Market Liquidity and Inventory Cycles

Ziqi Qiao *

November 21, 2024

Link to the latest version

Abstract

Using U.S. data, I document that inventory investment is pro-cyclical, the inventory-sales (I/S) ratio is counter-cyclical, and aggregate markup is pro-cyclical. However, existing models fail to fully explain these empirical patterns. To address this gap, I introduce a model where inventory holdings are driven by search friction and influenced by market liquidity, defined as the trade-off between markup and the speed of sales. In this model, sellers stock goods and post prices, while buyers select which sellers to visit. Due to a lack of coordination among buyers, sales are stochastic, often resulting in leftover inventory. This friction introduces a new incentive for sellers to hold inventory: to improve buyer experience by offering better terms of trade. However, this incentive leads to a negative externality. In the calibrated economy, sellers, on aggregate, overstock by 1.6%, equivalent to 0.2% of annual GDP, while the welfare cost amounts to 0.27% of annual GDP after accounting for distributional effects. Unlike the Walrasian framework, where prices are determined by market-clearing conditions, sellers in my model actively adjust both prices and quantities. This pricing flexibility allows the model to generate predictions that align with observed patterns of inventory cycles.

*University of Wisconsin – Madison. Email: zqiao7@wisc.edu. I am grateful to Randall Wright for his invaluable guidance and Guanming Shi for her support.

1 Introduction

Inventories represent approximately 13% of U.S. GDP, highlighting their crucial role in resource allocation within the economy. This raises two key questions: what role do inventories play, and is their substantial allocation efficient? Beyond their size, inventories are a fundamental driver of business cycle dynamics.[1] For instance, inventory investment is highly volatile, matching the fluctuations in output and moving closely with the business cycle. Several robust empirical regularities define this relationship: inventory investment is pro-cyclical, the inventory-sales (I/S) ratio is counter-cyclical, and aggregate markups are pro-cyclical. These patterns persist across various measures of the business cycle, emphasizing the need to understand inventory behavior to gain deeper insights into economic fluctuations. Despite their significant influence on resource allocation and business cycles, inventories remain underrepresented in most contemporary macroeconomic models.

This paper employs a structural VAR model to document the cyclicality of inventories and proposes a new framework to explain these patterns. I develop a dynamic general equilibrium model where inventory holdings are driven by search frictions and influenced by market liquidity, a trade-off between markup and sales speed. Model calibration indicates aggregate overstock equivalent to 0.2% of annual GDP, with an associated welfare cost of 0.27% of annual GDP. In addition to steady-state analysis, the model generates impulse responses that are consistent with empirical evidence.

Existing literature provides three primary explanations for inventory behavior: production smoothing, stock-out avoidance, and fixed restock costs. However, these models struggle to fully capture the observed cyclical patterns. Production smoothing models predict that production will be less variable than sales or a counter-cyclical markup, which contradicts the data (Blinder, 1986; West, 1990; Bils & Kahn, 2000; Kryvtsov & Midrigan, 2012; Nekarda & Ramey, 2013). Stock-out avoidance suggests that inventories fall when sales rise, contrary

[1] The importance of inventories in business cycles has been well-documented in the literature. See (Metzler, 1941; Blinder & Maccini, 1991; Ramey & West, 1999).

to the pro-cyclicality of inventory stock (Kahn, 1987; Coen-Pirani, 2004). Meanwhile, fixed restock cost models struggle to account for the persistence of the I/S ratio and often focus solely on intermediate inventory, neglecting final goods (Arrow *et al.*, 1951; Haltiwanger & Maccini, 1988; Khan & Thomas, 2007).

Beyond these explicit explanations, broader literature treats inventory in reduced form. Kydland & Prescott (1982) and Christiano (1988) model inventory as a production factor, while Wen (2011) and Auernheimer & Trupkin (2014) incorporate inventory into the utility functions. The empirical performance of these models isn't fully satisfactory: they often fail to predict correct signs and the great persistence of inventory-related responses.

This paper proposes a new micro foundation to explain inventory behavior. I incorporate search friction (Burdett *et al.*, 2001) into a heterogeneous-agent general equilibrium framework (Aiyagari, 1994). In the model, sellers post prices and inventories, while buyers select which sellers to visit based on a mixed strategy, given the lack of coordination. Consequently, sales are stochastic, but the process is endogenous. Buyers adjust their visiting strategies such that the expected utility of visiting any seller is the same in equilibrium. Thus, sellers in equilibrium offer a uniform shopping experience rather than a uniform price. This setup allows sellers to adjust prices and quantities flexibly. Sellers posting lower prices attract more buyers, increasing the likelihood of sales. At the same time, holding more inventory allows sellers to charge higher prices based on the buyer-seller ratio, directly affecting profit margins.

However, holding additional inventory incurs costs, making market liquidity a concern. If sellers wish to sell their inventory quickly, they must reduce prices. The model is realistic in two key ways: sales are stochastic, and sellers can choose the stochastic process. This involves risks, and sellers with greater capital holdings are better positioned to manage this risk. Sellers with more capital tend to post higher prices, meaning equilibrium features a joint distribution between capital and inventory holdings.

The proposed model uses a parsimonious set of parameters to capture inventory behaviors through clear economic mechanisms. In periods of economic growth, high aggregate capital reduces interest rates, prompting sellers to invest more in inventory, making inventory investment pro-cyclical. Unlike in a Walrasian market, where sellers take prices as given, the model allows sellers to adjust both prices and inventory. As a result, the quantity response of inventory is moderated. The inventory-sales ratio becomes counter-cyclical, while the markup remains pro-cyclical. Given market liquidity concerns, sellers optimize by charging higher prices and holding inventories longer, leading to greater persistence in the model's responses compared to standard models.

In addition to explaining the inventory cycle at the aggregate level, the model aligns with findings from recent micro studies. For example, Cavallo & Kryvtsov (2023) demonstrate that inventory stockouts have inflationary effects, while Kim (2021) shows that capital insufficiency can lead to price cuts and inventory liquidation. Moreover, Kryvtsov & Vincent (2021) find that the frequency of price cuts is counter-cyclical, a prediction supported by the model.

