NBER WORKING PAPER SERIES

PRICING AND PRODUCTION WITHOUT THE INVISIBLE HAND

Joel P. Flynn
George Nikolakoudis
Karthik Sastry

Working Paper 35390
http://www.nber.org/papers/w35390

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2026

We are grateful to Susanto Basu, Javier Bianchi, Saroj Bhattarai (Discussant), Vasco Carvalho, Andres Drenik, Christian Hellwig, Giuseppe Moscarini, Michael Peters, Pascual Restrepo, and Aleh Tsyvinski for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Pricing and Production Without the Invisible Hand
Joel P. Flynn, George Nikolakoudis, and Karthik Sastry
NBER Working Paper No. 35390
June 2026
JEL No. E23, E32, E40

ABSTRACT

Modern theories of the business cycle do not allow for the simultaneous rational choice of both prices and quantities, instead assuming that an "invisible hand" determines one of these variables to clear markets. In this paper, we develop a macroeconomic model in which both prices and quantities are chosen optimally by firms and exchange is both voluntary and efficient. As a consequence, individual markets will generically be in Walrasian disequilibrium: either slack (over-supplied) or rationed (under-supplied). The absence of market clearing changes pricing and production in qualitatively important ways: markups are governed by the probability of rationing rather than the elasticity of demand, and higher uncertainty reduces production and increases markups. Marrying the Old and New Keynesian traditions, we study general Walrasian disequilibrium with rational expectations and optimal firm decisions. On a technical level, we characterize cross-market spillovers arising from rationed demand with differentiated goods, overcoming the standard combinatorial problem that arises when studying multi-market disequilibrium. Unlike in New Keynesian economies, monetary shocks propagate by reducing product-market slack, raising aggregate labor productivity and consumption with muted effects on employment, while uncertainty shocks act as stagflationary cost-push shocks.

Joel P. Flynn
Yale University
Department of Economics
and NBER
joel.flynn@yale.edu

Karthik Sastry
Princeton University
Department of Economics
and NBER
ksastry@princeton.edu

George Nikolakoudis
The University of Chicago
Economics
gnikolakoudis@gmail.com

1 Introduction

We study the equilibrium determination of pricing *and* production in economies in which both of these decisions are made directly by economic agents. Although this may sound descriptive of standard general equilibrium models, that is far from the case. Indeed, Arrow (1959) describes the presence of "a logical gap in the usual formulations of the theory of the perfectly competitive economy, namely that there is no place for a rational decision with respect to prices as there is with respect to quantities." In modern models of imperfect competition in general equilibrium, such as the monetary model of Blanchard and Kiyotaki (1987) or the textbook New Keynesian model of Woodford (2003b), Arrow's critique can be inverted: there is a rational decision with respect to prices but not quantities. For example, the New Keynesian model requires agents to produce goods even when they would rather not do so (Huo and Ríos-Rull, 2020; Holden, 2025), violating the core economic principle of voluntary exchange. Thus, it is no new argument that the delegation of pricing or production to the invisible hand or the Walrasian auctioneer is potentially undesirable.

In this paper, to overcome the Arrow (1959) critique, we build a general equilibrium macroeconomic model in which firms make both pricing and production decisions to maximize expected profits under uncertainty and exchange is both voluntary and efficient. The model features markets that are rationed and slack—as in classical analyses of disequilibrium in the Old Keynesian tradition (see *e.g.,* Barro and Grossman, 1971)—while maintaining rational expectations equilibrium and optimal pricing in the New Keynesian tradition. Marrying these approaches requires developing a theory of optimal firm decisionmaking in the presence of cross-market demand spillovers that arise in disequilibrium models.

We find that severing the invisible hand has a number of microeconomic and macroeconomic implications. At the microeconomic level, firms' optimal markups are governed by the probability that demand is rationed and not merely the price elasticity of demand (*cf.* the formula of Lerner, 1934). Moreover, increases in volatility have *first-order* effects on firms' decisions, driving up prices, reducing labor inputs, and increasing markups. At the macroeconomic level, we characterize cross-market demand spillovers, finding that aggregate demand externalities can turn negative because increasing production in rationed markets reduces demand for non-rationed markets (*cf.* Blanchard and Kiyotaki, 1987). Uncertainty shocks act as stagflationary cost-push shocks. Monetary shocks propagate via the absorption or release of slack; because of this, their effects depend on the amount of uncertainty and market power in the economy. While our baseline model is intentionally as simple as possible, we also describe how it can be augmented to feature inventories, disequilibrium in product and labor markets, and dynamics with sticky decisions.

The Pricing and Production Problem. We first examine a single firm that must choose prices and inputs for production before it knows its demand and productivity. To focus on the novel economics of this problem, we consider the simplest possible model of this setting: the firm operates a constant returns-to-scale production function and faces a demand curve with constant price elasticity of demand. The firm is uncertain about the strength of its demand and productivity, which are jointly log-normally distributed. Problems of this variety lie at the core of virtually all standard macroeconomic models, including the New Keynesian model and the Real Business Cycle model with monopolistic firms.

The standard practice is to suppose that the firm picks either its price or its quantity, and then that the remaining variable is determined by market clearing. To overcome the Arrow (1959) critique of this assumption, we instead adopt an axiomatic approach (Bénassy, 1993). First, we assume that exchange must be voluntary: no one can be forced to trade more than they would be willing. Second, we assume that exchange is efficient: there are no trades that fail to take place that both sides of the market would be willing and able to make. These axioms of exchange imply that the transacted quantity is the minimum of what the firm produces and what the household demands at the price set by the firm. For this reason, the firm faces sharply different pricing and production incentives *ex post* depending on which of two states occurs: demand exceeds supply and the market is rationed, or supply exceeds demand and the market is slack. The firm's problem is to balance these trade-offs *ex ante*.

We first characterize the optimal pricing and input decisions in terms of a single fixed-point equation for the firm's optimal *tightness* $t = p^{-\eta}/L$, where p is the firm's price, η is the price elasticity of demand, and L is the firm's labor input. This equation requires that the firm's tightness is proportional to the ratio of the firm's expected productivity when the firm rations demand to the firm's expected demand when the firm over-produces. We derive this condition from the firm's indifference to locally changing prices. Intuitively, consider a firm raising its prices by 1 percent, a marginal amount. If the market ends up rationed, they sell all of their produced goods at the now higher price and make 1% higher revenues. If the market ends up slack, then the firm continues to earn 1% more per unit sold but also sells η% fewer units, where η is the elasticity of demand; the overall effect on revenues is a $(\eta - 1)$% reduction. The optimal choice of tightness balances these forces, weighted by their relative likelihood of occurring. It is a fixed point because the probability of rationing itself depends on the firm's choices through, and only through, tightness.

The firm's optimal price is pinned down by setting a markup equal to the inverse probability that demand is rationed on the expected marginal cost of production. We derive this expression from the firm's indifference to locally changing its labor input. Without market clearing, hiring an additional unit of labor is beneficial only whenever the market is rationed, while costs are always paid. Therefore, to ensure that hiring a marginal worker has zero profit (and that inframarginal workers are profitable), the markup is the inverse rationing probability. Thus, optimal pricing deviates from the classical Lerner (1934) formula, in which the optimal markup is solely determined by the elasticity of firms' residual demand. Instead, uncertainty about supply and demand forces now drives the optimal markup.

Indeed, uncertainty has unusually strong effects on prices and inputs in this model: the perceived standard deviations of supply and demand shocks have *first-order* effects on decisions. This is in contrast to most widely used business cycle frameworks, in which the standard deviations of these factors are only relevant to second-order or third-order (Schmitt-Grohé and Uribe, 2004). This feature arises from the discontinuity in firms' marginal revenues as a firm's product switches from being oversupplied to rationed. Moreover, the discontinuity of marginal revenues is a fundamental property of voluntary and efficient exchange.[1]

We characterize the effects of uncertainty and show that higher uncertainty increases firms' optimal prices and depresses their optimal usage of labor inputs. Intuitively, firms reduce their labor to insure themselves against low demand realizations in which production is wasted, and they raise prices so that purchasing labor continues to be profitable. Both this result and the economic argument underlying it require that firms jointly make pricing and production decisions, rather than delegating one margin to the invisible hand. Our framework thus implies that cyclical fluctuations in uncertainty can drive cyclical fluctuations in markups. Finally, we show that these qualitative insights translate to more general environments, such as those that feature inventories or disequilibrium in input markets.

Pricing and Production in General Equilibrium. We next embed this firm problem in an otherwise canonical and intentionally simple general equilibrium macroeconomic model. Following Blanchard and Kiyotaki (1987), households have constant elasticity of substitution preferences over a continuum of varieties with idiosyncratic taste shocks. Households further have Golosov and Lucas (2007) preferences over the consumption aggregate, money, and labor supply.

The household's problem in this model differs from that in standard analyses of monopolistic competition because the household cannot consume more of any good than has been produced. The resulting failure of market clearing in any one market modifies effective demand in other markets. We characterize these spillovers analytically and show that the presence of rationed markets increases effective demand in slack markets. Intuitively,

[1] To make this point precise, we contrast our model to one with a matching function that violates the principle of efficient exchange, as in the literature on matching in goods markets (see Petrongolo and Pissarides, 2001, for a review), and show that uncertainty only matters for firm decision-making to second-order.

consumers substitute to other varieties once a particular variety is out of stock. Thus, the standard logic that production decisions are strategic complements (Blanchard and Kiyotaki, 1987) is no longer universally true: expanding production of rationed varieties *reduces* the demand for non-rationed varieties, generating negative aggregate demand externalities.

Next, we study how these market interactions shape prices, production, and aggregate consumption in general equilibrium. As in our partial equilibrium framework, these variables respond to first order in uncertainty. In particular, an increase in uncertainty unambiguously raises firms' prices and reduces aggregate consumption. For this reason, uncertainty shocks have similar effects to markup shocks in general equilibrium. We contrast this case to the Walrasian benchmark, in which demand volatility has no first-order impact on firms' pricing decisions. Moreover, the channel for these effects is a reduction in aggregate labor productivity (GDP per labor hour), despite the fact that physical productivity is unchanged. This occurs because an increase in uncertainty raises the economy's "excess capacity"—that is, certain varieties are not consumed because there is no invisible hand to incentivize consumers to purchase those goods which are in excess supply.

We then study the monetary transmission mechanism when markets do not clear. Unanticipated changes in aggregate demand affect real consumption, even though production is predetermined, by giving households resources to consume goods in previously overproduced markets. This mechanism is therefore distinct from that of standard New Keynesian frameworks in which changes in consumption arise entirely from firms' market-clearing commitments to produce whatever the market demands (for a discussion, see Flynn et al., 2026). Instead, it is more reminiscent of the "Old Keynesian" notion that consumption increases by picking up slack in the economy (for a discussion, see Barro, 2025). To make this point formal, we show that monetary passthrough is discontinuous in the zero-uncertainty limit of our economy: when there is no uncertainty in the economy, there is no slack in any market, and unanticipated monetary shocks have no effect on output. We additionally characterize how passthrough is endogenous to—and ambiguously affected by—uncertainty and market power, based on the way that each force affects the spare capacity available to absorb marginal demand shocks.

Dynamics: A New Old Keynesian Model. Finally, we study the dynamic implications of our framework when pricing and production decisions are also "sticky" in the sense of Calvo (1983): only a fraction of firms can adjust pricing and production in any given period. This allows for realistic sluggishness in decision making and allows us to transparently compare our approach to that of the textbook New Keynesian model.

We first characterize the properties of our New Old Keynesian (NOK) model. First, as in the static analysis, uncertainty has first-order effects on allocations. These uncertainty shocks provide a microfoundation for stagflationary shocks ("cost-push shocks") that are essential to explain the business cycle (*e.g.*, Smets and Wouters, 2007). Second, we show that the first-order propagation of shocks for the price level and real consumption are the same as their New Keynesian counterparts: that is, conditional on the initial shock to the price level or consumption, these variables behave identically in the NK and NOK models. Third, we show that propagation for physical output, employment, and labor productivity is sharply different. The intuition is the dynamic analogue to what we found in the static analysis: consumption and sales can fluctuate due to absorbing and releasing product-market slack, even without corresponding movements in inputs. Thus, the NOK model generates dynamic propagation of shocks into inflation as per the New Keynesian tradition with physical effects on production and economic slack as per the Old Keynesian tradition.

We finally solve a calibrated version of this model nonlinearly to illustrate two key economic lessons. First, monetary expansions lead to consumption booms fueled by tightening product markets, with a more limited response of production and employment. This transmission mechanism is consistent with high-frequency empirical evidence of monetary transmission from Buda et al. (2025), who document rapid (week-to-week) responses of sales to monetary expansions without corresponding responses of production or inputs (including labor). Moreover, the prediction that monetary expansions increase labor productivity is consistent with the classic empirical findings of Christiano et al. (2005), who show that GDP expands more quickly than employment in response to shock monetary expansions.

Second, uncertainty shocks lead to sharp declines in consumption, increases in prices, and an increase in product-market tightness. As a result, they reduce labor productivity: for close to the same level of physical production, high prices dissuade consumers and lead to lower transaction volumes. These effects are completely absent in the nested NK model, in which volatility has no effect on allocations even in the nonlinear solution. In this way, our findings contribute to the literature studying how and why spikes in uncertainty, known to be concurrent with recessions (see, *e.g.*, Bloom et al., 2018), might translate into economic contractions (as shown empirically by Bloom, 2009) and inflation (as shown empirically by Mumtaz and Ruch, 2025). In aggregate data, our uncertainty shocks would manifest as contractionary and inflationary reductions in the Solow residual with ambiguous effects on labor, consistent with the findings of Basu et al. (2006).

Related Literature. The prevailing tradition in modern macroeconomics is to study models in which firms set prices (see *e.g.,* Woodford, 2003b) or quantities (see *e.g.,* Angeletos and La'O, 2010) and all markets clear *ex post*.[2] Our work is most related to a literature

[2] Reis (2006) and Flynn et al. (2024) consider settings in which firms can choose to set either a price or a quantity. Flynn et al. (2026) allow firms to choose supply functions, but maintain market clearing.

that explores equilibrium in contexts in which markets do not clear (see Bénassy, 1993, for a review). For example, Barro and Grossman (1971) and Michaillat and Saez (2015) study models in which there can be rationing and excess supply in the product and labor markets. A key challenge in this literature is the combinatorial nature of clearing across markets and accounting for all possible combinations of cross-market demand spillovers. Methodologically, we show that allowing for uncertainty and monopolistic competition (as in Dixit and Stiglitz, 1977) smooths over any discontinuities in the firm decision problem and gives rise to a well-defined distribution of markets that are below and above capacity in general equilibrium, allowing for an analytical characterization of these spillovers. This is a different approach than those based on tâtonnement, which postulate (Jaffé, 1967) or derive (Lorenzoni and Werning, 2026) an explicit process of price adjustment and study whether and how fast outcomes converge to Walrasian equilibrium. For example, in Lorenzoni and Werning (2026), while firms' prices are sticky, their quantities are not: they always produce the quantity that clears the market at their fixed price.

Our analysis relates to a long line of work that studies deeper microfoundations for the structure of product markets, including analyses of sequential transactions (*e.g.,* Lucas and Woodford, 1993; Eden, 1994) and search (*e.g.,* Diamond, 1982; Lagos and Wright, 2005; Rocheteau and Wright, 2005; Ghassibe and Zanetti, 2022; Bianchi et al., 2024; Bai et al., 2025). Within this vein, the papers most related to our analysis are Huo and Ríos-Rull (2020), Holden (2025), and Markauskas et al. (2026), which augment New Keynesian models with the ability of workers to not supply labor or firms to not produce goods if they do not wish to do so.[3] In both instances, this leads to the potential for labor or goods demand to be rationed. Our analysis differs from these in the critical respect that firms must make both pricing and production decisions before the realization of uncertainty. This leads to the important economic difference that markets can end up in a state of either excess supply or rationing, and is the feature that gives rise to the novel implications of our analysis for the propagation of monetary and uncertainty shocks.

Outline. The rest of the paper proceeds as follows. Section 2 introduces the decision problem of a monopolist that must choose both prices and inputs under uncertainty and characterizes its properties. Section 3 embeds this decision in a general equilibrium model and characterizes household behavior. Section 4 characterizes equilibrium prices and production and leverages this to study the propagation of uncertainty shocks and monetary shocks. Section 5 extends the static model to a dynamic setting and quantitatively illustrates the dynamic implications of the theory. Section 6 concludes.

[3] Liu (2026) embeds a similar rationing mechanism to that of Holden (2025) in a search-and-matching framework, allowing him to study the interaction of rationing with real rigidities.

2 Pricing and Production in Partial Equilibrium

We begin by studying the decision problem of a monopolist that makes pricing and production decisions before their demand is realized. We characterize the firm's optimal choices and, in particular, how they depend on uncertainty about payoff-relevant variables.

2.1 Model: The Firm's Problem

A firm sets a price $p \in \mathbb{R}_{++}$ and purchases labor $L \in \mathbb{R}_{+}$ at a cost $w \in \mathbb{R}_{++}$. Given these choices, the firm produces quantity $q^s \in \mathbb{R}_{+}$ according to a constant returns-to-scale production function given by:

$$q^s = AL \tag{1}$$

where $A \in \mathbb{R}_{++}$ is the firm's productivity. Moreover, the firm faces demand q^d for its product:

$$q^d = zp^{-\eta} \tag{2}$$

where $z \in \mathbb{R}_{++}$ is a demand shifter and $\eta > 1$ is the elasticity of demand. The firm values its profits according to a nominal stochastic discount factor $\Lambda \in \mathbb{R}_{++}$. We assume that the variables (Λ, A, z, w) are jointly log-normally distributed with mean μ, variance-covariance matrix Σ, and cumulative distribution function G. We use μ_X and σ_X^2 to represent, respectively, the mean and the variance of the logarithm of an arbitrary random variable X.

The amount of the good that is ultimately *sold* to consumers, q, derives from two basic and classic principles (see Bénassy, 1993, for a review). The first is voluntary exchange: no agent can be forced to trade more than that which they are able and/or willing, *i.e.,* $q \leq q^s$ and $q \leq q^d$. The second is efficient exchange: there are no trades that fail to take place that both sides of the market are willing and able to make, *i.e.,* it is never the case that both $q < q^s$ and $q < q^d$. These two principles imply that the ultimate transacted quantity is the minimum of what is supplied by the firm and what is demanded by consumers:

$$q = \min\{q^s, q^d\} \tag{3}$$

There are three possible outcomes in this market. First, the firm may supply more than what is demanded, $q^s > q^d$. In this case, we say that the market is *slack*. Second, the firm may supply less than what is demanded, $q^s < q^d$. In this case, we say that the market is *rationed*. Finally, the firm may supply exactly what is demanded, or $q^s = q^d$. In this case, we say that the market is *in Walrasian equilibrium*.

The firm makes both its pricing and input choices under uncertainty about the state of demand, its own productivity, factor costs, and the stochastic discount factor. This implies, crucially, that the firm is also unsure of the eventual state of the market: slack, rationed, or in equilibrium. As we elaborate on below, this can be equivalently interpreted as the uncertain outcome of the market over a single transaction period or the average over many transaction periods.

Formally, the firm solves the following problem:

$$\begin{aligned}
&\sup_{p,L} \int_{\mathbb{R}_4^{++}} \Lambda(pq - wL)\,\mathrm{d}G(\Lambda, A, z, w) \\
&\text{s.t. } q = \min\{q^s, q^d\},\ q^s = AL,\ q^d = zp^{-\eta}
\end{aligned} \tag{4}$$

This reveals the technically non-standard and economically important feature of this problem: the firm's revenue function has a kink at the point at which demand equals supply.

Why Study Disequilibrium? Supply and demand are actively equated by institutions in only a handful of real-world settings, such as centralized financial exchanges. More commonly, markets fluctuate between being rationed and being slack. Our analysis belongs in a long tradition of work that models this phenomenon explicitly rather than assuming that markets are cleared via the intervention of an auctioneer or "invisible hand" (Bénassy, 1993).

As an illustrative metaphor, consider how our model might describe the decisions of a hair salon. On certain occasions, the manager sets prices and chooses how many hairdressers to work in the salon. When making these choices, the manager knows that the salon will be unable to provide exactly the quantity of haircuts that happens to be demanded on any given day. Sometimes, the salon will be full and customers will be turned away. That is, the market may be rationed. Other times, hairdressers will sit idle. That is, the market may be slack. While, as we shall see, an optimizing manager will make these choices such that the salon's capacity equals demand "on average," the incentives for price and quantity choices will be markedly different than those in a market in which capacity always equals demand.[4] Viewed in this way, the key premises of our analysis are the familiar ideas that both prices and inputs are, for a variety of practical reasons, adjusted only infrequently, while underlying demand and cost conditions fluctuate.

As a practical matter, there is abundant evidence that many real-world markets are out of equilibrium at any given point in time—particularly for services, which constitute about 80% of the modern US economy, and for perishable goods. The rationing of goods often manifests

[4] It is worth contrasting this perspective to the standard approach familiar from New Keynesian models. A "New Keynesian hair salon" sets its prices and commits to paging hairdressers in real-time when demand is strong and sending them home when demand is low, paying them only for the time they spend actively cutting hair.

in *stockouts*, which are prevalent in microeconomic data and correlate with business cycle events (see *e.g.,* Cavallo and Kryvtsov, 2024; Holden, 2025). For staffed services—such as those of restaurants, barber shops, law firms, management consultants, or call services—there are less systematic data on rationing, but the experience of waiting for or being turned away from these products is familiar. Similarly, one easily measured consequence of slack is the outright wastage of goods. For example, the US Department of Agriculture estimates that 31% of the US food supply is wasted at the retail and consumer levels (USDA Economic Research Service, 2026). Toward a more systematic measure of slack, a long tradition in economics has sought to directly measure the extent of *capacity utilization*, or the fraction of installed inputs that currently sit idle rather than being reallocated to other purposes to clear markets (Corrado and Mattey, 1997). Direct survey evidence from establishments documents that firms often employ excess labor over the cycle (Fay and Medoff, 1985) and that employed workers often spend a substantial fraction of their work day idle, on average (Michaillat and Saez, 2015). Similarly, Liu (2026) provides evidence that a substantial fraction of products are produced below capacity.

Of course, these examples may also bring to mind certain features of real-world settings that are *not* in our intentionally simple model. Chief among them are inventories and failures of market clearing for inputs. We abstract from these forces in our main analysis for analytical clarity and consider them as extensions in Section 2.4.

2.2 Optimal Firm Behavior

We now characterize the firm's optimal price and input choices under uncertainty.

Theorem 1. *There exists a unique optimal price p^* and labor input L^*. Moreover, optimal tightness $t^* = \frac{p^{*-\eta}}{L^*}$ is the unique solution to:*

$$t = \frac{1}{\eta - 1} \frac{\mathbb{E}[\Lambda A \mathbb{I}[t \geq U]]}{\mathbb{E}[\Lambda z \mathbb{I}[t \leq U]]} \tag{5}$$

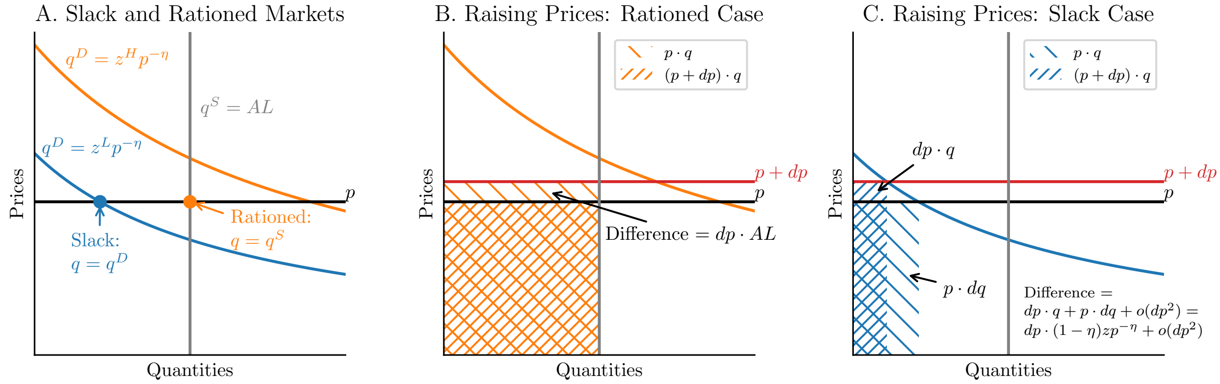
where $U = A/z$. Moreover, the optimal price is given by:

$$p^* = \frac{\mathbb{E}[\Lambda w]}{\mathbb{E}[\Lambda A \mathbb{I}[t^* \geq U]]} \tag{6}$$

and the optimal labor input is given by $L^* = \frac{p^{*-\eta}}{t^*}$

Proof. See Appendix A.1. □

Figure 1: The Incentives to Raise Prices



Note: This figure illustrates the incentives for price setting in the model. Panel A illustrates cases where the market is rationed or slack, depending on the realization of demand. Panels B and C show the marginal revenue effect of increasing prices from p to $p + \mathrm{d}p$ in states that are rationed versus slack, respectively. Equation 7 is obtained by setting the expected (risk-adjusted) marginal revenue effect to zero, averaging over slack and rationed cases under the stochastic discount factor, and Equation 5 of Theorem 1 is obtained by rearranging this condition.

Together, equations 5 and 6 describe the firm's choices of both inputs and prices. We now describe the derivation of both as well as the economic intuition underlying them.

Optimal Tightness. We derive the expression for firms' optimal tightness (Equation 5) from the firm's indifference about marginally changing prices. To derive this, we first recognize that the marginal effect of changing prices depends on the state of the market. Figure 1 illustrates this in a simple example with fixed productivity and two demand states, high ($z = z^H$) and low ($z = z^L < z^H$). The market is rationed (demand is greater than supply) whenever $q^d = zp^{-\eta} \geq AL = q^s$. Using our definitions of tightness, $t := p^{-\eta}/L$, and the shock to supply-demand imbalances, $U := A/z$, we can re-write the condition for rationing as $t \geq U$. Panel A of Figure 1 shows that, in this example, the market is rationed when demand is high and slack when demand is low. We finally observe that, since random variables are smoothly distributed in our problem, the event of exact Walrasian equilibrium has probability zero; therefore, it will suffice to consider the cases in which the market is rationed and slack.

If the market is rationed, a slightly higher price earns additional revenue $\mathrm{d}p$ for all units produced. That is, the change in revenues is $AL \cdot \mathrm{d}p$ (Panel B of Figure 1). The firm values these revenues in risk-adjusted terms, conditional on the event that the market is rationed. Thus, conditional on the rationed state, expected marginal benefits are $\mathbb{E}[\Lambda AL | t \geq U] \cdot \mathrm{d}p$.

If the market is slack, then raising prices affects the quantity of goods that will be sold as the firm moves up its demand curve. That is, its revenue at a given price is $p \times zp^{-\eta}$ and the first-order change from raising prices is $\mathrm{d}p \cdot q^D + p \cdot \mathrm{d}q^D$ (Panel C of Figure 1). This is readily calculated as $(1-\eta)zp^{-\eta} \cdot \mathrm{d}p$. Moreover, this state occurs when $t \leq U$. Thus, conditional on the slack state, expected marginal losses are $(\eta - 1)\mathbb{E}[\Lambda z p^{-\eta}|t \leq U] \cdot \mathrm{d}p$.

At the optimal price, the expected losses from raising prices when the market is slack must be equal to the firm's expected gains from raising prices whenever the market is rationed, multiplied by the respective probabilities of being in each state:

$$\underbrace{(\eta-1)\mathbb{E}[\Lambda z p^{-\eta}|t < U]}_{\text{Revenue losses when slack}} \times \underbrace{(1-\mathbb{P}[t \geq U])}_{\text{Slack probability}} = \underbrace{\mathbb{E}[\Lambda A L|t \geq U]}_{\text{Revenue gains when rationed}} \times \underbrace{\mathbb{P}[t \geq U]}_{\text{Rationing probability}} \tag{7}$$

Rearranging this equation then yields Equation 5.

