1 summarizes the five externally calibrated parameters.

[16] This assumption aligns with Foster et al. (2008), where $\sigma_{zn} = 0$ is necessary for their estimation strategy, which assumes orthogonality between demand and productivity. Using Colombian data and an alternative strategy that relaxes this assumption, Eslava et al. (2024) report a correlation of -0.07 between demand and productivity. We address this in Appendix D.2, showing that our results are robust even when allowing for correlated shocks.

4.2 Internally Calibrated Parameters

To be consistent with the firm dynamics evidence, we refer to Foster et al. (2008), who provide direct estimates of firm-level idiosyncratic productivity and demand processes using data on U.S. manufacturing firms from 1977 to 1997.[17] We rely on these estimates to discipline the parameters governing the AR(1) processes for idiosyncratic firm productivity (ρ_z, σ_z) and demand shocks (ρ_n, σ_n) in the model. In addition, we use two moments from the same study to calibrate the parameters governing the Kimball demand system (ω, ψ). In contrast to the conventional calibration strategy in the menu cost literature, we rely on only a single pricing moment – the frequency of price changes.[18]

We jointly calibrate these seven parameters to match the frequency of price changes and six firm-dynamics moments from Foster et al. (2008): the five-year autocorrelation and the annual variance of demand and TFP, the correlation between TFPQ and price, and the correlation between TFPQ and TFPR. We construct firm-dynamics moments from model simulated data in the exact same way as Foster et al. (2008) do in their empirical analysis.[19] Details on the empirical data targets as well as the calibration algorithm are delegated to Appendix C.

Before presenting the results, a discussion on the identification of parameters is in order. While all seven parameters influence the model's ability to match all calibration targets, some parameters are more responsible for matching specific target moments. Some of these relationships are intuitive. The fixed cost, f, plays a significant role in determining the model-implied frequency of price changes, while the shock-process parameters $(\rho_z, \sigma_z, \rho_n, \sigma_n)$ are closely related to the empirical persistence and cross-sectional dispersion in firm-level demand and productivity reported in Foster et al. (2008).

Less obvious may be the links between ψ, ω and the correlations of TFPQ with prices and TFPQ with TFPR. To understand these, it is instructive to start from a CES demand system (with $\psi = 0$). In the absence of pricing frictions, the profit-maximizing rule in that framework delivers a pricing strategy that sets price as a constant markup over marginal

[17]Appendix C.1.1 describes the procedure used by Foster et al. (2008), and Appendix C.2 provides details on how we replicate their estimation using model-simulated data.

[18]Appendix C.4 considers an alternative strategy where we target the average size of price changes instead of the frequency and show that the two strategies lead to very similar outcomes.

[19]One may be concerned that some of the underlying assumptions in Foster et al. (2008) are not satisfied in our model. Two noteworthy ones are flexible prices and CES demand. The former is less likely to have large effects on annual firm aggregates and five-year autocorrelations of firm idiosyncratic demand and productivity. In general, treating the model generated data in the same way as Foster et al. (2008) means that the model moments and empirical moments are affected by these two sources of misspecification equally.

cost. As a result, a firm's optimal price is inversely proportional to its productivity, that is, Corr$(P, TPFQ) = -1$. Moreover, in a CES demand system, TFPR is equalized across firms, so Corr$(TFPR, TPFQ) = 0$. This occurs because optimizing firms operate at the point where the marginal revenue product of labor ($p_{i,t}z_{i,t}$) equals the nominal market wage.

Under a Kimball demand system, both productivity and demand factors affect a firm's optimal price, as shown in Section 2.2. In particular, deviations from CES, controlled by the parameter ψ, reduce the influence of productivity on pricing relative to demand. Because some price changes are due to demand shocks, this weakens the perfect negative correlation between price and TFPQ. The parameter ω, on the other hand, governs the elasticity of substitution between varieties around a symmetric equilibrium, with a higher ω indicating less substitutability across varieties. In a Kimball demand system, more productive firms can charge higher prices if the elasticity of substitution is lower (ω is higher), leading to a higher correlation coefficient between TFPQ and TFPR.[20]

To demonstrate that the calibration targets are indeed informative for the respective parameters, we solve the model for a large set of quasi-random parameter vectors and show a strong relationship between each internally calibrated parameter and its corresponding empirical target moment. The results are reported in Appendix C.3.

The results of the internal calibration are reported in Table 2. The top panel shows the targeted moments, whose calculations are described in more detail in Appendices C.1.1 and C.1.2. The second panel presents the seven parameters calibrated jointly. The results indicate that all seven moments are matched very closely. The calibrated processes for idiosyncratic demand and productivity are highly persistent, with monthly autocorrelations of 0.997 and 0.98, respectively. While innovations to idiosyncratic productivity are larger, with a standard deviation of 0.06 compared to 0.02 for demand, the stationary distribution of idiosyncratic demand exhibits greater dispersion due to the highly persistent nature of the process.

Regarding the parameters governing the shape of the demand function, the calibrated value of $\psi = -1.27$ suggests that a substantial deviation from CES demand is necessary to match the data. The calibrated value of $\omega = 1.29$ implies a price elasticity of demand of 4.44 and a markup of 29% under a symmetric equilibrium. Appendix C.5 discusses calibrations and estimates from other studies with Kimball demand, showing that our calibration lies modestly within the range of values considered in the literature.

[20] Appendix A.2 provides supplementary derivations on the properties of the symmetric equilibrium.

Table 2: Internal Calibration

Moment	Data	Baseline Model
Frequency of price changes	0.11	0.12
5-year autocorrelation of z_t^i	0.32	0.32
Cross-sectional standard deviation of z_t^i	0.26	0.25
5-year 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	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.

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.

4.3 Untargeted Pricing Moments

Table 3 reports three important pricing moments that are not targeted in the calibration of the baseline model, alongside their data counterparts computed from U.S. CPI microdata.[21] These moments include 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

[21] We borrow these estimates from Vavra (2014), which we review in more detail in Appendix C.1.2. One may be concerned about whether pricing facts from CPI data are the correct benchmark. However, Nakamura and Steinsson (2008) show that the key pricing moments computed from PPI data do not differ significantly.

model matches all three moments very well.

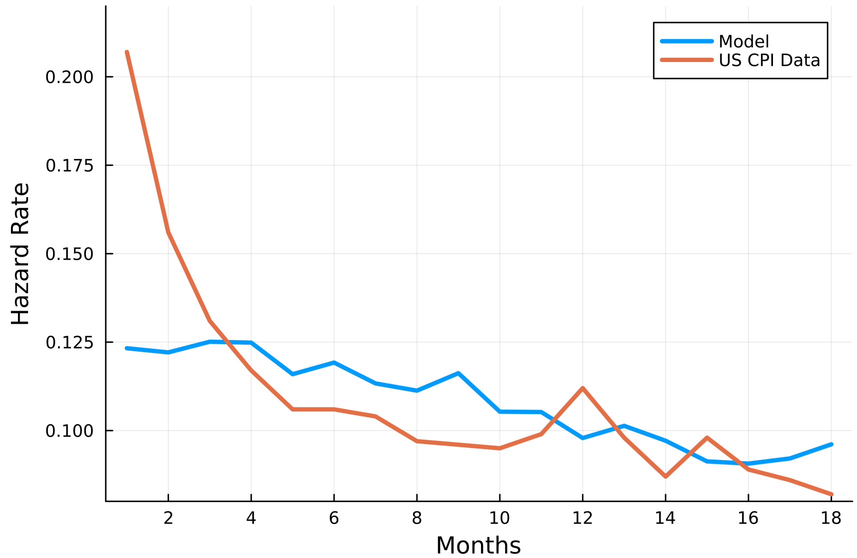
The success of the model addresses the skepticism in Klenow and Willis (2016) regarding the relevance of micro real rigidities by demonstrating that a menu cost model incorporating micro real rigidities can indeed be consistent with both direct estimates of idiosyncratic firm processes and the price-setting behavior of firms. This success is due to the interplay between micro real rigidities introduced through Kimball demand and the inclusion of demand shocks. Crucially, the precise calibration of these two components is essential for achieving the results. Figure A-5 in Appendix C.3 illustrates that the value of ψ is positively related to the average size of price changes, as stronger strategic complementarities reduce the dispersion between firms' prices. Furthermore, Table A-7 in Appendix D.4 demonstrates that an alternative calibration without demand shocks results in smaller price changes. Thus, it is remarkable that our model matches the pricing moments as well as it does given that the key parameters are informed by the firm dynamics moments in Foster et al. (2008).

We also examine the hazard function of price changes generated by the model. The price adjustment hazard is defined as the probability that a price will change h periods after the last adjustment, conditional on the price spell lasting h periods. Empirically, the hazard function is observed to be either downward-sloping (Nakamura and Steinsson, 2008; Baley and Blanco, 2019) or flat (Klenow and Kryvtsov, 2008).[22] As highlighted by Nakamura and Steinsson (2008), simple menu cost models are typically unable to produce hazard functions consistent with the empirical evidence. This limitation largely depends on the calibration of the idiosyncratic processes.

In models with trend inflation, the hazard function tends to be upward-sloping when idiosyncratic shocks are small. This occurs because, once a price changes, it takes time for trend inflation, the dominant source of price changes under small idiosyncratic shocks,to accumulate sufficiently to trigger another adjustment. Larger and more persistent idiosyncratic shocks flatten the hazard function by inducing temporary price changes, which are often quickly reversed, thus increasing the early part of the hazard function. Thus when idiosyncratic shocks become sufficiently large, a simple menu-cost model can produce a downward-sloping hazard function. However, Nakamura and Steinsson (2008) argue that such calibrations are unrealistic, as they conflict with micro-level pricing facts. This conclusion has motivated alternative pricing models that aim to rationalize a downward-sloping hazard. For instance, Baley and Blanco (2019) introduce firm-level uncertainty and learning

[22]While the literature generally finds a mildly negative slope, Alvarez et al. (2023a) show that controlling for product heterogeneity can reverse the sign of the slope.

Figure 2: Hazard Function of Price Change



Note: This figure plots the pricing hazard from model generated data against the empirical hazard estimated from US CPI in Baley and Blanco (2019).

to generate frequent price changes shortly after an adjustment.