This paper also contributes to the literature on decentralized market models, particularly in the study of pricing and trading mechanisms. [2] In the model, sellers face a trade-off between the price they post and the speed of sales. This mechanism not only produces price dispersion but also allows for fire sales, where some prices fall below production costs—something that occurs in reality but is absent from standard models.

The model adds to the search literature[3] by considering a setting with many-to-one matches, where the differentiation in available quantities influences the equilibrium terms of trade. While Geromichalos (2012, 2014) modeled many-to-one matches in a directed search environment, those models involved ex post contracts and did not account for inventory.

[2]For example, Duffie *et al.* (2007) modeled an over-the-counter market, Chang & Zhang (2018) modeled intermediation and Gale & Kariv (2007) modeled trading network.

[3]Rogerson *et al.* (2005) surveyed labor search and Lagos *et al.* (2017) surveyed money searches.

My model, by contrast, incorporates pre-committed inventory stocks, introducing greater matching friction. Some sellers are left with unsold inventory, while others may run out of stock. Although this approach is more realistic, it sacrifices analytical tractability, requiring the model to be solved numerically. Shevchenko (2004) also models inventory in a search environment, featuring random search and Nash bargaining. The hold-up problem in their model leads to understocking. In contrast, I model directed search, eliminating the hold-up problem, while the quantity margin introduces an externality that causes sellers to overstock on aggregate. This externality is absent in a competitive search equilibrium where only one-to-one matching is considered.

Finally, the model's stochastic process is endogenous, as sellers choose both the level of inventory to hold and the transition process, including pricing decisions and sales probabilities. This provides a specific type of micro-foundation for the idiosyncratic shocks present in heterogeneous-agent models (Huggett, 1993; Aiyagari, 1994) and offers new insights into agent mobility.

The rest of the paper is organized as follows. Section 2 documents the empirical behavior of aggregate inventory. Section 3 sets up the model and discusses the economic mechanisms. Section 4 calibrates the model and evaluates the model's empirical performance. Section 5 concludes.

2 Empirical regularities

One might expect inventory stock to decrease over time due to technological advancements, but this is not the case. Figure 1 shows the inventory-to-GDP ratio using U.S. quarterly data. On average, inventory stock represents about 13% of annual GDP, with no clear trend toward a decline. Given its size, inventory plays a significant role in resource allocation.

Beyond its substantial size, inventory is also closely correlated with business cycles. Fig-

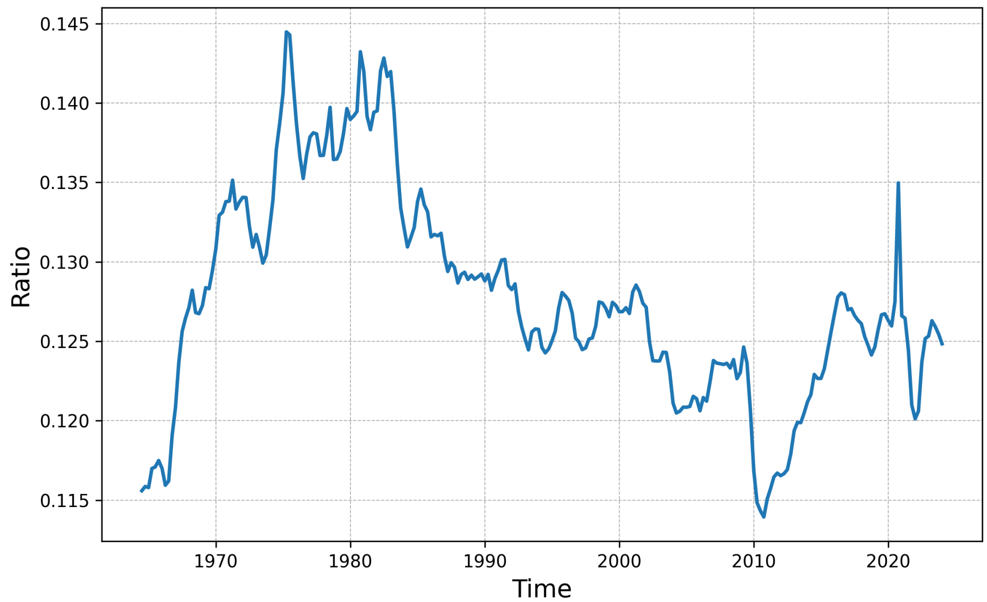

Figure 1: Inventory-GDP ratio, U.S. quarterly 1964–2023

Inventory is the real non-farm private inventory. GDP is the real GDP. Data source: U.S. Bureau of Economic Analysis.

ure 2 compares the growth rates of inventory and GDP. The two time series co-move closely, exhibiting both a positive correlation and similar magnitudes of fluctuation. Inventory is therefore a valuable indicator of business cycles.

To further assess the importance of inventory in business cycles, I compute the descriptive statistics shown in Table 1. The data, sourced from the U.S. Bureau of Economic Analysis and the U.S. Bureau of Labor Statistics, spans from 1964 to 2023 on a quarterly basis. Total output y is represented by real GDP, inventory i is the real non-farm private inventory, and sales s represent the real final sales of domestic products. Aggregate markup $mk*$ is calculated as the ratio of the Producer Price Index (PPI) to wages. [4] I use the log difference to measure cycles and, for comparison with the literature, I also report the HP-filtered cyclical components. The results from these two measures are qualitatively similar and quantitatively close.

[4]Specifically, PPI is the Producer Price Index by Commodity for Final Demand (Finished Goods). I weigh it by the Implicit Price Deflator for GDP to control for inflation. Wages data come from the Average Hourly Earnings of Production and Nonsupervisory Employees (Total Private).