Optimal Prices and Markups. We derive the firm's optimal price (Equation 6) from the firm's indifference about marginally changing inputs. Increasing inputs always has a cost, proportional to wages, but only sometimes has a benefit. If the market is slack (goods are oversupplied), then extra production generates no revenues. If the market is rationed (goods are undersupplied), then the firm can sell additional units at its chosen price. Putting these ideas together, the expected gains from marginally increasing inputs are given by $\mathbb{E}[\Lambda A p|t \geq U] \cdot \mathbb{P}[t \geq U] + 0 \cdot \mathbb{P}[t < U]$, while the expected losses are given by $\mathbb{E}[\Lambda w]$. At the optimum, these expected losses and gains must be equated. Re-arranging this expression yields Equation 6.

This pricing condition suggests a direct connection between the rationing probability and the markup. In particular, the condition can be rewritten in terms of the probability that demand is rationed, $r = \mathbb{P}[t \geq U]$:

$$p^* = \frac{1}{r}\frac{\mathbb{E}[\Lambda w]}{\mathbb{E}[\Lambda A|t^* \geq U]} \tag{8}$$

Thus, the firm optimally places a markup of the inverse rationing probability on an object that is related to marginal costs, with two twists. First, the firm risk-adjusts using the stochastic discount factor Λ. Second, the expectation of productivity corrects for the inference that when demand is rationed, productivity is likely to be low and marginal costs are likely to be high. In the natural but special case that shuts down both of these channels—the stochastic discount factor and the firm's productivity are known—the firm's price is exactly a markup of $1/r$ on its expected marginal costs: $p^* = \frac{1}{r}\frac{\mathbb{E}[w]}{A}$.

Conspicuously missing is any *direct* role for the elasticity of demand. In the textbook derivation of the Lerner (1934) formula in a market-clearing setting, markups are solely determined by the elasticity of demand because of the following logic: firms equate the marginal revenue and marginal cost of additional production, and the marginal revenue depends only on the price and the elasticity of demand. In our setting, marginal revenues do not *directly* depend on the shape of the demand curve, holding fixed tightness and, in turn, the probability of rationing.

The Role of Log Normality. The arguments above use two necessary conditions for optimality—the firm's indifference to small changes in p or L—and therefore do not directly require our additional assumptions that random variables follow a log normal distribution. In this sense, the economic intuition underlying the two necessary conditions does not rely on our functional form for uncertainty. In the proof of Theorem 1, we use the log-normality assumption only to establish the *sufficiency* of the conditions derived above and the uniqueness of the problem's solution.

2.3 The Effects of Uncertainty on Decisions

We now study the effects of uncertainty on firm behavior in this setting. To do this, we consider the random vector and we rescale the volatility of each variable about which firms are uncertain $(\ln \Lambda, \ln A, \ln z, \ln w)$ by $\delta \geq 0$.[5] Thus, δ scales the total uncertainty faced by the firm while preserving the covariance structure of the shocks.

The Complete Information Limit. The next result shows that the complete information limit corresponds to nothing more than the benchmark model of conduct by a monopolistic firm: set a price equal to a taste-based markup of $\frac{\eta}{\eta-1}$ over nominal marginal costs.

Proposition 1. *In the complete information limit, we have that:*

1. *Prices are a constant taste-based markup over nominal marginal costs:*

$$p^{FI} \equiv \lim_{\delta \to 0} p^* = \frac{\eta}{\eta - 1} \exp\{\mu_w - \mu_A\} \tag{9}$$

2. *Labor converges to its full information value:*

$$L^{FI} \equiv \lim_{\delta \to 0} L^* = \left(\frac{\eta}{\eta - 1}\right)^{-\eta} \exp\{\mu_z + (\eta - 1)\mu_A - \eta\mu_w\} \tag{10}$$

[5] That is, collapsing $\theta = (\ln \Lambda, \ln A, \ln z, \ln w)$, we consider a family of random vectors $\theta_\delta = \mu_\theta + \delta \Sigma^{\frac{1}{2}} Z$, where $Z \sim N(0, I)$ is a four-dimensional standard Gaussian noise vector.

3. *The supply-demand imbalance in choices is the fundamental supply-demand imbalance:*

$$t^{FI} \equiv \lim_{\delta \to 0} t^* = \exp\{\mu_A - \mu_z\} \tag{11}$$

4. *The rationing probability $r = \mathbb{P}[q^d \geq q^s]$ is the inverse taste-based markup:*

$$\lim_{\delta \to 0} r^* = \frac{\eta - 1}{\eta} \tag{12}$$

Proof. See Appendix A.2 □

Statements 1 and 2 imply that, as uncertainty approaches zero, the resulting prices and quantities coincide with those in the standard benchmark without uncertainty, where firms set prices and then produce to satisfy demand. Statement 3 establishes that any gaps between demand and supply vanish in this limit. Consequently, in the zero-uncertainty limit, the framework yields an underpinning for markets in Walrasian equilibrium: one can simply view firms as choosing prices and having quantities determined by market clearing (or vice versa). Moreover, this claim makes clear that any deviations of prices or quantities from their Walrasian levels must stem from the existence of uncertainty.

Statement 4 shows that the limiting rationing probability is non-zero and coincides with the inverse of the zero-uncertainty markup. Intuitively, small (but non-zero) uncertainty implies that demand may still be higher or lower than what the firm anticipates *ex ante*, leading to rationed or slack markets. Moreover, the precise form of the limit is *necessary* to ensure that prices coincide with the Lerner (1934) formula as uncertainty vanishes (Statement 1). To understand why this is the case, we consider again the firm's marginal incentives to raise prices (Equation 7). When uncertainty is vanishingly small, sales in slack and rationed markets are known and equal to each other, canceling out on both sides of Equation 7. Therefore, the firm equalizes the revenue elasticity in each state weighted by the rationing probabilities:

$$\underbrace{(\eta - 1)}_{\text{Revenue elast. when slack}} \times \underbrace{(1 - \mathbb{P}[t \geq U])}_{\text{Slack probability}} = \underbrace{1}_{\text{Revenue elast. when rationed}} \times \underbrace{\mathbb{P}[t \geq U]}_{\text{Rationing probability}} \tag{13}$$

Re-arranging and noting that $r = \mathbb{P}[t \geq U]$ gives the result. A further implication is that, as the market becomes perfectly competitive ($\eta \to \infty$), the rationing probability converges to one. Intuitively, more severe competition raises the losses from higher prices in the slack state while leaving constant the gains from higher prices in the rationed state; thus, the firm always wants to ration. The opposite logic implies that as monopoly power becomes very strong ($\eta \to 1$), the rationing probability converges to zero.

The First-Order Effects of Volatility. Having understood that uncertainty is the core element of this theory's departure from standard Walrasian equilibrium, we now characterize how uncertainty affects optimal behavior to first order. We do this by studying how choices depend to first order on uncertainty, parametrized by $\delta > 0$, around the point at which there is no uncertainty ($\delta = 0$). Crucially, this is a first-order perturbation of the volatility (standard deviation) of random variables, not of the variance.

The following proposition derives the first-order effects of volatility on the firm's optimal prices and labor input:

Theorem 2. *Up to a first-order approximation in volatility, the firm's optimal price and labor inputs are given by:*

$$\begin{aligned} \ln p^* &= \ln p^{FI} + \left(\phi(\kappa) - \frac{\kappa}{\eta}\right)\sigma_U \\ \ln L^* &= \ln L^{FI} - \eta\phi(\kappa)\sigma_U \end{aligned} \tag{14}$$

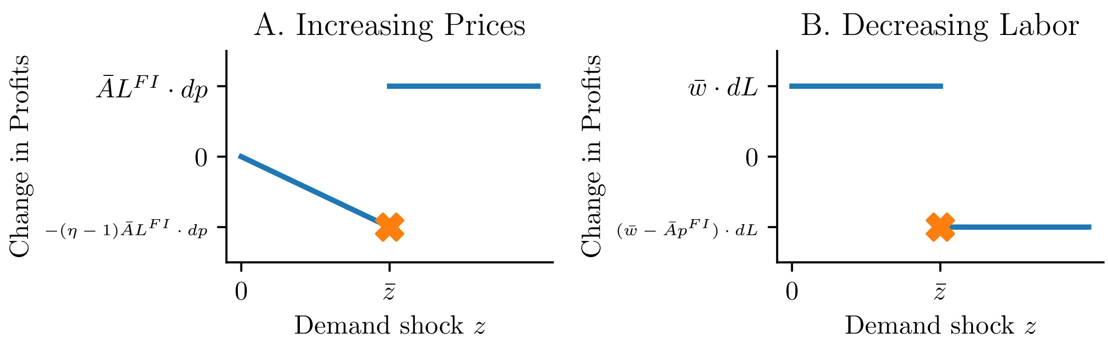
where $\kappa = \Phi^{-1}\left(\frac{\eta-1}{\eta}\right)$ and ϕ and Φ are the standard normal PDF and CDF, respectively. Moreover, to the first order, higher uncertainty leads firms to raise prices ($\Omega \equiv \phi(\kappa) - \kappa/\eta > 0$) and reduce labor ($-\eta\phi(\kappa) < 0$).

Proof. See Appendix A.3. □

This result has three implications. First, uncertainty has *first-order* effects on firms' choices. By contrast, in standard models in which firms maximize the expectation of smooth functions of choice variables and shocks, choices depend only (if at all) on *second-order* terms of volatility (first-order terms of variance). Second, σ_U, the standard deviation of $\log U = \log A - \log z$, is a sufficient statistic for the first-order effects of uncertainty. For reasons that will become clearer below, this arises because U summarizes how shocks affect whether the market is rationed versus slack. Finally, uncertainty moves choices in a particular direction: a more uncertain firm sets a higher price and hires less labor.

To understand the economic intuition behind these points, it is useful to consider how a small amount of demand uncertainty affects the firm's incentives in the neighborhood of their complete-information choices. Concretely, consider a full-information benchmark in which the stochastic discount factor is constant and productivity, wages, and demand are respectively fixed at $\bar{A}$, $\bar{w}$, and $\bar{z}$. The firm optimally chooses some p^{FI} and L^{FI} to maximize profits, which puts the market in Walrasian equilibrium. The firm would lose profits by perturbing either variable in either direction. In particular, if they were to raise prices by some amount $\mathrm{d}p$, the market would become slack and the firm would sell fewer goods at

Figure 2: Demand Shocks and Marginal Incentives



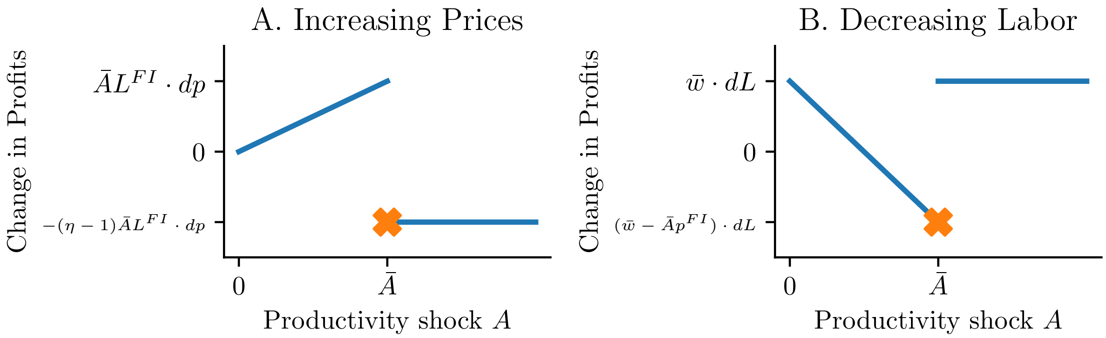
Note: This figure shows the marginal effects on profits of increasing prices by dp (panel A) and reducing labor by dL (panel B) as a function of the realized demand shock z, starting from choices (p^{FI}, L^{FI}) that maximize profits conditional on full knowledge that $z = \bar{z}$. The orange X's indicate marginal effects in this full information case. The blue lines correspond to marginal effects conditional on other possible realizations of z.

a higher price. The net effect would be to lose $(\eta - 1)\bar{A}L^{FI} \cdot \mathrm{d}p$ dollars by moving up the demand curve (recall Figure 1, Panel C). If the firm were to remove a marginal worker, then the market would become rationed. The firm would save $\bar{w} \cdot \mathrm{d}L$ in labor costs but lose $\bar{A}p^{FI} \cdot \mathrm{d}L$ in revenues. Because their markup is positive ($p^{FI} > \bar{w}/\bar{A}$), this entails a profit loss.

Now let us add a small amount of uncertainty about demand and analyze the firm's first-order incentives to deviate from the full-information optimal price. Under small uncertainty, the firm's marginal incentives are described by averaging the blue lines in Figure 2 over some distribution of z centered around $\bar{z}$. For prices, adding variability in z raises two new possibilities. First, demand could exceed $\bar{z}$. In this case, the market would be rationed and setting a higher price would earn a profit of $\bar{A}L \cdot \mathrm{d}p > 0$: that is, high prices allow the firm to capture the *upside* of high demand. Second, demand could be lower than $\bar{z}$. In this case, the market is slack, and price increases would lead to a smaller revenue loss than in the case where $z = \bar{z}$. By implication, for any demand realization $z \neq \bar{z}$, the marginal payoff to increasing prices (blue solid line) exceeds the marginal payoff evaluated under full information that $z = \bar{z}$ (orange X). It must also be true that the *expected* marginal profit has increased once uncertainty is added. Therefore, the firm has an incentive to raise prices.

For labor choice, the argument is similar. If demand is higher than $\bar{z}$, removing a worker is no worse than before; but if demand is lower than $\bar{z}$, removing a worker simply saves on costs. So, when taking an expectation over possible states, the firm now has an incentive to reduce the labor they hire (Panel B of Figure 2).

Figure 3: Productivity Shocks and Marginal Incentives



Note: This figure shows the marginal effects on profits of increasing prices by dp (panel A) and reducing labor by dL (panel B) as a function of the realized productivity A, starting from choices (p^{FI}, L^{FI}) that maximize profits conditional on full knowledge that $A = \bar{A}$. The orange X's indicate marginal effects in this full information case. The blue lines correspond to marginal effects conditional on other possible realizations of A.

We can also apply a closely related argument to understand the effects of uncertainty about productivity (Figure 3). The possibility that productivity may be *lower* than $\bar{A}$, which would make the market rationed, generates states with a strictly positive gain to raising prices. The possibility that productivity may be *higher* than $\bar{A}$, which would make the market slack, generates states with a strictly positive gain to reducing labor. The key theme, as above, is induced uncertainty about whether the market will be rationed or slack. This helps explain why the two random variables that affect market clearing enter Theorem 2 through the sufficient statistic $U = A/z$, and moreover why uncertainty about other random variables that do not affect market clearing does not enter Theorem 2.

The formal argument underpinning Theorem 2 involves directly perturbing the fixed-point equation derived in Theorem 1. This argument clarifies that the outcome of this informal intuition carries through once we properly account for two sets of indirect effects that this simple intuition ignores. The first set of effects relates to the fixed point between the arguments above: if the firm changes p, this affects the incentives to change L, and *vice versa*. The second set of effects relates to firms' *inference* about payoff-relevant variables given the state of the market. Our general argument shows that neither interferes with the basic argument above regarding the first-order effects of uncertainty on pricing and labor. In the process, we derive the exact responsiveness of these variables to uncertainty as well as the responses of tightness and the rationing probability.

Underlying both the basic intuition and the formal argument is the fact that firms' profits are kinked at the point at which the market clears. This kink itself arises from the basic technological fact that, under efficient exchange, trade switches from being on the demand curve to being on the supply curve at the exact point at which the market clears. Moreover, the presence of the kink is both why volatility has first-order effects and why the effects of volatility can be unambiguously signed. To make precise that these phenomena arise from the failure of interim market clearing, in Appendix B.1, we consider a model that smooths this kink via a *reduced-form* inefficient matching function (as is often studied in the literature on matching in goods markets, see *e.g.,* Petrongolo and Pissarides, 2001; Michaillat and Saez, 2015; Ghassibe and Zanetti, 2022; Bianchi et al., 2024; Bai et al., 2025). We show that when the kink is smoothed, volatility ceases to have any effect on firms' choices to first order (see Proposition 7), differentiating this theory from these approaches.

Uncertainty-Based Markups. Earlier, we noted that markups are not solely determined by the elasticity of demand in our model. We can now use Theorem 2 to say something more precise: to the first-order, uncertainty adds a second, strictly positive component to the mark-up over expected marginal costs, $\mathcal{M} = p/\exp(\mu_w - \mu_A)$:

$$\ln \mathcal{M} = \underbrace{\ln\left(\frac{\eta}{\eta-1}\right)}_{\text{Taste-Based Markup}} + \underbrace{\Omega\sigma_U}_{\text{Uncertainty-Based Markup}} \tag{15}$$

This has at least two concrete implications. First, our model can rationalize differences between markups estimated via the two most standard approaches: measuring the elasticity of demand on the "demand side" and combining this with assumptions about how firms compete (*e.g.*, Bresnahan, 1989; Berry et al., 1995), or measuring the ratio between prices and estimated marginal costs on the "supply side" (*e.g.*, De Loecker and Warzynski, 2012). Under our model, the standard translation from demand elasticity to markups would be incorrect. Even under correct specification of the model of monopolistic competition, this method would always understate markups.[6] Second, our framework can provide a rationale for markup *shocks* as a consequence of changes in firms' idiosyncratic uncertainty (*e.g.*, as measured by Bloom et al., 2018; Bachmann et al., 2026). Aggregate business-cycle models often rely on markup shocks to generate fluctuations in inflation (*e.g.*, as in Smets and Wouters, 2007); as we will elaborate on in our general-equilibrium model (Section 4) and quantification (Section 5), our model can rationalize fluctuations in markups from fluctuations in uncertainty, without potentially implausible changes in consumers' tastes.

[6] Studies applying the production approach to public firms in the US typically find markups in the range of 1.2-1.6 (*e.g.*, De Loecker et al., 2020; Demirer, 2025). By contrast, studies applying the demand approach to retail goods in the US find that the majority of markets have demand elasticities above 8 (*e.g.*, Hottman et al., 2016). This gap would be consistent with the existence of a positive uncertainty-based markup.

Our model's mechanism by which volatility affects markups is different from (and complementary to) existing approaches. In particular, Di Tella et al. (2025) show that markups may arise as a compensating differential for the exposure of entrepreneurial labor to risk, while a literature on precautionary pricing suggests that non-linearities in firms' profit functions can lead uncertainty to increase prices (see *e.g.,* Kimball, 1989). Our explanation is different economically from these and, moreover, is particularly strong in the precise sense that the effects of volatility are first-order in our framework, while they are second-order in these existing frameworks.

2.4 Extensions

Before embarking on our general equilibrium analysis, we first highlight two extensions of our framework to richer environments. For simplicity, we abstract from these elements in our subsequent general equilibrium model, but it is simple to see how the model could be extended to feature them.

Inventories. In this simple model, unsold production has no value to the firm. For services, this is a reasonable assumption: if a hairdresser sits in the salon and nobody shows up, there is no way to recoup the lost time. Moreover, for perishable goods, such as food, a similar principle applies. However, for durable goods, there is naturally a non-zero value of unsold production. To model this value, let us suppose abstractly that each unsold unit of production becomes inventory $I = \max\{AL - zp^{-\eta}, 0\}$, each unit of which has future value $\Lambda' \in \mathbb{R}_{++}$, where Λ' is jointly log-normal with the other random variables that enter the producer's problem. In Appendix B.2, we solve this problem. We show that uncertainty continues to have first-order effects on firm behavior. Moreover, the economics of the problem are very similar. The primary qualitative effect of introducing inventories is a lowering of the optimal rationing probability, as slack is now less costly for the firm.

Disequilibrium in Multiple Markets. In our model, we relaxed the imposition of market clearing in the goods market but maintained market clearing in the input (labor) market. Thus, we abstract from the interactions between non-Walrasian labor and product markets that are central to the disequilibrium analyses of Barro and Grossman (1971) and Michaillat and Saez (2015). In Appendix B.3, we show how an extension of our model to allow for input-market disequilibrium retains many of the qualitative properties highlighted above. In particular, we study an extension of our model in which firms make three choices: prices p, labor *demand* L^d, and a posted wage w. The firm faces a labor supply curve of the form $L^s = vw^\chi$, where χ is a known elasticity and v is an unknown shock, and the labor market is also described by disequilibrium with voluntary and efficient exchange (*i.e.,* the labor ultimately hired is $L = \min\{L^s, L^d\}$). Thus, the firm operates as a monopolist in product markets and a monopsonist in input markets, but nonetheless must make all decisions in advance. We show in Proposition 9 that the firm's optimal choice is described by a system of equations determining product-market tightness ($t = p^{-\eta}/L^d$), labor-market tightness ($t^L = L^d/w^{\chi}$), and prices. Both price markups and wage markdowns are tightly linked to the expected probability of rationing, for reasons that echo our discussion of Theorem 1. Moreover, the discontinuities in the marginal profit functions, now spanning four different states (rationing or slack in product and/or labor markets), open the door to first-order effects of uncertainty on allocations, in line with our discussion of Theorem 2.

3 A General Equilibrium Model

We now embed this partial equilibrium model in a general equilibrium setting to understand how uncertainty shocks and monetary shocks propagate. The goal of this analysis is to consider the simplest possible and most standard general equilibrium closure to illustrate the new economic mechanisms that emerge from this model of firm decision-making. For this reason, we largely follow the microfoundations of Blanchard and Kiyotaki (1987).

3.1 Primitives and Timing

Households. There is a continuum of families $\ell \in [0,1]$, each comprised of a continuum of individuals $j \in [0,1]$, all of whom have identical preferences and endowments. Each individual has constant elasticity of substitution preferences over a continuum of consumption varieties, indexed by $i \in [0,1]$:

$$C_{j\ell} = \left(\int_0^1 \theta_i^{\frac{1}{\eta}} c_{ij\ell}^{\frac{\eta-1}{\eta}} \, \mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \tag{16}$$

where $\eta > 1$ is the elasticity of substitution between varieties and θ_i is a variety-specific quality shock, which we assume to be IID log-normal with mean μ_θ and variance σ_θ^2 and common across the heterogeneous individuals. Each individual has Golosov and Lucas (2007) preferences over the consumption aggregate, money, and labor supply given by:

$$\ln C_{j\ell} + \ln M_{j\ell} - \alpha N_{j\ell} \tag{17}$$

where $\alpha > 0$ indexes disutility from labor. Each individual supplies labor to the firms, receives the profits from the firms, and receives transfers from the government. They can purchase a share $s_{j\ell} \in \mathbb{R}_+$ of the family's shopping of each variety $c_{i\ell}$, *i.e.*, $c_{ij\ell} = s_{j\ell} c_{i\ell}$. Shares are sold at the zero-profit price:

$$P_\ell = \frac{\int_0^1 p_i c_{il} \, \mathrm{d}i}{C_\ell} \tag{18}$$

where C_ℓ is the CES aggregate of the family's purchases of each variety $c_{i\ell}$. Thus, each household faces the following budget constraint:

$$P_\ell \, s_{j\ell} C_\ell + M_{j\ell} = w N_{j\ell} + \Pi + T \tag{19}$$

where T is a monetary transfer from the government and Π are aggregate profits. As all households are symmetric, we drop (j, ℓ) subscripts when there is no ambiguity.

Firms. There is a continuum of firms indexed by $i \in [0, 1]$, each of which is a monopolist producing good i. As in Section 2, each firm operates a constant returns to scale production technology $\bar{c}_i = A_i L_i$, where $\bar{c}_i$ is their quantity produced, A_i is their (stochastic) productivity (which is jointly log-normal with θ_i), and L_i is the amount of labor that they hire. The firms are owned in equal shares by the households.

The Government. There is a government that issues money in quantity M. They finance money creation via lump-sum taxation of households, such that $M = T$.

Timing. The model is split into three periods: the morning, the afternoon, and the evening. In the morning, the government announces the money supply $\bar{M}$, the firms set their prices p_i and hire labor L_i, and each individual chooses its labor supply N and its share of the family's shopping s. In the afternoon, productivity shocks $(A_i)_{i \in [0,1]}$ are realized and production takes place, $\bar{c}_i = A_i L_i$. In the evening, demand shocks $(\theta_i)_{i \in [0,1]}$ are realized, each individual chooses its monetary holdings M, and each family makes purchases of individual varieties, $(c_i)_{i \in [0,1]}$. In the evening, goods markets are supply-constrained: the families cannot purchase more of any variety than what has actually been produced, $\bar{c}_i$. To model this, we follow the standard approach in disequilibrium of imposing an anonymous and non-manipulable rationing rule (see Bénassy, 1993, for a discussion) of the form $c_{i\ell} = \min\{c_{i\ell}^d, \bar{c}_i\}$, where $c_{i\ell}^d$ is the demand for good i submitted by family ℓ and $\bar{c}_i$ is the available supply of good i. Adopting this rule, as opposed to a uniform rationing rule, has the desirable property that it is incentive compatible for families to submit demands which coincide with their desired demands.[7] Finally, to focus solely on the failure of market clearing in the goods markets, we

[7] As is well known (see *e.g.*, Bénassy, 1993), failures of incentive compatibility generally lead to a failure of equilibrium existence. As a simple example, suppose instead that a proportional rule is used that fills only a fraction α of submitted demands. Internalizing this, households would optimally inflate their demands by

will assume that the labor market and money market clear ($N = \int L_i \, \mathrm{d}i$, $M = \bar{M}$). Both of these assumptions could be relaxed along the lines of our multi-market extension from Section 2.4.

3.2 Optimal Household Behavior

Before fully defining our rational expectations equilibrium concept, it is useful to first characterize optimal household choices in both the morning and the evening.

To do this, we first describe the optimal allocation of family spending across markets i conditional on some (to-be-determined) level of total expenditure E. Under the rationing rule $c_{i\ell} = \min\{c_{i\ell}^d, \bar{c}_i\}$, where $c_{i\ell}^d$ is the demand for good i submitted by household ℓ and $\bar{c}_i$ is the available supply of good i, it is always as if the household chooses its optimal consumption level subject to the constraint that $c_{i\ell} \le \bar{c}_i$. As all households are identical, we henceforth again suppress ℓ subscripts and write each household's problem as follows:[8]

$$\max_{\{c_i\}_{i\in[0,1]}} \left(\int_0^1 \theta_i^{\frac{1}{\eta}} c_i^{\frac{\eta-1}{\eta}} \, \mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \quad \text{s.t.} \quad \int_0^1 p_i c_i \, \mathrm{d}i = E, \quad c_i \le \bar{c}_i \text{ for all } i \in [0,1] \tag{20}$$

To solve this problem, we define for each good i the statistic:

$$\iota_i = \frac{\frac{p_i^{-\eta}}{L_i}}{\frac{A_i}{\theta_i}} \tag{21}$$

Intuitively, this statistic represents the ultimate imbalance between supply and demand for each variety. Without loss of generality, we re-index i such that i is ordered by ι_i. That is, $i < j$ if and only if $\iota_i < \iota_j$. The following result characterizes household demand for varieties in this setting.

Proposition 2. *There exists a unique index threshold $k \in [0,1]$ such that:*

1. *For all varieties $j \in [0,k)$, the household's demand is given by:*

$$c_j = \frac{E - \int_k^1 p_i \bar{c}_i \, \mathrm{d}i}{\int_0^k \theta_i p_i^{1-\eta} \, \mathrm{d}i} \theta_j p_j^{-\eta} \tag{22}$$

2. *For all varieties $j \in (k,1]$, the household demands $c_j = \bar{c}_j$.*

---

a fraction of $1/\alpha$. However, to allocate the fixed resources, the fraction of submitted demands that is filled must fall to some $\alpha' < \alpha$. This logic repeats *ad infinitum*, leading to non-existence.

[8] Here, we are implicitly assuming there is sufficient production to allow the household to spend E. We later discuss what happens when a monetary injection is so large as to make this infeasible.