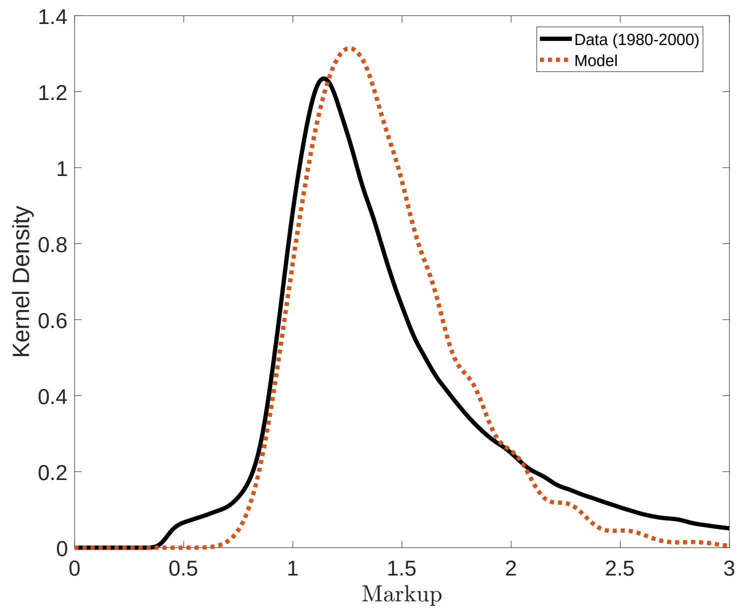
Figure 2 shows the mildly downward-sloping hazard curve produced by our model, compared with the empirical hazard function estimated by Baley and Blanco (2019) using U.S. CPI data.[23] While the model-implied hazard does not exhibit a steep decline in the first few months, it successfully captures the overall downward-sloping pattern observed in the data. This is consistent with the logic explained earlier. The presence of idiosyncratic shocks increases the hazard rate, as having two orthogonal shocks provides more reasons for price changes shortly after an adjustment. Furthermore, the inclusion of both types of shocks reduces the need for each shock to be particularly large, addressing the criticism of Nakamura and Steinsson (2008) noted above.

4.4 External Validity: Markups, Pass-through, and the Size Distribution

The work of Foster et al. (2008) is the only study for the U.S. that systematically estimates productivity and demand shocks at the firm level across different industries. This estimation relies on price and quantity data at the product level, alongside other information such as

[23]The shape of the hazard function reported by Baley and Blanco (2019) aligns with that reported by Nakamura and Steinsson (2008) for processed food, which they suggest is representative of the hazard function for many other product groups.

Figure 3: 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.

inputs. By focusing on a carefully selected set of firms operating in industries that produce homogeneous products, they are able to separately estimate the processes for productivity and demand shocks.

However, one might question the external validity of the estimations provided by Foster et al. (2008) when applied to the broader economy. We address these concerns by comparing our model to additional statistics computed from different samples. Despite being calibrated to a limited set of manufacturing industries, our model successfully matches untargeted moments for the broader economy across several dimensions—namely, price-setting evidence, the cross-sectional markup distribution, and the overall dispersion of firm growth rates.

4.4.1 Markups

One prominent application of Kimball demand, among other non-CES systems, is in modeling variable markups. Because a firm's desired markup depends on both its idiosyncratic productivity and demand, the cross-sectional distribution of productivity and demand, combined with pricing frictions, results in a non-degenerate markup distribution in the model.

Figure 3 plots the kernel density of the cross-sectional distribution of gross markups from both the model and the data. The empirical distribution of markups is computed using data on public firms between 1980 and 2000, following the method of De Loecker et al. (2020).[24] We find that the markup distribution generated by the model closely resembles the untargeted empirical distribution. The median gross markup in the model is 1.35, compared to 1.33 in the data. However, the model-implied markup distribution exhibits lower variance than the empirical distribution, particularly in the tails. Specifically, our model generates fewer firms with gross markups exceeding 2 and does not produce markups significantly below 1. Matching these extreme tails would require additional features, such as non-Gaussian demand shocks, monopolistic firms (large positive markups), customer capital, or firm exit (negative markups).

The success of our model in replicating the empirical markup distribution is perhaps unsurprising given the findings of Edmond et al. (2023), among others, who use Kimball demand systems to model firm markups. Nonetheless, it is noteworthy that while our model is calibrated to six moments related to demand, supply, and productivity estimated from selected manufacturing industries, it provides a remarkable fit to the untargeted distribution of estimated markups across a broad set of industries. This result lends external validity to the generalizability of the estimates from Foster et al. (2008).

4.4.2 Cost Pass-through

We turn to comparing the degree of cost pass-through implied by our calibration to the empirical estimates from the literature. In the model, as shown in (14), the strength of cost pass-through is determined by ω and ψ. Given the calibrated values of these two parameters, the model implies a cost pass-through of 38%. The literature in both open-economy and closed-economy macroeconomics, reviewed in Appendix C.1.3, has estimated cost pass-through using various datasets from different countries and consistently finds pass-through rates in the range of 20% to 40%. The alignment of the model's implied cost pass-through with this extensive empirical evidence provides additional external validity to our calibration strategy.

[24] Despite the potential limitations of using revenue data to estimate markups, De Ridder et al. (2024) show that while the estimated levels of markups can be noisy, this method provides reliable estimates of the dispersion and shape of the markup distribution. More details on the empirical markup estimation can be found in Appendix C.1.3.

4.4.3 Comparison with Other Countries

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, employing the same methodology as Foster et al. (2008). They report similar values for Corr($TFPQ, TFPR$) and Corr($TFPQ, P$), which are crucial for pinning down the Kimball demand system parameters. Specifically, they report Corr($TFPQ, TFPR$) = 0.69 and Corr($TFPQ, P$) = –0.65, compared to the Foster et al. (2008) values of 0.75 and –0.54, respectively.

In more recent work, Eslava et al. (2024) apply a different approach to the same Colombian manufacturing firms data, which allows them to relax the orthogonality assumption between idiosyncratic supply and demand. In Appendix D.1, we report an alternative calibration of our model to the moments reported in Eslava et al. (2024). We observe that the properties of idiosyncratic demand and productivity among Colombian manufacturing firms exhibit similar patterns to those studied in Foster et al. (2008). Furthermore, we find that under the alternative calibration, the model is also able to match the average size of price changes in Colombia and cost pass-through, even though these are untargeted moments.

4.4.4 Firm Dynamics

Because 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 whether the estimates from Foster et al. (2008) can be generalized beyond the eleven industries they study. We benchmark our estimates to Davis et al. (2006), a study that utilizes the Longitudinal Business Database, providing a comprehensive measure of U.S. business dynamics. They estimate the cross-sectional standard deviation of firm revenue growth rates to be 0.39 over the period 1982–1997, while in our simulated data, this untargeted moment is 0.41.

4.5 Robustness

In this section, we explore a range of robustness exercises to our quantitative analysis.

4.5.1 CES Demand with Decreasing Returns to Scale

The baseline analysis focuses on Kimball demand as a source of micro real rigidity. Alternatively, micro real rigidity and strategic complementarity in pricing can arise from an upward-sloping marginal cost curve. We explore this setup with supply-side micro real rigidity, featuring CES demand and decreasing returns to scale at the firm level, similar to Burstein and Hellwig (2007). By applying the same calibration strategy, we show in Appendix D.3 that this model setup can also match both targeted and untargeted moments of the data, similar to the baseline model. While supply-side micro real rigidity can jointly match pricing and firm dynamics evidence, it requires an extremely low returns to scale of 0.21 and does not replicate the untargeted markup distribution well. Nonetheless, while we focus solely on the demand side in our analysis for simplicity, there may be potential advantages for combining both supply and demand-side micro real rigidities in future work.

4.5.2 Correlated Shocks

The estimation strategy of Foster et al. (2008) assumes orthogonality between shocks to idiosyncratic demand and productivity. The approach ofEslava et al. (2024) relaxes this assumption and finds that demand and supply shocks estimated using data on Colombian manufacturing firms exhibit a weak negative correlation of -0.07. Appendix D.2 summarizes two alternative calibrations where we allow σ_{zn} to be non-zero while holding all other parameters fixed at their baseline levels. The results show that even with a moderately high correlation coefficients of 0.4 or -0.4, the key model moments do not differ significantly. The degree of correlation does affect the average size of price changes implied by the model. When productivity shocks are positively correlated with demand shocks, the average size of price changes is smaller. This occurs because firms that receive favorable productivity shocks—which put downward pressure on prices—tend to receive negative demand shocks, that dampen the firm's desire to lower prices.

4.5.3 No Idiosyncratic Demand

Appendix D.4 analyzes the baseline model in the absence of demand shocks. In doing so, we must 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. Furthermore, the average price change is somewhat smaller, and the markup distribution is less dispersed. These results arise from the absence of demand shocks that prompt firms to change their prices, thus highlighting the role of demand shocks in delivering a better match with untargeted pricing moments.

To further scrutinize the role of demand shocks in helping the model match micro pricing data, we solve the model while turning off supply shocks, holding all other parameters fixed at the baseline values. Without either demand or supply shocks, the model would fall short in both the frequency and size of price changes, showing that having both shocks is crucial in explaining micro pricing facts in the data.

4.5.4 Leptokurtic Shocks

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 median estimated kurtosis of 4.5 in the literature (Alvarez et al., 2016). In Appendix C.4, we introduce leptokurtic demand shocks following Midrigan (2011) to generate more realistic kurtosis of non-zero price changes. This extension does not change the conclusions of the baseline model.

4.5.5 Klenow and Willis (2016)

We revisit the original critique of micro real rigidity from Klenow and Willis (2016) by assessing the fit of their quantitative model in light of firm dynamics evidence. Appendix D.5 summarizes the model using their calibration, which targets a frequency of price change of 0.09.[25] The key observation is that their calibration does not fit the firm dynamics in Foster et al. (2008) well. In fact, once TFP and demand shocks are properly disciplined by the evidence, a model with Kimball demand turns out to generate an average size of price changes that is just right.

4.5.6 Translog Demand

The literature has used other demand systems with variable demand elasticity, besides Kimball demand, which also lead to strategic complementarity in pricing. In Appendix D.6, we

[25]The baseline calibration in Klenow and Willis (2016) ($\theta = 5, \epsilon = 10$ using their specification) translates 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 a smaller pass-through of cost shocks.

consider a translog demand system that is frequently used in trade and industrial organization literature as an alternative non-CES demand system. 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.

The model featuring a translog demand system matches the four moments derived 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 can be attributed to the properties of translog that restrict the cost pass-through to be between 50% and 100%. Relatedly, given the size of the underlying shocks, the size of price changes turns out to be 12%, which is considerably larger than observed in the data. Due to having one additional degree of freedom and no restrictions on the cost pass-through, a Kimball demand system is more flexible than a translog system and thus more desirable to work with. Nonetheless, the translog demand system performs better than models with CES demand.