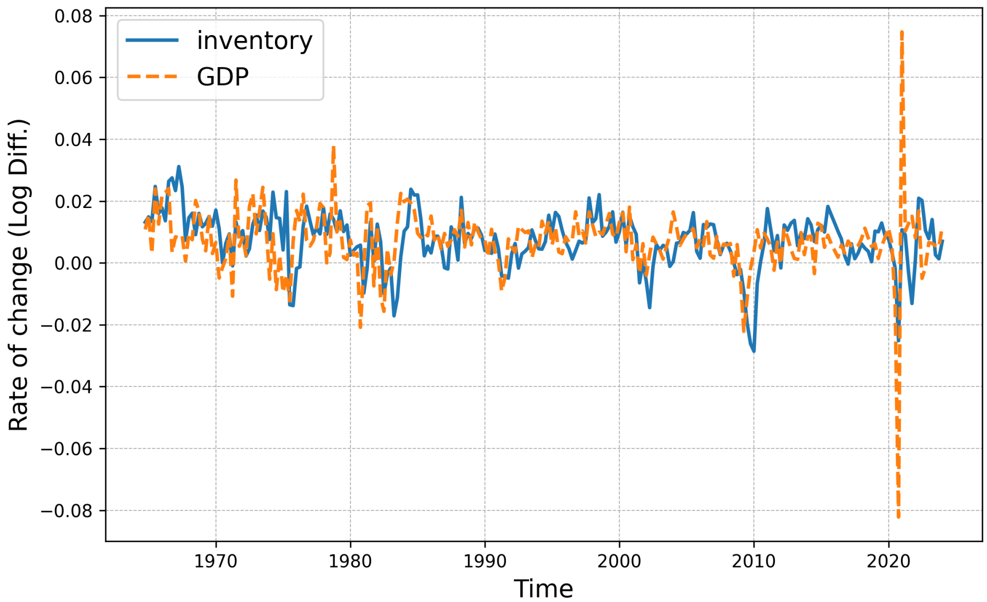

Figure 2: Inventory-GDP comovement, U.S. quarterly 1964–2023.
Both variables are in log difference. Inventory is the real non-farm private inventory. GDP is the real GDP. Data source: U.S. Bureau of Economic Analysis.

The top panel of Table 1 compares the volatility of output, sales, and inventory. Inventory is nearly as volatile as output, while sales are 83% as volatile as output. This observation rejects the production smoothing hypothesis as the sole explanation for inventory holding. If agents were holding inventory solely to smooth production against volatile sales, inventory investment would be less volatile than sales. The bottom panel of Table 1 shows the correlations between variables. The correlation between inventory and output is 0.41, which does not support stock-out avoidance as the sole reason for holding inventory. If stock-out avoidance were the primary motive, inventory levels would decrease as sales increase. Additionally, the inventory-sales (I/S) ratio has a negative correlation of -0.11 with GDP growth. Although inventory is pro-cyclical, it does not rise proportionately with sales, which is puzzling within standard models where agents only adjust quantities.

Lastly, the autocorrelation of the I/S ratio is approximately 0.97, indicating strong persistence. This persistence makes it unlikely that fixed restocking costs alone explain inventory holding. Under a fixed restocking cost model, agents would restock inventory only

Table 1: Descriptive statistics, U.S. quarterly data 1964-2023

		Cycle measures	
Variable	Description	Log diff.	HP-filtered
	Volatility		
$\sigma(y)$	GDP	0.01	0.02
$\sigma(s)/\sigma(y)$	sales to GDP	0.83	0.85
$\sigma(i)/\sigma(y)$	inventory to GDP	0.83	1.18
	Correlation		
$\rho(s, y)$	sales and GDP	0.88	0.96
$\rho(i, y)$	inventory and GDP	0.41	0.55
$\rho(is, y)$	I/S and GDP	-0.11	-0.16
$\rho(is_t, is_{t-1})$	I/S autocorrelation	0.97	0.76
$\rho(mk^*, y)$	markup* and GDP	0.13	0.32

Markup* = PPI/wage

when it falls below a certain threshold and gradually deplete it over time, leading to a less persistent I/S ratio. These raw moments suggest that current inventory models fail to fully account for these dynamics. To examine their interactions further, I employ a structural VAR model.

In the VAR model, I include three variables: sales, inventory, and aggregate markup. All variables are unfiltered and represented in log form to account for potential unit root issues. Since the data is quarterly, I use eight lags and include both constant and trend terms. Structural identification is based on Cholesky decomposition. To analyze the impulse response to a sales shock, I order the variables as sales, inventory, and markup, under the assumption that a sales shock immediately impacts inventory size, while prices and production costs remain fixed for the period. Changing the order of inventory and markup does not qualitatively alter the impulse responses.

Figure 3 plots the impulse response of sales, inventory, I/S ratio, and aggregate markup