Moreover, the threshold k is the unique solution to Equation 22 when we replace j by k. The aggregate level of consumption is given by:

$$C = \left(\left(E - \int_k^1 p_i \bar{c}_i \,\mathrm{d}i \right)^{\frac{\eta-1}{\eta}} \left(\int_0^k \theta_i p_i^{1-\eta} \,\mathrm{d}i \right)^{\frac{1}{\eta}} + \int_k^1 \theta_i^{\frac{1}{\eta}} \bar{c}_i^{\frac{\eta-1}{\eta}} \,\mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \tag{23}$$

Proof. See Appendix A.4. □

The outcomes are as if a representative agent sequentially visited markets in the following way. The household starts by consuming everything produced in market $i = 1$, which is the most under-supplied. The household then goes to all subsequent, better-supplied markets and consumes everything that is produced until they hit a threshold market k at which they are exactly indifferent between consuming everything and just a little bit less. At this point, we have:

$$\bar{c}_k = \frac{E - \int_k^1 p_i \bar{c}_i \,\mathrm{d}i}{\int_0^k \theta_i p_i^{1-\eta} \,\mathrm{d}i} \theta_k p_k^{-\eta} \tag{24}$$

Then, for all subsequent markets, the household's demand is of the iso-elastic form:

$$c_j = z_j p_j^{-\eta} \tag{25}$$

but where:

$$z_j = D\theta_j \quad \text{with} \quad D = \frac{E - \int_k^1 p_i \bar{c}_i \,\mathrm{d}i}{\int_0^k \theta_i p_i^{1-\eta} \,\mathrm{d}i} \tag{26}$$

Thus, for all markets in which the household is not rationed, outcomes lie on the demand curve. A critical upshot of this analysis is that it generates the exact demand curves for the firm that we assumed in the partial equilibrium analysis; the amount that the firm actually sells is given by:

$$c_j = \min\{z_j p_j^{-\eta}, \bar{c}_j\} \tag{27}$$

Moreover, as $z_j = D\theta_j$, D is deterministic, and θ_j is log-normal, we also obtain that z_j is endogenously log-normally distributed, as we assumed in Section 2.

Cross-Substitution Without the Invisible Hand. In the limiting case in which all markets are slack, we recover the demand system of Dixit and Stiglitz (1977):

$$c_j = \theta_j \frac{E}{P} \left(\frac{p_j}{P} \right)^{-\eta}, \quad P = \left(\int_0^1 \theta_i p_i^{1-\eta} \,\mathrm{d}i \right)^{\frac{1}{1-\eta}} \tag{28}$$

In contrast, with rationing, we may write the demand system for all slack markets as

$$c_j = \theta_j \frac{E - \int_k^1 p_i \bar{c}_i \, \mathrm{d}i}{P_k} \left(\frac{p_j}{P_k} \right)^{-\eta}, \quad P_k = \left(\int_0^k \theta_i p_i^{1-\eta} \, \mathrm{d}i \right)^{\frac{1}{1-\eta}} \tag{29}$$

Hence, relative to the Dixit and Stiglitz (1977) benchmark, the demand shifter with rationing needs to allocate all income that is not spent on markets that stock out, $E - \int_k^1 p_i \bar{c}_i \, \mathrm{d}i$, on the remaining markets, which feature a price index of P_k. This also reveals that increasing the tightness of the economy by lowering $\bar{c}_i$ acts as a positive demand shifter for slack markets.[9] In other words, if consumers are unable to purchase certain goods because they are rationed, this will increase the demand for goods that are available. Thus, there are negative aggregate demand externalities from production in rationed markets on demand in slack markets. These *spillovers* link the aggregate tightness of the economy with the effective demand experienced by producers (and thus their chosen prices and quantities). Other studies have shown that the absence of market clearing in one market can modify the supply or demand of other markets, most notably in the presence of a homogeneous good and the labor market (Barro and Grossman, 1971; Bénassy, 1993). However, to the best of our knowledge, the characterization of these spillovers in the presence of imperfectly substitutable goods is novel to our analysis.

Labor Supply and Expenditure Decisions. The remainder of the household block yields standard conditions for labor supply and aggregate demand. In the morning, individuals supply labor and buy shares. Letting Λ denote the Lagrange multiplier on the household's budget constraint, we have from the first-order conditions for optimal labor choice that $w\Lambda = \alpha$, for optimal money holdings that $\frac{1}{M} = \Lambda$, and for optimal shareholding that $\frac{1}{s} = \Lambda P_\ell C_\ell$. As all households are symmetric, in equilibrium, $s = 1$, $P_\ell = P$, and $C_\ell = C$. Thus, we obtain the familiar solution under Golosov and Lucas (2007) preferences that wages are proportional to the money supply, nominal expenditures equal the money supply, and the marginal value of income is the inverse of nominal expenditures:[10]

$$w = \alpha M, \qquad E = PC = M, \qquad \Lambda = \frac{1}{PC} = \frac{1}{M} \tag{30}$$

[9] This can be proven formally by considering a small deviation $\bar{c}_i + \delta$ for a non-zero measure of goods $i \in [k, 1]$ and differentiating D with respect to δ.

[10] The assumption that families buy consumption varieties has only one, purely technical role in our analysis, which is ensuring that aggregate nominal expenditures are exogenous, in the tradition of Golosov and Lucas (2007). If instead each individual household purchased varieties, then the marginal value of wealth Λ would depend not on the average price P but instead the *marginal* cost of increasing C. The household's money demand would pin down the product of C and this marginal price. Under this structure, Proposition 2 would hold as written and the essential results of Theorems 3 and 4 would extend, but with different exact expressions. The results under this alternative assumption are omitted for brevity.

3.3 Rational Expectations Equilibrium

We study an equilibrium with the following key characteristics. When making choices in the morning, firms do not know the realizations of idiosyncratic shocks to productivity (A_i) and consumer tastes (θ_i). Firms and households form rational expectations about all variables. We formally describe equilibrium below:

Definition 1 (Equilibrium)**.** *An equilibrium is a collection of variables:*

$$\{\{t_i, p_i, L_i, c_i, A_i, \theta_i, z_i\}_{i\in[0,1]}, C, M, P, E, \Pi, T, N, k, w, \bar{M}, \Lambda\} \tag{31}$$

such that:

1. *Each firm $i \in [0, 1]$ chooses prices, labor, and tightness (p_i, L_i, t_i) according to Equations 5 and 6.*
2. *Households choose consumption of each variety $\{c_i\}_{i\in[0,1]}$ according to Equation 27 where the demand shifter $\{z_i\}_{i\in[0,1]}$ is given by Equation 26 and the threshold k satisfies Equation 24.*
3. *Households choose expenditure, consumption, money balances, and labor supply (E, C, M, N) according to equations 23 and 30.*
4. *The money market clears, the labor market clears, Π is the aggregate profit, the price index is $P = (\int p_i c_i \,\mathrm{d}i)/C$, and $\Lambda = 1/(PC)$ is the stochastic discount factor.*

When we study an unanticipated monetary shock, we consider a scenario in which the government changes the money supply from their announcement in the morning of $\bar{M}$ to some value $\bar{M}\Delta$ in the afternoon. Unanticipated monetary shocks differ from anticipated monetary shocks in the sole respect that all decisions made in the morning (production, pricing, and labor supply) cannot respond.

4 Pricing and Production in General Equilibrium

We now combine the firm and household problem to characterize equilibrium outcomes. We then use this model to study the propagation of monetary shocks and uncertainty shocks.

4.1 Characterizing Equilibrium

Firms' production and pricing choices generate spillovers in other markets through effective demand, D (defined in Equation 26). Anticipating this in rational expectations equilibrium, firms adjust their pricing and production decisions. We now characterize this two-way interaction between effective demand and firms' choices to describe aggregate output.

To understand how firms' choices matter, we first trace their effects through household expenditure. From the household's budget constraint and their first-order conditions for optimal choice, we have that total expenditure $M = PC$ is allocated across slack and rationed markets according to:

$$PC = \int_0^k p_i c_i \,\mathrm{d}i + \int_k^1 p_i c_i \,\mathrm{d}i = p \int_0^k \theta_i D p^{-\eta} \,\mathrm{d}i + p \int_k^1 A_i L \,\mathrm{d}i \tag{32}$$

where, in the second equality, we use that spending in slack markets lies on the demand curve ($c_i = \theta_i D p^{-\eta}$) while spending in rationed markets coincides with production ($c_i = \bar{c}_i = A_i L$).

We next observe that the firm's first-order condition for pricing relates exactly to these two expenditure terms. In the morning, firms equalize the expected marginal revenue losses if the market is slack with the gains if the market is rationed (recall Equation 7). Using the law of large numbers, we can express this in terms of average *ex post* values across markets:

$$(\eta - 1) \int_0^k \theta_i D p^{-\eta} \,\mathrm{d}i = \int_k^1 A_i L \,\mathrm{d}i \tag{33}$$

Moreover, if we multiply both sides by p, we obtain a relationship between average *sales* in slack and rationed markets.

Combining these conditions for optimal household and firm behavior, we obtain that:

$$PC = \frac{p}{\eta - 1} \int_k^1 A_i L \,\mathrm{d}i + p \int_k^1 A_i L \,\mathrm{d}i = \frac{\eta}{\eta - 1} p L \int_k^1 A_i \,\mathrm{d}i \tag{34}$$

Solving for real consumption C, and using the fact that the fraction of rationed markets is $r = 1 - k$, we immediately obtain the following result:

Proposition 3. *Equilibrium real output follows $C = \mathcal{A}L$, where*

$$\mathcal{A} = \frac{\eta}{\eta - 1} \times \underbrace{\frac{p}{P}}_{\textit{Misallocation}} \times \underbrace{r}_{\textit{Rationing}} \times \underbrace{\mathbb{E}[A_i | t \geq U_i]}_{\textit{Physical Productivity}} \tag{35}$$

where $\mathbb{I}[t \geq U_i]$ is the event that market i is rationed.

The variable $\mathcal{A}$ is the economy's Solow residual or aggregate productivity, measuring the wedge between inputs and *consumed* (or purchased) outputs. This, in turn, is determined by three endogenous forces. The first is *misallocation*, captured by the wedge $p/P < 1$. In this economy, the price of goods does not correspond to the real consumption deflator P. Intuitively, if all goods were available in slack markets, then households could consume one unit of C by buying an additional unit of each good; therefore, it would be the case that $P = p$. Instead, because some markets are rationed and, therefore, the economy is misallocated, purchasing C is more expensive. The second term is the rationing probability r. When more markets are rationed, a higher fraction of output produced is actually consumed, increasing the ratio of consumption to inputs (labor). The final term is *physical* productivity, but with a twist: holding fixed the other two terms, it is only productivity in rationed (supply-determined) markets that matters. Intuitively, these are likely to be markets in which productivity is relatively *low*.

Turning toward shock propagation, Proposition 3 makes clear that shocks can affect the economy via two channels: changing aggregate labor L and changing aggregate productivity $\mathcal{A}$.[11] In our simple model, we have adopted Golosov and Lucas (2007) preferences to shut down the more familiar channel that shocks affect equilibrium labor to focus on the more novel economics of how shocks affect the endogenous productivity of labor. Of course, with more general preferences, a similar analysis can be conducted.

4.2 The Effects of Uncertainty Shocks

As we showed earlier, our core point of departure from the standard, Walrasian model is the combination of *ex ante* firm choices and uncertainty. Therefore, we next describe how uncertainty affects general-equilibrium outcomes. We once more use the approach of studying a first-order perturbation of allocations in a linear scaling of the standard deviation of shocks.[12] The following theorem characterizes the first-order response of these variables to uncertainty (recalling that $\Omega = \phi(\kappa) - \kappa/\eta > 0$):

Theorem 3. *Up to a first-order approximation in volatility, the equilibrium price level and aggregate consumption are uniquely given by:*

$$\ln P = \ln P^{*,FI} + \Omega\sigma_U \quad \textit{and} \quad \ln C = \ln C^{*,FI} - \Omega\sigma_U, \tag{36}$$

where the constants $(\ln P^{*,FI}, \ln C^{*,FI})$ *are the full-information (Walrasian) price level and consumption level, which we provide in the appendix.*

Proof. See Appendix A.5. □

The first-order impact of uncertainty on the price level coincides with the partial equilibrium impact on a single firm's price derived in Theorem 2. In our model, equilibrium effects

[11] Of course, ours is not the only model in which shocks to demand and other factors propagate by changing aggregate productivity. However, as we will see, the precise mechanism in our model is different than those in existing approaches (see *e.g.,* Qiu and Ríos-Rull, 2022; Bai et al., 2025).

[12] The relevant shocks are A_i and θ_i. We therefore parametrize uncertainty by a scalar δ, letting $\vartheta_i = (\ln A_i, \ln \theta_i) \sim N(\mu, \delta^2\Sigma)$, and we consider a first-order expansion of allocations in δ around $\delta = 0$.

on firms' behavior operate through the *mean* level of demand; however, as shown earlier, these means do not affect the firm's desired probability to ration and, in turn, their desired prices. Moreover, *first-order* movements in the price level P are identical to first-order movements in the posted prices p. Finally, the equal and opposite response of aggregate consumption follows from the fact that, in equilibrium, household expenditure equals the total circulation of money in the economy ($PC = M$; Equation 30).

An implication of this analysis is that uncertainty shocks propagate as stagflationary or "cost-push" shocks into prices and consumption. In standard New Keynesian models, such shocks are typically microfounded via time-variation in the household's elasticity of substitution between different consumption varieties (see *e.g.,* Ireland, 2004), which are hard to relate to business cycle events in concrete economic terms. It is potentially appealing that uncertainty shocks, which are well-documented to be cyclical and large (see *e.g.,* Bloom, 2009; Bloom et al., 2018), act as cost-push shocks. Thus, our theory provides an explanation of the microeconomic origins of cost-push shocks, which are central to accounting for the business cycle (see *e.g.,* Smets and Wouters, 2007).

As we foreshadowed, the channel by which uncertainty shocks propagate is via changing aggregate labor productivity:

$$\mathrm{d}\ln\frac{C}{L} = \mathrm{d}\ln\mathcal{A} = -\Omega\,\mathrm{d}\sigma_U \tag{37}$$

Intuitively, because there is no invisible hand to clear markets, the consequence of higher prices is that goods are produced but go unsold.

4.3 The Effects of Monetary Shocks

We now study the propagation of monetary shocks. Anticipated increases in the money supply that are known to the firms in the morning are fully neutral for standard reasons, regardless of whether firms are uncertain about idiosyncratic shocks. An *un*anticipated increase in the evening would also be neutral in the full-information case of our model. Intuitively, firms produce just enough in the afternoon to clear markets in the evening at their predetermined prices, having anticipated some level of aggregate demand. Thus, when the monetary shock hits in the evening, consumers want to spend their additional money, but all goods are out of stock. By contrast, in our model *with* uncertainty, unanticipated monetary shocks can affect the real economy purely through the absorption of slack, even though (by construction) all goods' prices and physical production are unaffected.

To show this, we define the extent of monetary non-neutrality of an unanticipated monetary shock as:

$$\mathscr{M} = \left.\frac{\partial \ln C}{\partial \ln \Delta}\right|_{\ln \Delta = 0} \tag{38}$$

We also define the right and left derivatives of the passthrough as $\mathscr{M}^+$ and $\mathscr{M}^-$. These terms capture how much output expands to a small positive and a small negative change in demand, respectively.

Monetary Non-Neutrality. We begin by characterizing the extent of monetary non-neutrality and its relationship with underlying firm-level uncertainty:

Theorem 4. *Up to a first-order approximation in volatility, the extent of monetary non-neutrality is given by:*

$$\mathscr{M} = 1 - \frac{1}{\eta}\left(\phi(\kappa) + \frac{\eta - 1}{\eta}\kappa\right)\sigma_U \tag{39}$$

Proof. See Appendix A.6. □

Theorem 4 is intimately related to how much slack there is in the economy. To see this, consider for exposition the case with only productivity uncertainty. We can write the elasticity of consumption with respect to an unanticipated change in money balances as

$$\frac{\partial \ln C}{\partial \ln \Delta} = \frac{M \times C^{1/\eta} \times \int_0^k c_i^{-1/\eta}\,\mathrm{d}i \times \left(\frac{\partial}{\partial M} z p^{-\eta}\right)}{C} \tag{40}$$

The last two terms reflect the fact that consumption can only increase in slack markets, where sales are demand-determined. Furthermore, by the definition of the demand shifter, an increase in one unit of money raises the quantity demanded in slack markets by $(1/\int_0^k p^{1-\eta}) \times p^{-\eta} = \eta/(kp)$. Thus, since $1/\eta$ fraction of markets are slack in equilibrium in the zero-uncertainty limit, households spend η more units of money in each slack market and increase their consumption of each variety in excess supply by η/p. Using the definition of the price index $P = \int p_i c_i\,\mathrm{d}i\,/C$ and the fact that nominal expenditures are equal to money holdings, $PC = M$, we can then rewrite Equation 40 as

$$\frac{\partial \ln C}{\partial \ln \Delta} = \underbrace{\left(\frac{\int_0^1 c_i\,\mathrm{d}i}{C}\right)^{\frac{\eta-1}{\eta}}}_{\text{consumption misallocation}} \times \underbrace{\eta \int_0^k \left(\frac{\int_0^1 c_i\,\mathrm{d}i}{c_i}\right)^{\frac{1}{\eta}} \mathrm{d}i}_{\text{average excess consumption in slack markets}} \tag{41}$$

This makes it clear that the passthrough of an aggregate demand shock is determined by two features. First, as reflected by the first term, it depends on the extent of misallocation in the economy, holding fixed the average consumption of each variety. This term is unity to first order in uncertainty, which reflects the well-known result that "small" distortions on consumption only affect welfare to second-order (Harberger, 1964). In contrast, the second term differs from unity if uncertainty is non-zero and reflects how much more consumption is present in slack markets relative to the average market. Since goods are imperfectly substitutable and households cannot increase consumption of rationed varieties, this term reduces the passthrough of an aggregate demand shock. The term $\phi(\kappa) + \frac{\eta-1}{\eta}\kappa$ reflects this average consumption gap and is strictly positive for all values of η.

Beyond these first-order effects of monetary shocks, there are also interesting non-linearities that arise for two reasons. First, as households consume goods in previously slack markets, they consume goods of progressively lower quality (which is why these goods were not being sold to begin with). Thus, each progressive dollar of monetary stimulus has a progressively smaller effect on real consumption. Second, at some level of money injection, the economy runs out of slack altogether, and further monetary stimulus ceases to have any real effects. Of course, in a dynamic setting (as we will see in Section 5), the economy could produce in future periods to absorb any (persistent) change in money balances.

The Role of Uncertainty. This argument also reveals that uncertainty and slack lie at the heart of monetary non-neutrality. Concretely, as uncertainty vanishes, so does the average consumption gap between the average market and slack markets. Thus, in the zero uncertainty limit, the passthrough of the money shock to consumption is equal to unity. Moreover, uncertainty is critical to ensure that the passthrough of monetary policy is well-defined (even in the limit). Indeed, as discussed previously, if there is no uncertainty, markets are in a state of Walrasian equilibrium and there is no slack. Thus, an unanticipated increase in the money supply renders money neutral. Conversely, an unanticipated contraction implies that money reduces consumption one-to-one, because households reduce the consumption of all goods uniformly. This discontinuity arises because of the kink in the aggregate supply curve: as uncertainty vanishes, optimal producer behavior ensures that output lies exactly on the kink. We now write $\mathscr{M}(\delta)$ as the extent of monetary non-neutrality when uncertainty about supply-demand imbalances is given by δ. We formalize this discussion in the following corollary.

Proposition 4. *The extent of monetary non-neutrality is discontinuous in the zero-uncertainty limit. That is, we have that:*

$$\lim_{\delta \to 0} \mathscr{M}^+(\delta) = 1, \ \lim_{\delta \to 0} \mathscr{M}^-(\delta) = 1 \quad \textit{and} \quad \mathscr{M}^+(0) = 0, \ \mathscr{M}^-(0) = 1 \tag{42}$$

Proof. See Appendix A.7. □

The Role of Market Power. Finally, given uncertainty σ_U, we observe that the degree of monetary neutrality is determined solely by market power, as captured by the demand elasticity η. The following proposition identifies the value of η for which monetary passthrough is dampened as much as possible.

Proposition 5. *The extent of monetary non-neutrality, $\mathscr{M}$, is non-monotonically related to market power. In particular, to first order in volatility, $\mathscr{M}$ is strictly decreasing in η if $\eta < \eta^*$ and strictly increasing in η if $\eta > \eta^*$, where $\eta^* = \frac{1}{1-\Phi(\kappa^*)} \approx 2.83$ and κ^* is the unique solution to the equation:*

$$1 = \frac{\phi(x)}{1-\Phi(x)}\left(x + \frac{\phi(x)}{\Phi(x)}\right) \tag{43}$$

Proof. See Appendix A.8. □

The non-monotonic nature of this result can be explained as follows. As established in Theorem 4, the degree of monetary non-neutrality falls as the consumption gap between the average slack market and the average market in the entire economy widens. The magnitude of this mechanism weakens with higher firm market power: when market power is lower, the relative cost of overproduction rises and the likelihood that a market becomes rationed increases (Proposition 1). At the same time, a reduction in market power makes goods more substitutable, which diminishes the relative welfare losses from consuming "overproduced" varieties. The cutoff value $\eta^* \approx 2.83$ pins down the point at which these two opposing effects offset each other. Interestingly, this cutoff value does not depend on any model parameters.

5 Dynamics: A New Old Keynesian Model

To most transparently discuss the novel economics of our setting, our baseline model was static (in the style of Blanchard and Kiyotaki, 1987). To make the parallel to the New Keynesian model as tight as possible, we now extend our baseline model of the firm to a dynamic setting and subject firms to Calvo (1983) stickiness. The key departure of our model from the standard New Keynesian model is that firms' prices and quantities are *both* sticky, and so goods markets do not necessarily clear in each period. In this sense, this model marries elements of the New Keynesian model of price setting with an Old Keynesian model of Walrasian disequilibrium. Our model differs relative to the Old Keynesian tradition in the critical respect that Walrasian disequilibrium is not assumed; it is instead an outcome of a rational expectations equilibrium. For these reasons, we refer to this model as the New Old Keynesian (NOK) Model.

5.1 The NOK Model

Time is discrete and infinite, indexed by $\tau \in \mathbb{N}$ to avoid confusion with market tightness t.

Households. The household side of the model is identical to our baseline model except that they now act dynamically. That is, we continue to assume that:

$$C_\tau = \left(\int \theta_{i,\tau}^{\frac{1}{\eta}} c_{i,\tau}^{\frac{\eta-1}{\eta}} \, \mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \tag{44}$$

where $\theta_{i,\tau}$ is IID lognormal with mean μ_θ and variance $\sigma^2_{\theta,\tau}$. The household has preferences over streams of aggregate consumption, money balances $\tilde{M}_\tau$, and labor supply given by:

$$\sum_{j=0}^{\infty} \beta^j \left(\ln C_{\tau+j} + \ln \tilde{M}_{\tau+j} - \alpha N_{\tau+j} \right) \tag{45}$$

where $\beta \in [0,1)$ is the discount factor and the ideal price index is given by $P_\tau = (\int p_{i,\tau} c_{i,\tau} \, \mathrm{d}i)/C_\tau$. As before, it is as if each household faces the flow budget constraint:

$$P_\tau C_\tau + \tilde{M}_\tau + B_\tau = w_\tau N_\tau + \Pi_\tau + T_\tau + \tilde{M}_{\tau-1} + (1 + i_{\tau-1}) B_{\tau-1} \tag{46}$$

where the only new ingredient is the presence of nominal bonds B_τ, which have nominal interest rate i_τ. Thus, the household's optimality conditions imply in equilibrium that:[13]

$$w_\tau = \alpha P_\tau C_\tau \quad \text{and} \quad P_\tau C_\tau = \frac{i_\tau}{1 + i_\tau} \tilde{M}_\tau \tag{47}$$

Thus, as in Golosov and Lucas (2007), monetary policy controls nominal GDP and wages. It is not essential that the monetary authority does this through changing the money supply or interest rates. In what follows, we denote nominal aggregate demand by $M_\tau \equiv \frac{i_\tau}{1+i_\tau} \tilde{M}_\tau$ and treat it directly as an exogenous shock process.[14] We also emphasize that it is simple to supplement the household side of this model with a dynamic IS equation, as in the standard New Keynesian model.

Importantly, household demand over varieties continues to follow Proposition 2:

$$c_{i,\tau} = \min\{ z_{i,\tau} p_{i,\tau}^{-\eta}, \bar{c}_{i,\tau} \} \tag{48}$$

[13]See, *e.g.*, Proposition 1 of Flynn et al. (2026).

[14]A cash-in-advance constraint would yield similar equilibrium conditions (see Woodford, 2003a, for a discussion of these points).

where:

$$z_{i,\tau} = D_\tau \theta_{i,\tau} \quad \text{with} \quad D_\tau = \frac{M_\tau - \int_{k_\tau}^1 p_{i,\tau}\bar{c}_{i,\tau}\,\mathrm{d}i}{\int_0^{k_\tau} \theta_{i,\tau} p_{i,\tau}^{1-\eta}\,\mathrm{d}i} \tag{49}$$

and k_τ solves $\bar{c}_{k_\tau,\tau} = z_{k_\tau,\tau} p_{k_\tau,\tau}^{-\eta}$, where we are re-indexing i in each period by $\iota_{i,\tau} = t_{i,\tau}/U_{i,\tau}$, as we did in the static model.

Firms. The departure of this model from our baseline framework is that firms' decisions are sticky: there is a constant probability $\omega \in [0, 1)$ that the firm cannot reset its decision over both prices and labor in each period. Thus, this model extends the standard Calvo (1983) model to also incorporate stickiness along the input margin. The rest of our model is as before. The firm produces $A_{i,\tau} L_{i,\tau}$, where $A_{i,\tau}$ is IID log-normal with mean μ_A and variance $\sigma^2_{A,\tau}$, and independent of $\theta_{i,\tau}$. Moreover, the nominal SDF is $\Lambda_\tau = (P_\tau C_\tau)^{-1}$. Given these assumptions, the firm's problem at time τ is:

$$\sup_{p,L} \sum_{j=0}^{\infty} (\beta\omega)^j \mathbb{E}_{i,\tau}[\Lambda_{\tau+j} \left(\min\{z_{i,\tau+j} p^{1-\eta}, p A_{i,\tau+j} L\} - w_{\tau+j} L\right)] \tag{50}$$

where $\mathbb{E}_{i,\tau}$ denotes the conditional expectation given knowledge of all shocks that realized in periods $\hat{\tau} \leq \tau - 1$. Applying the same arguments as the proof of Theorem 1, firms' optimal tightness and prices must satisfy the following conditions:

$$\begin{aligned} t_{i,\tau} &= \frac{1}{\eta - 1} \frac{\sum_{j=0}^{\infty} (\beta\omega)^j \mathbb{E}_{i,\tau}[\Lambda_{\tau+j} A_{i,\tau+j} \mathbb{I}[t_{i,\tau} \geq U_{i,\tau+j}]]}{\sum_{j=0}^{\infty} (\beta\omega)^j \mathbb{E}_{i,\tau}[\Lambda_{\tau+j} z_{i,\tau+j} \mathbb{I}[t_{i,\tau} \leq U_{i,\tau+j}]]} \\ p_{i,\tau} &= \frac{\sum_{j=0}^{\infty} (\beta\omega)^j \mathbb{E}_{i,\tau}[\Lambda_{\tau+j} w_{\tau+j}]}{\sum_{j=0}^{\infty} (\beta\omega)^j \mathbb{E}_{i,\tau}[\Lambda_{\tau+j} A_{i,\tau+j} \mathbb{I}[t_{i,\tau} \geq U_{i,\tau+j}]]} \end{aligned} \tag{51}$$

and labor inputs are determined residually as $L_{i,\tau} = p_{i,\tau}^{-\eta}/t_{i,\tau}$. Critically, these equations just involve present discounted values of the same conditional expectations that characterized the firm's optimal behavior in the static setting. Moreover, they retain the block-recursive structure; once the fixed point for $t_{i,\tau}$ has been solved, the optimal choices for $p_{i,\tau}$ and $L_{i,\tau}$ can be determined residually.

Monetary Policy. We consider a situation in which money and interest rates are initially held fixed such that nominal demand in the steady-state is $\bar{M}$. At date $\tau = 0$, it is announced that nominal aggregate demand will follow $M_\tau = \bar{M}\Delta_\tau$ for some sequence $\{\Delta_\tau\}_{\tau\in\mathbb{N}}$ such that $\Delta_\tau \to \Delta \in \mathbb{R}_{++}$. As mentioned above, these aggregate demand shocks can be equivalently driven by changes to the money supply, interest rates, or both.