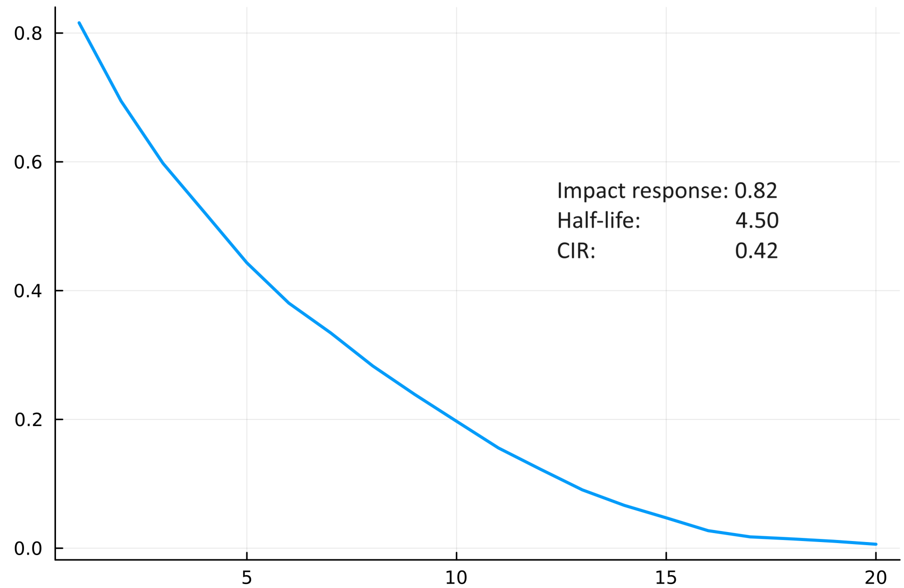
5 Monetary Non-Neutrality

Having validated the model's ability to reconcile firm-level shocks with pricing dynamics, we now assess the degree of monetary non-neutrality that the calibrated model can generate. In particular, we examine how a nominal expenditure shock, – ϵ_t in (25) – which can also be interpreted as a monetary policy shock, affects real output in the model. Monetary policy is said to be neutral if the nominal shock is fully transmitted to prices leaving real aggregate output and consumption unaffected. If prices do not respond or only partially respond to the shock, monetary policy will affect real output in the economy.

We consider four measures of non-neutrality. The first measure is the unconditional standard deviation of consumption in a long model simulation. Greater volatility in real consumption indicates less monetary neutrality and larger real effects of nominal shocks in the model because there are no other aggregate shocks that can cause fluctuations in aggregate consumption. The other three measures pertain to the Impulse Response Function (IRF) of real output in response to a monetary innovation. Specifically, we examine the response of real output to a positive nominal expenditure shock of size 0.2%, which doubles the monthly growth rate of aggregate expenditure.[26] From this IRF, we report three key

[26] As the model is solved non-linearly, response of output also display some non-linearity with respect to the size of the shock. We leave a more complete analysis of this nonlinearity to future work.

Figure 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.

outcomes. The first is the peak response of output as a fraction of the shock size, which, in a model without internal propagation, occurs on impact. The second is the half-life of the impulse response, measuring the persistence of the shock's effect. Lastly, these two measures are synthesized in the cumulative impulse response (CIR), which sums the output response as a fraction of the shock over the period it is non-zero and normalizes by the number of periods in a year—12 in our case.

Our calibrated model generates substantial volatility in aggregate consumption, with a monthly standard deviation of 0.52%.[27] Turning to output responsiveness to monetary shocks, Figure 4 displays the impulse response of real output as a fraction of the shock size for the baseline model. On impact, approximately 82% of the nominal expenditure increase translates into higher real output. The real effects of the shock diminish over time, with a half-life of 4.5 months, eventually dissipating after 20 months. The cumulative impulse response over this horizon is 0.42.

The degree of monetary non-neutrality generated by the calibrated model lies in the upper

[27]The standard deviation of detrended U.S. quarterly consumption during 1989 to 1998 is 1.5% (Golosov and Lucas, 2007). Thus, a back-of-the-envelope calculation suggests that monetary shocks in our model account for roughly a third of the consumption volatility observed in the data.

Table 4: Monetary Non-neutrality across Models

	Baseline	CES	No Demand	CES+DRS	Translog
Impact Response	0.82	0.56	0.80	0.77	0.73
CIR	0.42	0.11	0.39	0.25	0.29

The first column refers to the baseline calibrated model. The second column refers to a model with CES and only productivity shocks calibrated to the frequency and average size of price changes, which is named CES I in Appendix D.8. The third column refers to the baseline model without idiosyncratic demand shocks (Appendix D.4). The fourth column refers to the calibrated model with CES demand and decreasing returns to scale (Appendix D.3). Lastly, the fifth column refers to the model with translog demand in Appendix D.6.

range of values reported in the literature, as summarized by Mongey (2021). Moreover, the peak output response to nominal shocks in our model is comparable to results from studies that incorporate macro real rigidity as an alternative source of real transmission. For example, Nakamura and Steinsson (2010) develop a multi-sector model with production networks, introducing macro real rigidity via sticky marginal costs, and obtain a peak impulse response of 0.80. This comparison highlights that micro real rigidity, as represented in our model, can generate a degree of non-neutrality comparable to models with macro real rigidities, while simultaneously aligning with micro-data on firm dynamics and pricing.[28]

It is useful to highlight the two key sources of amplification in our setup relative to a simple menu cost model (Golosov and Lucas, 2007). The first source is micro real rigidities, which introduce strategic complementarities in pricing via the Kimball demand system. With micro real rigidities, firms adjusting their prices tend to make smaller adjustments than they would under a CES demand system. This behavior reflects their desire to remain closer to the prices of competitors who are not adjusting. This real rigidity amplifies the degree of monetary non-neutrality.[29]

To assess the quantitative role of micro real rigidities in generating monetary non-neutrality, we compare the baseline model to a CES menu cost model similar to Golosov and Lucas (2007). The second column in Table 4 presents the impact response and CIR for a menu cost model with CES demand and only idiosyncratic productivity shocks, cali-

[28]Table A.1 in Mongey (2021) shows that menu-cost models without real rigidities that are calibrated to the US economy, typically generate a peak output response in the range of 0.35 to 0.50. Richer models that include alternative sources of real rigidities find higher values in the 0.7-0.8 range.

[29]This intuition is formalized in Alvarez et al. (2023b), who derive analytic results in a menu cost model with strategic complementarities framed as a Mean Field Game. Their findings show that complementarity consistently increases the impulse response of output to a nominal shock at every horizon. They also show that their theoretical result holds under a Calvo model of pricing.

brated exclusively to pricing moments, without accounting for any firm dynamics moments.[30] Deviating from CES demand has substantial effects on the impact and persistence of the monetary shock. Specifically, the impact response is 45% larger, and the CIR measure is nearly four times larger.

The second factor that affects the degree of monetary non-neutrality is the strength of selection in price adjustment. The real response to nominal shocks depends not only on how many prices adjust but also on which prices adjust. In menu cost models, selection is strong: the firms that incur the fixed costs of price adjustment are precisely those with the largest desired price changes. In the absence of idiosyncratic shocks, prices respond solely to aggregate shocks, and only those prices most misaligned with the aggregate shock adjust. However, the introduction of idiosyncratic shocks weakens this selection effect, as firms now respond to both aggregate shocks and disturbances in their individual states. In fact, Nakamura and Steinsson (2010) demonstrate that, holding the frequency of price changes constant, more volatile idiosyncratic shocks lead to weaker selection and stronger non-neutrality.

In our baseline model, firms are subject to two orthogonal idiosyncratic shocks, a feature that potentially weakens the selection of price changes. To quantify the impact of this second force, the third column in Table 4 examines a version of the baseline model without idiosyncratic demand shocks. Consistent with the presence of a selection effect, we observe a slightly smaller degree of monetary non-neutrality. A back-of-the-envelope calculation reveals that less than 10% of the amplification when moving from the standard CES model to the baseline model can be attributed to the inclusion of idiosyncratic demand shocks, as disciplined by the data. The remaining 90% arises from micro real rigidity in the form of the Kimball demand system.[31]

Given the centrality of micro real rigidities in amplifying monetary non-neutrality, it is natural to ask whether this channel could be as important in other sticky price models. To explore this, Appendix D.7 uses a Calvo (1983) pricing framework in which firms receive i.i.d. shocks that determine when they can adjust their prices. This setup serves as a useful benchmark, as it eliminates the selection margin by construction. Specifically, the randomness of price adjustments means that the probability of price adjustment is independent of the size of desired price change. Therefore, any difference observed under the Kimball demand system must be attributed to micro real rigidities. In fact, we find that under our

[30] Further details on the calibration are provided in Appendix D.8, referring to the CES I model.

[31] Table A-5 shows that a model with with correlated demand and supply shocks exhibit less non-neutrality, lending support to the potential relevance of the selection channel.

calibrated idiosyncratic processes, transitioning from CES demand to the calibrated Kimball demand system nearly doubles the CIR non-neutrality measure in the Calvo model.

Finally, the last two columns of Table 4 explore alternative strategies for generating micro real rigidities. The fourth column examines a model where the concavity in the profit function arises from increasing marginal costs. Specifically, we report the degree of non-neutrality generated by the model described in Appendix D.3, which features CES demand and decreasing returns to scale at the firm level, similar to Burstein and Hellwig (2007). The fifth column presents another demand-based micro real rigidity, stemming from the translog model described in Appendix D.6. Interestingly, while both models increase the degree of non-neutrality relative to the CES framework in the second column, the amplification is noticeably milder in these alternative setups compared to the benchmark model.

6 Conclusion

We re-investigate the importance of micro real rigidities as a source of monetary non-neutrality. While prior studies, such as Golosov and Lucas (2007), demonstrate that nominal frictions alone cannot account for the observed non-neutrality, Ball and Romer (1990) highlights that adding real rigidities can amplify these effects. Among real rigidities, micro real rigidities in the form of strategic complementarities have been proposed as a key mechanism. However, the literature summarized by Nakamura and Steinsson (2010) concludes that models with micro real rigidities fail to replicate observed pricing moments for plausible parameter values.

We address these limitations by introducing a simple menu-cost model featuring a Kimball demand system with non-constant elasticity and idiosyncratic productivity and demand shocks. Crucially, our calibration draws directly from empirical estimates of firm-level productivity and demand processes that incorporate both prices and quantities. This avoids relying on ad hoc calibrations to match pricing moments and allows for a richer investigation of strategic complementarities.

Our calibrated model successfully replicates key empirical moments across the firm dynamics and pricing literatures. It matches firm-level dynamics from Foster et al. (2008), including price, quantity, and revenue responses, while also aligning with pricing facts such as the frequency, size, direction, and dispersion of price changes. Importantly, the degree of monetary non-neutrality generated by our model lies at the upper range reported in the literature, overcoming prior negative results.

The inclusion of idiosyncratic demand shocks, in conjunction with a Kimball demand system, dampens the selection effect in price adjustments, leading to stronger real output responses to aggregate shocks. This highlights the importance of jointly modeling productivity and demand shocks, addressing critiques such as those raised by Klenow and Willis (2016). Additionally, our model produces realistic pass-through rates and a markup distribution consistent with data. Moreover, our calibration of Kimball demand system offers a tractable and empirically grounded approach that DSGE models can adopt to generate levels of monetary non-neutrality consistent with U.S. microdata.

By bridging pricing, markup, and firm dynamics within a unified framework, our model provides a foundation for future research. This work opens the door to integrating real (e.g., investment, employment, entry/exit), nominal (e.g., price setting), and other firm decisions (e.g., markup adjustments, pass-through) into a cohesive modeling structure, paving the way for deeper insights into the interplay between monetary policy, firm dynamics, and macroeconomic outcomes.

References

Alvarez, Fernando E, Katarína Borovičková, and Robert Shimer, "Consistent Evidence on Duration Dependence of Price Changes," Technical Report 2023.

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

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, December 2008, *98* (5), 1998–2031.

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.