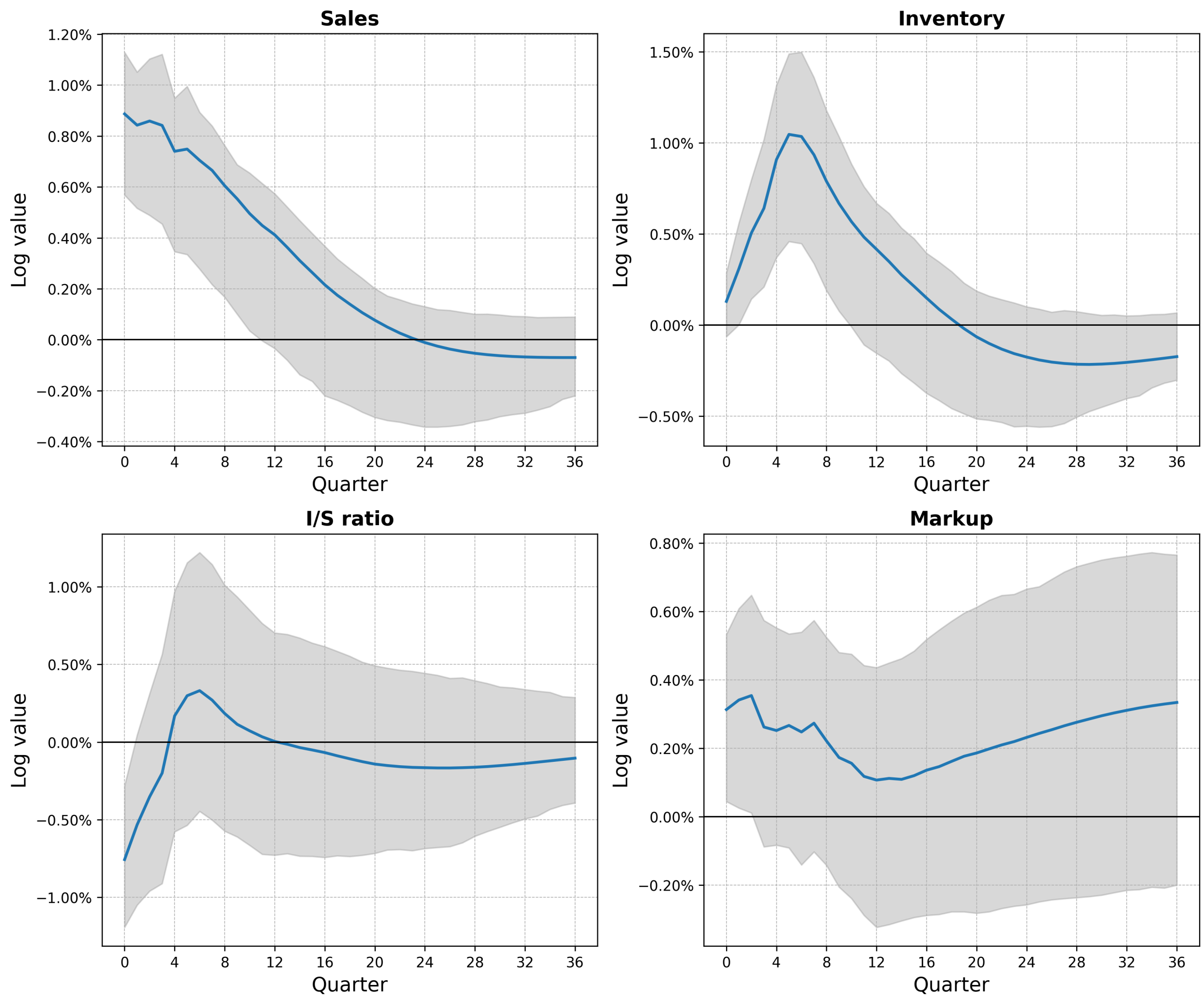

Figure 3: Impulse responses to sales shock, U.S. quarterly 1964–2023

Structural VAR using Cholesky decomposition. Vector order: (sales, inventory, markup). VAR specification: 8 lags with both constant and trend. All variables are in log form and unfiltered.

to the sales shock. The shaded areas depict the 95% confidence intervals. [5] The signs of the responses illustrate the cyclicality to be explained: (1) inventory is pro-cyclical, (2) the I/S ratio is counter-cyclical [6], and (3) aggregate markup is pro-cyclical. Meanwhile, the impulse responses are also very persistent. The signs and general patterns of the impulse responses are robust to HP-filtered cycles. [7] The confidence intervals for the HP-filtered variables are

[5]The confidence intervals are non-cumulative and calculated by bootstraps with 500 runs.

[6]To address the concern of cointegration between inventory and sales, I conduct the Johansen test and rerun the VAR with the constructed inventory-sales relation. The counter-cyclicality persists, as shown in Appendix B.

[7]Appendix A reports the results.

tighter, as the filter removes the smooth trend.

As explained above, existing models have difficulty capturing either the signs of the movements or their persistence. The main goal of this paper is to build a model that explains these trends. The next section proposes the model, and I then calibrate it to data in Section 4. The primary exercise is to compare the impulse responses from the model to those in Figure 3.

3 The model

This section introduces a dynamic general equilibrium model with heterogeneous agents and endogenous idiosyncratic risks.

In this model, time is discrete and continues indefinitely. There are three types of agents: a continuum of sellers with measure 1, a continuum of buyers with measure μ, and a representative firm. In each period, two markets operate sequentially: a Walrasian market (WM) and a search market (SM). The representative firm plays a role only in WM, while the sellers and buyers make decisions in both markets. Theoretically these two markets can operate simultaneously, as we do not degenerate the distributions of capital and inventory. However, sequential operation greatly reduces the computational burden. To further simplify the problem, let the buyers have linear utility, so they don't save and make only static decisions. [8] This simplification eliminates the heterogeneity among buyers and permits the application of market utility, which is beneficial for the many-to-one matching problem introduced later.

The WM operates like a standard Walrasian general equilibrium. Specifically, the representative firm rents capital K_t from the sellers, hires labor L_t from the sellers and buyers, and produces output Y_t with a time-invariant production technology $Y_t = f(K_t, L_t)$ that

[8] The result doesn't change when the buyers make intertemporal decisions with a quasi-linear preference. The point here is to homogenize buyers' capital holdings before they enter the search market.

has constant returns to scale. The markets for Y, K, and L are competitive, so the return to capital is $r_t = f_K(K_t, L_t)$ and the wage rate is $w_t = f_L(K_t, L_t)$. Since $f(\cdot)$ has constant returns to scale, the representative firm makes zero profit in equilibrium. The sellers and buyers then purchase the WM outputs as the numeraire.

The sellers have a technology to convert the WM output Y into the SM goods x. They do not directly consume x but can sell them to buyers for WM output. Before entering SM, the sellers make all the quantity decisions–consumption c_t, direct saving $\hat{k}$, and inventory holding $\hat{x}$ for SM.

After all the quantity decisions in WM, the sellers set the prices in SM. Due to a lack of coordination, the meetings between sellers and buyers are subject to search friction. The sellers decide what price to charge while taking into account the buyers' visiting responses. In the price posting problem, each seller posts a tuple $(\hat{x}, p, n)$, where $\hat{x}$ is the available inventory for sale, p is the price, and n is the buyer-seller ratio. After observing all the posts, each buyer chooses which seller to visit. Note that at this point, the buyers are still homogeneous, allowing market utility to apply. Specifically, the ex ante utility of visiting each seller should be same across all buyers. Denote market utility as J, an equilibrium object. Knowing the processes, we just need to work out the meeting and consumption probabilities to establish the dynamic programming problem.

An advancement of this model is the consideration of many-to-one matching: each buyer only meets a single seller, while each seller can be visited by multiple buyers. In each submarket $(\hat{x}, p, n)$, each buyer visits each seller with equal probability. [9] It follows that the expected number of visits is n for all sellers in submarket $(\hat{x}, p, n)$. Therefore, the probability

[9] This holds true in the unique non-coordinate equilibrium (Galenianos & Kircher, 2012).