5.2 The Linearized NOK Model

To analyze how this dynamic model behaves, we log-linearize it around its steady state. In the steady state, the previous equations imply that $P_\tau \equiv P$, $C_\tau \equiv C$, $t_{i,\tau} \equiv t$, $p_{i,\tau} \equiv p$. The solution of this model reduces to that of our static model. Indeed, the firm's optimality conditions reduce to:

$$tD = \frac{1}{\eta - 1} \frac{\mathbb{E}[A\mathbb{I}[tD \geq A/\theta]]}{\mathbb{E}[\theta\mathbb{I}[tD \leq A/\theta]]} \quad \text{and} \quad p = \frac{\alpha \bar{M}}{\mathbb{E}[A\mathbb{I}[tD \geq A/\theta]]} \tag{52}$$

Moreover, D solves:

$$D = \frac{\bar{M} - pL \int_k^1 A_i \, \mathrm{d}i}{p^{1-\eta} \int_0^k \theta_i \, \mathrm{d}i} \quad \text{where} \quad k = 1 - r = 1 - \Phi\left(\frac{\ln(Dt) - (\mu_A - \mu_\theta)}{\sqrt{\sigma_A^2 + \sigma_\theta^2}}\right) \tag{53}$$

completing the characterization of the steady state.

The shocks in this model are those to nominal aggregate demand $\hat{M}_\tau$ and those to the volatility of the supply-demand imbalance $\hat{\sigma}_{U,\tau}$. For any variable Z with steady-state value $\bar{Z}$, we let $\hat{Z} = \ln\big(Z/\bar{Z}\big)$ denote the log-deviation from the steady state. We also define $\bar{t}_\tau$ and $\bar{p}_\tau$ as the average log-deviation from the steady state in firms' tightness and prices, and we denote price inflation by $\pi_\tau = \Delta\bar{p}_\tau$ and tightness inflation by $\upsilon_\tau = \Delta\bar{t}_\tau$. We have that:

Proposition 6. *When log-linearized around the steady state and given a process for aggregate demand shocks and volatility shocks $\{\hat{M}_\tau, \hat{\sigma}_{U,\tau}\}_{\tau\in\mathbb{N}}$, the dynamic equilibria of the model solve the following three equations:*

1. *The New Keynesian Phillips Curve:*

$$\pi_\tau = \Gamma(\hat{M}_\tau - \bar{p}_\tau) + \beta\mathbb{E}_\tau[\pi_{\tau+1}] + \zeta\hat{\sigma}_{U,\tau} \tag{54}$$

2. *The New Old Keynesian Tightness Curve:*

$$\upsilon_\tau = -\Gamma(\hat{D}_\tau + \bar{t}_\tau) + \beta\mathbb{E}_\tau[\upsilon_{\tau+1}] + \psi\hat{\sigma}_{U,\tau} \tag{55}$$

3. *The Demand Spillovers Curve:*

$$\hat{D}_\tau = \gamma_M\hat{M}_\tau + \gamma_p\bar{p}_\tau + \gamma_t\bar{t}_\tau + \gamma_U\hat{\sigma}_{U,\tau} \tag{56}$$

where $\Gamma = \frac{(1-\omega)(1-\beta\omega)}{\omega}$ and the remaining constants $(\zeta, \psi, \gamma_M, \gamma_p, \gamma_t, \gamma_U)$ are derived in the proof of the result.

Proof. See Appendix A.9. □

In this representation, we track the three key endogenous state variables in the economy: inflation, market tightness, and demand spillovers. Using these variables, we can calculate the equilibrium behavior of consumption, labor, and other aggregates.

Consistent with our results in static equilibrium (Theorem 3), uncertainty shocks continue to have first-order effects on allocations. In particular, a spike in uncertainty appears as if a "cost push shock" in the New Keynesian Phillips curve, increasing inflation holding fixed marginal cost. Uncertainty also has first-order effects on tightness inflation and demand spillovers for largely the same reasons as in the static analysis.

Very strikingly, the dynamics of the New Keynesian Phillips curve in Equation 54 show that the first-order *propagation* of shocks is just as in the standard New Keynesian model. Intuitively, this arises because the effects of (current and future) rationing on desired markups have no first-order effects *in deviations from the steady state*: that is, while they affect steady-state prices and markups themselves, they do not affect conditional price dynamics. This suggests, as we will soon verify in a (nonlinear) numerical example below, that the broad dynamic behavior of inflation—a variable held fixed in our intentionally stylized static analysis (Section 3)—is as in the standard benchmark analysis. We also remark that, due to the tight link between the price level, nominal demand, and consumption ($C_\tau = M_\tau/P_\tau$, for all time periods τ), this implies similar dynamics for consumption across the models.

Nonetheless, the *mechanics* of shock transmission—as well as the productivity and input market effects—will prove to be quite different, and much more in line with the Old Keynesian tradition. In particular, shock transmission is governed by the dynamic coevolution of market tightness and cross-market demand spillovers. Below, we show how this leads to a qualitatively different picture for *which* shocks are relevant for the business cycle and *why* shocks have the effect they do on inflation and consumption.

5.3 Calibration

Having described the intuition for the linearized dynamics, we now study the fully nonlinear shock responses in the New Old Keynesian Model. We compare these dynamics to their equivalent under the standard New Keynesian model under the same microfoundations and parameter values. Thus, concretely, the two models considered differ only in their assumptions about the markets for intermediate goods. In the NOK model, firms choose prices and inputs in advance, so individual markets may be slack or rationed; whereas in the comparable NK model, inputs are fully adjustable *ex post* to ensure market clearing.

We choose intentionally standard parameters for illustration: a quarterly discount rate of $\beta = 0.99$; an elasticity of substitution of $\eta = 8$ (consistent with firm-level evidence from

Hottman et al., 2016); and a stickiness parameter (probability decisions are fixed quarter-to-quarter) of $\omega = 0.75$, implying an average duration of prices and labor of 4 quarters. In Appendix B.4, we show how the model can be used to compute microeconomic volatility using the reported volatilities of physical TFP (TFTQ) and revenue TFP (TFPR) from Foster et al. (2008) along with information from Bloom et al. (2018) about the unforecastability of firm-level shocks. Using this method, we calibrate $\sigma_A = 0.10$ and $\sigma_z = 0.21$, so $\sigma_U = 0.23$.

5.4 The Propagation of Monetary Shocks

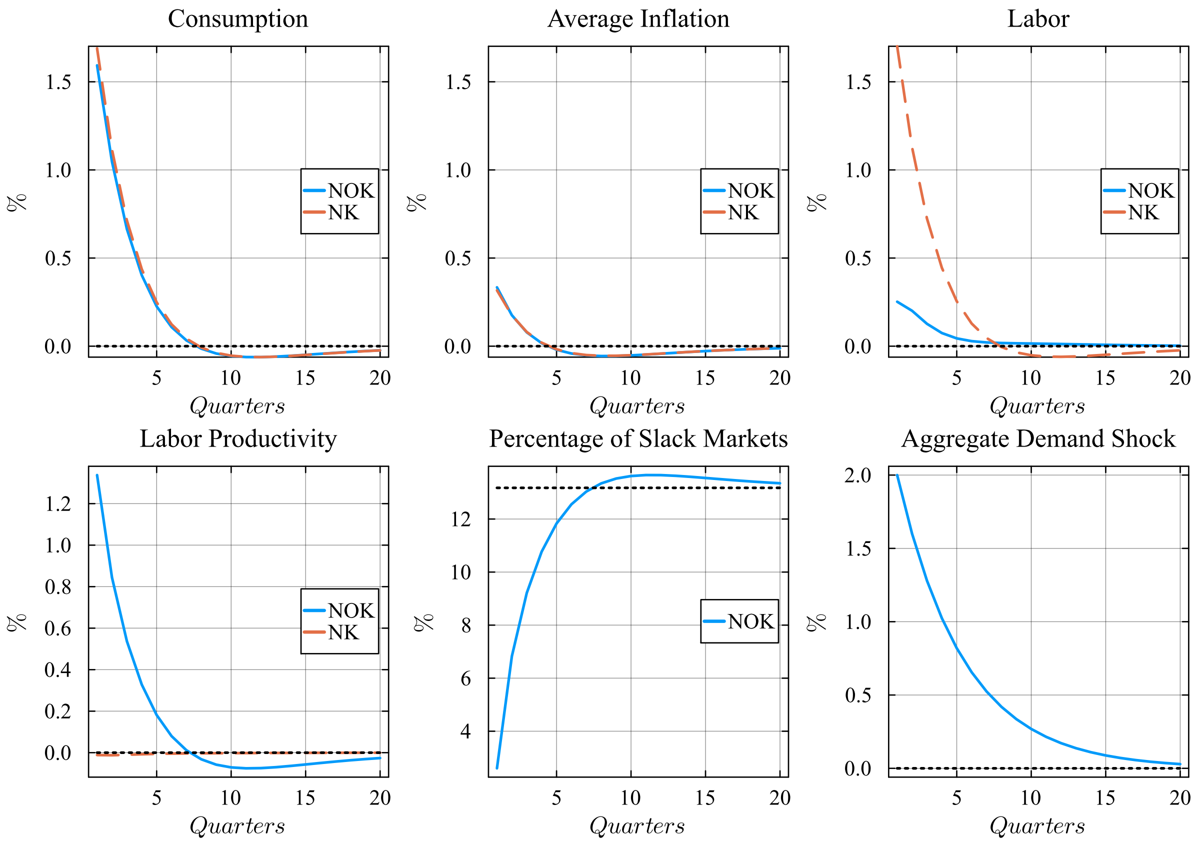
We first study a temporary monetary expansion, which is the dynamic analogue of the analysis in Section 4.3. We plot the non-linear dynamics of the NK and NOK models following an unexpected and temporary 2% increase in nominal aggregate demand M_τ in Figure 4.

Consider first the standard New Keynesian (NK) framework, where both aggregate consumption and inflation rise. The increase in inflation is driven by changes in firms' marginal costs via wages, as summarized by the canonical New Keynesian Phillips Curve in Equation 54. Output then changes because firms commit to hiring the labor necessary to meet the corresponding changes in aggregate demand. Consequently, the change in aggregate labor closely tracks the change in aggregate consumption. Furthermore, labor productivity falls slightly, since price dispersion lowers the marginal utility derived from consuming each individual variety. Finally, note that in the NK model, there is no slack as all markets are in Walrasian equilibrium.

Next, consider the NOK model. The response of inflation is nearly identical in the NOK and NK models. As previously explained, this is because changes in firms' rationing probabilities are only second-order in determining firms' reset prices. Hence, to a very accurate first-order approximation, inflation dynamics are determined by the same New Keynesian Phillips Curve in both models. Since aggregate consumption is simply total nominal demand divided by the price level, the path of aggregate consumption must also be similar in both models.

However, the monetary transmission mechanism in the NOK model is fundamentally different. In the NK setting, higher consumption is generated by a potentially implausible commitment assumption in which firms hire more labor in input markets to meet demand while holding prices fixed. In contrast, in the NOK model, the rise in consumption is primarily due to a reduction in slack in the economy. For instance, on impact, the share of slack markets decreases from 12.5% to 1%. Notably, this does not mean that total production is unchanged in the NOK model. Instead, firms that can adjust their decisions find it

Figure 4: Impulse Responses to an Aggregate Demand Shock



Note: This figure shows the response of macroeconomic variables to a temporary 2% increase in the money supply at $\tau = 0$, beginning from a steady state and decaying at 20% per quarter. Each panel contrasts the responses in the New Old Keynesian model (blue solid line) with those in the New Keynesian model (orange dashed line) with the same parametrization. The bottom right plot shows the path of the money supply M_τ. "Consumption" is the consumption quantity index (Equation 44); "Average Inflation" is the growth in the ideal price index, P_τ; "Labor" is total labor hours; "Labor productivity" is GDP (Consumption) per labor hour; and the "Percentage of Slack Markets" is the percentage of markets for intermediate goods in which supply exceeds demand. All variables except inflation and the percentage of slack markets are expressed in percent deviations from the steady state.

optimal to raise prices and expand labor input in order to take advantage of the temporarily higher aggregate demand. Critically, the increase in consumption arises from both higher production *and* lower slack. This is reflected in an increase in measured labor productivity, as shown in the bottom-left panel.

This slack-absorbing transmission mechanism for monetary expansions (or slack-creating mechanism for contractions) is consistent with evidence from Buda et al. (2025) on the perhaps surprisingly immediate effects of monetary policy shocks that are visible in data that aggregate economic transactions at a high frequency. In particular, these authors show that a one-standard-deviation monetary policy shock in their data affects sales, consumption, and investment on the order of 1% within only 1 to 2 months, while employment is reduced by a tenth of this (0.1%). This is hard to reconcile with the New Keynesian benchmark studied here, in which output, production, and employment must move one-for-one with one another; it is also likely inconsistent with any alternative view that incorporates capital in production, given reasonable values for the elasticity of the capital stock to investment and the elasticity of output to the capital stock. The findings are, however, consistent with the monetary transmission mechanism in the NOK model, which operates through changes in product market slack: in the short run, as aggregate demand increases, consumers flood into previously slack markets; in the longer run, firms slowly begin to respond by also ramping up production.

We finally observe that both the high-frequency analysis of Buda et al. (2025) as well as more classical, business-cycle frequency analysis in the US (Christiano et al., 2005) imply a significant response of labor productivity to monetary policy shocks. In the New Keynesian benchmark we study, this is impossible as labor productivity is purely determined by (exogenous) technology and quantitatively negligible second-order effects of misallocation. In the New Old Keynesian Model, as markets tighten and transactions increase, the economy produces more GDP for the same amount of labor inputs despite there being no change in underlying production technology.

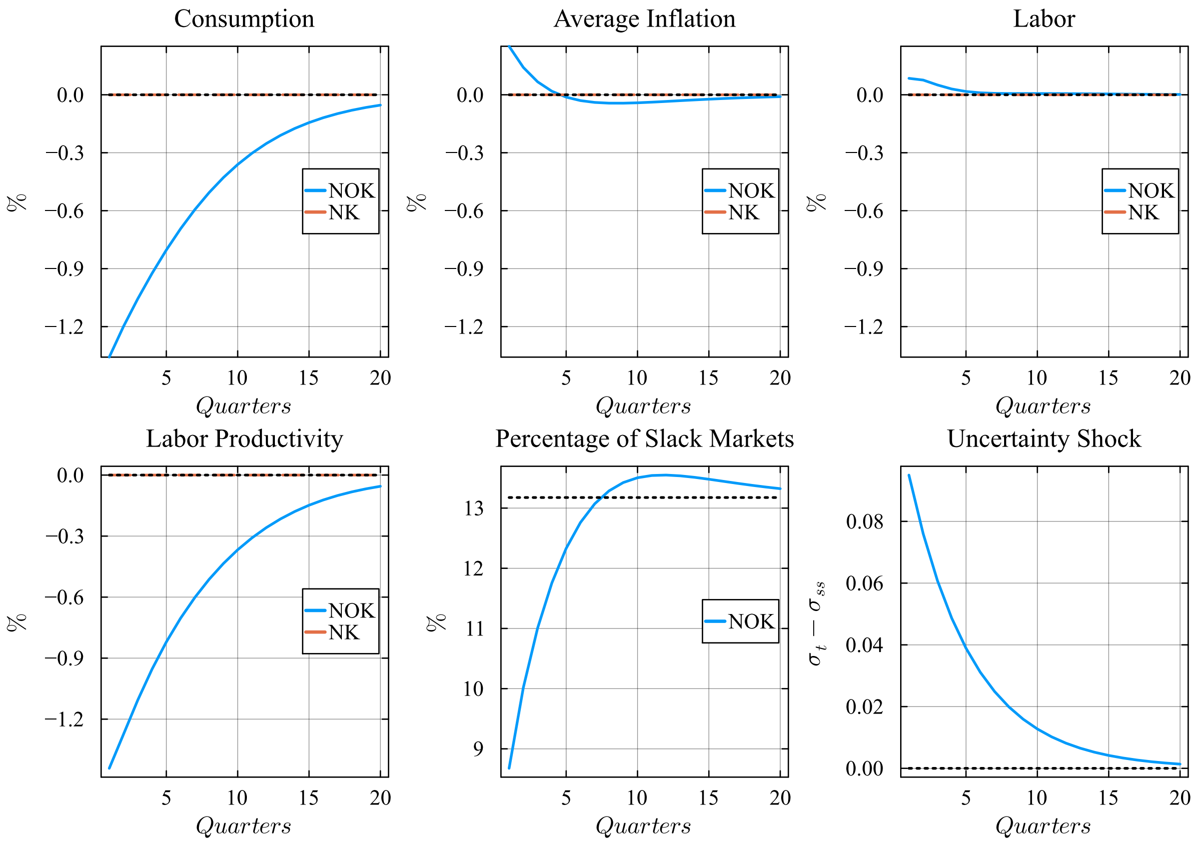
5.5 The Propagation of Uncertainty Shocks

We next consider a temporary increase in the uncertainty that firms face about idiosyncratic demand and productivity, which is the dynamic analogue of the analysis in Section 4.2. We think of this as a stylized representation of the "uncertainty shocks" uncovered by previous work using firm-level microdata. For instance, Bloom et al. (2018) estimate that the standard deviation of firm-level shocks (specifically, to revenue-based TFP) can more than quadruple in a recession versus a normal state. We study a much smaller shock in which uncertainty regarding the relevant variable, the supply-demand shifter $U = A/z$, increases by 0.1, or 43% of its steady-state value.[15] We plot the non-linearly solved response of macroeconomic variables to this shock in Figure 5.

In the standard New Keynesian model, uncertainty shocks have zero effect on economic

[15] We note that our definition of an uncertainty shock, while consistent with the main empirical findings of Bloom et al. (2018), differs from the notion of an "uncertainty shock" studied by Basu and Bundick (2017) in a variant of the New Keynesian model. These authors study a change in the volatility of an aggregate preference shock for the representative consumer, calibrated to generate a realistic spike in the forward volatility of aggregate equity returns.

Figure 5: Impulse Responses to an Uncertainty Shock



Note: This figure shows the response of macroeconomic variables to a transitory increase in uncertainty, beginning from a steady state and decaying at 20% per quarter. Each panel contrasts the responses in the New Old Keynesian model (blue solid line) with those in the New Keynesian model (orange dashed line) with the same parametrization. The shock, shown in the bottom right panel, corresponds to a temporary increase in uncertainty about the idiosyncratic demand shock θ and the idiosyncratic productivity shock A, such that $\sigma_U = \mathrm{StdDev}[\log A - \log \theta]$ rises by 0.1 on impact, and the ratio of uncertainty regarding θ and A remains constant. "Consumption" is the consumption quantity index (Equation 44); "Average Inflation" is the growth in the ideal price index, P_τ; "Labor" is total labor hours; "Labor productivity" is GDP (Consumption) per labor hour; and the "Percentage of Slack Markets" is the percentage of markets for intermediate goods in which supply exceeds demand. All variables except inflation, the percentage of slack markets, and uncertainty are expressed in percent deviations from steady state.

allocations. Firms' optimal reset prices are invariant to the level of product demand as well as uncertainty thereof. Because the "invisible hand" determines quantities produced, there is also no higher-order effect on aggregates via misallocation in the economy. As such, the shocks have no effects.

In the New Old Keynesian model, uncertainty shocks lead to sharp contractions in consumption and labor productivity along with a temporary spike in inflation. Echoing our findings in the static model (Theorem 3), adjusting firms increase prices and slightly *increase* labor. Nonetheless, significantly more markets end up rationed and total consumption declines on impact. This shows up as a large decline in labor productivity (GDP per worker), despite constant physical productivity and almost constant labor inputs. In essence, the economy has contracted because of price increases and widespread "stockouts" (rationing). This broad pattern of transmission resembles that of a markup shock in the standard New Keynesian model, but for the fact that transmission is largely through changes in market tightness rather than changes in physical production.

Our findings offer one possible explanation for empirical findings that increases in economic uncertainty, broadly defined, lead to economic contractions and inflation (*e.g.*, Bloom, 2009; Basu et al., 2021; Mumtaz and Ruch, 2025). This happens for a particularly stark reason in our model: even though physical output is close to constant, firms' high prices discourage purchases. Thus, it appears as if the economy has lost productivity in the aggregate, even though technology has not changed.

By the same token, our findings can also be interpreted as an alternative take on what econometricians identify as a "productivity shock" in aggregate data. Provided that standard approaches to adjusting total factor productivity (Fernald, 2014) for utilization rates are imperfect, especially for measuring wasted nondurable output or idle time in services, our uncertainty shock manifests as an immediate decline in aggregate productivity. In the empirical literature, such shocks have negative effects on output and positive effects on inflation, but ambiguous effects on labor hours and investment (Gali, 1999; Basu et al., 2006). Moreover, it remains contested whether such shocks can be reasonably linked back to underlying changes in production technology (see, *e.g.*, the discussion in Alexopoulos, 2011). Our model allows for aggregate productivity shocks to arise from the interaction of uncertainty and the malfunctioning of the invisible hand, potentially helping to resolve this tension.

6 Conclusion

To overcome the Arrow (1959) critique and its modern inversion, we have developed a general equilibrium model in which both prices and quantities are determined by rational economic actors, not the invisible hand. The resulting model generates a number of interesting economic predictions that standard theories do not.

At the microeconomic level, the presence of slack and rationing changes production and pricing decisions in qualitatively important ways. For instance, optimal monopoly pricing no longer obeys the Lerner (1934) formula and uncertainty has first-order effects on firms' decisions, increasing prices and depressing labor demand. At the macroeconomic level, we illustrate that the presence of uncertainty gives rise to a well-defined distribution of slack and rationed markets. This allows us to overcome the combinatorial challenge of market clearing, which has inhibited research on previous models of Walrasian disequilibrium (Bénassy, 1993) and characterize cross-market spillovers in a classic Dixit and Stiglitz (1977) framework. Our results reveal that the presence of rationed markets substantially alters the standard logic under monopolistic competition that there are positive aggregate demand externalities (Blanchard and Kiyotaki, 1987): increasing production in rationed markets induces negative demand spillovers in slack markets. Moreover, the propagation of demand and uncertainty shocks is qualitatively different from standard theories. In particular, demand shocks raise aggregate consumption by reducing economic slack, generating increases in aggregate labor productivity while microeconomic productivity remains unchanged. Moreover, uncertainty shocks matter to first order and raise prices while depressing output, akin to a cost-push or markup shock.

Finally, we have shown how this framework can be tractably combined with sticky decisions (Calvo, 1983) in a dynamic setting: one must simply add the dynamic tightness curve and the demand spillovers curve to a standard New Keynesian Phillips curve. Indeed, as our framework is modular, it could be tractably embedded in any existing equilibrium business cycle model, including larger-scale DSGE models. The simple New Old Keynesian (NOK) model that we introduce has the feature that demand and uncertainty shocks—as suggested by the static theory—propagate very differently than in New Keynesian economies. Beyond potential conceptual appeal, the NOK model can account for three important facts that simple New Keynesian models struggle to rationalize: (i) uncertainty shocks have large contractionary effects on output while generating inflation (Bloom, 2009; Mumtaz and Ruch, 2025), (ii) labor productivity increases in response to demand shocks (Christiano et al., 2005), and (iii) the response of consumption to demand shocks is much larger than the response of labor to demand shocks (Buda et al., 2025). We therefore argue that the propagation mechanisms revealed by the NOK model are both microeconomically appealing and capable of matching important facts. The incorporation of further realistic features, such as disequilibrium in labor markets and investment, would be an interesting avenue for future research.

References

Alexopoulos, M. (2011). Read all about it!! what happens following a technology shock? *American Economic Review*, 101(4):1144–79.

Angeletos, G.-M. and La'O, J. (2010). Noisy business cycles. *NBER Macroeconomics Annual*, 24(1):319–378.

Arrow, K. J. (1959). *Toward a theory of price adjustment*, volume 1959. Stanford University Press Stanford.

Bachmann, R., Carstensen, K., Lautenbacher, S., and Schneider, M. (2026). Uncertainty and change: Survey evidence of firms' subjective beliefs. *American Economic Review*. Forthcoming.

Bai, Y., Rios-Rull, J.-V., and Storesletten, K. (2025). Demand shocks as technology shocks. *Review of Economic Studies*.

Barro, R. J. (2025). The old keynesian model. Technical report, National Bureau of Economic Research.

Barro, R. J. and Grossman, H. I. (1971). A general disequilibrium model of income and employment. *The American Economic Review*, 61(1):82–93.

Basu, S. and Bundick, B. (2017). Uncertainty shocks in a model of effective demand. *Econometrica*, 85(3):937–958.

Basu, S., Candian, G., Chahrour, R., and Valchev, R. (2021). Risky business cycles. Working Paper 28693, National Bureau of Economic Research.

Basu, S., Fernald, J. G., and Kimball, M. S. (2006). Are technology improvements contractionary? *American Economic Review*, 96(5):1418–1448.

Bénassy, J.-P. (1993). Nonclearing markets: microeconomic concepts and macroeconomic applications. *Journal of Economic Literature*, 31(2):732–761.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bianchi, J., McKay, A., and Mehrotra, N. (2024). How should monetary policy respond to housing inflation? Technical report, Federal Reserve Bank of Minneapolis.

Birnbaum, Z. W. (1942). An inequality for mill's ratio. *The Annals of Mathematical Statistics*, 13(2):245–246.

Blanchard, O. J. and Kiyotaki, N. (1987). Monopolistic competition and the effects of aggregate demand. *The american economic review*, pages 647–666.

Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, 77(3):623–685.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., and Terry, S. J. (2018). Really uncertain business cycles. *Econometrica*, 86(3):1031–1065.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. *Handbook of industrial organization*, 2:1011–1057.

Buda, G., Carvalho, V. M., Corsetti, G., Duarte, J. B., Hansen, S., Moura, A. S., Ortiz, Á., Rodrigo, T., Rodríguez Mora, J., and Alves da Silva, G. (2025). The short lags of monetary policy. Working paper, Janeway Institute.

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

Cavallo, A. and Kryvtsov, O. (2024). Price discounts and cheapflation during the post-pandemic inflation surge. *Journal of Monetary Economics*, 148:103644.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45.

Corrado, C. and Mattey, J. (1997). Capacity utilization. *Journal of Economic Perspectives*, 11(1):151–167.

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

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American economic review*, 102(6):2437–2471.

Demirer, M. (2025). Production function estimation with factor-augmenting technology: An application to markups. Working paper, MIT Sloan.

Di Tella, S., Malgieri, C., and Tonetti, C. (2025). Risk markups. Technical report, National Bureau of Economic Research.

Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *Journal of political Economy*, 90(5):881–894.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, 67(3):297–308.

Eden, B. (1994). The adjustment of prices to monetary shocks when trade is uncertain and sequential. *Journal of Political Economy*, 102(3):493–509.

Fay, J. A. and Medoff, J. L. (1985). Labor and output over the business cycle: Some direct evidence. *The American Economic Review*, 75(4):638–655.

Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. Working Paper 2012-19, Federal Reserve Bank of San Francisco.

Flynn, J. P., Nikolakoudis, G., and Sastry, K. A. (2024). Prices vs. quantities: A macroeconomic analysis.

Flynn, J. P., Nikolakoudis, G., and Sastry, K. A. (2026). A theory of supply function choice and aggregate supply. *American Economic Review*, 116(2):710–748.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425.

Gali, J. (1999). Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations? *American Economic Review*, 89(1):249–271.

Ghassibe, M. and Zanetti, F. (2022). State dependence of fiscal multipliers: the source of fluctuations matters. *Journal of Monetary Economics*, 132:1–23.

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

Harberger, A. (1964). Taxation, resource allocation, and welfare. In *The role of direct and indirect taxes in the Federal Reserve System*, pages 25–80. Princeton University Press.

Holden, T. D. (2025). Rationing under sticky prices. *Working Paper*.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Huo, Z. and Ríos-Rull, J.-V. (2020). Sticky wage models and labor supply constraints. *American Economic Journal: Macroeconomics*, 12(3):284–318.

Ireland, P. N. (2004). Technology shocks in the new keynesian model. *Review of Economics and Statistics*, 86(4):923–936.

Jaffé, W. (1967). Walras' theory of tatonnement: a critique of recent interpretations. *Journal of Political Economy*, 75(1):1–19.

Kimball, M. S. (1989). The effect of demand uncertainty on a precommitted monopoly price. *Economics Letters*, 30(1):1–5.

Lagos, R. and Wright, R. (2005). A unified framework for monetary theory and policy analysis. *Journal of Political Economy*, 113(3):463–484.

Lerner, A. P. (1934). The concept of monopoly and the measurement of monopoly power. *The review of economic studies*, 1(3):157–175.

Liu, Y. (2026). Inflation and competition in an old keynesian model. Working paper, SSRN.