Basu, Susanto, "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare," *American Economic Review*, 1995, *85* (3), 512–531.

Beck, Günter W and Sarah M Lein, "Price elasticities and demand-side real rigidities in micro data and in macro models," *Journal of Monetary Economics*, 2020, *115*, 200–212.

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.

Blanchard, Olivier and Jordi Galí, "Real wage rigidities and the New Keynesian model," *Journal of money, credit and banking*, 2007, *39*, 35–65.

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

Calvo, Guillermo A, "Staggered prices in a utility-maximizing framework," *Journal of monetary Economics*, 1983, *12* (3), 383–398.

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 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.

Chari, Varadarajan V, Patrick J Kehoe, and Ellen R McGrattan, "Sticky price models of the business cycle: can the contract multiplier solve the persistence problem?," *Econometrica*, 2000, *68* (5), 1151–1179.

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.

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," *Review of Economic Studies*, 2024, *91* (1), 259–300.

Fisher, Jonas and Martin Eichenbaum, "Evaluating models of sticky prices," Technical Report, Society for Computational Economics 2005.

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.

Gagliardone, Luca, Mark Gertler, Simone Lenzu, and Joris Tielens, "Anatomy of the Phillips Curve: Micro Evidence and Macro Implications," Technical Report, National Bureau of Economic Research 2023.

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

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 _ , "In search of real rigidities," *NBER Macroeconomics Annual*, 2011, *25* (1), 261–310.

_ , _ , 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.

Julio, Juan Manuel, Héctor Manuel Zárate, and Manuel Darío Hernández, "Rigideces de precios al consumidor en Colombia," 2011.

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.

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.

Ridder, Maarten De, Basile Grassi, and Giovanni Morzenti, “The Hitchhiker’s Guide to Markup Estimation: Assessing Estimates from Financial Data,” 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.

Woodford, Michael, *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2003.

Internet Appendix (For Online Publication)

A Derivations

A.1 Demand Elasticity

For our specification of Kimball demand, the price elasticity of demand depends on p/P, n, and $\frac{ny}{Y}$ in the following ways

$$\begin{aligned}
\frac{\partial \epsilon}{\partial\left(p/P\right)} &= \frac{\psi\omega^2}{\lambda n\left(\frac{ny}{Y}\right)^2\left(1-\omega\right)^2}\left(\frac{p}{\lambda nP}\right)^{\frac{\omega(1+\psi)}{1-\omega}-1} \\
\frac{\partial \epsilon}{\partial n} &= -\frac{\psi\omega Y}{n^2 y\left(\omega-1\right)} \\
\frac{\partial \epsilon}{\partial(ny/Y)} &= -\frac{\psi\omega}{\left(\frac{ny}{Y}\right)^2\left(\omega-1\right)}
\end{aligned}$$

A.2 Symmetric Equilibrium

In a symmetric equilibrium, all firms have identical demand ($n^i = n$), productivity ($z^i = z$), and prices ($p^i = p$). As a result, the aggregate price index under our specification of $G(.)$ reduces to $P = \frac{1}{1+\psi}\left[\int\limits_0^1 \left(\frac{p}{n}\right)^{\frac{1+\omega\psi}{1-\omega}} di\right]^{\frac{1-\omega}{1+\omega\psi}} + \frac{\psi}{1+\psi}\int\limits_0^1 \frac{p}{n} di = \frac{p}{n}$.

Substituting $P = \frac{p}{n}$ into the expression for $\lambda = \left[\int\limits_0^1 \left(\frac{p}{nP}\right)^{\frac{1+\omega\psi}{1-\omega}} di\right]^{\frac{1-\omega}{1+\omega\psi}}$ yields $\lambda = 1$ in a symmetric equilibrium.

Finally, $\int_0^1 \frac{ny}{Y} di = \frac{ny}{Y} = 1$ holds in the symmetric equilibrium because all firms have the same effective market share.

A.3 Pass-through

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

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

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

$$\frac{1}{p_i - \frac{W}{z_i}} + \frac{\frac{\omega(1+\psi)}{1-\omega}\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}-1}\frac{1}{\lambda n_i P}}{\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}}+\psi} \quad = \quad 0 \tag{A-2}$$

$$\left(\frac{p_i}{\lambda n_i P}\right)^{\frac{\omega(1+\psi)}{1-\omega}}\left[1-\left(\frac{W}{z_i}-p_i\right)\left(\frac{\omega\left(1+\psi\right)}{1-\omega}\frac{1}{p_i}\right)\right] \quad = \quad -\psi \tag{A-3}$$

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.4 Π_{13} and Π_{11}

Let $\Pi_{pz} = \frac{\partial^2 \Pi}{\partial(p_i/P)\partial n_i}$ and $\Pi_{pn} = \frac{\partial^2 \Pi}{\partial(p_i/P)\partial n_i}$.

$$\frac{\partial \Pi_{pz}}{\partial \psi} = \frac{WY\omega^2 \log\left(\frac{p_i}{P\lambda n_i}\right)}{\left(\lambda n_i^2 z_i^2 \left(\omega - 1\right)^2\right)\left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(1+\psi)}{\omega-1}-1}} \tag{A-6}$$

$$
\begin{aligned}
\frac{\partial \Pi_{pn}}{\partial \psi} =& \frac{Y\left(\psi + \frac{1}{\left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(\psi+1)}{\omega-1}}}\right)}{n_i^2(\psi+1)^2} + \frac{Y\left(\frac{\omega \log\left(\frac{p_i}{P\lambda n_i}\right)}{(\omega-1)\left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(\psi+1)}{\omega-1}}} - 1\right)}{n_i^2(\psi+1)} \\
&- \frac{Y\omega^2 \frac{p_i}{P}\left(\frac{p_i}{P} - \frac{W}{z_i}\right)}{\lambda^2 n^4 (\omega-1)^2 \left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(\psi+1)}{\omega-1}+2}} - \frac{2Y\omega^2 \log\left(\frac{p_i}{P\lambda n_i}\right)\left(\frac{p_i}{P} - \frac{W}{z_i}\right)}{\lambda n^3 (\omega-1)^2 \left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(\psi+1)}{\omega-1}+1}} \\
&- \frac{Y\omega^2 \frac{p_i}{P} \log\left(\frac{p_i}{P\lambda n_i}\right)}{\lambda n^3 (\omega-1)^2 \left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(\psi+1)}{\omega-1}+1}} + \frac{Y\omega^2 \frac{p_i}{P} \log\left(\frac{p_i}{P\lambda n_i}\right)\left(\frac{\omega(\psi+1)}{\omega-1}+1\right)\left(\frac{p_i}{P} - \frac{W}{z_i}\right)}{\lambda^2 n^4 (\omega-1)^2 \left(\frac{p_i}{P\lambda n_i}\right)^{\frac{\omega(\psi+1)}{\omega-1}+2}}
\end{aligned} \tag{A-7}
$$

As we show in Appendix A.2, $\lambda = 1$ and $\frac{p}{P} = n$ in a symmetric equilibrium. When this is the case, the condition $\frac{\partial \Pi_{pz}}{\partial \psi} = 0$ holds (solid blue line in Figure A-2b). Consequently, as ψ becomes more negative, $|\Pi_{11}|$ increases (Figure A-2a) and the responsiveness of prices to productivity shocks ($\zeta = -\frac{\Pi_{13}}{\Pi_{11}}$) is increasingly muted.

Meanwhile for demand shocks, the cross-derivative Π_{pn} varies with ψ. In particular, $\frac{\partial \Pi_{pn}}{\partial \psi} = \frac{-Y\omega^2 \frac{p_i}{P}\left(\frac{p_i}{P} - \frac{W}{z_i}\right)}{\lambda^2 n^4 (\omega-1)^2} < 0$. For more negative values of ψ, both Π_{pn} and $|\Pi_{11}|$ increases. As illustrated in Figure A-2, Π_{pn} increases at a faster rate than $|\Pi_{11}|$ as ψ decreases, so ζ in Equation (4) increases signifying larger pass-through of demand shocks to desired prices. This highlights that while stronger micro real rigidities mute the response of prices to productivity, they raise price responsiveness to idiosyncratic demand.

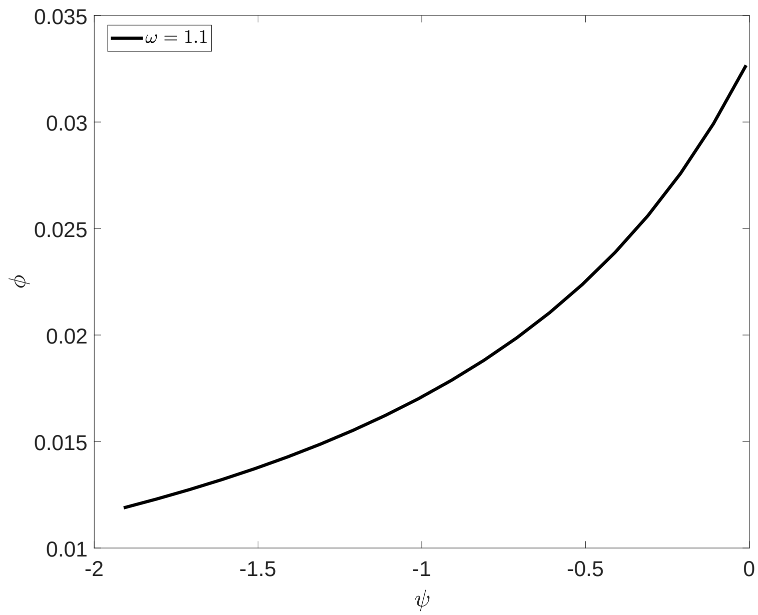
B Quantitative Model

B.1 Rewriting the Problem

Firms need to observe $\mathcal{S}$ and understand its law of motion 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.

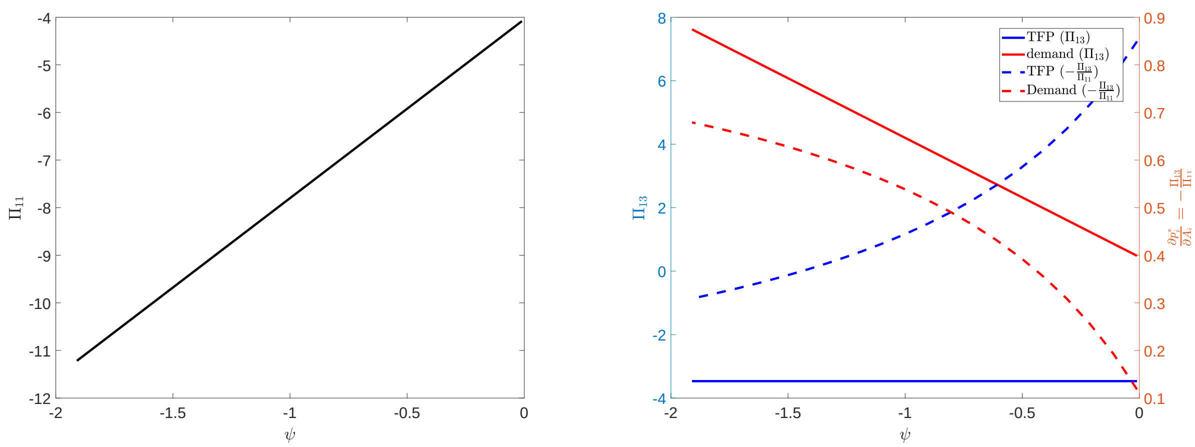
Because money supply $S_t = P_t C_t$ exhibits positive growth on average, nominal prices will increase over time. To ensure that the state variables remain stationary, we normalize all

Figure A-1: Degree of real rigidity ϕ



Note: This plots ϕ in Equation (3) against ψ under a symmetric equilibrium where all firms are identical.