π_s of making sales s follows a truncated Poisson distribution.

$$\pi_s(\hat{x}, n) = \begin{cases} \frac{n^s e^{-n}}{s!} & \text{if } s < \hat{x} \\ 1 - \sum_{i=0}^{\hat{x}-1} \frac{n^i e^{-n}}{i!} & \text{if } s = \hat{x} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Note that stock-out avoidance plays a role here, as the probability of sales is truncated beyond the inventory level. Assume the buyers have equal probabilities of consumption if the number of buyers exceeds the inventory holdings. The probability α of a buyer consuming the SM goods is then

$$\begin{aligned} \alpha(\hat{x}, n) &= \sum_{i=0}^{\hat{x}-1} \frac{n^i e^{-n}}{i!} + \sum_{i=\hat{x}}^{\infty} \frac{n^i e^{-n}}{i!} \frac{\hat{x}}{i+1} \\ &= \sum_{i=0}^{\hat{x}-1} \frac{n^i e^{-n}}{i!} + \frac{\hat{x}}{n} \left(1 - \sum_{i=0}^{\hat{x}} \frac{n^i e^{-n}}{i!} \right) \end{aligned} \tag{2}$$

Note that a single buyer's consumption probability depends on the probabilities of how many other buyers show up, so the threshold is at $\hat{x} - 1$. Knowing these key probabilities, we now turn to the preferences and technology before formally defining the equilibrium.

The representative firm has a Cobb-Douglas production function $Y = AK^\gamma L^{1-\gamma}$, where A is the time-invariant TFP and γ is the capital share. In WM, buyers make a labor decision by solving the following optimization problem:

$$\max_{l \in [0,1]} \; wl - \zeta \frac{l^\epsilon}{\epsilon} \tag{3}$$

where w is the wage rate, l is the labor supply, ϵ measures the labor supply elasticity, and ζ alters the level of disutility from working. It follows that the buyers' optimal labor supply is $l^* = \left(\frac{w}{\zeta}\right)^{1/(\epsilon-1)}$ and their WM income is wl^*. In SM, buyers derive utility η from consuming x and gain linear utility from consuming Y. Their optimization problem entails deciding which submarket $(\hat{x}, p, n)$ to search. Since they do not save, buyers consume all remaining income. Their market utility can then be written as

$$J = \max_{(\hat{x},p,n)} \ \alpha(\hat{x}, n)(\eta + wl^* - p) + [1 - \alpha(\hat{x}, n)]wl^* \tag{4}$$

The sellers' problem is dynamic, with two state variables: capital k and inventory x. In WM, they supply capital and earn rk. At the same time, each unsold unit of inventory incurs a cost of δ. To ensure that sellers have enough resources to cover the holding costs, I assume that sellers also supply labor l in WM and earn wl income. The disutility from labor for sellers is the same as that for buyer's. The total available resources for a seller at the beginning of WM is $(1+r)k+wl-\delta x$. In terms of spending, they choose their consumption c and investments in capital $\hat{k}$ and inventory $\hat{x}$ that can be sold in the subsequent SM. In SM, they post the price p and the buyer-seller ratio n, along with the inventory size $\hat{x}$. Assume free disposal and let the cost of producing $\hat{x}$ given x be $\max\{\hat{x} - x, 0\}^\kappa/a$. Note that for $\kappa > 1$, the production cost of SM goods is convex, giving sellers an incentive to smooth production.

Taking into account the buyers' expected utility J, we can write the sellers' SM value $V(\cdot)$ function as

$$V(\hat{k}, \hat{x}; w, J) = \max_{p,n} \ \sum_{s=0}^{\hat{x}} \pi_s(\hat{x}, n) \cdot W(\hat{k} + sp, \hat{x} - s; r', w', J') \tag{5}$$

$$\text{s.t.} \quad J = \alpha(\hat{x}, n)(\eta - p) + wl^* \tag{6}$$

$$p \leq \min\{wl^*, \eta\}$$

where $W(\cdot)$ is the WM value function calculated below.

$$W(k,x;r,w,J) = \max_{c,\hat{k},\hat{x}} \frac{c^{1-\sigma}}{1-\sigma} - \zeta\frac{l^{\epsilon}}{\epsilon} + \beta V(\hat{k},\hat{x};w,J) \tag{7}$$

$$\text{s.t.} \quad c + \hat{k} + C(\hat{x};x) = (1+r)k + wl$$

$$C(\hat{x};x) = \frac{1}{a}\left(\max\{\hat{x}-x,0\}^{\kappa} + \delta x\right)$$

The posted price has two potential upper bounds: buyers' utility η from consuming SM goods and their budget constraint wl^*. With the dynamic programming problem established, we are now ready to define the equilibrium formally.

Definition 1. *A stationary equilibrium is the value functions W and V, market values (r,w,J), aggregate quantities (K,X), policy functions $(\hat{k},\hat{x},p^*,n^*)$, and measure $F_{K,X}$, such that*

1. *Optimality: given (r,w,J), $(p^*,n^*,\hat{k},\hat{x})$, W and V solve (5) and (7).*
2. *Clearing: $r = f_K(K,\mu l^* + \bar{l})$, $w = f_L(K,\mu l^* + \bar{l})$, $K = \int k \,\mathrm{d}F_{K,X}$ and $\mu = \int n^* \,\mathrm{d}F_{K,X}$.*
3. *Stationarity: $F_{K,X}(k,x) = \int \sum_{s=0}^{\hat{x}} \pi_s(\hat{x},n^*) \mathbb{1}\{\hat{k} + sp^* \leq k\} \cdot \mathbb{1}\{\hat{x} - s \leq x\} \,\mathrm{d}F_{K,X}$*

Note that when SM production is costly, it may shut down, resulting in no inventory. For our purposes, we only consider the parameter range when SM is open. Unlike the standard heterogeneous-agent framework, the stochastic process here is endogenous; that is, the sellers choose the supports—-how much inventory to hold—-and the transition probabilities—-what price to charge and the mean sales. This process isn't necessarily ergodic, so the standard proofs for the uniqueness of the stationary distribution do not apply here. As an analytical solution is not available, I solve the model numerically and present the details in the following section.