Lorenzoni, G. and Werning, I. (2026). Tâtonnement and price setting in general equilibrium. *Working Paper*.

Lucas, R. E. and Woodford, M. (1993). Real effects of monetary shocks in an economy with sequential purchases. Technical Report 4250, National Bureau of Economic Research.

Markauskas, J., Ravn, M. O., and Rios-Rull, J.-V. (2026). Sticky prices and constraints to produce. Technical report, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Michaillat, P. and Saez, E. (2015). Aggregate demand, idle time, and unemployment. *The Quarterly Journal of Economics*, 130(2):507–569.

Mumtaz, H. and Ruch, F. U. (2025). Policy uncertainty and aggregate fluctuations: Evidence from emerging and developed economies. *The World Bank Economic Review*.

Petrongolo, B. and Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic literature*, 39(2):390–431.

Qiu, Z. and Ríos-Rull, J.-V. (2022). Procyclical productivity in new keynesian models. Technical report, National Bureau of Economic Research.

Reis, R. (2006). Inattentive producers. *The Review of Economic Studies*, 73(3):793–821.

Rocheteau, G. and Wright, R. (2005). Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium. *Econometrica*, 73(1):175–202.

Schmitt-Grohé, S. and Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of economic dynamics and control*, 28(4):755–775.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American economic review*, 97(3):586–606.

USDA Economic Research Service (2026). Estimates of food loss at the retail and consumer levels. Technical report, US Department of Agriculture Economic Research Service. Available at: https://www.ers.usda.gov/sites/default/files/images/foodloss_infographicaa.pdf.

Woodford, M. (2003a). Imperfect common knowledge and the effects of monetary policy. In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*. Princeton University Press, Princeton.

Woodford, M. (2003b). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Appendices

A Omitted Derivations and Proofs

A.1 Proof of Theorem 1

Proof. We begin by establishing that an optimum exists. Define the firm's payoff from an arbitrary pair of price and labor input (p, L) as:

$$J(p, L) = \mathbb{E}[\Lambda \min\{zp^{1-\eta}, ApL\} - \Lambda wL] \tag{57}$$

Now consider a sequence of policies $\{p_n, L_n\}_{n\in\mathbb{N}}$ such that either $p_n \to \infty$ or $L_n \to \infty$. We observe that $\min\{x, y\} \le x$ and so we may write:

$$J(p_n, L_n) \le \mathbb{E}[\Lambda z]p_n^{1-\eta} - \mathbb{E}[\Lambda w]L_n \tag{58}$$

From this inequality, if $p_n \to \infty$, then $\limsup_n J(p_n, L_n) \le 0$. Moreover, if $L_n \to \infty$ and $p_n \not\to 0$, then $\limsup_n J(p_n, L_n) \le 0$. If $L_n \to \infty$ and $p_n \to 0$, it is immediate that $\limsup_n J(p_n, L_n) \le 0$ (as it is $-\infty$). Writing J in integral form, we have that:

$$J(p, L) = \int_0^\infty \int_0^\infty \left[\int_0^{Ap^\eta L} \Lambda zp^{1-\eta} g(z, A, \Lambda)\,\mathrm{d}z + \int_{Ap^\eta L}^\infty \Lambda pALg(z, A, \Lambda)\,\mathrm{d}z\right] \mathrm{d}A\,\mathrm{d}\Lambda - \mathbb{E}[\Lambda w]L \tag{59}$$

It is immediate that this is a continuous function of (p, L) over $\mathbb{R}^2_+$ (while the integrand in the first integral is not well-behaved, observe that it is bounded above by $\Lambda ApL \sup_z g(z, A, \Lambda)$).

We now show that there is a $(\bar{p}, \bar{L})$ such that $J(\bar{p}, \bar{L}) > 0$. We observe that:

$$J(p, L) = \mathbb{E}[\Lambda \min\{zt, A\}p - \Lambda w]L \tag{60}$$

Thus, the firm's payoff is strictly positive at (p, L) if and only if:

$$p\mathbb{E}[\Lambda \min\{zt, A\}] > \mathbb{E}[\Lambda w] \tag{61}$$

To derive the existence of such a $(\bar{p}, \bar{L})$, consider a sequence $\{p_n, L_n\}_{n\in\mathbb{N}}$ such that $L_n p_n^\eta \to 0$ and $p_n > \frac{\mathbb{E}[\Lambda w]}{\mathbb{E}[\Lambda A]} + \varepsilon$ for all $n \in \mathbb{N}$ and some $\varepsilon > 0$. Under this construction, we have that $t_n \to \infty$ and so $\mathbb{E}[\Lambda \min\{zt_n, A\}] \to \mathbb{E}[\Lambda A]$. Thus, there exists an N such that for all $n > N$:

$$p_n > \frac{\mathbb{E}[\Lambda w]}{\mathbb{E}[\Lambda A]} + \varepsilon > \frac{\mathbb{E}[\Lambda w]}{\mathbb{E}[\Lambda \min\{zt_n, A\}]} \tag{62}$$

For any such n, we then have that $J(p_n, L_n) > 0$. Setting $\bar{p} = p_n$ and $\bar{L} = L_n$ for any such n suffices.

Pick $\varepsilon \in (0, J(\bar{p}, \bar{L}))$. We have shown that $\limsup_n J(p_n, L_n) \leq 0$ for all sequences $\{p_n, L_n\}_{n\in\mathbb{N}}$ such that $\max\{p_n, L_n\} \to \infty$. Thus, there exists an R such that $J(p, L) \leq \varepsilon$ whenever $\max\{p, L\} \geq R$. Hence, we have that:

$$\sup_{(p,L)\in\mathbb{R}^2_+} J(p, L) = \sup_{(p,L)\in[0,R]^2} J(p, L) \tag{63}$$

An optimum therefore exists as $[0, R]^2$ is compact and J is continuous, by Weierstrass' theorem.

We now show that $(0, 0)$ is not optimal. First, observe that if $p = 0$, then it must be the case that $L = 0$ (otherwise reducing L would yield a strict improvement). Second, if $L = 0$, then the firm's payoff is zero. We therefore have that $0 = J(0, 0) < J(\bar{p}, \bar{L})$.

Thus, an optimum exists (p^*, L^*) and all optima must lie in $\mathbb{R}^2_{++}$. As J is continuously differentiable, the first-order conditions $J_p(p^*, L^*) = 0$ and $J_L(p^*, L^*) = 0$ must hold at an optimum. By applying Leibniz's rule, we compute the derivative in prices:

$$\begin{aligned}
J_p(p, L) = \int_0^\infty \int_0^\infty \Bigg[\int_0^{Ap^\eta L} &\Lambda(1-\eta) z p^{-\eta} g(z, A, \Lambda)\,\mathrm{d}z + \Lambda \left(\eta A p^{\eta-1} L\right) \left(ApLg(Ap^\eta L, A, \Lambda)\right) \\
&+ \int_{Ap^\eta L}^\infty \Lambda A L g(z, A, \Lambda)\,\mathrm{d}z - \Lambda \left(\eta A p^{\eta-1} L\right) \left(pALg(Ap^\eta L, A, \Lambda)\right) \Bigg] \mathrm{d}A\,\mathrm{d}\Lambda \\
= \int_0^\infty \int_0^\infty &\left[\int_0^{Ap^\eta L} \Lambda(1-\eta) z p^{-\eta} g(z, A, \Lambda)\,\mathrm{d}z + \int_{Ap^\eta L}^\infty \Lambda A L g(z, A, \Lambda)\,\mathrm{d}z \right] \mathrm{d}A\,\mathrm{d}\Lambda \\
= (1-\eta)\mathbb{E}&\left[\Lambda z p^{-\eta} \mathbb{I}[z \leq Ap^\eta L]\right] + \mathbb{E}\left[\Lambda A L \mathbb{I}[z \geq Ap^\eta L]\right]
\end{aligned} \tag{64}$$

Thus, dividing all terms by L, we can re-write the first-order condition for prices as:

$$0 = (1-\eta)\mathbb{E}[\Lambda z \mathbb{I}[t \leq A/z]]t + \mathbb{E}[\Lambda A \mathbb{I}[t \geq A/z]] \tag{65}$$

Which implies that:

$$t = \frac{1}{\eta - 1} \frac{\mathbb{E}[\Lambda A \mathbb{I}[t \geq A/z]]}{\mathbb{E}[\Lambda z \mathbb{I}[t \leq A/z]]} = \frac{1}{\eta - 1} \frac{\mathbb{P}[A/z \leq t]\mathbb{E}[\Lambda A | A/z \leq t]}{(1 - \mathbb{P}[A/z \leq t])\,\mathbb{E}[\Lambda z | A/z \geq t]} \tag{66}$$

We now compute the derivative in labor:

$$
\begin{aligned}
J_L(p, L) = &\int_0^\infty \int_0^\infty \Bigg[\Lambda \left(Ap^\eta\right) \left(ApLg(Ap^\eta L, A, \Lambda)\right) + \int_{Ap^\eta L}^\infty \Lambda A p g(z, A, \Lambda) \,\mathrm{d}z \\
&- \Lambda \left(Ap^\eta\right) \left(ApLg(Ap^\eta L, A, \Lambda)\right) \Bigg] \mathrm{d}A \,\mathrm{d}\Lambda - \mathbb{E}[w\Lambda] \\
&= \mathbb{E}[\Lambda A \mathbb{I}[z \geq Ap^\eta L]]p - \mathbb{E}[\Lambda w]
\end{aligned}
\tag{67}
$$

Thus, the first-order condition for labor can be expressed as:

$$
p = \frac{\mathbb{E}[\Lambda w]}{\mathbb{E}[\Lambda A \mathbb{I}[t \geq A/z]]} \tag{68}
$$

We have now shown that any optimal (p, L) must solve Equations 66 and 68. Moreover, we observe that Equation 66 is a one-dimensional fixed-point equation. Moreover, for any solution t^* of Equation 66, we can find p^* by substitution into Equation 68 and labor is then given by $L^* = \frac{1}{t^* p^{*\eta}}$. Thus, the set of solutions to the pair of first-order conditions is pinned down by the set of solutions to Equation 66.

We finally show there is a unique such solution. To do so, we define $x = \ln t$, and we write the Equation 66 as $x = f(x)$, where $f : \mathbb{R} \to \mathbb{R}$ is given by:

$$
\begin{aligned}
f(x) = &\ln\left(\frac{1}{\eta - 1}\right) + \ln\left(\mathbb{P}[\ln A - \ln z \leq x] \mathbb{E}[\exp\{\ln \Lambda + \ln A\} | \ln A - \ln z \leq x]\right) \\
&- \ln\left(\left(1 - \mathbb{P}[\ln A - \ln z \leq x]\right) \mathbb{E}[\exp\{\ln \Lambda + \ln z\} | \ln A - \ln z \geq x]\right)
\end{aligned}
\tag{69}
$$

To analyze this equation, we first state and prove a lemma for the conditional expectations of linear combinations of Gaussian random variables when they are conditioned on lying in a half-space.

Lemma 1. *Consider an n-dimensional normal random variable $X \sim N(\mu, \Sigma)$ and for two vectors $a, b \in \mathbb{R}^n$, define $S = a'X$ and $T = b'X$. The following statements are true:*

$$
\begin{aligned}
\mathbb{P}[S \leq x]\mathbb{E}[\exp\{T\}|S \leq x] &= \exp\left\{\mu_T + \frac{1}{2}\sigma_T^2\right\} \Phi\left(\frac{x - \mu_S}{\sigma_S} - \frac{\sigma_{S,T}}{\sigma_S}\right) \\
\mathbb{P}[S \geq x]\mathbb{E}[\exp\{T\}|S \geq x] &= \exp\left\{\mu_T + \frac{1}{2}\sigma_T^2\right\} \left(1 - \Phi\left(\frac{x - \mu_S}{\sigma_S} - \frac{\sigma_{S,T}}{\sigma_S}\right)\right)
\end{aligned}
\tag{70}
$$

where μ_Z, σ_Z, $\sigma_{W,Z}$ denote the mean, standard deviation, and covariance for generic random variables Z and W.

Proof. We derive the first equality, the second follows by an entirely analogous argument. First, we apply the standard multivariate normal conditional expectation formula to derive:

$$T|S = s \sim N\left(\mu_T + \frac{\sigma_{S,T}}{\sigma_S^2}(s - \mu_S), \sigma_T^2 - \frac{\sigma_{S,T}^2}{\sigma_S^2}\right) \tag{71}$$

Applying the formula for the moment generating function of a Gaussian random variable, we have that:

$$\mathbb{E}[\exp\{T\}|S = s] = \exp\left\{\mu_T + \frac{1}{2}\sigma_T^2 - \frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\exp\left\{\frac{\sigma_{S,T}}{\sigma_S}\frac{s - \mu_S}{\sigma_S}\right\} \tag{72}$$

We then have that:

$$\mathbb{E}[\exp\{T\}|S \leq x] = \exp\left\{\mu_T + \frac{1}{2}\sigma_T^2 - \frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\mathbb{E}\left[\exp\left\{\frac{\sigma_{S,T}}{\sigma_S}\frac{S - \mu_S}{\sigma_S}\right\}|S \leq x\right] \tag{73}$$

We now compute:

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{\frac{\sigma_{S,T}}{\sigma_S}\frac{S - \mu_S}{\sigma_S}\right\}|S \leq x\right] \\
&\quad = \frac{1}{\mathbb{P}[S \leq x]}\int_{-\infty}^{x}\exp\left\{\frac{\sigma_{S,T}}{\sigma_S}\frac{z - \mu_S}{\sigma_S}\right\}\frac{1}{\sigma_S\sqrt{2\pi}}\exp\left\{-\frac{(z - \mu_S)^2}{2\sigma_S^2}\right\}\mathrm{d}z \\
&\quad = \frac{1}{\mathbb{P}[S \leq x]}\int_{-\infty}^{x}\frac{1}{\sigma_S\sqrt{2\pi}}\exp\left\{\frac{\sigma_{S,T}}{\sigma_S}\frac{z - \mu_S}{\sigma_S} - \frac{(z - \mu_S)^2}{2\sigma_S^2}\right\}\mathrm{d}z \\
&\quad = \frac{1}{\mathbb{P}[S \leq x]}\int_{-\infty}^{x}\frac{1}{\sigma_S\sqrt{2\pi}}\exp\left\{-\frac{(z - \mu_S - \sigma_{S,T})^2}{2\sigma_S^2} + \frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\mathrm{d}z \\
&= \frac{1}{\mathbb{P}[S \leq x]}\exp\left\{\frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\int_{-\infty}^{x}\frac{1}{\sigma_S\sqrt{2\pi}}\exp\left\{-\frac{(z - \mu_S - \sigma_{S,T})^2}{2\sigma_S^2}\right\}\mathrm{d}z \\
&= \frac{1}{\mathbb{P}[S \leq x]}\exp\left\{\frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\int_{-\infty}^{x}\frac{1}{\sigma_S\sqrt{2\pi}}\exp\left\{-\frac{(z - \mu_S - \sigma_{S,T})^2}{2\sigma_S^2}\right\}\mathrm{d}z \\
&= \frac{1}{\mathbb{P}[S \leq x]}\exp\left\{\frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\int_{-\infty}^{\frac{x-\mu_S}{\sigma_S} - \frac{\sigma_{S,T}}{\sigma_S}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}y^2\right\}\mathrm{d}y \\
&= \frac{1}{\mathbb{P}[S \leq x]}\exp\left\{\frac{1}{2}\frac{\sigma_{S,T}^2}{\sigma_S^2}\right\}\Phi\left(\frac{x - \mu_S}{\sigma_S} - \frac{\sigma_{S,T}}{\sigma_S}\right)
\end{aligned} \tag{74}$$

Combining these steps, we have that:

$$\mathbb{P}[S \leq x]\mathbb{E}[\exp\{T\}|S \leq x] = \exp\left\{\mu_T + \frac{1}{2}\sigma_T^2\right\}\Phi\left(\frac{x - \mu_S}{\sigma_S} - \frac{\sigma_{S,T}}{\sigma_S}\right) \tag{75}$$

As we claimed. $\square$

We now apply Lemma 1 to re-express f as:

$$\begin{aligned}
f(x) &= \ln\left(\frac{1}{\eta-1}\right) + \mu_\Lambda + \mu_A + \frac{1}{2}\left(\sigma_\Lambda^2 + \sigma_A^2 + 2\sigma_{\Lambda,A}\right) + \ln\Phi\left(\frac{x-\mu_U}{\sigma_U} - \beta^S\sigma_U\right) \\
&\quad - \left(\mu_\Lambda + \mu_z + \frac{1}{2}\left(\sigma_\Lambda^2 + \sigma_z^2 + 2\sigma_{\Lambda,z}\right)\right) - \ln\left(1 - \Phi\left(\frac{x-\mu_U}{\sigma_U} - \beta^D\sigma_U\right)\right) \\
&= \ln\left(\frac{1}{\eta-1}\right) + \mu_A - \mu_z + \frac{1}{2}\left(\sigma_A^2 - \sigma_z^2 + 2\sigma_{\Lambda,A} - 2\sigma_{\Lambda,z}\right) + \ln\frac{\Phi\left(\frac{x-\mu_U}{\sigma_U} - \beta^S\sigma_U\right)}{1-\Phi\left(\frac{x-\mu_U}{\sigma_U} - \beta^D\sigma_U\right)}
\end{aligned} \tag{76}$$

where $\mu_U = \mathbb{E}[\ln A - \ln z]$, $\sigma_U^2 = \mathbb{V}[\ln A - \ln z]$, $\beta^S = \mathbb{C}[\ln\Lambda + \ln A, \ln A - \ln z]/\sigma_U^2$, $\beta^D = \mathbb{C}[\ln\Lambda + \ln z, \ln A - \ln z]/\sigma_U^2$. We can further define the following transformations:

$$y = \frac{x-\mu_U}{\sigma_U}, \alpha = \beta^S\sigma_U, \beta = \beta^D\sigma_U, K = \frac{1}{\sigma_U}\left(\ln\left(\frac{1}{\eta-1}\right) + \frac{1}{2}\left(\sigma_A^2 - \sigma_z^2 + 2\sigma_{\Lambda,A} - 2\sigma_{\Lambda,z}\right)\right) \tag{77}$$

and then we can write the fixed-point equation as $y = g(y)$, where:

$$g(y) = K + \frac{1}{\sigma_U}\ln\frac{\Phi(y-\alpha)}{1-\Phi(y-\beta)} \tag{78}$$

We observe that $g(y)$ is strictly increasing and it has derivative:

$$g'(y) = \frac{1}{\sigma_U}\left(n(y-\alpha) + m(y-\beta)\right) \tag{79}$$

where $n(z) = \phi(z)/\Phi(z)$ and $m(z) = \phi(z)/(1-\Phi(z))$ are the left and right inverse Mills ratios, respectively. To find sufficient conditions for a unique fixed point, we use the following fact about Mills ratios: $n(z) > \max\{0, -z\}$ and $m(z) > \max\{0, z\}$. Thus, we have that:

$$\begin{aligned}
\inf_y g'(y) &= \inf_y \frac{1}{\sigma_U}\left(n(y-\alpha) + m(y-\beta)\right) > \inf_y \frac{1}{\sigma_U}\left(\max\{0, \alpha - y\} + \max\{0, y-\beta\}\right) \\
&\geq \frac{\max\{0, \alpha-\beta\}}{\sigma_U}
\end{aligned} \tag{80}$$

Thus, if $\alpha - \beta \geq \sigma_U$, then $\inf_y g'(y) > 1$ and there is a unique fixed point. This condition reduces to:

$$1 \leq \frac{\alpha-\beta}{\sigma_U} = \beta^S - \beta^D = 1 \tag{81}$$

Thus, there is always a unique fixed point.

We have thus shown that: (i) an optimum exists, (ii) any optimum must lie in $\mathbb{R}^2_{++}$, (iii) any optimum in $\mathbb{R}^2_{++}$ must satisfy Equations 66 and 68, and (iv) there is a unique solution to Equations 66 and 68. This completes the proof. $\square$

A.2 Proof of Proposition 1

Proof. We prove Claim 3, then Claim 4, and finally Claims 1 and 2.

Claim 3. Defining $x(\delta) = \ln t(\delta)$, we first prove that $\lim_{\delta\to 0} x(\delta) = \mu_A - \mu_z = x(0)$. Consider the fixed-point equation (Equation 5) and let the numerator and denominator on the right-hand side be:

$$\begin{aligned} N_\delta(t) &= \mathbb{E}[\Lambda_\delta A_\delta \mathbb{I}[t \geq A_\delta/z_\delta]] = \mathbb{P}[t \geq A_\delta/z_\delta]\mathbb{E}[\Lambda_\delta A_\delta | t \geq A_\delta/z_\delta] \\ D_\delta(t) &= \mathbb{E}[\Lambda_\delta z_\delta \mathbb{I}[t \leq A_\delta/z_\delta]] = (1 - \mathbb{P}[t \geq A_\delta/z_\delta])\mathbb{E}[\Lambda_\delta z_\delta | t \leq A_\delta/z_\delta] \end{aligned} \tag{82}$$

By Theorem 1, we have that $t(\delta) = \exp\{x(\delta)\}$ is unique for all $\delta > 0$ and moreover that $t(\delta)$ solves:

$$(\eta - 1)t(\delta)D_\delta(t(\delta)) = N_\delta(t(\delta)) \tag{83}$$

We also have that $t(0) = \exp\{\mu_A - \mu_z\}$. Because $\frac{A_\delta}{z_\delta} = \exp\{\ln A_\delta - \ln z_\delta\} \to^{a.s.} \exp\{\mu_A - \mu_z\}$ as $\delta \to 0$, we have that:

$$\lim_{\delta\to 0} \mathbb{P}[t \geq A_\delta/z_\delta] = \begin{cases} 1, t > t(0), \\ 0, t < t(0). \end{cases} \tag{84}$$

Thus, if $\lim_\delta t(\delta) < t(0)$, then $\lim_{\delta\to 0} N_\delta(t(\delta)) = 0$ and $\lim_{\delta\to 0} D_\delta(t(\delta)) = \exp\{\mu_\Lambda + \mu_z\} > 0$, so Equation 83 cannot hold, which is a contradiction. Conversely, if $\lim_{\delta\to 0} t(\delta) > t(0)$, then $\lim_{\delta\to 0} N_\delta(t(\delta)) = \exp\{\mu_\Lambda + \mu_A\} > 0$ and $\lim_{\delta\to 0} D_\delta(t(\delta)) = 0$, so Equation 83 cannot hold, which is a contradiction. Thus, we have proven that $\lim_{\delta\to 0} t(\delta) = t(0)$.

Claim 4. Rewriting the fixed-point equation in Equation 83, we have that:

$$(\eta - 1)t(\delta)(1 - \mathbb{P}[t(\delta) \geq A_\delta/z_\delta])\mathbb{E}[\Lambda_\delta z_\delta | t(\delta) \leq A_\delta/z_\delta] = \mathbb{P}[t(\delta) \geq A_\delta/z_\delta]\mathbb{E}[\Lambda_\delta A_\delta | t(\delta) \geq A_\delta/z_\delta] \tag{85}$$

which implies that:

$$\frac{\mathbb{P}[t(\delta) \geq A_\delta/z_\delta]}{1 - \mathbb{P}[t(\delta) \geq A_\delta/z_\delta]} = (\eta - 1)t(\delta)\frac{\mathbb{E}[\Lambda_\delta z_\delta | t(\delta) \leq A_\delta/z_\delta]}{\mathbb{E}[\Lambda_\delta A_\delta | t(\delta) \geq A_\delta/z_\delta]} \tag{86}$$

Taking the $\delta \to 0$ limit of both sides and using Claim 3, we have that:

$$\lim_{\delta\to 0} \frac{\mathbb{P}[t(\delta) \geq A_\delta/z_\delta]}{1 - \mathbb{P}[t(\delta) \geq A_\delta/z_\delta]} = (\eta - 1)\exp\{\mu_A - \mu_z\}\exp\{\mu_z - \mu_A\} = \eta - 1 \tag{87}$$

Thus, we have that:

$$\lim_{\delta\to 0}\mathbb{P}[t(\delta)\geq A_\delta/z_\delta]=\frac{\eta-1}{\eta} \tag{88}$$

as claimed.

Claims 1 and 2. By Theorem 1, we have that for all $\delta>0$, the optimal price is uniquely given by:

$$p(\delta)=\frac{\mathbb{E}[\Lambda_\delta w_\delta]}{\mathbb{P}[t(\delta)\geq A_\delta/z_\delta]\mathbb{E}[\Lambda_\delta A_\delta|t(\delta)\geq A_\delta/z_\delta]} \tag{89}$$

Taking the $\delta\to 0$ limit of both sides and applying Claims 3 and 4, we have that:

$$\lim_{\delta\to 0}p(\delta)=\frac{\eta}{\eta-1}\exp\{\mu_w-\mu_A\} \tag{90}$$

Finally, we use the definition of t alongside Claims 1 and 3 to show

$$\lim_{\delta\to 0}L(\delta)=\lim_{\delta\to 0}\frac{p(\delta)^{-\eta}}{t(\delta)}=\left(\frac{\eta}{\eta-1}\right)^{-\eta}\exp\{\mu_z+(\eta-1)\mu_A-\eta\mu_w\} \tag{91}$$

as claimed. □

A.3 Proof of Theorem 2

Proof. We break the proof of this result into three lemmas. First, we characterize the first-order effects of volatility on the logarithm of the optimal supply-demand imbalance.