Figure A-2: Π_{11} and Π_{13}



(a) Π_{11} on ψ

(b) Π_{13} and ζ for TFP and demand shocks

Note: The left panel plots Π_{11} against ψ under a symmetric equilibrium. The right panel plots Π_{13} and ζ in Equation (4) against ψ for productivity and demand separately.

nominal variables by S_t. Consequently, we can rewrite the firm's profit function

$$\pi\left(\frac{p_t^i}{S_t}; n_t^i, z_t^i, \frac{P_t}{S_t}, \lambda_t\right) = \left(\frac{p_t^i/S_t}{P_t/S_t} - \frac{1}{z_t^i P_t/S_t}\right)\frac{(P_t/S_t)^{-1}}{n_t^i(1+\psi)}\left[\left(\frac{p_t^i/S_t}{\lambda_t n_t^i (P_t/S_t)}\right)^{\frac{\omega(1+\psi)}{1-\omega}} + \psi\right] \quad \text{(A-8)}$$

where we 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

$$\log\left(\frac{P_t}{S_t}\right) = F\left(\frac{P_{t-1}}{S_t}\right) = \alpha_0 + \alpha_1 \log\left(\frac{P_{t-1}}{S_t}\right)$$
$$\log\left(\lambda_t\right) = G\left(\frac{P_{t-1}}{S_t}\right) = \beta_0 + \beta_1 \log\left(\frac{P_{t-1}}{S_t}\right)$$

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

$$\begin{aligned}\log\left(\frac{P_t}{S_{t+1}}\right) &= \log\left(\frac{P_t}{S_t}\right) + \log\left(\frac{S_t}{S_{t+1}}\right)\\ &= \alpha_0 + \alpha_1 \log\left(\frac{P_{t-1}}{S_t}\right) - \left(\mu + \sigma_S \epsilon_{t+1}\right)\end{aligned}$$

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

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

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

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

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

$$V\left(\frac{p_{t-1}^i}{S_t}; n_t^i, z_t^i, \frac{P_{t-1}}{S_t}\right) = \max\left[V_N\left(\frac{p_{t-1}^i}{S_t}; n_t^i, z_t^i, \frac{P_{t-1}}{S_t}\right), V_A\left(n_t^i, z_t^i, \frac{P_{t-1}}{S_t}\right)\right]$$

B.2 Computational Strategy

A sketch of the computation algorithm is as follows. We first make guesses of the coefficients $(\alpha_0^0, \alpha_1^0, \beta_0^0, \beta_1^0)$ 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 $(\alpha_0^1, \alpha_1^1, \beta_0^1, \beta_1^1)$ which is then used to 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

C.1 Empirical Moments

C.1.1 Firm-Level Productivity and Demand Processes

Using the quinquennial Census of Manufactures between 1977 to 1997, Foster et al. (2008) estimate firm-level productivity and demand for eleven product markets with minimal vertical differentiation.[32] With 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:

$$TFPQ_{it} = \ln q_{it} - \alpha_l \ln l_{it} - \alpha_k \ln k_{it} - \alpha_m \ln m_{it} - \alpha_e \ln e_{it}, \tag{A-9}$$

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

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

revenue-based TFP, which is derived similarly but replaces the quantity produced with the firm's revenue:,

$$TFPR_{it} = \ln p_{it}q_{it} - \alpha_l \ln l_{it} - \alpha_k \ln k_{it} - \alpha_m \ln m_{it} - \alpha_e \ln e_{it}. \tag{A-10}$$

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

$$\ln q_{it} = \alpha_0 + \alpha_1 \widehat{\ln p_{it}} + \sum_t \alpha_t \text{YEAR}_t + \alpha_2 \ln(\text{INCOME})_{mt} + n_t^i, \tag{A-11}$$

using an instrumental variable regression. Here, the log-price, $\ln p_{it}$, is instrumented by the estimate of TFPQ from (A-9), which acts as a supply shifter. The regression includes time fixed effects and the average income in a plant's local market, m, defined using the Bureau of Economic Analysis' Economic Areas. The residual from this equation is interpreted as a pure demand shifter for the firm.

For the eleven products analyzed, Foster et al. (2008) report average five-year autocorrelations of 0.32 for idiosyncratic TFPQ and 0.62 for demand. The cross-sectional dispersion of TFPQ and demand are 0.26 and 1.16, respectively. This indicates that demand shocks are more persistent and more dispersed across firms. Additionally, 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.

C.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.[33] 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.

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

C.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 time period 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.[34] 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. Amiti et al. (2019) find a 60% cost pass-through – in the higher end among estimates in the literature – and a 40% pass-through of competitors' price changes using Belgian manufacturing data. Recent studies using merged data on both costs and prices recover cost pass-through estimates that are similar to the 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. Gagliardone et al. (2023) uses a structural model and Belgian manufacturing microdata and find the cost pass-through in the range of 0.35 to 0.47 across different model specifications. 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%.

[34] 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.

C.2 Calibration Details

The model-based moments we need for calibration are 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,

$$
\begin{aligned}
TFPQ_{it} &= \ln q_{it} - \ln l_{it}, && \text{(A-12)}\\
TFPR_{it} &= \ln (p_{it} q_{it}) - \ln l_{it}. && \text{(A-13)}
\end{aligned}
$$

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),

$$
\ln q_{it} = \alpha_0 + \alpha_1 \widehat{\ln p_{it}} + \text{Time FE} + \eta_{it}, \tag{A-14}
$$

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.

C.3 Identification of Model Parameters

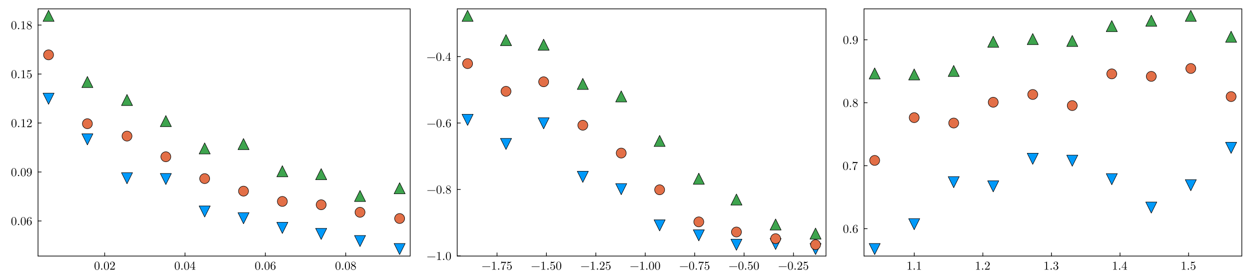
In order to demonstrate that the calibration targets are indeed informative for the respective parameters we borrow an exercise from Daruich (2022), which mimics the first stage of the multistart global optimization proposed by Arnoud et al. (2022).

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.[35]

[35] A uniform Sobol sequence (Sobol, 1967) is a sequence of points that spans the n-dimensional hypercube

Figure A-3: 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.

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.

Figure A-3 plots the values of three key model-implied target moment against the values of the parameter it is assigned to.[36] 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

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.

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

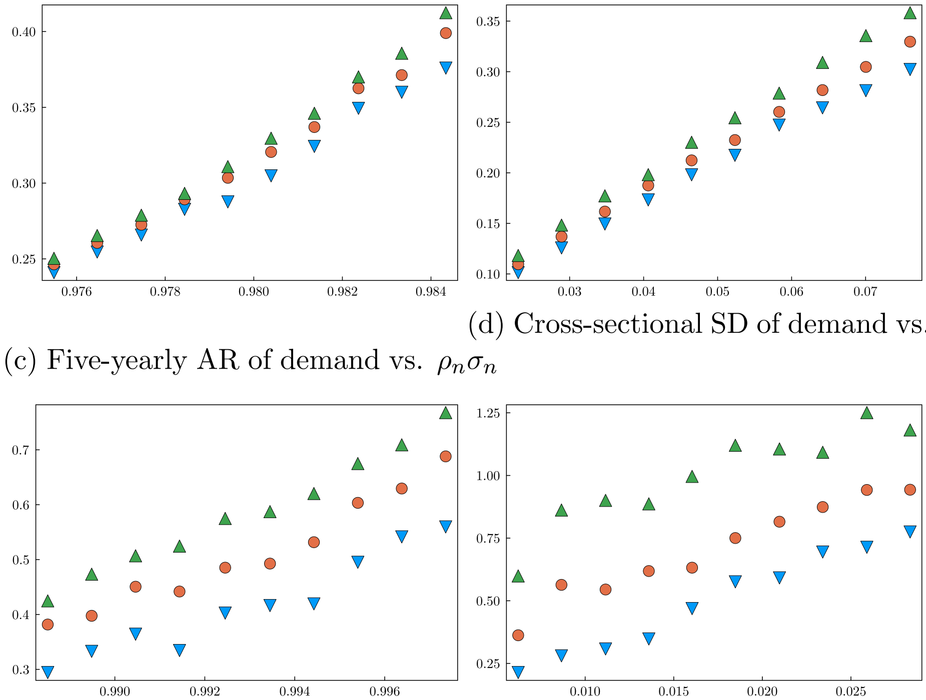
Figure A-4: Identification of Internally-Calibrated Parameters

(b) Cross-sectional SD of TFPQ vs.

(a) Five-yearly AR of TFPQ vs. ρ_z σ_z

(d) Cross-sectional SD of demand vs.

(c) Five-yearly AR of demand vs. $\rho_n \sigma_n$



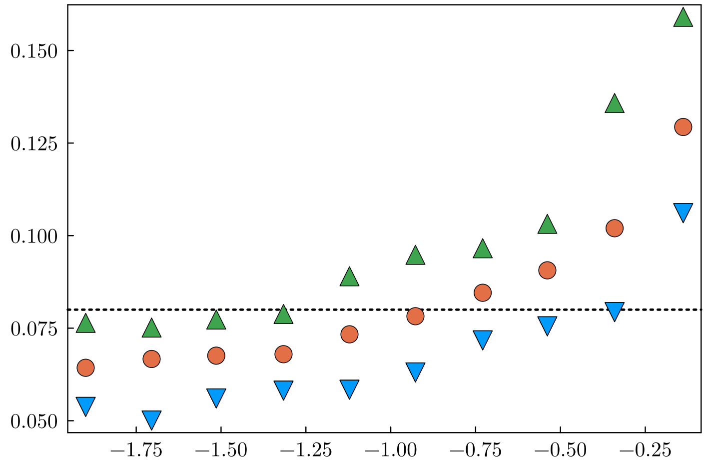
Note: For each decile of a given parameter plotted on the horizontal axis, the red dot shows the median of the moment that is assigned to the parameter. The blue down-pointing triangles and green up-pointing triangles show the 25^{th} and 75^{th} percentiles respectively.