Though the full characterization of the equilibrium relies on computation, we can observe the key economic forces analytically. The endogenous stochastic process reflects the trade-off between the probability of sales and markup. Hence, sellers face market liquidity when making decisions; to sell faster, they must lower their prices. This relationship can be formally captured by the tightness elasticity of price derived from (6):

$$\lambda(\hat{x}, p, n) \equiv \frac{\mathrm{d}\log p}{\mathrm{d}\log n}\Big|_{(J,w)} = \frac{\alpha_n(\hat{x}, n)n}{\alpha(\hat{x}, n)} \left(\frac{\eta}{p} - 1\right) < 0 \tag{8}$$

Market liquidity λ is defined as the elasticity described above. It can be grouped into two parts. The first part, $\frac{\alpha_n(\hat{x},n)n}{\alpha(\hat{x},n)}$, represents the elasticity of the consumption probability, while the second part, $\left(\frac{\eta}{p} - 1\right)$, reflects the consumer surplus. When the chance of consumption is very elastic or when the surplus is high, sellers must reduce prices drastically to attract more consumers and liquidate inventories quickly. This creates a new incentive to hold inventory, as the elasticity of the consumption probability depends on the inventory holding.

Proposition 1. *In the case of overstock, $\hat{x} \geq n$,*

$$\left|\frac{\alpha_n(\hat{x}+1, n)n}{\alpha(\hat{x}+1, n)}\right| < \left|\frac{\alpha_n(\hat{x}, n)n}{\alpha(\hat{x}, n)}\right| \tag{9}$$

It follows immediately that

$$|\lambda(\hat{x}+1, p, n)| < |\lambda(\hat{x}, p, n)| \tag{10}$$

Holding more inventory improves market liquidity.

Proof. See Appendix C. □

Proposition 1 shows that holding more inventory makes the consumption probability less elastic. Hence, to attract more buyers, the magnitude of the price cut is smaller when inventory holdings are greater. In other words, sellers have an additional incentive to hold inventory, which allows them to post more profitable terms of trade. Note that this incentive arises in the overstock case $\hat{x} \geq n$, where inventory exceeds the expected number of consumers. In the understock case $\hat{x} < n$, the sign can be ambiguous. Nonetheless, the calibrated model shows that sellers overstock in equilibrium, which aligns with our daily observations.

On the other hand, we can immediately see that $\frac{\partial|\lambda(\hat{x},p,n)|}{\partial p} < 0$. The concern for market liquidity diminishes when prices are higher, so sellers who can charge higher prices worry less about market liquidity at the margin. Such complementarity has equilibrium consequences. In the model, sellers make portfolio decisions between capital and inventory. Sellers with more capital can tolerate a higher risk of low sales and thus tend to post higher prices. A higher price alleviates the market liquidity concern, which encourages them to post even higher prices. Therefore, the wealth level affects individual profitability; wealthier agents are also more profitable. This mechanism can generate a heavy tail in the wealth distribution, which poses an empirical puzzle for the Walrasian framework. In the Walrasian framework, the law of one price governs profitability for all agents. Given the uniform profitability across all wealth levels, wealth accumulation solely depends on the diminishing marginal benefit of consumption, and agents have no incentive to maintain a high wealth level. In my model, agents' profitability is wealth-dependent: the richer have better profitability. Even though the agents are ex ante homogeneous, in the stationary equilibrium, some agents may cluster at a high-wealth level in the distribution, as seen in the numerical solution.

Before turning to the numerical exercise, let me elucidate how the model helps to explain the cyclical behavior of inventory.

Proposition 2. *Given the distribution of posting strategies $G(x,p)$ that delivers market utility J, a change in buyer utility $\tilde{\eta} > \eta$ leads to $\tilde{J} > J$.*

Moreover, the new equilibrium tightness satisfies $\frac{\alpha(x_1,\tilde{n}_1)}{\alpha(x_1,n_1)} < \frac{\alpha(x_2,\tilde{n}_2)}{\alpha(x_2,n_2)}$ for any $p_1 > p_2$ for all x_1, x_2 in the support.

It follows that $\int p\mathbb{E}_s[s\pi_s(x, \tilde{n}_{x,p})]\,\mathrm{d}G(x,p) > \int p\mathbb{E}_s[s\pi_s(x, n_{x,p})]\,\mathrm{d}G(x,p)$.

Proof. See Appendix D. □

Proposition 2 considers a sales shock generated by a utility change. When buyers experience a positive utility shock, they adjust their visiting behavior so that the new market utility condition holds, i.e., all submarkets deliver the same expected utility. Since inventories and prices are pre-committed by the sellers, the only adjustable margin is market tightness. As buyers derive more utility from consumption, they tolerate higher prices and gravitate toward submarkets that provide a higher consumption probability. Consequently, aggregate sales increase.

The sales shock essentially increases the total capital in the economy. Hand-to-mouth buyers would consume the output if they do not receive the SM goods. The increasing sales convert otherwise consumed output into sellers' capital stock. With more capital available, the interest rate decreases. Capital investment becomes less attractive, while inventory investment increases. Thus, inventory investment is pro-cyclical.

At the same time, sellers also consider the price margin. Higher inventory improves market liquidity, as shown in Proposition 1. Therefore, sellers charge higher prices, which reduces buyers' visits and, in turn, hinders the incentive to hold excessive inventory. This leads to two immediate implications. Although inventory increases, it does not keep pace with the rise in sales, resulting in a counter-cyclical inventory-sales ratio. On the other hand, posting higher prices implies that markup is pro-cyclical.

In short, the portfolio decision induces pro-cyclical inventory, while the additional price margin leads to a counter-cyclical I/S ratio and pro-cyclical markup. In the numerical exercise below, I compare the model's responses to the VAR responses.

4 A numerical solution

I first describe the numerical algorithm used to find a stationary equilibrium. I then calibrate the model parameters to match the long-run average moments. Following this, I report on the steady state and dynamic responses.

4.1 Algorithm

The numerical algorithm benefits from the sequential solutions of (9) and (10). We define a grid on the space of state variables (k, x), where k is continuous and x is discrete by the model setup. The computation loop is as follows.

(I) Guess a pair (r^0, J^0). Use the CRS property of $f(K, L)$ to obtain w^0 and K^0.