Lemma 2. *We have that:*

$$x(\delta)=\mu_A-\mu_z+\kappa\sigma_U\delta+O(\delta^2) \tag{92}$$

Proof. We showed in the proof of Theorem 1 that $f'_\delta(x)>1$ for all $x\in\mathbb{R}$ and $\delta>0$. Thus, by application of the implicit function theorem, we have that $x(\delta)$ is continuously differentiable at all points $\delta>0$. We define its derivative at $\delta=0$ by continuity: $x'(0)=\lim_{\delta\to 0}x'(\delta)$. We now make use of the change of variable:

$$y(\delta)=\frac{x(\delta)-\mu_U}{\delta\sigma_U} \tag{93}$$

And we observe that:

$$\lim_{\delta\to 0}y(\delta)=\lim_{\delta\to 0}\frac{x(\delta)-\mu_U}{\delta\sigma_U}=\frac{1}{\sigma_U}\lim_{\delta\to 0}\frac{x(\delta)-x(0)}{\delta}=\frac{1}{\sigma_U}x'(0) \tag{94}$$

We moreover have that:

$$\delta\sigma_U y(\delta) = \ln\left(\frac{1}{\eta-1}\right) + \frac{1}{2}\delta^2(\sigma_A^2 - \sigma_z^2 + 2\sigma_{\Lambda,A} - 2\sigma_{\Lambda,z}) + \ln\left(\frac{\Phi(y(\delta)-\delta\beta^S\sigma_U)}{1-\Phi(y(\delta)-\delta\beta^D\sigma_U)}\right) \quad (95)$$

Thus, taking the $\delta \to 0$ limit of both sides, we have that:

$$0 = \ln\left(\frac{1}{\eta-1}\right) + \ln\left(\frac{\Phi(\frac{1}{\sigma_U}x'(0))}{1-\Phi(\frac{1}{\sigma_U}x'(0))}\right) \quad (96)$$

And so we obtain that:

$$x'(0) = \sigma_U \Phi^{-1}\left(\frac{\eta-1}{\eta}\right) \quad (97)$$

Application of Taylor's remainder theorem then yields the claimed formula. □

Second, we use this Lemma to derive the effects of volatility on prices and inputs:

Lemma 3. *We have that:*

$$(\ln p)(\delta) = \ln\left(\frac{\eta}{\eta-1}\right) + \mu_w - \mu_A + \left(\phi(\kappa) - \frac{\kappa}{\eta}\right)\sigma_U\delta + O(\delta^2) \quad (98)$$

And:

$$(\ln L)(\delta) = \mu_z + (\eta-1)\mu_A - \eta\mu_w - \eta\ln\left(\frac{\eta}{\eta-1}\right) - \eta\phi(\kappa)\sigma_U\delta + O(\delta^2) \quad (99)$$

Proof. By Theorem 1, we have that:

$$(\ln p)(\delta) = \mu_w - \mu_A + \frac{1}{2}\delta^2\left(\sigma_w^2 - \sigma_A^2 + 2\sigma_{\Lambda,w} - 2\sigma_{\Lambda,A}\right) - \ln\Phi\left(y(\delta) - \delta\beta^S\sigma_U\right) \quad (100)$$

Using the fact from Lemma 2 that $y(0) = \Phi^{-1}\left(\frac{\eta-1}{\eta}\right)$, we have that:

$$(\ln p)(0) = \mu_w - \mu_A - \ln\Phi(y(0)) = \ln\left(\frac{\eta}{\eta-1}\right) + \mu_w - \mu_A \quad (101)$$

Moreover, as x is continuously differentiable for all $\delta > 0$, we also have that y is continuously differentiable for all $\delta > 0$ and, therefore, that p is continuously differentiable for all $\delta > 0$. As before, we define derivatives at zero by continuity. Differentiating both sides of Equation 95 and evaluating at $\delta = 0$, we obtain that:

$$\sigma_U y(0) = \left(y'(0) - \beta^S\sigma_U\right) n(y(0)) + \left(y'(0) - \beta^D\sigma_U\right) m(y(0)) \quad (102)$$

where we remind that $n(z) = \phi(z)/\Phi(z)$ and $m(z) = \phi(z)/(1-\Phi(z))$ are the left and right inverse Mills ratios, respectively. Hence,

$$y'(0) = \frac{\sigma_U y(0) + \beta^S \sigma_U n(y(0)) + \beta^D \sigma_U m(y(0))}{n(y(0)) + m(y(0))} \tag{103}$$

Differentiating both sides of Equation 100 and evaluating at $\delta = 0$, we obtain:

$$(\ln p)'(0) = -\left(y'(0) - \beta^S \sigma_U\right) n(y(0)) \tag{104}$$

Combining the previous two equations, we have that:

$$(\ln p)'(0) = -\left(\frac{\sigma_U y(0) + \beta^S \sigma_U n(y(0)) + \beta^D \sigma_U m(y(0))}{n(y(0)) + m(y(0))} - \beta^S \sigma_U\right) n(y(0)) \tag{105}$$

We now make use of the following facts:

$$y(0) = \Phi^{-1}\left(\frac{\eta-1}{\eta}\right),\ n(y(0)) = \frac{\phi(y(0))}{\frac{\eta-1}{\eta}},\ m(y(0)) = \frac{\phi(y(0))}{1-\frac{\eta-1}{\eta}}, \tag{106}$$

After some algebra, one obtains that:

$$\begin{aligned}(\ln p)'(0) &= (\beta^S - \beta^D)\sigma_U \phi\left(\Phi^{-1}\left(\frac{\eta-1}{\eta}\right)\right) - \frac{\sigma_U}{\eta}\Phi^{-1}\left(\frac{\eta-1}{\eta}\right) \\ &= \sigma_U \phi\left(\Phi^{-1}\left(\frac{\eta-1}{\eta}\right)\right) - \frac{\sigma_U}{\eta}\Phi^{-1}\left(\frac{\eta-1}{\eta}\right)\end{aligned} \tag{107}$$

where the final equality exploits the fact that $\beta^S - \beta^D = 1$. Application of Taylor's remainder theorem then yields the claimed formula for $\ln p$. Using the fact that $\ln L = -\eta \ln p - x$ then yields the formula for labor. □

We finally show that L decreases and p increases in uncertainty. The first claim is immediate as $\phi(\kappa) > 0$. The second claim holds if and only if $\phi(\kappa) - \kappa/\eta > 0$. Since $\kappa = \Phi^{-1}((\eta-1)/\eta)$, this inequality can be re-written as:

$$\phi(\kappa) - \kappa(1 - \Phi(\kappa)) > 0 \tag{108}$$

which holds if and only if $m(\kappa) > \kappa$, where $m(\cdot)$ is the right inverse Mills ratio. This is a global property of the Mills ratio, establishing the inequality. □

A.4 Proof of Proposition 2

Proof. To characterize the solution to Equation 20, we attach a Lagrange multiplier λ to the expenditure constraint and μ_i to each $c_i \leq \bar{c}_i$ constraint. The first-order condition for each c_i is then:

$$\theta_i^{\frac{1}{\eta}} c_i^{-\frac{1}{\eta}} C^{\frac{1}{\eta}} - \lambda p_i - \mu_i = 0 \tag{109}$$

Thus, we have that:

$$c_i = \theta_i C \left(\lambda p_i + \mu_i\right)^{-\eta} \tag{110}$$

Taking the ratio of this condition for two varieties i and j yields:

$$c_i = \frac{\theta_i}{\theta_j} \frac{\left(\lambda p_i + \mu_i\right)^{-\eta}}{\left(\lambda p_j + \mu_j\right)^{-\eta}} c_j \tag{111}$$

Substituting this into the expenditure constraint, we have that:

$$E = \frac{c_j}{\theta_j \left(\lambda p_j + \mu_j\right)^{-\eta}} \int p_i \theta_i \left(\lambda p_i + \mu_i\right)^{-\eta} \mathrm{d}i \tag{112}$$

And so we have that:

$$c_j = E \theta_j \left(\lambda p_j + \mu_j\right)^{-\eta} \left(\int p_i \theta_i \left(\lambda p_i + \mu_i\right)^{-\eta} \mathrm{d}i\right)^{-1} \tag{113}$$

Moreover, $c_j \leq \bar{c}_j$. Thus, exploiting complementary slackness ($c_j < \bar{c}_j \implies \mu_j = 0$), we have:

$$c_j = \bar{c}_j \min\left\{1, \frac{\theta_j}{A_j} \frac{p_j^{-\eta}}{L_j} D\right\} \tag{114}$$

where $D = E\lambda^{-\eta} \left(\int p_i \theta_i \left(\lambda p_i + \mu_i\right)^{-\eta} \mathrm{d}i\right)^{-1}$, which is independent of j. Thus, re-ordering markets according to ι, we have that there exists a unique $k \in [0, 1]$ such that all markets with $j < k$ have $c_j < \bar{c}_j$ and all markets with $j > k$ have $c_j = \bar{c}_j$. Thus, we can write the expenditure constraint as:

$$E = D \int_0^k \theta_i p_i^{1-\eta} \mathrm{d}i + \int_k^1 p_i \bar{c}_i \mathrm{d}i \tag{115}$$

which implies that:

$$D = \frac{E - \int_k^1 p_i \bar{c}_i \mathrm{d}i}{\int_0^k \theta_i p_i^{1-\eta} \mathrm{d}i} \tag{116}$$

yielding Equation 22. To complete the proof, we plug these demands directly into the consumption aggregator, yielding Equation 23. □

A.5 Proof of Theorem 3

Proof. We begin by solving for the mean value of the demand shock faced by each firm i in equilibrium. From Proposition 2, we have that the demand shock z_i is:

$$z_i = \frac{M - \int_k^1 p_i \bar{c}_i \, \mathrm{d}i}{\int_0^k \theta_i p_i^{1-\eta} \, \mathrm{d}i} \theta_i \tag{117}$$

where here we have used the fact that $E = M$ in equilibrium. Thus, we have that:

$$\mathbb{E}[z_i] = \frac{M - \int_k^1 p_i \bar{c}_i \, \mathrm{d}i}{\int_0^k \theta_i p_i^{1-\eta} \, \mathrm{d}i} \exp\left\{\mu_\theta + \frac{1}{2}\sigma_\theta^2\right\} \tag{118}$$

where $\mu_z = \ln \mathbb{E}[z_i] - \frac{1}{2}\sigma_\theta^2$. To simplify further, we observe that (i) firms are *ex ante* symmetric and (ii) have a uniquely optimal solution for prices and labor. Thus in any equilibrium, all firms choose the same price p and the same labor input L, and they ration with the same *ex ante* probability r. With this, we state the following formula for mean demand.

Lemma 4. *We have that:*

$$\exp\{\mu_z\} = \frac{M}{p} \frac{1}{p^{-\eta} \exp\left\{\frac{1}{2}\sigma_\theta^2\right\} \left(1 - \Phi\left(\Phi^{-1}(r) - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\right) + \bar{L} \exp\{\mu_A + \frac{1}{2}\sigma_A^2\} \Phi\left(\Phi^{-1}(r) - \frac{\sigma_{A,U}}{\sigma_U}\right)} \tag{119}$$

where $\bar{L} = L \exp\{-\mu_z\}$ is invariant to μ_z.

Proof. We first solve for the integrals in the numerator and denominator of Equation 118:

$$\begin{aligned} \int_k^1 p_i \bar{c}_i \, \mathrm{d}i &= pL\mathbb{E}\left[\exp\{\ln A\} \mathbb{I}\left[\ln A - \ln \theta \leq \ln \frac{A_k}{\theta_k}\right]\right] \\ &= pL \exp\left\{\mu_A + \frac{1}{2}\sigma_A^2\right\} \Phi\left(\frac{\ln \frac{A_k}{\theta_k} - \mathbb{E}[\ln A - \ln \theta]}{\sigma_U} - \frac{\sigma_{A,U}}{\sigma_U}\right) \end{aligned} \tag{120}$$

$$\begin{aligned} \int_0^k \theta_i p_i^{1-\eta} \, \mathrm{d}i &= p^{1-\eta}\mathbb{E}\left[\exp\{\ln \theta\} \mathbb{I}\left[\ln A_i - \ln \theta_i \geq \ln \frac{A_k}{\theta_k}\right]\right] \\ &= p^{1-\eta} \exp\left\{\mu_\theta + \frac{1}{2}\sigma_\theta^2\right\} \left(1 - \Phi\left(\frac{\ln \frac{A_k}{\theta_k} - \mathbb{E}[\ln A - \ln \theta]}{\sigma_U} - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\right) \end{aligned} \tag{121}$$

Combining these expressions with Equation 118, we obtain that:

$$\begin{aligned}
&\mathbb{E}[z_i]p^{1-\eta}\left(1-\Phi\left(\frac{\ln\frac{A_k}{\theta_k}-\mathbb{E}[\ln A-\ln\theta]}{\sigma_U}-\frac{\sigma_{\theta,U}}{\sigma_U}\right)\right)\\
&\qquad = M - pL\exp\left\{\mu_A+\frac{1}{2}\sigma_A^2\right\}\Phi\left(\frac{\ln\frac{A_k}{\theta_k}-\mathbb{E}[\ln A-\ln\theta]}{\sigma_U}-\frac{\sigma_{A,U}}{\sigma_U}\right)
\end{aligned} \tag{122}$$

We now simplify to remove the term $\ln\frac{A_k}{\theta_k}$. To do this, we observe that k and $\ln\frac{A_k}{\theta_k}$ satisfy the following relationship:

$$\begin{aligned}
k &= \mathbb{P}\left[\ln\theta_i-\ln A_i\le\ln\frac{\theta_k}{A_k}\right] = 1-\mathbb{P}\left[\ln A_i-\ln\theta_i\le\ln\frac{A_k}{\theta_k}\right]\\
&= 1-\mathbb{P}\left[\frac{\ln A_i-\ln\theta_i-\mathbb{E}[\ln A_i-\ln\theta_i]}{\sigma_U}\le\frac{\ln\frac{A_k}{\theta_k}-\mathbb{E}[\ln A_i-\ln\theta_i]}{\sigma_U}\right]\\
&= 1-\Phi\left(\frac{\ln\frac{A_k}{\theta_k}-\mathbb{E}[\ln A_i-\ln\theta_i]}{\sigma_U}\right)
\end{aligned} \tag{123}$$

We can now also note that the threshold index k is related to the *ex ante* rationing probability, since all firms are *ex ante* identical: $k=1-r$. From this, we can solve that $\ln\frac{A_k}{\theta_k}$ in terms of the rationing probability:

$$r=\Phi\left(\frac{\ln\frac{A_k}{\theta_k}-\mathbb{E}[\ln A_i-\ln\theta_i]}{\sigma_U}\right)\implies\ln\frac{A_k}{\theta_k}=\sigma_U\Phi^{-1}(r)+\mathbb{E}[\ln A_i-\ln\theta_i] \tag{124}$$

Combining this with Equation 122 yields the relationship:

$$\mathbb{E}[z_i]p^{1-\eta}\left(1-\Phi\left(\Phi^{-1}(r)-\frac{\sigma_{\theta,U}}{\sigma_U}\right)\right)=M-pL\exp\{\mu_A+\frac{1}{2}\sigma_A^2\}\Phi\left(\Phi^{-1}(r)-\frac{\sigma_{A,U}}{\sigma_U}\right) \tag{125}$$

We now return to the firm's problem to establish the general relationship between p, L, and r with μ_z, noting that all other terms in this expression are parameters. From the expressions in Theorem 1, it is immediate to verify that $\exp\{\mu_z\}t$ is invariant to μ_z and p is invariant to μ_z. These facts imply that L is log-linear in μ_z: $L=\bar{L}\exp\{\mu_z\}$ and r is invariant to μ_z. Exploiting these facts and rearranging, we obtain the claimed expression. □

We now compute the first-order expansion of μ_z in volatility.

Lemma 5. *We have that:*

$$\mu_z=(\eta-1)\ln\frac{\eta}{\eta-1}+(\eta-1)(\mu_w-\mu_A)+\ln M+\eta\phi(\kappa)\sigma_U\delta+O(\delta^2) \tag{126}$$

Proof. We first compute μ_z in the $\delta \to 0$ limit. In this limit, by Equation 119, we have that:

$$
\begin{aligned}
&\exp\{\mu_z(0)\} = \lim_{\delta\to 0} \exp\{\mu_z(\delta)\} = \frac{M}{p(0)} \frac{1}{p(0)^{-\eta}\frac{1}{\eta} + \bar{L}(0)\exp\{\mu_A\}\frac{\eta-1}{\eta}} \\
&= \frac{M}{\frac{\eta}{\eta-1}\exp\{\mu_w - \mu_A\}} \\
&\times \frac{1}{\left(\frac{\eta}{\eta-1}\right)^{-\eta}\exp\{-\eta\mu_w + \eta\mu_A\}\frac{1}{\eta} + \exp\{(\eta-1)\mu_A - \eta\mu_w\}\left(\frac{\eta}{\eta-1}\right)^{-\eta}\exp\{\mu_A\}\frac{\eta-1}{\eta}} \\
&= \frac{M}{\left(\frac{1}{\eta-1}+1\right)\left(\frac{\eta}{\eta-1}\right)^{-\eta}\exp\{(\eta-1)(\mu_A-\mu_w)\}} \\
&= \left(\frac{\eta}{\eta-1}\right)^{\eta-1} \frac{M}{\exp\{(\eta-1)(\mu_A - \ln\alpha - \ln M)\}}
\end{aligned}
\tag{127}
$$

where the final equality uses the fact that $w = \alpha M$.

We now proceed to find the slope of μ_z in δ at $\delta = 0$. We have that:

$$
\begin{aligned}
\mu_z = \ln M - \ln p - \ln \bar{L} - \ln\Bigg(&\exp\{\mu_z\}t\exp\left\{\frac{1}{2}\sigma_\theta^2\right\}\left(1-\Phi\left(\Phi^{-1}(r) - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\right) \\
&+ \exp\{\mu_A + \frac{1}{2}\sigma_A^2\}\Phi\left(\Phi^{-1}(r) - \frac{\sigma_{A,U}}{\sigma_U}\right)\Bigg)
\end{aligned}
\tag{128}
$$

We now define:

$$
\begin{aligned}
A(\delta) &= (\exp\{\mu_z\}t)(\delta)\exp\left\{\frac{1}{2}\delta^2\sigma_\theta^2\right\}\left(1-\Phi\left(\Phi^{-1}(r(\delta)) - \delta\frac{\sigma_{\theta,U}}{\sigma_U}\right)\right) \\
B(\delta) &= \exp\{\mu_A + \frac{1}{2}\delta^2\sigma_A^2\}\Phi\left(\Phi^{-1}(r(\delta)) - \delta\frac{\sigma_{A,U}}{\sigma_U}\right)
\end{aligned}
\tag{129}
$$

And so we have that:

$$
\mu_z'(0) = -(\ln p)'(0) - (\ln \bar{L})'(0) - \frac{A'(0) + B'(0)}{A(0) + B(0)}
\tag{130}
$$

From Theorem 2, we have that $(\ln p)'(0) = \left(\phi(\kappa) - \frac{\kappa}{\eta}\right)\sigma_U$ and $(\ln \bar{L})'(0) = -\eta\phi(\kappa)\sigma_U$. We also have that:

$$
A(0) = \exp\{\mu_A\}\frac{1}{\eta}, \quad B(0) = \exp\{\mu_A\}\frac{\eta-1}{\eta}
\tag{131}
$$

and so $A(0) + B(0) = \exp\{\mu_A\}$. Moreover, we calculate:

$$
\begin{aligned}
A'(0) &= (\exp\{\mu_z\}t)(\delta)'(0)\frac{1}{\eta} - \exp\{\mu_A\}\left(r'(0)\Phi^{-1\prime}(r(0)) - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\phi(\Phi^{-1}(r(0))) \\
&= \exp\{\mu_A\}\frac{\kappa}{\eta}\sigma_U - \exp\{\mu_A\}\left(r'(0)\frac{1}{\phi(\kappa)} - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\phi(\kappa) \\
B'(0) &= \exp\{\mu_A\}\left(r'(0)\frac{1}{\phi(\kappa)} - \frac{\sigma_{A,U}}{\sigma_U}\right)\phi(\kappa)
\end{aligned} \tag{132}
$$

Combining this information we have that:

$$
\begin{aligned}
\frac{A'(0) + B'(0)}{A(0) + B(0)} &= \frac{\kappa}{\eta}\sigma_U - \left(r'(0)\frac{1}{\phi(\kappa)} - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\phi(\kappa) + \left(r'(0)\frac{1}{\phi(\kappa)} - \frac{\sigma_{A,U}}{\sigma_U}\right)\phi(\kappa) \\
&= \frac{\kappa}{\eta}\sigma_U + \frac{\sigma_{\theta,U} - \sigma_{A,U}}{\sigma_U}\phi(\kappa)
\end{aligned} \tag{133}
$$

And so we obtain:

$$
\begin{aligned}
\mu_z'(0) &= -\left(\phi(\kappa) - \frac{\kappa}{\eta}\right)\sigma_U + \eta\phi(\kappa)\sigma_U - \frac{\kappa}{\eta}\sigma_U + \frac{\sigma_{A,U} - \sigma_{\theta,U}}{\sigma_U}\phi(\kappa) \\
&= \left((\eta - 1) + \frac{\sigma_{A,U} - \sigma_{\theta,U}}{\sigma_U^2}\right)\phi(\kappa)\sigma_U \\
&= \eta\phi(\kappa)\sigma_U
\end{aligned} \tag{134}
$$

As claimed. □

To obtain equilibrium labor inputs, we combine the formula for μ_z with the formula for $\ln L$ from Theorem 2 (see, in particular, Lemma 3):

$$
\begin{aligned}
\ln L &= \mu_z + (\eta - 1)\mu_A - \eta\mu_w - \eta\ln\left(\frac{\eta}{\eta - 1}\right) - \eta\phi(\kappa)\sigma_U \\
&= (\eta - 1)\ln\frac{\eta}{\eta - 1} + (\eta - 1)(\mu_w - \mu_A) + \ln M + \eta\phi(\kappa)\sigma_U \\
&\quad + (\eta - 1)\mu_A - \eta\mu_w - \eta\ln\left(\frac{\eta}{\eta - 1}\right) - \eta\phi(\kappa)\sigma_U \\
&= -\ln\frac{\eta}{\eta - 1} - \mu_w + \ln M = -\ln\frac{\eta}{\eta - 1} - \ln\alpha
\end{aligned} \tag{135}
$$

To solve for aggregate consumption, we use Equation 23 from Proposition 2 along with the fact that $E = M$:

$$
C = \left(\left(M - \int_k^1 p_i\bar{c}_i \,\mathrm{d}i\right)^{\frac{\eta-1}{\eta}}\left(\int_0^k \theta_i p_i^{1-\eta}\right)^{\frac{1}{\eta}} + \int_k^1 \theta_i^{\frac{1}{\eta}}\bar{c}_i^{\frac{\eta-1}{\eta}}\,\mathrm{d}i\right)^{\frac{\eta}{\eta-1}} \tag{136}
$$

We moreover re-arrange Equation 118 to obtain that:

$$M - \int_k^1 p_i \bar{c}_i \,\mathrm{d}i = \exp\{\mu_z - \mu_\theta\} \left(\int_0^k \theta_i p_i^{1-\eta} \,\mathrm{d}i \right) \tag{137}$$

Thus, we have that:

$$\ln C = \frac{\eta}{\eta - 1} \ln \left(\exp\left\{ \frac{\eta - 1}{\eta} (\mu_z - \mu_\theta) \right\} \left(\int_0^k \theta_i p_i^{1-\eta} \,\mathrm{d}i \right) + \int_k^1 \theta_i^{\frac{1}{\eta}} \bar{c}_i^{\frac{\eta-1}{\eta}} \,\mathrm{d}i \right) \tag{138}$$

We already computed the first integral on the right-hand side in the proof of Lemma 4:

$$\int_0^k \theta_i p_i^{1-\eta} \,\mathrm{d}i = \exp\left\{ \mu_\theta + \frac{1}{2}\sigma_\theta^2 \right\} p^{1-\eta} \left(1 - \Phi\left(\Phi^{-1}(r) - \frac{\sigma_{\theta,U}}{\sigma_U} \right) \right) \tag{139}$$

We can further compute the second integral by defining the random variable $\zeta_i = \ln \left(\theta_i^{\frac{1}{\eta}} A_i^{\frac{\eta-1}{\eta}} \right)$ and following similar steps to those in Lemma 4:

$$\begin{aligned}
\int_k^1 \theta_i^{\frac{1}{\eta}} \bar{c}_i^{\frac{\eta-1}{\eta}} \,\mathrm{d}i &= L^{\frac{\eta-1}{\eta}} \int_k^1 \theta_i^{\frac{1}{\eta}} A_i^{\frac{\eta-1}{\eta}} \,\mathrm{d}i \\
&= L^{\frac{\eta-1}{\eta}} \mathbb{E}\left[\exp\{\zeta\} \mathbb{I}\left[\ln A - \ln\theta \le \ln A_k - \ln\theta_k \right] \right] \\
&= L^{\frac{\eta-1}{\eta}} \exp\{\mu_\zeta + \frac{1}{2}\sigma_\zeta^2\} \Phi\left(\Phi^{-1}(r) - \frac{\sigma_{\zeta,U}}{\sigma_U} \right)
\end{aligned} \tag{140}$$

Putting all of this together, we have that:

$$\begin{aligned}
\ln C = \frac{\eta}{\eta - 1} \ln \Bigg(& \exp\left\{ \frac{\eta - 1}{\eta}\mu_z + \frac{1}{\eta}\mu_\theta + \frac{1}{2}\sigma_\theta^2 \right\} p^{1-\eta} \left(1 - \Phi\left(\Phi^{-1}(r) - \frac{\sigma_{\theta,U}}{\sigma_U} \right) \right) \\
& + L^{\frac{\eta-1}{\eta}} \exp\{\mu_\zeta + \frac{1}{2}\sigma_\zeta^2\} \Phi\left(\Phi^{-1}(r) - \frac{\sigma_{\zeta,U}}{\sigma_U} \right) \Bigg)
\end{aligned} \tag{141}$$

and we note that we have already solved for all quantities on the right-hand side. We now parameterize volatility by δ and compute consumption to first order in δ. That is, we wish to compute:

$$(\ln C)(\delta) = (\ln C)(0) + (\ln C)'(0)\delta + O(\delta^2) \tag{142}$$

We can first compute:

$$
\begin{aligned}
(\ln C)(0) &= \frac{\eta}{\eta-1}\ln\left(\exp\left\{\frac{\eta-1}{\eta}\mu_z(0)+\frac{1}{\eta}\mu_\theta\right\}p(0)^{1-\eta}\left(1-r(0)\right)+L(0)^{\frac{\eta-1}{\eta}}\exp\{\mu_\zeta\}r(0)\right)\\
&= \frac{\eta}{\eta-1}\ln\left(\frac{1}{\eta}\exp\left\{\frac{\eta-1}{\eta}\mu_z(0)+\frac{1}{\eta}\mu_\theta+(1-\eta)\ln p(0)\right\}+\frac{\eta-1}{\eta}\exp\left\{\mu_\zeta+\frac{\eta-1}{\eta}\ln L(0)\right\}\right)\\
&= \frac{\eta}{\eta-1}\ln\left(\frac{1}{\eta}\exp\left\{\frac{\eta-1}{\eta}\left((\eta-1)\ln\frac{\eta}{\eta-1}+(\eta-1)(\ln\alpha-\mu_A)+\eta\ln M\right)+\frac{1}{\eta}\mu_\theta\right.\right.\\
&\quad\left.+(1-\eta)\left(\ln\left(\frac{\eta}{\eta-1}\right)+\ln\alpha+\ln M-\mu_A\right)\right\}\\
&\quad\left.+\frac{\eta-1}{\eta}\exp\left\{\frac{1}{\eta}\mu_\theta+\frac{\eta-1}{\eta}\mu_A-\frac{\eta-1}{\eta}\left(\ln\left(\frac{\eta}{\eta-1}\right)+\ln\alpha\right)\right\}\right)\\
&= \mu_A+\frac{\mu_\theta}{\eta-1}-\ln\alpha-\ln\left(\frac{\eta}{\eta-1}\right)
\end{aligned}
\tag{143}
$$

We now compute the derivative, first noting that:

$$
(\ln C)'(0) = \frac{\eta}{\eta-1}C(0)^{\frac{1}{\eta}-1}\left(A+B\right) \tag{144}
$$

where:

$$
\begin{aligned}
A &= \left(\frac{\eta-1}{\eta}\mu_z'(0)+(1-\eta)(\ln p)'(0)\right)\frac{1}{\eta}\exp\left\{\frac{\eta-1}{\eta}\mu_z(0)+\frac{1}{\eta}\mu_\theta+(1-\eta)\ln p(0)\right\}\\
&\quad-\phi(\Phi^{-1}(r(0)))\left(\frac{r'(0)}{\phi(\Phi^{-1}(r(0)))}-\frac{\sigma_{\theta,U}}{\sigma_U}\right)\exp\left\{\frac{\eta-1}{\eta}\mu_z(0)+\frac{1}{\eta}\mu_\theta+(1-\eta)\ln p(0)\right\}\\
&= C(0)^{\frac{\eta-1}{\eta}}\left(\frac{1}{\eta}\frac{\eta-1}{\eta}\kappa\sigma_U-\phi(\kappa)\left(\frac{r'(0)}{\phi(\kappa)}-\frac{\sigma_{\theta,U}}{\sigma_U}\right)\right)
\end{aligned}
\tag{145}
$$

and:

$$
\begin{aligned}
B &= \frac{\eta-1}{\eta}(\ln L)'(0)\exp\left\{\mu_\zeta+\frac{1}{2}\sigma_\zeta^2+\frac{\eta-1}{\eta}(\ln L)(0)\right\}\frac{\eta-1}{\eta}\\
&\quad\phi(\Phi^{-1}(r(0)))\left(\frac{r'(0)}{\phi(\Phi^{-1}(r(0)))}-\frac{\sigma_{\zeta,U}}{\sigma_U}\right)\exp\left\{\mu_\zeta+\frac{1}{2}\sigma_\zeta^2+\frac{\eta-1}{\eta}(\ln L)(0)\right\}\\
&= C(0)^{\frac{\eta-1}{\eta}}\phi(\kappa)\left(\frac{r'(0)}{\phi(\kappa)}-\frac{1}{\eta}\frac{\sigma_{\theta,U}}{\sigma_U}-\frac{\eta-1}{\eta}\frac{\sigma_{A,U}}{\sigma_U}\right)
\end{aligned}
\tag{146}
$$

Combining all of this, we have that:

$$
\begin{aligned}
(\ln C)'(0) &= \frac{\eta}{\eta-1}\left(\frac{1}{\eta}\frac{\eta-1}{\eta}\kappa\sigma_U - \frac{\eta-1}{\eta}\phi(\kappa)\left(\frac{\sigma_{A,U}}{\sigma_U} - \frac{\sigma_{\theta,U}}{\sigma_U}\right)\right) \\
&= -\left(\phi(\kappa) - \frac{1}{\eta}\kappa\right)\sigma_U
\end{aligned}
\tag{147}
$$