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.

Figure A-4 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-year 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.

The key takeaway from this exercise is that the links between the parameters and mo-

Figure A-5: 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 A-3. Specifically, they are random draws from a hypercube of parameter space.

ments are quite tight. While it is too computationally intensive, if one were to consider a 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.

C.4 Alternative Calibration Strategies

We explore two alternative calibration strategies. In the first one, we target the size rather than the frequency of price changes as we did in the baseline calibration. Second, as it is well known that the distribution of price changes implied by a standard menu-cost model with Gaussian shocks exhibits negative excess kurtosis in contrast to the positive excess kurtosis found in the empirical distribution, we consider an alternative calibration with leptokurtic demand shocks which can better match the empirical kurtosis.[37]

The results are presented in Table A-1. 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. The third column reports the results from the calibration where we target the size of non-zero price changes. This alternative calibration delivers a similar fit to the data and the calibrated parameters do not differ much from the baseline. At the same

[37] 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).

Table A-1: Baseline Calibration and Alternative Calibrations

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-year 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-year 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.

time, it is able to match both the frequency of price changes and fraction of price changes that are price increases – both untargeted – well. The frequency is price changes is is lower than the baseline model at 9% in order to generate larger price adjustments on average.

The last column in Table A-1 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 (20) 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.[38] This calibration delivers

[38]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.

Source	Elasticity	Super-elasticity
Bergin and Feenstra (2000)	3	1.3
Chari et al. (2000)	10	385
Fisher and Eichenbaum (2005)	11	10
Gopinath and Itskhoki (2010)	5	4
Kimball (1995)	11	471
Klenow and Willis (2016)	5	10
Woodford (2003)	7.8	6.7
Beck and Lein (2020)	3.2	1.93
Harding et al. (2022)	11	138.1
Harding et al. (2023)	2.6	26.4
This paper	4.4	5.7

Table A-2: Elasticity and Super-elasticity Specifications from the Literature

Note: This table summarizes the demand elasticity and super-elasticity implied by the calibration of non-CES demand systems in the literature. The numbers reported for Beck and Lein (2020) are their median empirical estimates.

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.

C.5 Comparison with Literature

Our baseline calibration of ψ and ω implies a demand elasticity of 4.4 and super-elasticity of 5.7. Table (A-2) summarizes the implied elasticities and super-elasticities from the literature. Among the specifications surveyed, demand elasticities range from 3 to 11. The super-elasticity features much more variation from as low as 1.3 to as high as 471, while the majority are in the range from 2 to 10. In fact, the implied demand curvature in our baseline calibration is broadly consistent with the literature and does not represent an outlier.

D Robustness

D.1 Colombian Data

Drawing upon empirical firm dynamics evidence based on narrowly defined manufacturing industries in the United States (Foster et al., 2008), our baseline analysis arrives at the conclusion that a menu cost model featuring micro real rigidities and carefully calibrated supply and demand processes can indeed be consistent with non-targeted firm-level pricing facts. To test the robustness of this result, we apply our model to data from Colombian manufacturing firms as reported in Eslava et al. (2024). Like Foster et al. (2008), Eslava et al. (2024) uses price and quantity data of manufacturing firms inputs and outputs to estimate firm-level productivity and demand processes. However, instead of separately estimating firms' production and demand function using an instrumental variable strategy, Eslava et al. (2024) utilizes a Generalized Method of Moments (GMM) approach to jointly estimate productivity and demand. This method allows for a more flexible analysis by relaxing the orthogonality assumption between productivity and demand shocks. Furthermore, the analysis is performed using data drawn from a broad set of manufacturing industries and products rather than from selected industries with homogeneous products.

Using the empirical estimates reported in Eslava et al. (2024) for Colombia, we recalibrate our baseline model following a similar calibration strategy as in the main analysis. Specifically, we choose the parameters $\{f, \rho_z, \sigma_z, \rho_n, \sigma_n, \omega, \psi, \sigma_{zn}\}$ to match seven empirical moments in Eslava et al. (2024), as well as the price adjustment frequency in Colombia. The firm dynamics moments that we use to discipline the model are largely the same as those we use in the main analysis with one exception. Because Eslava et al. (2024) do not report statistics on TFPR, we replace $corr(TFPR, TFPQ)$ with the correlation between firm output and markup as a data target. For the price adjustment frequency in Colombia, we reference statistics from Julio et al. (2011) who compute micro pricing moments using Colombian CPI data. Between 1999 and 2008, the average monthly price adjustment frequency for goods in the CPI basket is approximately 10% to 15%.

Table A-4 reports the internal calibration to Colombian data. Overall, the model is able to fit Colombia firm dynamics statistics well. In terms of firm-level idiosyncratic processes, we observe a number of similarities between the calibration to Colombian data and the calibration in our main analysis using US data. First of all, the calibrated process for idiosyncratic productivity and demand are both highly persistent, with a monthly autocor-

Table A-3: Externally Calibrated Parameters with Colombian Data

Parameter	Description	Value	Source
β	Discount Factor	0.9966	Annual discount rate of 4%
χ	Labor disutility	1	Normalization
μ	Growth rate of S	0.0056	Annual inflation rate of 7%
σ_S	SD of shocks to nom. expenditure	0.0104	Vavra (2014)

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

relation of 0.994 for demand and 0.987 for productivity.[39] In both samples, idiosyncratic demand exhibits more persistence compared to idiosyncratic productivity. Also, in both calibrations, the standard deviation of shocks to productivity is larger than that of shocks to demand. In terms of the curvature of the demand system, the calibrated values of ψ and ω imply a high cost passthrough to price of 70%, as compared to 38% in our baseline calibration.

In the main analysis, the model calibrated only to firm dynamics moments and price adjustment frequency is able to match untargeted pricing moments well. Interestingly, we find that this also holds in the alternative calibration using Colombian data. In the calibrated model, the average size of price adjustment is 20% – which is roughly in line with the size distribution reported in Figure 8 of Julio et al. (2011). In terms of real response to nominal shocks, we find that the model calibrated to Colombian data exhibit less monetary non-neutrality due to smaller degree of strategic complementarity. While the impact response is high at 0.75, the cumulative impact response is 0.18 and the half-life is 3 months.

D.2 Correlated Shocks

In this section, we relax the assumption of $\rho_{zn} = \frac{\sigma_{nz}}{\sigma_z \sigma_n} = 0$ and allow stochastic innovations to firm demand and productivity to be correlated. In particular, we solve the model with two levels of correlation between demand and productivity shocks $\rho_{zn} = \{-0.4, 0.4\}$. In doing so, we keep all other parameters fixed at their baseline values but vary the menu cost f to keep the frequency of price changes identical across all specifications. The results are summarized in Table (A-5). They show that even with the mildly large correlations assumed, our conclusions in the main text are robust.

[39] Table A-4 report the annual autocorrelations of productivity and demand.

Table A-4: Internal Calibration with Colombian Data

Moment	**Data**	**Model**
Frequency of price changes	0.10 - 0.15	0.13
Yearly autocorrelation of z_t^i	0.91	0.91
Cross-sectional standard deviation of z_t^i	0.75	0.76
Yearly autocorrelation of n_t^i	0.98	0.98
Cross-sectional standard deviation of n_t^i	0.89	0.87
Corr between demand and TFPQ	–0.07	–0.12
Corr between price and TFPQ	–0.73	–0.75
Corr between output and markup	0.45	–0.50
Parameter	**Description**	**Value**
ψ	Super-elasticity	–0.30
ω	Elasticity	1.69
ρ_z	Persistence of z_t^i	0.987
σ_z	Standard deviation of z_t^i	0.11
ρ_n	Persistence of n_t^i	0.994
σ_n	Standard deviation of n_t^i	0.08
f	Menu cost	0.025
ρ_{zn}	Correlation between ϵ_n and ϵ_z	–0.07

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

Table A-5: Model Summary with Correlated Supply and Demand Shocks

Moment	**Data**	$\rho_{zn} = 0$	$\rho_{zn} = 0.40$	$\rho_{zn} = -0.40$
5-year autocorr of z_t^i	0.32	**0.32**	0.32	0.32
Cross-sectional SD of z_t^i	0.26	**0.25**	0.28	0.29
5-year autocorr of n_t^i	0.62	**0.62**	0.82	0.76
Cross-sectional SD of n_t^i	1.16	**1.05**	1.28	0.99
Corr b/w TFPR and TFPQ	0.75	**0.74**	0.72	0.67
Corr b/w price and TFPQ	–0.54	**–0.57**	–0.42	–0.63
Frequency of Price Changes	0.11	**0.11**	**0.11**	**0.11**
Average Size of Price Changes	0.08	0.07	0.05	0.08
Average Markup	1.56	1.42	1.53	1.46
Cross-sectional SD of Markup	0.72	0.39	0.51	0.40
Impact Response		0.83	0.77	0.77
CIR		0.43	0.43	0.36

D.3 Model with CES Demand and Decreasing Returns to Scale

In this section, we consider an alternative source of micro real rigidity arising from decreasing returns to scale technology. To do this, we make two changes to our baseline model. Instead of Kimball demand, intermediate goods producers face CES demand with a demand shifter

$$y_{it} = n_{it} Y_{it} \left(\frac{p_i}{P}\right)^{-\theta} \tag{A-15}$$

In addition, the production technology exhibits decreasing returns to scale ($\alpha < 1$).[40]

$$y = z_{it} l_{it}^{\alpha} \tag{A-16}$$

Besides these two modifications, all other aspects of the model remain unchanged and the resulting structure is identical to the baseline model in Burstein and Hellwig (2007).

This version of the model share many properties as our baseline model, which is unsurprising given the discussion in Section 2.2 that both curvature in the demand function and marginal cost function can raise the concavity of the profit function, which is central to micro real rigidity. In particular, the two model specifications deliver similar implications for the pass-through of productivity and demand shocks to prices.

To see this, consider again the static price-setting problem of a firm under flexible prices. The first-order condition to the static profit-maximization problem is

$$(1-\theta)\, p_{it}^{-\theta} n_{it} Y_t P_t^{\theta-1} + \frac{\theta}{\alpha} \frac{W_t}{P_t} \left(\frac{n_{it}}{z_{it}} Y_t P_t^{\theta}\right)^{\frac{1}{\alpha}} p_{it}^{\frac{-\theta}{\alpha}-1} = 0 \tag{A-17}$$

Log-linearizing (A-17) around a symmetric steady state yields the following expression for the optimal price:

$$\hat{p}^* = \frac{1-\alpha}{\alpha\theta-\theta-\alpha}\left(\hat{n}+\hat{Y}\right) + \frac{\alpha\theta-\theta}{\alpha\theta-\theta-\alpha}\hat{P} - \frac{\alpha}{\alpha\theta-\theta-\alpha}\hat{W} + \frac{1}{\alpha\theta-\theta-\alpha}\hat{z} \tag{A-18}$$

where hatted variables denote log-deviations from the steady state.