(II) Guess an initial value function V^0 on the grid of (k, x)

  (i) For each $(\hat{k}, \hat{x})$, grid search for optimal (p, n). Note that given J, p and n are bijective, so we only search over n. This provides us the optimal value $\hat{V}^0$ over $(\hat{k}, \hat{x})$, along with the policy functions for p and n.

  (ii) For each (k, x), grid search for optimal $(\hat{k}, \hat{x})$ using $\hat{V}^0$. This gives a new value function V^1 and the policy functions for $(\hat{k}, \hat{x})$.

  (iii) If V^0 and V^1 are distant, replace V^0 with V^1 and iterate to (i). If V^0 and V^1 are close, exit the inner loop.

(III) State an initial distribution F^0 over grids (k, x).

  – Use the policy functions from (II) to iterate the distribution until the convergence reaches some F^1.

(IV) Use F^1 to compute the average market tightness $\hat{\mu} = \int n^* \, \mathrm{d}F$. If $\hat{\mu} > \mu$ $(\hat{\mu} < \mu)$, increase (decrease) J^0 and return to (II). If $\hat{\mu}$ and μ are close, go to the next step.

(V) Use F^1 to compute the aggregate capital K. If $K > K^0$ ($K < K^0$), decrease (increase) r^0 and go to (II), unless K and K^0 are close.

4.2 Calibration

One concern when solving the model numerically is that different initial distributions might converge to different steady states. [10] Although different initial distributions do not affect the impulse responses qualitatively, the non-uniqueness adds an element of ad-hocness to the calibration. The following results initiate the iteration with a uniform distribution.

Table 2: Model calibration

Parameter		Target	Data	Model
		Non-equilibrium object		
A	0.03	normalization	-	-
σ	2.0	risk aversion	-	-
γ	0.33	capital share	0.3 - 0.4	0.33
ϵ	1.53	labor supply elasticity	1.90	1.90
		Equilibrium object		
β	0.992	annual interest rate	0.03	0.03
ζ	0.033	labor hour	0.51	0.51
κ	1.6	inventory/GDP	0.14	0.16
δ	0.03	sales/GDP	0.99	0.99
η	0.04	consumption/GDP	0.65	0.68
μ	2.6	wage/GDP	0.46	0.47
a	220	markup	1.2 - 2.3	1.47

Table 2 reports the calibrated parameters and compares the model moments to the target moments. The parameters in the top panel are set without computing the equilibrium. I normalize the TFP A to 0.03 and choose a preference parameter σ of 2.0, which is a typical value in the business cycle literature. The parameters γ and ϵ are chosen to directly match the capital share and labor supply elasticity, respectively. The parameters in the bottom panel target equilibrium objects and require a joint search. Overall, the model matches the

[10] The steady states are not dense, so a marginal change in the initial distribution does not affect the outcome.

equilibrium objects to the data moments quite well.

One thing to notice is that the calibration of β differs from the typical macro calibration. Usually β alone pins down the steady-state interest rate, $1 = \beta(1+r)$, because saving capital is the only technology for moving resources across time. Here, the sellers can invest in another technology: the SM goods. The optimal portfolio decision equates the marginal return of capital to the marginal return of the SM goods, with adjustments for sales risks. In terms of the equilibrium interest rate, the advancement of the new technology is significant. As a result, $1 \neq \beta(1 + r)$ in general. The more advanced the SM technology is, the higher the steady-state interest rate will be. In the calibrated steady state, $1 > \beta(1 + r)$.

Most of the calibration targets are straightforward, except for the markup. The magnitude of aggregate markup is still understudied. Traditionally, estimation assumes that the source of markup is market power (Berry *et al.*, 1995; De Loecker *et al.*, 2020). However, the model here operates in a competitive environment where search friction generates the markup. Different estimates of markup can affect the model parameters; nonetheless, the impulse responses are robust to different sets of parameters.

4.3 The steady state

I report the steady state for the calibrated parameters below. As shown in Figure 4, the value functions are increasing in both capital and inventory, given the free disposal. Figure 5 plots the WM policy functions. In the left panel, the policy functions for inventory are increasing in capital holdings, meaning that larger sellers produce more SM goods. In the right panel, the policy functions for capital are not monotonically increasing, as the sellers invest in inventory once they accumulate enough capital.

What's new in this paper is the policy functions of price posting, displayed in Figure 6. In equilibrium, all sellers provide the same expected utility to the buyers. Given a buyer-seller ratio, the more inventory one seller holds, the more likely a buyer is to consume, and

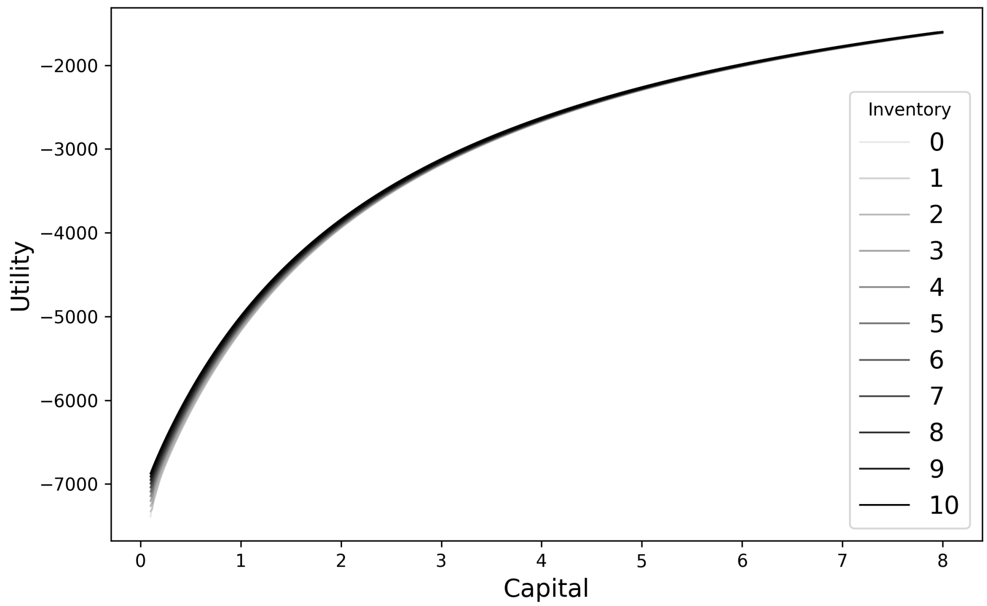

Figure 4: Value function

The value functions across different inventories are close. Lower inventory yields lower value function.