The claimed expression follows from using the definition $\Omega = \phi(\kappa) - \frac{\kappa}{\eta}$. Finally, the claimed expression for the price level follows immediately from observing that $\ln P = \ln M - \ln C$. $\square$

A.6 Proof of Theorem 4

Proof. We now exploit the fact from Equation 27 that aggregate consumption is given by:

$$
\ln C = \frac{\eta}{\eta-1}\ln\left(\left(\exp\{\ln M\} - \int_k^1 p_i\bar{c}_i\,\mathrm{d}i\right)^{\frac{\eta-1}{\eta}}\left(\int_0^k \theta_i p_i^{1-\eta}\,\mathrm{d}i\right)^{\frac{1}{\eta}} + \int_k^1 \theta_i^{\frac{1}{\eta}}\bar{c}_i^{\frac{\eta-1}{\eta}}\,\mathrm{d}i\right) \tag{148}
$$

where $\ln M = \ln \bar{M} + \ln\Delta$, and $p_i = p$ and $L_i = L$ are invariant to Δ. Differentiating this expression with respect to Δ, we have that:

$$
\frac{\partial \ln C}{\partial \ln \Delta} = \frac{\partial \ln C}{\partial \ln M}\frac{\partial \ln M}{\partial \ln \Delta} = \frac{\partial \ln C}{\partial \ln M} \tag{149}
$$

Moreover, we can compute:

$$
\begin{aligned}
\frac{\partial \ln C}{\partial \ln M} &= \frac{\eta}{\eta-1}C^{-\frac{\eta-1}{\eta}}\frac{\partial}{\partial \ln M}\left(\left(\exp\{\ln M\} - \int_k^1 p_i\bar{c}_i\,\mathrm{d}i\right)^{\frac{\eta-1}{\eta}}\left(\int_0^k \theta_i p_i^{1-\eta}\,\mathrm{d}i\right)^{\frac{1}{\eta}} + \int_k^1 \theta_i^{\frac{1}{\eta}}\bar{c}_i^{\frac{\eta-1}{\eta}}\,\mathrm{d}i\right) \\
&= \frac{\eta}{\eta-1}C^{-\frac{\eta-1}{\eta}}\frac{\partial}{\partial \ln M}\left(\left(\exp\{\ln M\} - pL\int_k^1 A_i\,\mathrm{d}i\right)^{\frac{\eta-1}{\eta}}\left(p^{1-\eta}\int_0^k \theta_i\,\mathrm{d}i\right)^{\frac{1}{\eta}} + L^{\frac{\eta-1}{\eta}}\int_k^1 \theta_i^{\frac{1}{\eta}}A_i^{\frac{\eta-1}{\eta}}\,\mathrm{d}i\right) \\
&= \frac{\eta}{\eta-1}C^{-\frac{\eta-1}{\eta}}\left[\frac{\eta-1}{\eta}\left(M + pLA_k\frac{\partial k}{\partial M}\right)\left(\frac{\exp\{\ln M\} - pL\int_k^1 A_i\,\mathrm{d}i}{p^{1-\eta}\int_0^k \theta_i\,\mathrm{d}i}\right)^{-\frac{1}{\eta}}\right. \\
&\qquad \left. + \frac{1}{\eta}p^{1-\eta}\theta_k\frac{\partial k}{\partial M}\left(\frac{\exp\{\ln M\} - pL\int_k^1 A_i\,\mathrm{d}i}{p^{1-\eta}\int_0^k \theta_i\,\mathrm{d}i}\right)^{\frac{\eta-1}{\eta}} + L^{\frac{\eta-1}{\eta}}\theta_k^{\frac{1}{\eta}}A_k^{\frac{\eta-1}{\eta}}\frac{\partial k}{\partial M}\right]
\end{aligned}
\tag{150}
$$

We now evaluate at $\ln\Delta = 0$. This allows us to make use of the fact that:

$$
D = \frac{\exp\{\ln \bar{M}\} - pL\int_k^1 A_i\,\mathrm{d}i}{p^{1-\eta}\int_0^k \theta_i\,\mathrm{d}i} = \exp\{\mu_z - \mu_\theta\} \tag{151}
$$

which gives us that:

$$
\begin{aligned}
\left.\frac{\partial \ln C}{\partial \ln M}\right|_{M=\bar{M}} &= C^{-\frac{\eta-1}{\eta}}\bar{M}\exp\left\{-\frac{1}{\eta}\left(\mu_z-\mu_\theta\right)\right\} \\
&+ \frac{\eta}{\eta-1}C^{-\frac{\eta-1}{\eta}}\left.\frac{\partial k}{\partial M}\right|_{M=\bar{M}}\left[\left(\frac{\eta-1}{\eta}A_kL+\frac{1}{\eta}\theta_kDp^{-\eta}\right)pD^{-\frac{1}{\eta}}-(A_kL)^{\frac{\eta-1}{\eta}}(\theta_kDp^{-\eta})^{\frac{1}{\eta}}pD^{-\frac{1}{\eta}}\right] \\
&= C^{-\frac{\eta-1}{\eta}}\bar{M}\exp\left\{-\frac{1}{\eta}\left(\mu_z-\mu_\theta\right)\right\} \\
&+ \frac{\eta}{\eta-1}C^{-\frac{\eta-1}{\eta}}pD^{-\frac{1}{\eta}}\left.\frac{\partial k}{\partial M}\right|_{M=\bar{M}}\left[\frac{\eta-1}{\eta}A_kL+\frac{1}{\eta}\theta_kDp^{-\eta}-(A_kL)^{\frac{\eta-1}{\eta}}(\theta_kDp^{-\eta})^{\frac{1}{\eta}}\right] \\
&= C^{-\frac{\eta-1}{\eta}}\bar{M}\exp\left\{-\frac{1}{\eta}\left(\mu_z-\mu_\theta\right)\right\}+\frac{\eta}{\eta-1}C^{-\frac{\eta-1}{\eta}}\left.\frac{\partial k}{\partial M}\right|_{M=\bar{M}}pD^{-\frac{1}{\eta}}\left[A_kL-A_kL\right] \\
&= \bar{M}\exp\left\{-\frac{\eta-1}{\eta}\ln C-\frac{1}{\eta}\left(\mu_z-\mu_\theta\right)\right\}
\end{aligned}
\tag{152}
$$

where we exploited the fact that $A_kL=\theta_kDp^{-\eta}$ for the threshold market k. We now compute the small volatility expansion. We first compute:

$$
\begin{aligned}
\lim_{\delta\to 0}\mathscr{M} &= \lim_{\delta\to 0}\left.\frac{\partial \ln C}{\partial \ln M}\right|_{M=\bar{M}} = \bar{M}\exp\left\{-\frac{\eta-1}{\eta}\ln C(0)-\frac{1}{\eta}\mu_z(0)+\frac{1}{\eta}\mu_\theta\right\} \\
&= \bar{M}\exp\{-\ln\bar{M}\}=1
\end{aligned}
\tag{153}
$$

where we have used the expressions from Theorem 3. We now compute:

$$
\begin{aligned}
\lim_{\delta\to 0}\frac{\partial}{\partial\delta}\left[\left.\frac{\partial \ln C}{\partial \ln M}\right|_{M=\bar{M}}\right] &= \left(-\frac{\eta-1}{\eta}(\ln C)'(0)-\frac{1}{\eta}\mu_z'(0)\right)\bar{M}\exp\left\{-\frac{\eta-1}{\eta}\ln C(0)-\frac{1}{\eta}\mu_z(0)+\frac{1}{\eta}\mu_\theta\right\} \\
&= \left(\frac{\eta-1}{\eta}\left(\phi(\kappa)-\frac{\kappa}{\eta}\right)\sigma_U-\frac{1}{\eta}\eta\phi(\kappa)\sigma_U\right)\times 1 \\
&= -\frac{1}{\eta}\left(\phi(\kappa)+\frac{\eta-1}{\eta}\kappa\right)\sigma_U
\end{aligned}
\tag{154}
$$

Completing the proof. □

A.7 Proof of Proposition 4

Proof. In Theorem 4, we proved that $\lim_{\delta\to 0}\mathscr{M}=1$. Moreover, the arguments showed that $\ln C$ is continuous in $\ln\Delta$, and hence $\mathscr{M}^+=\mathscr{M}^-$ for any δ, showing that $\lim_{\delta\to 0}\mathscr{M}^+=\lim_{\delta\to 0}\mathscr{M}^-=\lim_{\delta\to 0}\mathscr{M}=1$.

Now consider the economy in which $\delta = 0$. Let $\bar{c}_i$ denote production in each market i and $\bar{C}$ denote the consumption aggregator of these values. If $\ln \Delta = 0$, then demand equals supply in all markets. That is, $c_i = \bar{c}_i$ for all i and $C = \bar{C}$. If $\ln \Delta > 0$, then demand (Equation 28) exceeds supply in all markets. Hence, all markets are rationed, $c_i = \bar{c}_i$ for all i, and $C = \bar{C}$. Therefore, $\mathscr{M}^+ = 0$. If $\ln \Delta < 0$, then, by a symmetric argument, demand in each market i is $\Delta \bar{c}_i < \bar{c}_i$. Therefore, since C is constant returns to scale, $C = \Delta \bar{C}$ and, as a result, $\mathscr{M}^- = 1$.

□

A.8 Proof of Proposition 5

Proof. From Theorem 4, we have that:

$$\operatorname{sgn}\left(\frac{\partial \mathscr{M}}{\partial \eta}\right) = \operatorname{sgn}\left(-\frac{\partial}{\partial \eta}\left[\frac{1}{\eta}\left(\phi(\kappa) + \frac{\eta - 1}{\eta}\kappa\right)\right]\right) \tag{155}$$

We now make use of the fact that $\eta = 1/(1 - \Phi(\kappa))$ to write:

$$\frac{1}{\eta}\left(\phi(\kappa) + \frac{\eta - 1}{\eta}\kappa\right) = (1 - \Phi(\kappa))(\phi(\kappa) + \kappa\Phi(\kappa)) \equiv g(\kappa) \tag{156}$$

Observing that κ is a strictly increasing function of η, it suffices to sign the derivative of g, which is given by:

$$\begin{aligned} g'(\kappa) &= -\phi(\kappa)(\phi(\kappa) + \kappa\Phi(\kappa)) + (1 - \Phi(\kappa))(\phi'(\kappa) + \Phi(\kappa) + \kappa\phi(\kappa)) \\ &= \Phi(\kappa)(1 - \Phi(\kappa)) - \phi(\kappa)(\phi(\kappa) + \kappa\Phi(\kappa)) \end{aligned} \tag{157}$$

where we have used the fact that $\phi'(\kappa) = -\kappa\phi(\kappa)$. The roots of g' must therefore solve:

$$1 = m(\kappa)(\kappa + n(\kappa)) \equiv R(\kappa) \tag{158}$$

where $m(\cdot)$ and $n(\cdot)$ are respectively the right and left inverse Mills ratios. Moreover, we know that m is a strictly increasing function. Thus, if $\kappa + n(\kappa)$ is an increasing function, then $g(\kappa)$ has at most one root. This is an increasing function if:

$$0 \leq 1 + r'(\kappa) = 1 - (\kappa + n(\kappa))n(\kappa) \iff n(\kappa)(\kappa + n(\kappa)) \leq 1 \tag{159}$$

We now use some facts about Mills ratios. First, as $\Phi(-x) = 1 - \Phi(x)$ and $\phi(-x) = \phi(x)$, we have that $n(\kappa) = m(-\kappa)$. Thus, we require that $m(-\kappa)(\kappa + m(-\kappa)) \leq 1$. Equivalently, we must prove that $m(x)(-x + m(x)) \leq 1$ for all $x \in \mathbb{R}$. But this is equivalent to the inequality $m(x) \leq \frac{1}{2}(x + \sqrt{x^2+4})$, which holds for all $x \in \mathbb{R}$ by the theorem of Birnbaum (1942). We further have that $R(0) = m(0)(0 + n(0)) = 4\phi(0)^2 = 4(1/\sqrt{2\pi})^2 = 2/\pi < 1$. Moreover, $\lim_{\kappa\to\infty} R(\kappa) = \infty$ (as $m(\kappa) \geq \kappa$ and $n(\kappa) \geq 0$). Thus, we have that g' changes sign exactly once, at the unique solution κ^* of the equation $R(\kappa) = 1$. Moreover, when $\kappa < \kappa^*$, we have that $g'(\kappa) > 0$ and when $\kappa > \kappa^*$, we have that $g'(\kappa) < 0$. Hence, the extent of monetary non-neutrality is decreasing in η when $\eta < \eta^*$ and increasing in η when $\eta > \eta^*$, where $\eta^* = 1/(1 - \Phi(\kappa^*))$. Numerical calculations yield that $\kappa^* \approx 0.376$ and $\eta^* \approx 2.83$. $\square$

A.9 Proof of Proposition 6

Proof. We start by log-linearizing the firm's dynamic tightness and pricing decisions.

Lemma 6. *When log-linearized around the steady state, the optimal reset tightness and reset prices follow:*

$$\hat{t}_\tau = -(1-\beta\omega)\left(\hat{D}_\tau + \tilde{\psi}\hat{\sigma}_{U,\tau}\right) + \beta\omega\mathbb{E}_\tau[\hat{t}_{\tau+1}] \tag{160}$$

$$\hat{p}_\tau = (1-\beta\omega)\left(\hat{M}_\tau - \tilde{\zeta}\hat{\sigma}_{U,\tau}\right) + \beta\omega\mathbb{E}_\tau[\hat{p}_{\tau+1}] \tag{161}$$

where $\tilde{\psi}$ and $\tilde{\zeta}$ are constants given in the proof. Moreover, the implied optimal reset level of labor inputs follows:

$$\hat{L}_\tau = (1-\beta\omega)(\hat{D}_\tau - \eta\hat{M}_\tau + (\tilde{\psi} + \eta\tilde{\zeta})\hat{\sigma}_{U,\tau}) + \beta\omega\mathbb{E}_\tau[\hat{L}_{\tau+1}] \tag{162}$$

Proof. We begin by log-linearizing the forward-looking decision rules of the firm that we derived in Equation 51:

$$\begin{aligned} t_\tau &= \frac{1}{\eta-1}\frac{\sum_{j=0}^\infty (\beta\omega)^j \frac{1}{M_{\tau+j}}\mathbb{E}_\tau[A_{\tau+j}\mathbb{I}[t_\tau D_{\tau+j} \geq A_{\tau+j}/\theta_{\tau+j}]]}{\sum_{j=0}^\infty (\beta\omega)^j \frac{1}{M_{\tau+j}} D_{\tau+j}\mathbb{E}_\tau[\theta_{\tau+j}\mathbb{I}[t_\tau D_{\tau+j} \leq A_{\tau+j}/\theta_{\tau+j}]]} \\ p_\tau &= \frac{\frac{\alpha}{1-\beta\omega}}{\sum_{j=0}^\infty (\beta\omega)^j \frac{1}{M_{\tau+j}}\mathbb{E}_\tau[A_{\tau+j}\mathbb{I}[t_\tau D_{\tau+j} \geq A_{\tau+j}/\theta_{\tau+j}]]} \end{aligned} \tag{163}$$

We define $\mathcal{A}_{\tau+j} = \frac{1}{M_{\tau+j}}\mathbb{E}_\tau[A_{\tau+j}\mathbb{I}[t_\tau D_{\tau+j} \geq A_{\tau+j}/\theta_{\tau+j}]]$ and $\mathcal{B}_{\tau+j} = \frac{1}{M_{\tau+j}} D_{\tau+j}\mathbb{E}_\tau[\theta_{\tau+j}\mathbb{I}[t_\tau D_{\tau+j} \leq A_{\tau+j}/\theta_{\tau+j}]]$. We therefore have that:

$$\begin{aligned} \hat{t}_\tau &= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^\infty (\beta\omega)^j \left[\hat{\mathcal{A}}_{\tau+j} - \hat{\mathcal{B}}_{\tau+j}\right]\right] \\ \hat{p}_\tau &= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^\infty (\beta\omega)^j \left[-\hat{\mathcal{A}}_{\tau+j}\right]\right] \end{aligned} \tag{164}$$

We have already computed that:

$$\begin{aligned}
\mathcal{A}_{\tau+j} &= \frac{1}{M_{\tau+j}} \exp\left\{\mu_{A,\tau+j} + \frac{1}{2}\sigma^2_{A,\tau+j}\right\} \Phi\left(\frac{\ln\left(t_\tau D_{\tau+j}\right) - \mu_{U,\tau+j}}{\sigma_{U,\tau+j}} - \beta^S \sigma_{U,\tau+j}\right) \\
\mathcal{B}_{\tau+j} &= \frac{D_{\tau+j}}{M_{\tau+j}} \exp\left\{\mu_{\theta,\tau+j} + \frac{1}{2}\sigma^2_{\theta,\tau+j}\right\} \left(1 - \Phi\left(\frac{\ln\left(t_\tau D_{\tau+j}\right) - \mu_{U,\tau+j}}{\sigma_{U,\tau+j}} - \beta^D \sigma_{U,\tau+j}\right)\right)
\end{aligned} \tag{165}$$

and so we have that (as we have assumed that first moments of A and θ do not change):

$$\begin{aligned}
\hat{\mathcal{A}}_{\tau+j} &= -\hat{M}_{\tau+j} + \lambda_a \left(\hat{t}_\tau + \hat{D}_{\tau+j} + \phi_a \hat{\sigma}_{U,\tau+j}\right) \\
\hat{\mathcal{B}}_{\tau+j} &= -\hat{M}_{\tau+j} + \hat{D}_{\tau+j} - \lambda_b \left(\hat{t}_\tau + \hat{D}_{\tau+j} + \phi_b \hat{\sigma}_{U,\tau+j}\right)
\end{aligned} \tag{166}$$

where:

$$\begin{aligned}
\lambda_a &= \frac{1}{\bar{\sigma}_U} n \left(\frac{\ln\left(\bar{t}\bar{D}\right) - \mu_U}{\bar{\sigma}_U} - \beta^S \bar{\sigma}_U\right) > 0 \\
\lambda_b &= \frac{1}{\bar{\sigma}_U} m \left(\frac{\ln\left(\bar{t}\bar{D}\right) - \mu_U}{\bar{\sigma}_U} - \beta^D \bar{\sigma}_U\right) > 0 \\
\phi_a &= -\left(\left(\frac{\ln\left(\bar{t}\bar{D}\right) - \mu_U}{\bar{\sigma}_U} - \beta^S \bar{\sigma}_U\right) + 2\beta^S \bar{\sigma}_U\right) \bar{\sigma}_U \\
\phi_b &= -\left(\left(\frac{\ln\left(\bar{t}\bar{D}\right) - \mu_U}{\bar{\sigma}_U} - \beta^D \bar{\sigma}_U\right) + 2\beta^D \bar{\sigma}_U\right) \bar{\sigma}_U
\end{aligned} \tag{167}$$

Combining this information, we have that reset tightness follows (so long as $\lambda_a + \lambda_b \neq 1$, which we established in the proof of Theorem 1):

$$\begin{aligned}
\hat{t}_\tau &= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty} (\beta\omega)^j \left[(\lambda_a + \lambda_b)\hat{t}_\tau + (\lambda_a + \lambda_b - 1)\hat{D}_{\tau+j} + \left(\lambda_a \phi_a + \lambda_b \phi_b\right)\hat{\sigma}_{U,\tau+j}\right]\right] \\
&= (\lambda_a + \lambda_b)\hat{t}_\tau + (\lambda_a + \lambda_b - 1)(1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty} (\beta\omega)^j \left(\hat{D}_{\tau+j} + \tilde{\psi}\hat{\sigma}_{U,\tau+j}\right)\right]
\end{aligned} \tag{168}$$

where $\tilde{\psi} = (\lambda_a \phi_a + \lambda_b \phi_b)/(\lambda_a + \lambda_b - 1)$. This yields that:

$$\begin{aligned}
\hat{t}_\tau &= -(1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty} (\beta\omega)^j \left(\hat{D}_{\tau+j} + \tilde{\psi}\hat{\sigma}_{U,\tau+j}\right)\right] \\
&= -(1-\beta\omega)\left(\hat{D}_\tau + \tilde{\psi}\hat{\sigma}_{U,\tau}\right) + \beta\omega\mathbb{E}_\tau[\hat{t}_{\tau+1}]
\end{aligned} \tag{169}$$

Turning to reset prices, we have that:

$$
\begin{aligned}
\hat{p}_\tau &= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty}(\beta\omega)^j\left[\hat{M}_{\tau+j} - \lambda_a\left(\hat{t}_\tau + \hat{D}_{\tau+j} + \phi_a\hat{\sigma}_{U,\tau+j}\right)\right]\right] \\
&= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty}(\beta\omega)^j\hat{M}_{\tau+j}\right] - \lambda_a\left[\hat{t}_\tau + (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty}(\beta\omega)^j\left(\hat{D}_{\tau+j} + \phi_a\hat{\sigma}_{U,\tau+j}\right)\right]\right] \\
&= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty}(\beta\omega)^j\hat{M}_{\tau+j}\right] \\
&\qquad - \lambda_a\left[\hat{t}_\tau + (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty}(\beta\omega)^j\left(\hat{D}_{\tau+j} + \tilde{\psi}\hat{\sigma}_{U,\tau+j} + (-\tilde{\psi}+\phi_a)\hat{\sigma}_{U,\tau+j}\right)\right]\right] \\
&= (1-\beta\omega)\mathbb{E}_\tau\left[\sum_{j=0}^{\infty}(\beta\omega)^j\left(\hat{M}_{\tau+j} - \lambda_a(-\tilde{\psi}+\phi_a)\hat{\sigma}_{U,\tau+j}\right)\right] \\
&= (1-\beta\omega)\left(\hat{M}_\tau - \lambda_a(-\tilde{\psi}+\phi_a)\hat{\sigma}_{U,\tau}\right) + \beta\omega\mathbb{E}_\tau[\hat{p}_{\tau+1}]
\end{aligned}
\tag{170}
$$

Letting $\tilde{\zeta} = \lambda_a(-\tilde{\psi}+\phi_a)$ then yields the claimed expression. The formula for labor inputs follows from the identity that $\hat{L}_\tau = -\eta\hat{p}_\tau - \hat{t}_\tau$. $\square$

We next aggregate these variables and use the Calvo assumption to derive equations in terms of the average tightness, $\bar{t}_\tau$, and the average price level $\bar{p}_\tau$.

Lemma 7. *When log-linearized around the steady state, we have that price inflation π_τ and tightness inflation υ_τ obey the following equations:*

$$\pi_\tau = \Gamma(\hat{M}_\tau - \bar{p}_\tau) + \beta\mathbb{E}_\tau[\pi_{\tau+1}] + \zeta\hat{\sigma}_{U,\tau} \tag{171}$$

$$\upsilon_\tau = -\Gamma(\hat{D}_\tau + \bar{t}_\tau) + \beta\mathbb{E}_\tau[\upsilon_{\tau+1}] + \psi\hat{\sigma}_{U,\tau} \tag{172}$$

where $\Gamma = \frac{(1-\omega)(1-\beta\omega)}{\omega}$, $\zeta = -\Gamma\tilde{\zeta}$, and $\psi = -\Gamma\tilde{\psi}$.

Proof. We begin by noting that, under the Calvo assumption, $\bar{p}_\tau = \omega\bar{p}_{t-1} + (1-\omega)\hat{p}_\tau$, which implies that:

$$\hat{p}_\tau = \frac{\bar{p}_\tau - \omega\bar{p}_{\tau-1}}{1-\omega} = \frac{\bar{p}_\tau - \omega(\bar{p}_\tau - \Delta\bar{p}_\tau)}{1-\omega} = \bar{p}_\tau + \frac{\omega}{1-\omega}\Delta\bar{p}_\tau \tag{173}$$

Plugging this into the reset equation from Lemma 6, we obtain that:

$$\begin{aligned}\bar{p}_\tau + \frac{\omega}{1-\omega}\Delta\bar{p}_\tau &= (1-\beta\omega)(\hat{M}_\tau - \tilde{\zeta}\hat{\sigma}_{U,\tau}) + \beta\omega\mathbb{E}_\tau\left[\bar{p}_{\tau+1} + \frac{\omega}{1-\omega}\Delta\bar{p}_{\tau+1}\right] \\ &= (1-\beta\omega)(\hat{M}_\tau - \tilde{\zeta}\hat{\sigma}_{U,\tau}) + \beta\omega\mathbb{E}_\tau\left[\bar{p}_\tau + \Delta\bar{p}_{\tau+1} + \frac{\omega}{1-\omega}\Delta\bar{p}_{\tau+1}\right] \\ &= (1-\beta\omega)(\hat{M}_\tau - \tilde{\zeta}\hat{\sigma}_{U,\tau}) + \beta\omega\mathbb{E}_\tau\left[\bar{p}_\tau + \frac{1}{1-\omega}\Delta\bar{p}_{\tau+1}\right]\end{aligned} \tag{174}$$

Which we can re-arrange as:

$$\Delta\bar{p}_\tau = \Gamma(\hat{M}_\tau - \bar{p}_\tau) + \beta\mathbb{E}_\tau[\Delta\bar{p}_{\tau+1}] + \zeta\hat{\sigma}_{U,\tau} \tag{175}$$

We can now apply the same principle for tightness. In particular, we plug the fact that $\hat{t}_\tau = \bar{t}_\tau + \frac{\omega}{1-\omega}\Delta\bar{t}_\tau$ into the reset equation from Lemma 6 to obtain that:

$$\bar{t}_\tau + \frac{\omega}{1-\omega}\Delta\bar{t}_\tau = -(1-\beta\omega)(\hat{D}_\tau + \tilde{\psi}\hat{\sigma}_{U,\tau}) + \beta\omega\mathbb{E}_\tau\left[\bar{t}_{\tau+1} + \frac{\omega}{1-\omega}\Delta\bar{t}_{\tau+1}\right] \tag{176}$$

Identical algebraic steps yield the result. □

Finally, we log-linearize to obtain the demand shifter in slack markets.

Lemma 8. *When log-linearized around the steady state, the household's demand in slack markets follows:*

$$\hat{D}_\tau = \gamma_M\hat{M}_\tau + \gamma_p\bar{p}_\tau + \gamma_t\bar{t}_\tau + \gamma_U\hat{\sigma}_{U,\tau} \tag{177}$$

where γ_M, γ_p, and γ_t are coefficients derived in the proof.