The cost and demand pass-throughs under this specification are then given by

$$\frac{\partial \hat{p}_i^*}{\partial \widehat{mc}} = \frac{-1}{\alpha\theta-\theta-\alpha} \tag{A-19}$$

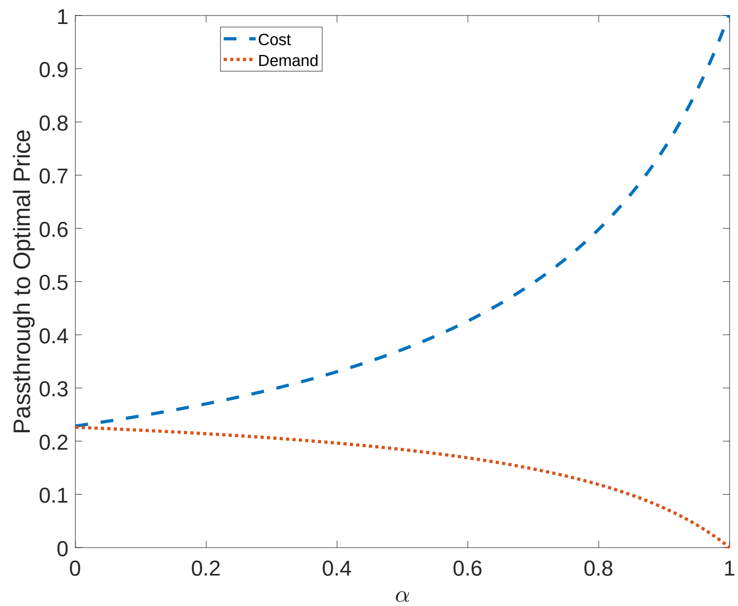
$$\frac{\partial \hat{p}_i^*}{\partial \hat{n}_i} = \frac{1-\alpha}{\alpha\theta-\theta-\alpha}, \tag{A-20}$$

Notice that with constant returns to scale ($\alpha = 1$), the cost pass-through to price is complete and the demand pass-through to price is zero. When the returns to scale parameter

[40] Note that a setup with CES demand and increasing returns to scale leads to strategic substitution in pricing. Even though we do not impose $\alpha < 1$, it is needed to match the model to data.

α falls below one so that technology exhibits decreasing returns to scale, the cost pass-through becomes smaller while demand begins to matter for optimal pricing as shown in Figure (A-6).

Figure A-6: Pass-through of Demand and Cost Shocks to Price



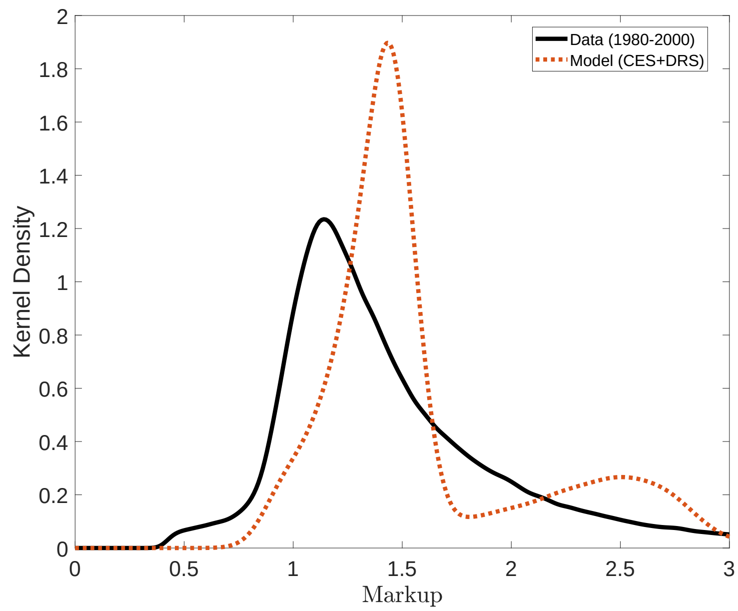
This figure plots the cost and demand pass-through in a model with CES demand with varying levels of returns to scale α, holding θ fixed.

Next, we explore if this alternative setup would have been successful at matching untargeted pricing moments if calibrated using our strategy to the same empirical data. In doing so, we replace the two parameters pertaining to the Kimball demand system (ω, ψ) with the elasticity of demand θ and return to scale parameter α while targeting the exact same set of data moments used in our main analysis.[41] We first note that a model with CES demand, idiosyncratic demand shifters, and decreasing returns to scale technology can be calibrated to the targeted moments as well as our baseline model, as shown in Table (A-6). It is worth pointing out that the correlations between TFPR and TFPQ, and between price and TFPQ, turn out to be informative about the elasticity of substitution θ and returns to scale parameter α.

With respect to the untargeted pricing moments, the calibrated model is able to match the average size of price changes, the fraction of changes that are upward, and the dispersion

[41]Externally calibrated parameters are held fixed at the values used in the main analysis.

Figure A-7: 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 with CES demand and decreasing returns to scale technology. Both kernel densities are computed using the optimal bandwidth for normal densities.

of price changes in the data almost exactly as shown in the bottom panel of Table (A-6). Furthermore, it is interesting that the pass-through of supply shocks implied by the calibration turns out to be 38%, which is identical to the pass-through implied by our baseline model with Kimball demand. Furthermore, this number is close to the value of 40% in the preferred calibration of Burstein and Hellwig (2007) even though they use a different calibration strategy than ours.

However, this model specification has two shortcomings compared to our baseline model. Firstly, although the average and cross-sectional dispersion of the model-implied markup distribution compare well with the data, the shape of the model-implied distribution exhibits a large hump near the frictionless desired markup which is counterfactual to the data.[42] Also, the calibrated returns to scale parameter of 0.21 appears to be implausibly low given that the literature typically finds constant returns to scale among manufacturing firms (Foster et al., 2008; Eslava et al., 2024). Even if one includes fixed or quasi-fixed inputs in the production function, a labor-share of 0.21 is fairly low.

[42]The frictionless desired gross markup under CES is given by $1 + \frac{\theta}{\theta-1} = 1.48$.

Table A-6: Internal Calibration: CES Demand with DRS

Parameter	Description	Value
θ	Elasticity of Substitution	3.1
α	Returns to scale	0.21
ρ_z	Persistence of z_t^i	0.98
σ_z	Standard deviation of z_t^i	0.053
ρ_n	Persistence of n_t^i	0.992
σ_n	Standard deviation of n_t^i	0.062
f	Menu cost	0.12
Targeted Moment	**Data**	**Model: CES+DRS**
Frequency of price changes	0.11	0.11
5-year autocorrelation of z_t^i	0.32	0.32
Cross-sectional standard deviation of z_t^i	0.26	0.26
5-year autocorrelation of n_t^i	0.62	0.64
Cross-sectional standard deviation of n_t^i	1.16	1.01
Corr between TFPR and TFPQ	0.75	0.75
Corr between price and TFPQ	–0.54	–0.55
Unargeted Moment		
Average Size of Price Changes	0.08	0.08
Fraction Up	0.65	0.64
SD(Δp)	0.08	0.08
Pass-through of Supply Shocks	20%-40%	38%
Average Markup	1.56	1.59
Cross-sectional SD of Markup	0.72	0.53
Impact Response	NA	0.77
CIR	NA	0.38

This table reports the calibration of the model with CES demand and decreasing returns to scale technology. The top panel reports the values of the calibrated parameters. The middle panel reports the targeted moments in the data and the model. The bottom panel reports various untargeted moments in the data and their model counterparts.

D.4 Baseline Model without Demand Shocks

Table (A-7) considers a modification to the baseline model without demand shocks. In this version of the model, σ_n is set to zero while all other parameters are fixed at the baseline calibration with the exception of f which is recalibrated to maintain the same frequency of price adjustment as in the baseline model.

The third and fourth columns in Table (A-8) report the pricing moments of two versions of the model with idiosyncratic demand or idiosyncratic productivity turned off, holding all other parameters fixed at the baseline value. Without demand shocks, frequency of price

Table A-7: Baseline Model without Demand

Moment	**Data**	**Baseline**	**No Demand**
5-year autocorr of z_t^i	0.32	**0.32**	0.32
Cross-sectional SD of z_t^i	0.26	**0.25**	0.26
5-year autocorr of n_t^i	0.62	**0.62**	0.04
Cross-sectional SD of n_t^i	1.16	**1.05**	0.42
Corr b/w TFPR and TFPQ	0.75	**0.75**	0.99
Corr b/w price and TFPQ	–0.54	**–0.57**	–0.98
Frequency of Price Changes	0.11	**0.12**	**0.12**
Average Size of Price Changes	0.08	0.07	0.06
Pass-through of Supply Shocks	20%-40%	38%	38%
Average Markup	1.56	1.42	1.38
Cross-sectional SD of Markup	0.79	0.39	0.28
Impact Response	NA	0.82	0.80
CIR	NA	0.42	0.39

Note: Boldface denotes calibration targets for each model.

Table A-8: Untargeted Pricing Moments

Moments	Data	Baseline	No Demand	No TFP
Frequency	0.11	0.12	0.10	0.06
Average Size	0.08	0.07	0.06	0.06
Fraction Up	0.65	0.58	0.66	0.73
SD(Δp)	0.08	0.07	0.06	0.05

Note: This table shows the four pricing moments: frequency of price adjustments, 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. The first column refers to the empirical data, the second column refers to the baseline model, and the third and last column refer to the model without idiosyncratic demand and productivity shocks respectively.

changes decreases slightly from 0.12 to 0.1 while the size of price changes drops from 0.07 to 0.06. Without productivity shocks, frequency goes down by half – due to the persistent nature of the demand process – while size of price changes only decreases slightly to 0.06. Of course we should emphasize that here we hold the parameters constant based on our baseline calibration, which was only feasible because we had demand shocks.

Table A-9: Alternative Model: Klenow-Willis

Moment	Data	Baseline	KW (2016)
5-year autocorr of z_t^i	0.32	**0.32**	0.00
Cross-sectional SD of z_t^i	0.26	**0.25**	0.27
5-year autocorr of n_t^i	0.62	**0.62**	0.00
Cross-sectional SD of n_t^i	1.16	**1.05**	0.23
Corr b/w TFPR and TFPQ	0.75	**0.75**	0.98
Corr b/w price and TFPQ	–0.54	**–0.57**	–0.77
Frequency of Price Changes	0.11	**0.12**	**0.09**
Average Size of Price Changes	0.08	0.07	0.14
Pass-through of Supply Shocks	20%-40%	38%	28%
Average Markup	1.56	1.42	1.64
Cross-sectional SD of Markup	0.79	0.39	0.57
Impact Response	NA	0.82	0.85

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.

D.5 Klenow and Willis (2016)

We solve our model using the calibration of Klenow and Willis (2016) in which they target a frequency of price change of 0.09.[43] The results are summarized in Table (A-9).