Proof. We start from the fact that:

$$D_\tau = \frac{M_\tau - R_\tau}{S_\tau} \tag{178}$$

where $R_\tau = \int_0^1 p_{i,\tau}L_{i,\tau}A_{i,\tau}\mathbb{I}[t_{i,\tau}D_\tau \geq A_{i,\tau}/\theta_{i,\tau}]\,\mathrm{d}i$ and $S_\tau = \int_0^1 p_{i,\tau}^{1-\eta}\theta_{i,\tau}\mathbb{I}[t_{i,\tau}D_\tau \leq A_{i,\tau}/\theta_{i,\tau}]\,\mathrm{d}i$. Thus, we can log-linearize this as:

$$\hat{D}_\tau = \frac{\bar{M}}{\bar{M}-\bar{R}}\hat{M}_\tau - \frac{\bar{R}}{\bar{M}-\bar{R}}\hat{R}_\tau - \hat{S}_\tau \tag{179}$$

Or, defining the steady-state expenditure share on slack markets as $s = (\bar{M}-\bar{R})/\bar{M}$, we have that:

$$\hat{D}_\tau = \frac{1}{s}\hat{M}_\tau - \hat{S}_\tau - \frac{1-s}{s}\hat{R}_\tau \tag{180}$$

From the arguments of Lemma 4, we have that:

$$\begin{aligned} R_\tau &= \exp\left\{\mu_{A,\tau} + \frac{1}{2}\sigma_{A,\tau}^2\right\} \int_0^1 p_{i,\tau} L_{i,\tau} \Phi\left(\frac{\ln(t_{i,\tau} D_\tau) - \mu_{U,\tau}}{\sigma_{U,\tau}} - \beta^S \sigma_{U,\tau}\right) \mathrm{d}i \\ S_\tau &= \exp\left\{\mu_{\theta,\tau} + \frac{1}{2}\sigma_{\theta,\tau}^2\right\} \int_0^1 p_{i,\tau}^{1-\eta} \left(1 - \Phi\left(\frac{\ln(t_{i,\tau} D_\tau) - \mu_{U,\tau}}{\sigma_{U,\tau}} - \beta^D \sigma_{U,\tau}\right)\right) \mathrm{d}i \end{aligned} \tag{181}$$

Log-linearizing these expressions around the steady state (and noting that all firms' decisions are symmetric in the steady state), we obtain that:

$$\begin{aligned} \hat{R}_\tau &= \bar{p}_\tau + \bar{L}_\tau + \lambda_a(\bar{t}_\tau + \hat{D}_\tau + \phi_a \hat{\sigma}_{U,\tau}) \\ \hat{S}_\tau &= (1-\eta)\bar{p}_\tau - \lambda_b(\bar{t}_\tau + \hat{D}_\tau + \phi_b \hat{\sigma}_{U,\tau}) \end{aligned} \tag{182}$$

Using the fact that $\bar{L}_\tau = -\eta \bar{p}_\tau - \bar{t}_\tau$, we then obtain that:

$$\begin{aligned} \hat{D}_\tau &= \frac{1}{s}\hat{M}_\tau - (1-\eta)\bar{p}_\tau + \lambda_b(\bar{t}_\tau + \hat{D}_\tau + \phi_b \hat{\sigma}_{U,\tau}) \\ &\quad - \frac{1-s}{s}\Big((1-\eta)\bar{p}_\tau - \bar{t}_\tau + \lambda_a(\bar{t}_\tau + \hat{D}_\tau + \phi_a \hat{\sigma}_{U,\tau})\Big) \\ &= \frac{1}{s}\hat{M}_\tau - \frac{1}{s}(1-\eta)\bar{p}_\tau + \left(\lambda_b - \frac{1-s}{s}(\lambda_a - 1)\right)\bar{t}_\tau \\ &\quad + \left(\lambda_b - \frac{1-s}{s}\lambda_a\right)\hat{D}_\tau + \left(\lambda_b\phi_b - \frac{1-s}{s}\lambda_a\phi_a\right)\hat{\sigma}_{U,\tau} \end{aligned} \tag{183}$$

Solving this equation for $\hat{D}_\tau$ yields the claim. □

This completes the proof of the claims in the result. □

B Additional Analysis

B.1 Matching and the Importance of Efficiency

Uncertainty in our model has large effects, in the sense that decision problems depend on volatility to first order. To understand what drives this result, it is instructive to extend the model to imagine that realized sales are determined by a reduced-form matching function $m : \mathbb{R}_+^2 \to \mathbb{R}_+$ such that $q = m(q^s, q^d)$. Some macroeconomic studies have explored the implications of such matching technologies for goods markets. For example, using different methods, Bianchi et al. (2024), Ghassibe and Zanetti (2022) and Bai et al. (2025) all study how to integrate search frictions in product markets with benchmark models of economic fluctuations.[16]

The principle of voluntary exchange requires that any matching function satisfy $m(x,y) \leq \min\{x,y\}$. However, if exchange is inefficient, it can be the case that $m(x,y) < \min\{x,y\}$. Thus, the qualitative property of efficiency necessarily generates a kink in the matching function and thereby the firm's profit function.

To show the role of efficiency in our analysis, we now revisit our analysis of a single firm's problem when there is a smooth, inefficient matching function. As before, we will scale the volatility that the firm perceives by δ. To define what we mean by smooth, observe that the firm's objective is now:

$$\tilde{J}(p, L; \delta) = \mathbb{E}[\Lambda \left(pm(AL, zp^{-\eta}) - wL\right)] \tag{184}$$

We say that the matching function is smooth if $\tilde{J}$ is three times continuously differentiable and for all $\delta \geq 0$ there exists a unique optimal $(p^*, L^*) \gg 0$ that solves the first-order conditions $0 = \tilde{J}_p(p^*, L^*; \delta) = \tilde{J}_L(p^*, L^*; \delta)$ and the second-order condition that the Hessian of $\tilde{J}$ is negative definite at this solution.

The following result shows that, under a smooth matching function, the main qualitative property of Theorem 2 disappears: prices and labor inputs are invariant to uncertainty to the first order.

Proposition 7. *If the matching function is smooth, then, up to a first-order approximation in volatility:*

$$p^* = p^{FI} \quad \textit{and} \quad L^* = L^{FI} \tag{185}$$

Proof. We define the function $F : \mathbb{R}^3_+ \to \mathbb{R}^2$ as:

$$F(\delta, \xi) = \begin{pmatrix} \tilde{J}_p(\xi; \delta) \\ \tilde{J}_L(\xi; \delta) \end{pmatrix} \tag{186}$$

where $\xi = (p, L)$. If the matching function is smooth, then F is well-defined and, moreover, we have that $\xi^*(\delta) = (p^*(\delta), L^*(\delta))$ is the unique solution to $F(\delta, \xi^*(\delta)) = 0$. Application of the implicit function theorem to ξ^* yields that:

$$\xi^{*\prime}(\delta) = -F_\xi(0, \xi^*(0))^{-1} F_\delta(0, \xi^*(0)) \tag{187}$$

and ξ^* is twice continuously differentiable. This is because we assumed that F is three times continuously differentiable and the derivative of F, which is the Hessian of $\tilde{J}$, is negative

[16] Bianchi et al. (2024) discuss how smooth matching functions relate to efficient matching functions in Section 2.3 of their paper, but use a smooth matching function for their main analysis.

definite and therefore invertible. Moreover, we define:

$$\begin{aligned} h(\theta_\delta, \xi) &= \Lambda_\delta \left(pm(A_\delta L, z_\delta p^{-\eta}) - w_\delta L \right) \\ \text{where} \quad \theta_\delta &= (\ln \Lambda_\delta, \ln A_\delta, \ln z_\delta, \ln w_\delta) = \mu_\theta + \delta \Sigma^{\frac{1}{2}} Z \end{aligned} \tag{188}$$

for $Z \sim N(0, I)$. We then have that for $i \in \{p, L\}$:

$$\frac{\partial}{\partial \delta} F_i(\delta, \xi) = \mathbb{E}[\nabla_\theta h_i(\theta_\delta, \xi) \Sigma^{\frac{1}{2}} Z] \implies \frac{\partial}{\partial \delta} F_i(0, \xi^*(0)) = \mathbb{E}[\nabla_\theta h_i(\mu_\theta, \xi^*(0)) \Sigma^{\frac{1}{2}} Z] = 0 \tag{189}$$

Combining these facts implies that ${\xi^*}'(0) = 0$. Application of Taylor's remainder theorem to ξ^* then yields the result. □

With inefficient exchange and a smooth matching function, the optimum at $\delta = 0$ is described by a first-order condition that is, itself, a smooth function of parameters. Therefore, adding small amounts of randomness to parameters has zero effect on the first-order-condition (Equation 189), which translates into a zero effect on the optimizer (Equation 187). This contrasts to the case of efficient matching in which first-order conditions *do not* hold when $\delta = 0$ and marginal payoffs are not a smooth function of parameters (recall Figures 2 and 3). This logic explains the sharp difference between Proposition 7 and Theorem 2.

B.2 Inventories

To model inventories, we assume that each unsold unit of production becomes inventory $I = \max\{AL - zp^{-\eta}, 0\}$, each unit of which has future value $\Lambda' \in \mathbb{R}_{++}$, where Λ' is jointly log-normal with the other random variables that enter the producer's problem. We assume that the retention value of inventory is such that it would be better not to produce today than to produce fully for inventory (in expectation), *i.e.*, $\mathbb{E}[\Lambda' A] < \mathbb{E}[\Lambda w]$. The following proposition characterizes the salient properties of the model when augmented with inventories:

Proposition 8. *If the firm's behavior is optimal, then tightness must solve the following fixed-point equation:*

$$t = \frac{\mathbb{E}[\Lambda A \mathbb{I}[t \geq U]]}{(\eta - 1)\mathbb{E}[\Lambda z \mathbb{I}[t \leq U]] - \eta \mathbb{E}[\Lambda' z \mathbb{I}[t \leq U]] \frac{\mathbb{E}[\Lambda A \mathbb{I}[t \geq U]]}{\mathbb{E}[\Lambda w] - \mathbb{E}[\Lambda' A \mathbb{I}[t \leq U]]}} \tag{190}$$

Moreover, given a solution for tightness t^*, any optimal price must be given by:

$$p^* = \frac{\mathbb{E}[\Lambda w] - \mathbb{E}[\Lambda' A \mathbb{I}[t^* \leq U]]}{\mathbb{E}[\Lambda A \mathbb{I}[t^* \geq U]]} \tag{191}$$

and any optimal labor input quantity must be given by $L^* = p^{*-\eta}/t^*$. Moreover, we have that the rationing probability converges in the zero uncertainty limit to:

$$r_I = \frac{\eta - 1}{\eta + \frac{1}{\exp\{(\mu_\Lambda + \mu_w) - (\mu_{\Lambda'} + \mu_A)\} - 1}} \tag{192}$$

and, up to a first-order approximation in volatility, that:

$$\ln t = \mu_U + \kappa_I \sigma_U \tag{193}$$

where $\kappa_I = \Phi^{-1}(r_I)$.

Proof. Applying similar arguments to Theorem 1, we have that any optimum must be interior. Thus, again following the same steps as Theorem 1, we now have that any optimal choices of p and L must satisfy the first-order condition with respect to labor:

$$0 = p\mathbb{E}[\Lambda A \mathbb{I}[t \geq U]] - \mathbb{E}[\Lambda w] + \mathbb{E}[\Lambda' A \mathbb{I}[t \leq U]] \tag{194}$$

and with respect to the price:

$$0 = (1 - \eta)\mathbb{E}[\Lambda z \mathbb{I}[t \leq U]]p^{-\eta} + \mathbb{E}[\Lambda A \mathbb{I}[t \geq U]]L + \eta \mathbb{E}[\Lambda' z \mathbb{I}[t \leq U]]p^{-\eta-1} \tag{195}$$

Equation 194 immediately rearranges into Equation 191. After some algebra, Equation 195 reduces to:

$$t = \frac{\mathbb{E}[\Lambda A \mathbb{I}[t \geq U]]}{(\eta - 1)\mathbb{E}[\Lambda z \mathbb{I}[t \leq U]] - \eta \mathbb{E}[\Lambda' z \mathbb{I}[t \leq U]]p^{-1}} \tag{196}$$

Combining this with the formula for p from Equation 191 then yields Equation 190.

We now solve for the first-order effects of volatility in the firms' choices. To do this, observe that we may write:

$$x(\delta) = \ln\left(\frac{1}{\eta - 1}\right) + \ln N(\delta) - \ln D(\delta) - \ln(1 - K(\delta)) \tag{197}$$

where:

$$K(\delta) = \frac{\eta}{\eta - 1} \frac{\hat{D}(\delta)}{D(\delta)} \frac{N(\delta)}{\mathbb{E}[\Lambda_\delta w_\delta] - S(\delta)} \tag{198}$$

and:

$$\hat{D}(\delta) = \mathbb{E}[\Lambda'_\delta z_\delta \mathbb{I}[t(\delta) \leq U_\delta]] = \mathbb{E}[\Lambda'_\delta z_\delta]\left(1 - \Phi(y(\delta) - \delta \hat{\beta}^D \sigma_U)\right) \tag{199}$$

$$S(\delta) = \mathbb{E}[\Lambda'_\delta A_\delta \mathbb{I}[t(\delta) \leq U_\delta]] = \mathbb{E}[\Lambda'_\delta A_\delta]\left(1 - \Phi(y(\delta) - \delta \hat{\beta}^S \sigma_U)\right) \tag{200}$$

and $\hat{\beta}^D = \mathrm{Cov}[\ln \Lambda' + \ln z, U]/\mathrm{Var}[U]$, $\hat{\beta}^S = \mathrm{Cov}[\ln \Lambda' + \ln A, U]/\mathrm{Var}[U]$. Thus, exploiting the same arguments as those in Theorem 2, we have that:

$$\delta\sigma_U y(\delta) = \ln\left(\frac{1}{\eta-1}\right) + \ln\left(\frac{\Phi(y(\delta) - \delta\beta^S\sigma_U)}{1 - \Phi(y(\delta) - \delta\beta^D\sigma_U)}\right) - \ln(1 - K(\delta)) + O(\delta^2) \tag{201}$$

Passing to the $\delta \to 0$ limit, we obtain that:

$$K(0) = 1 - \frac{1}{\eta-1}\frac{\Phi(y(0))}{1 - \Phi(y(0))} \tag{202}$$

Moreover, we can compute that:

$$K(0) = \frac{\eta}{\eta-1}\exp\{\mu_{\Lambda'} - \mu_\Lambda\}\frac{\exp\{\mu_\Lambda + \mu_A\}\Phi(y(0))}{\exp\{\mu_\Lambda + \mu_w\} - \exp\{\mu_{\Lambda'} + \mu_A\}(1 - \Phi(y(0)))} \tag{203}$$

We also recall that the rationing probability is given by $r(\delta) = \Phi(y(\delta))$. Thus, $r(0) = \Phi(y(0))$. We call $r(0) = r_I$ and we define $\kappa_I = \Phi^{-1}(r_I)$. We now combine the previous two equations to solve for r_I:

$$\frac{\eta}{\eta-1}\frac{\exp\{\mu_{\Lambda'} + \mu_A\}r_I}{\exp\{\mu_\Lambda + \mu_w\} - \exp\{\mu_{\Lambda'} + \mu_A\}(1 - r_I)} = 1 - \frac{1}{\eta-1}\frac{r_I}{1 - r_I} \tag{204}$$

This reduces to Equation 192. Moreover, we can observe that as $\mu_{\Lambda'} \to -\infty$ (*i.e.*, $\Lambda' \to^{a.s.} 0$), that $r_I \to \frac{\eta-1}{\eta}$. Moreover, we know from Theorem 2 that $x'(0) = \sigma_U y(0) = \sigma_U \kappa_I$. Thus, we have that:

$$(\ln t)(\delta) = \mu_U + \kappa_I\sigma_U\delta + O(\delta^2) \tag{205}$$

Completing the proof. □

Intuitively, the presence of inventories leads the firm to be less averse to over-production. For this reason, the presence of inventories decreases the rationing probability $r^I < \frac{\eta-1}{\eta}$ and also reduces the responsiveness of tightness to volatility $\kappa_I < \kappa$. However, the qualitative properties of the model remain much the same: rationing probabilities govern markups, and volatility continues to have first-order effects.[17]

B.3 Walrasian Disequilibrium in Multiple Markets

In our baseline model, for analytical simplicity, we maintained market clearing in the labor market. In this extension, we show how to allow for disequilibrium in both the product

[17] The first-order expansions for prices and labor inputs, although not reported for brevity, can be readily computed by combining the arguments in Theorem 2 and Proposition 8.

and labor markets within our framework, in the spirit of Barro and Grossman (1971) and Michaillat and Saez (2015).

We augment our baseline model, in which the firm took the wage as given, by supposing that the firm must also set wages $w \in \mathbb{R}_+$ alongside the price p and labor demand L^d. When the firm sets a wage w, it faces the labor supply curve $L^s = vw^\chi$, where $\chi > 0$ is the elasticity of labor supply and $v \in \mathbb{R}_{++}$ is a labor supply shifter. We assume that the variables (Λ, A, z, v) are jointly log-normally distributed. We impose the same voluntary and efficient exchange assumptions on the labor market, so the realized labor quantity is $L = \min\{L^s, L^d\}$. The firm now solves the following problem:

$$\sup_{p,w,L^d} \mathbb{E}\left[\Lambda(pq - wL)\right] \tag{206}$$

subject to constraints that describe the state of both the product and labor markets. We note that our baseline model would be recovered if v were known to the firm and labor supply were perfectly elastic ($\chi \to \infty$). In this case, the firm will choose for the labor market to be in Walrasian equilibrium, and moreover has no labor market power.

The two linked markets could be in one of four states, ignoring the knife-edge cases of Walrasian equilibrium. The labor market is rationed if and only if $t^L > v$, where we introduce the new variable $t^L = \frac{L^d}{w^\chi}$ to represent labor-market tightness or the imbalance between labor demand and supply. The product market is rationed if demand exceeds production, but the precise condition under which this occurs depends on the state of the labor market. If the labor market is slack, the condition remains $t = \frac{p^{-\eta}}{L^d} > A/z$. If the labor market is rationed, then labor is supply-determined, and the relevant condition is $\frac{p^{-\eta}}{L^s} > A/z$, or $t \cdot t^L > Av/z$. We summarize this by defining the indicator functions

$$R^L = \mathbb{I}[\{t^L > v\}],\ R^P = \mathbb{I}[\{t^L < v, t > U\} \cup \{t^L > v, t \cdot t^L > Uv\}] \tag{207}$$

to denote the states in which the labor and product markets, respectively, are rationed.

The solution to the firm's problem can be characterized by a pair of fixed-point equations that describe the firm's optimal desired tightness in both markets:

Proposition 9. *Any optimal desired tightness in the product and labor markets solves the following system of equations:*

$$\begin{aligned} t &= \frac{1}{\eta - 1} \frac{\mathbb{E}[\Lambda A(1 - R^L)R^P] + \frac{1}{t^L}\mathbb{E}[\Lambda A v R^L R^P]}{\mathbb{E}[\Lambda z(1 - R^P)]} \\ t^L &= \chi \frac{\mathbb{E}[\Lambda A v R^L R^P]}{\mathbb{E}[\Lambda A(1 - R^L)R^P]} - (1 + \chi)\frac{\mathbb{E}[\Lambda v R^L]}{\mathbb{E}[\Lambda(1 - R^L)]} \end{aligned} \tag{208}$$

Moreover, the firm's optimal price is

$$p = w\frac{\mathbb{E}[\Lambda(1-R^L)]}{\mathbb{E}[\Lambda A(1-R^L)R^P]} \tag{209}$$

Proof. We define tightness in the product market as $t = \frac{p^{-\eta}}{L^D}$ and tightness in the labor market as $t^L = \frac{L^D}{w^\chi}$. There are four possible cases for the two markets, ignoring the zero-probability events of Walrasian equilibrium in either or both markets:

1. Slack product and labor markets. This event occurs if $t^L < v$ and $t < U$. In this case, the firm's realized objective is

$$\Lambda(pq^d - wL^d) = \Lambda(zp^{1-\eta} - wL^d) \tag{210}$$

2. Rationed product market and slack labor market. This event occurs if $t^L < v$ and $t > U$. In this case, the firm's realized objective is

$$\Lambda(pq^s - wL^d) = \Lambda(pAL^d - wL^d) \tag{211}$$

3. Slack product market and rationed labor market. This event occurs if $t^L > v$ and $p^{-\eta}/w^\chi = tt^L < Uv$. In this case, the firm's realized objective is

$$\Lambda(pq^d - wL^s) = \Lambda(zp^{1-\eta} - vw^{1+\chi}) \tag{212}$$

4. Rationed product market and labor market. This event occurs if $t^L > v$ and $p^{-\eta}/w^\chi = tt^L > Uv$. In this case, the firm's realized objective is

$$\Lambda(pq^s - wL^s) = \Lambda(pAvw^\chi - vw^{1+\chi}) \tag{213}$$

We furthermore define the events

$$R^L = \{t^L > v\}, \;\; R^P = \{t^L < v, t > U\} \cup \{t^L > v, tt^L > Uv\} \tag{214}$$

to denote the states in which the labor and product markets, respectively, are rationed.

Using these observations, we can write the firm's objective as an expectation over these four events

$$\begin{aligned} &\mathbb{E}[\Lambda(zp^{1-\eta} - wL^d)(1-R^L)(1-R^P)] + \mathbb{E}[\Lambda(pAL^d - wL^d)(1-R^L)R^P] \\ &+ \mathbb{E}[\Lambda(zp^{1-\eta} - vw^{1+\chi})R^L(1-R^P)] + \mathbb{E}[\Lambda(pAvw^\chi - vw^{1+\chi})R^L R^P] \end{aligned} \tag{215}$$

Analogous arguments to those of Theorem 1 establish that the problem of maximizing J has a solution and that any solution must be interior. Necessary conditions for optimality are that the first-order gain to changing each of p, L^D, and w is zero.

We first take the first-order condition in L^d, observing that this has first-order effects on the firms' payoffs only when the labor market is slack ($t^L < v$):

$$0 = -w\mathbb{E}[\Lambda(1 - R^L)] + p\mathbb{E}[\Lambda A(1 - R^L)R^P] \tag{216}$$

Re-arranging,

$$p = w\frac{\mathbb{E}[\Lambda(1 - R^L)]}{\mathbb{E}[\Lambda A(1 - R^L)R^P]} \tag{217}$$

Next, we consider the first-order condition with respect to p

$$0 = (1 - \eta)p^{-\eta}\mathbb{E}[\Lambda z(1 - R^P)] + L^d\mathbb{E}[\Lambda A(1 - R^L)R^P] + w^\chi\mathbb{E}[\Lambda A v R^L R^P] \tag{218}$$

Dividing each term by L^d and re-arranging yields

$$t = \frac{1}{\eta - 1}\frac{\mathbb{E}[\Lambda A(1 - R^L)R^P] + \frac{1}{t^L}\mathbb{E}[\Lambda A v R^L R^P]}{\mathbb{E}[\Lambda z(1 - R^P)]} \tag{219}$$

We finally derive the first-order condition with respect to w:

$$0 = -L^d\mathbb{E}[\Lambda(1 - R^L)] - (1 + \chi)w^\chi\mathbb{E}[\Lambda v R^L] + \chi w^{\chi-1}p\mathbb{E}[\Lambda A v R^L R^P] \tag{220}$$

Dividing all terms by w^χ, and substituting p/w from Equation 217:

$$0 = -t^L\mathbb{E}[\Lambda(1 - R^L)] - (1 + \chi)\mathbb{E}[\Lambda v R^L] + \chi\mathbb{E}[\Lambda A v R^L R^P]\frac{\mathbb{E}[\Lambda(1 - R^L)]}{\mathbb{E}[\Lambda A(1 - R^L)R^P]} \tag{221}$$

Re-arranging to isolate t^L,

$$t^L = \chi\frac{\mathbb{E}[\Lambda A v R^L R^P]}{\mathbb{E}[\Lambda A(1 - R^L)R^P]} - (1 + \chi)\frac{\mathbb{E}[\Lambda v R^L]}{\mathbb{E}[\Lambda(1 - R^L)]} \tag{222}$$

□

The core economics of Equation 208 come from the producer's local indifference about changing prices and posted wages. The logic of the former is unchanged from before (see Theorem 1 and Figure 1), with only one twist: when the product market is rationed, the marginal revenue from raising prices could be proportional either to labor demand (slack labor markets) or labor supply (rationed labor markets). The logic of the latter is new to the two-market analysis and reflects three forces. If labor markets are slack, raising wages increases costs in proportion to labor demand. If labor markets are rationed, they increase costs by both raising inframarginal wages and moving up the labor supply curve. And if product markets are *also* rationed, then increasing wages leads to additional marginal production and sales.

Like its twin in Theorem 1, this characterization opens the door for uncertainty to have first-order effects on outcomes. A new implication, relative to the earlier analysis, is that uncertainty may have first-order effects on the excess supply of labor, which can be interpreted as involuntary unemployment.

We close by describing the condition relating wages and prices (Equation 209) and describing its relationship in terms of markups and markdowns. This condition is derived from the firm's indifference condition for labor demand, observing that this choice affects payoffs *only* if the labor market is slack. For intuition, it is easiest to consider the case in which Λ and A are constant, in which case

$$p = \frac{1}{r^C}\frac{w}{A} \geq \frac{w}{A}, \qquad\qquad w = r^C pA \leq pA \tag{223}$$

where $r^C \in [0,1]$ is the *conditional* probability that the product market is rationed given a slack labor market. These formulas imply that both markups and markdowns, standard measures respectively of monopoly and monopsony power, are governed by the probability of rationing. This, in turn, opens the possibility that both depend (to the first-order) on firms' uncertainty.

B.4 Calibration of σ_U

In this appendix, we describe a strategy for separately calibrating the volatility of productivity (A) and idiosyncratic demand shocks (z) faced for firms. This allows us to calibrate the volatility of the composite parameter $U = A/z$, which in turn is critical for calibrating the likelihood of rationing and slack. This calculation is especially subtle in our model for two reasons. First, productivity and demand shocks—which often have a symmetric role in standard market-clearing models, as shifters of the value marginal product of inputs—have sharply *asymmetric* roles in our theory, because they have different effects on the state of the market (rationed vs. slack). Second, as we will soon make clear, optimizing firms' choices of tightness complicates the mapping from fundamental shocks to the randomness an econometrician would observe in "value-based" measures of productivity.

Measurement. We base our calibration strategy on the availability of two measurements in any firm-level microdata with separate observations of inputs, quantities produced, and the value of output sold.

The first measurement is "physical TFP," or TFPQ. In the data, this is defined as plant output in *physical units* (*e.g.*, widgets) divided by an index of inputs. That is, in the data,

$$\ln \text{TFPQ} = \ln \text{Physical Output} - \ln f(\text{Inputs}) \tag{224}$$

where $f(\text{Inputs})$ is an identified production function (*e.g.*, under Cobb-Douglas, a sum of inputs in logarithmic units times their output elasticities). Thus, viewed through the lens of our model:

$$\ln \text{TFPQ} = \ln q^S - \ln L = \ln A \tag{225}$$

where A is the physical productivity of the firm.

The second measurement is revenue-based productivity, or TFPR. In the data, this is analogous to TFPQ but with *sales* (price times transacted quantity) replacing physical output:

$$\ln \text{TFPR} = \ln \text{Sales} - \ln f(\text{Inputs}) \tag{226}$$

Through the lens of the model, we interpret TFPR thusly. Products are sold at a constant unit price p, chosen in advance. Quantities sold equal $q^D = zp^{-\eta}$ if the market is slack, or $q^S = AL$ if the market is rationed. Thus,

$$\begin{aligned} \ln \text{TFPR} &= \ln p + \ln \min\{zp^{-\eta}, AL\} - \ln L \\ &= \ln p + \ln \min\{zt, A\} \end{aligned} \tag{227}$$

where we remind that the definition of tightness is $t = p^{-\eta}/L$. Moreover, in our model, optimal tightness depends on the variances of random variables as described in Theorem 1.

From Data to Theory. Foster et al. (2008) conduct a detailed comparison of quantity- and revenue-based measures of TFP using US Census of Manufacturing microdata from particular manufacturing sectors for which physical output is particularly homogeneous, such as cardboard boxes, white bread, and plywood. As the authors note, these sample restrictions are helpful to make cross-firm comparisons of physical productivity meaningful given the impracticality of constructing firm-specific measures of the quality of output. Their broad conclusion, across measurement methods, is that the volatility in quantity- and revenue-based TFP are comparable: that is, $\sigma_A \approx \sigma_{\text{TFPR}}$. Next, we observe that Bloom et al. (2018) use plant-level US Census of Manufacturing microdata *across sectors* to construct detailed measures of one-quarter-ahead firm uncertainty regarding revenue-based TFP. This incorporates three corrections that are useful for our analysis, versus directly using the estimates of Foster et al. (2008): (i) isolating the statistically unforecastable component of firm-level shocks, (ii) further estimating the fraction of that statistical uncertainty that is plausibly faced by the firm (given measurement error as well as unobservable information), and (iii) translating the estimate to a decision period of one quarter. This results in an estimated one-quarter ahead standard deviation of $\sigma_{\text{TFPR}} = 0.10$.[18]

What remains is to calibrate σ_z. We do this numerically to match the estimates $\sigma_A = \sigma_{\text{TFPR}} = 0.10$, simulating the distribution of $\ln \text{TFPR}$ using Equation 227 and, for each possible value of σ_z, also calculating the firm's optimal tightness using Theorem 1.[19] Doing this results in an estimate of $\sigma_z = 0.21$, and thus $\sigma_U = \sqrt{\sigma_A^2 + \sigma_z^2} = 0.23$.

[18]These calculations are based on the calibrated values of Table V of Bloom et al. (2018). Specifically, these authors consider an environment with two volatility states that evolve via a discrete Markov process. In the authors' calibration, the "low" and "high" volatility state respectively occur with probability 0.68 and 0.32. In the former state, the standard deviation of unforecastable quarterly TFPR shocks is $0.051 = 5.1\%$, and in the latter, it is $0.209 = 20.9\%$. Averaging these yields the estimate $0.68 \times 0.051 + 0.32 \times 0.209 = 0.10$.

[19]We assume, for this calculation, that all uncertainty about A and z comes from idiosyncratic shocks and moreover that this uncertainty is orthogonal to that in the stochastic discount factor Λ.