This model has three 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-year 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

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

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.

D.6 Translog Demand System

In addition to Kimball demand, we explore the translog demand system which is frequently used in the trade and 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

$$\ln P_t = \sum_{i=1}^{n} \alpha_i \ln P_{it} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \ln P_{it} \ln P_{jt} \tag{A-21}$$

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 at the back of this section 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 A-10 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

Table A-10: Alternative Model with Translog Demand

Moment	Data	Baseline	Translog
5-year autocorr of z_t^i	0.32	**0.32**	**0.29**
Cross-sectional SD of z_t^i	0.26	**0.25**	**0.27**
5-year autocorr of n_t^i	0.62	**0.62**	**0.68**
Cross-sectional SD of n_t^i	1.16	**1.05**	**1.11**
Corr b/w TFPR and TFPQ	0.75	**0.75**	**0.55**
Corr b/w price and TFPQ	–0.54	**–0.57**	**–0.76**
Frequency of Price Changes	0.11	**0.12**	**0.10**
Average Size of Price Changes	0.08	0.07	0.12
Pass-through of Supply Shocks	20%-40%	38%	60%
Average Markup	1.56	1.42	1.46
Cross-sectional SD of Markup	0.79	0.39	0.31
Impact Response	NA	0.82	0.73

Note: Boldface denotes calibration targets.

out to be 60%, well outside the relevant range. In fact, we show in Appendix D.6 that cost pass-through in the translog model is constrained to be between 50% and 100%.[44] Relatedly, given the size 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.

[44] 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%.

D.6.1 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

$$\epsilon(x) \equiv -\frac{p(x)}{xp'(x)} > 0 \text{ and } \rho(x) \equiv -\frac{xp''(x)}{p'(x)} \tag{A-22}$$

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

$$\rho^T(\epsilon) = \frac{3\epsilon - 1}{\epsilon^2} \tag{A-23}$$

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

$$\rho^K(\epsilon, b) = \frac{(1-b)\epsilon + 1}{\epsilon} \tag{A-24}$$

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

$$b(\epsilon) = \left(\frac{\epsilon - 1}{\epsilon}\right)^2 \tag{A-25}$$

So far we used the notation of Mrázová and Neary (2017). To convert the restriction in (A-25) to our notation, based on the expressions of the elasticity of demand and super-elasticity for our Kimball aggregator specification, we get

$$\epsilon = \frac{\omega}{\omega - 1} \text{ and } \psi = -b \tag{A-26}$$

Using the relationships in (A-26) in (A-25) we obtain

$$\psi = -\frac{1}{\omega^2} \tag{A-27}$$

which shows that given ω, we can find a value of 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.6.2 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

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

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-23), this simplifies to

$$\frac{d\log p}{d\log mc} = \frac{\epsilon}{2\epsilon - 1} \tag{A-29}$$

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

As we derived in Appendix A.3, our Kimball specification yields a cost pass-through of $\frac{1}{1-\psi\omega}$ in the symmetric steady state. Using (A-24) and (A-26), the formula in (A-28) simplifies exactly to the same formula.

D.7 Calvo Models

In this section, we consider the degree of non-neutrality in Calvo models, which serve as a good benchmark due to its prevalance as a modelling tool of nominal rigidities in the macroeconomics literature. Also, it features no state-dependency and hence no selection in price adjustment, and therefore can be used as an upper bound to the real effect of nominal shocks.

We first study a CES model with Calvo pricing and no idiosyncratic shocks to firms. 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

$$P_t = \left[\alpha X_t^{1-\theta} + (1-\alpha) P_{t-1}^{1-\theta}\right]^{\frac{1}{1-\theta}} \tag{A-30}$$

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

$$\hat{P}_t^{1-\theta} = \alpha \hat{X}_t^{1-\theta} + (1-\alpha) \hat{P}_{t-1}^{1-\theta} \tag{A-31}$$

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

$$\hat{P}_1 = \alpha\mu \tag{A-32}$$

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

$$\hat{P}_h = \left(\sum_{i=1}^{h} (1-\alpha)^{i-1}\right) \alpha\mu \tag{A-33}$$

The response of real output is therefore given by

$$\hat{C}_h = \hat{S} - \hat{P}_h \tag{A-34}$$

$$= \mu - \left(\sum_{i=1}^{h} (1-\alpha)^{i-1}\right) \alpha\mu \tag{A-35}$$

$$= (1-\alpha)^h \mu \tag{A-36}$$

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

Given the empirical price adjustment frequency, α would be set to 0.11. This implies an initial (and peak) response of output equalling 89% of the shock and a cumulative impulse response of 0.76.[45]

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

Table A-11: Monetary Non-neutrality: Calvo

	Baseline	Calvo I	Calvo II	Calvo III
Impact Response	0.82	0.89	0.86	0.98
CIR	0.42	0.76	0.67	1.31

Calvo I refers to the theoretical impulse response from a Calvo model with CES demand and no idiosyncratic shocks. Calvo II refers to a Calvo model with CES demand $\psi = 0$ and all other parameters held at the baseline calibration. Calvo III refers to a Calvo model with the same calibration as the baseline model which includes Kimball demand. All three models set the probability of price adjustment to be 0.11.

In Table (A-11), we also report non-neutrality results of Calvo models with idiosyncratic TFP and demand shocks with CES and Kimball demand. We do this by replacing state-dependent pricing with Calvo pricing, setting the probability of price adjustment to 0.11. In the CES version, we set $\psi = 0$ while holding all other parameters at the baseline level and find an impact response of 0.86 and CIR of 0.67. In the Kimball version, we keep all parameters related to demand curvature and idiosyncratic processes at the baseline calibration. The model with Kimball demand greatly amplifies the real effects of nominal shocks with an impact response of 0.98 and CIR of 1.31. This demonstrates that the role of micro real rigidities in generating monetary non-neutrality applies both to sticky price models with and without state-dependent pricing. Meanwhile, comparing the baseline menu cost model and the Calvo model with Kimball demand illustrate the role of selection in price adjustment. The Calvo model without selection exhibits substantially more non-neutrality relative to an otherwise identical model with state-dependent pricing due to menu costs.

D.8 Models with CES Demand

We present three 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-year autocorrelation and cross-sectional standard deviation of TFP . We still calibrate f to match the frequency of price changes. In CES III, we turn on the idiosyncratic demand on top of productivity and calibrate (ρ_n, σ_n)

According to their formula, the Calvo model has a CIR of 0.76 whereas a menu cost model à la Golosov and Lucas (2007) has a CIR of 0.13.

Table A-12: Internal Calibration

Moment	Data	CES I	CES II	CES III	Baseline
Frequency of price changes	0.11	**0.11**	**0.11**	**0.11**	**0.12**
Fraction of price increases	0.65	**0.64**	0.61	0.58	0.58
Size of price changes	0.08	**0.08**	0.15	0.14	0.07
5-year autocorr of z_t^i	0.32	0.00	**0.32**	**0.32**	**0.32**
Cross-sectional SD of z_t^i	0.26	0.03	**0.26**	**0.26**	**0.25**
5-year autocorr of n_t^i	0.62	0.01	0.00	**0.62**	**0.62**
Cross-sectional SD of n_t^i	1.16	0.01	0.04	**1.18**	**1.05**
Corr b/w TFPR and TFPQ	0.75	0.00	0.00	0.00	**0.74**
Corr b/w price and TFPQ	–0.54	–1.00	–1.00	–1.00	**–0.57**
Parameter	**Description**				
ψ	Super-elasticity	0	0	0	**–1.27**
ω	Elasticity	1.33	1.33	1.33	**1.29**
ρ_z	Persistence of z_t^i	**0.66**	**0.98**	**0.98**	**0.98**
σ_z	Standard deviation of z_t^i	**0.04**	**0.05**	**0.05**	**0.06**
ρ_n	Persistence of n_t^i	–	–	**0.992**	**0.997**
σ_n	Standard deviation of n_t^i	–	–	**0.05**	**0.02**
f	Menu cost	**0.01**	**0.06**	**0.03**	**0.03**

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.

to match the five-year autocorrelation and standard deviation of demand. In all 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 demand is CES.

Results are presented in Table A-12. CES I, which uses (ρ_z, σ_z, f) to match the first three pricing moments in the first panel,is able to match those moments very well. However, it completely misses the firm-dynamics moments. In addition, the size and persistence of the z process used to match pricing moments implies very small and nearly transitory movements in idiosyncratic productivity that is counterfactual to empirical estimates of firm productivity processes. Due to the absence of demand shocks, this version is, by definition, unable to address the remaining firm-dynamics moments. Turning to CES II, we focus on matching the first two firm-dynamics moments using (ρ_z, σ_z), while only targeting the frequency of price changes among pricing 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 average size of price changes completely. In an effort to match the more volatile and persistent idiosyncratic productivity process, the model

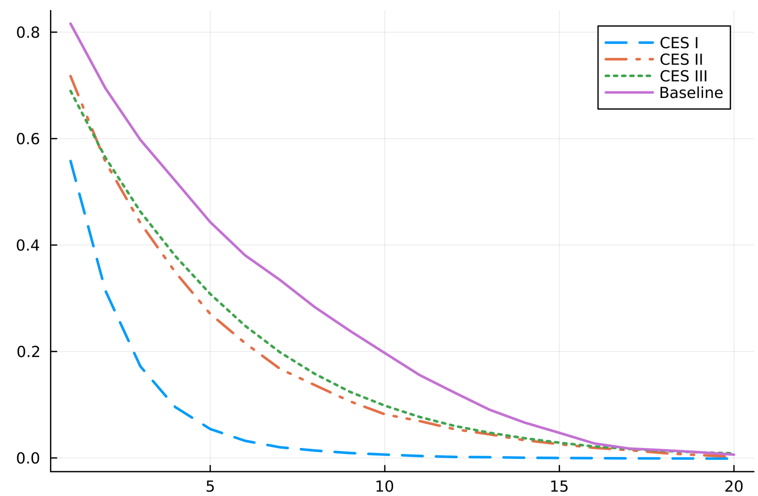


Figure A-8: Impulse Response of Real Output to a Nominal Expenditure Shock

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.

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. CES III attempts to match additional firm dynamics moments by including a idiosyncratic demand shock in addition to a idiosyncratic productivity shock, once again leaving all pricing moments except for the frequency of price changes as untargeted. This version is able to match the four firm-dynamics moments by picking appropriate parameters for the aforementioned shocks. Due to the CES structure, as explained above, the last two moments are still elusive for this version. And just like CES II, it fails to deliver on the untargeted pricing moments.

Figure A-8 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. Table A-13 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

Table A-13: 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.

larger CIR of 0.28 compared to CES I. Neither of these versions delivers a substantial level of non-neutrality, which is the key result of Golosov and Lucas (2007). CES III shows an impact response of 0.69, a half-life of 3.5 months, and a CIR of 0.30. Monetary policy in this version exhibits slightly greater non-neutrality due to the fact that idiosyncratic demand shocks act as random menu costs.[46] 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 thereby raising monetary non-neutrality.

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