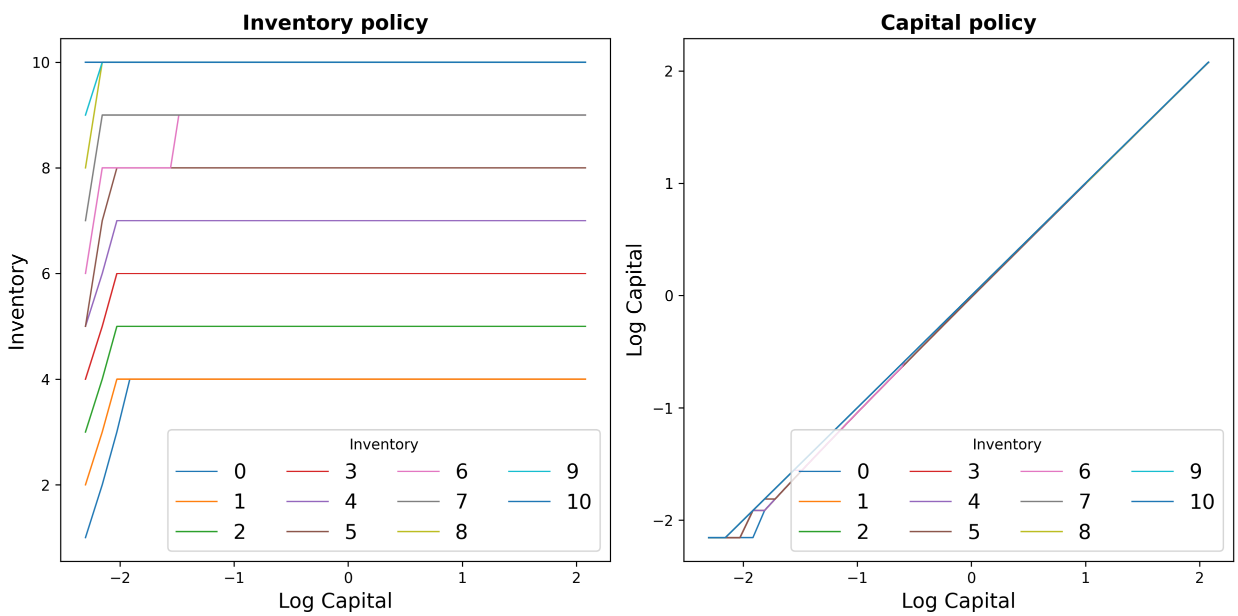

Figure 5: WM policies: inventory and capital

The differences across different inventory levels concentrate at the low capital level. Therefore, the x-axis is log transformed.

the higher the price should be. As Proposition 1 suggests, higher inventory improves market liquidity, as shown by a flatter indifference curve for greater inventory. The right panel of

Figure 6 displays the actual price-tightness pairs that have positive mass in equilibrium.

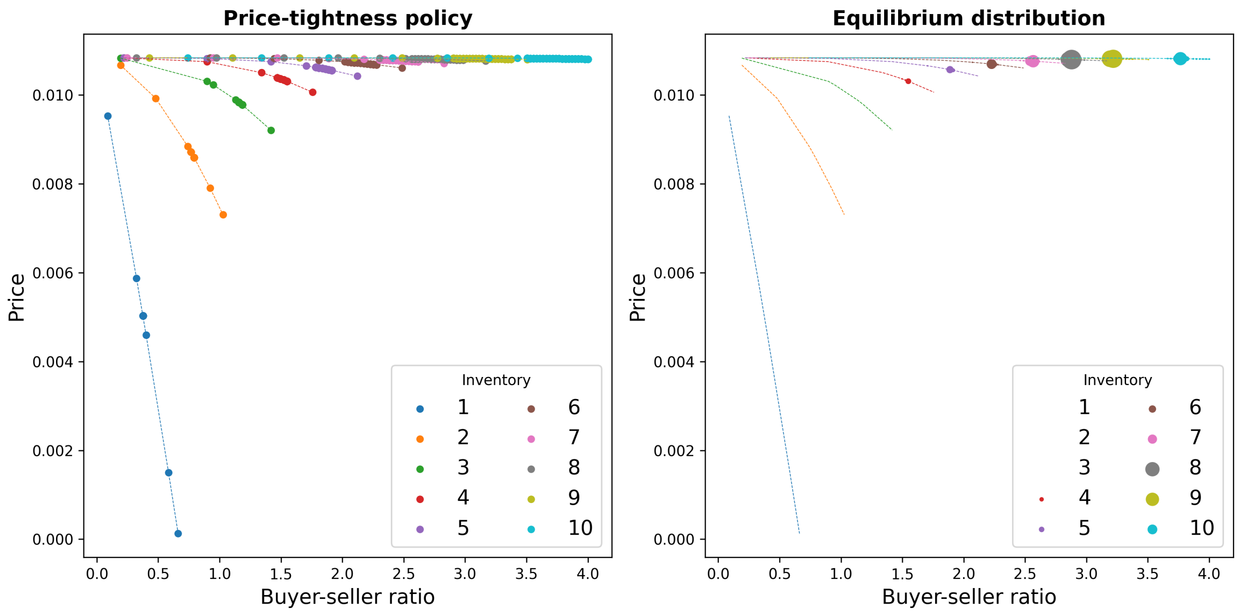

Figure 6: SM policies: price and tightness

The dots present the optimal responses. The dashed lines are the isoquants for the market utility condition. The left panel plots the policies for all states. The right panel plots the policies with positive mass in equilibrium. Bigger dots represent greater equilibrium mass.

The stationary distribution of inventory and capital are shown in Figure 7. The pre-trade distributions depict the state heterogeneity at the beginning of the SM. After the trades in SM happen, the inventory distribution shifts to the left and the capital distribution shifts to the right, shown as the post-trade distributions.

With the ex post heterogeneity in inventory and capital, the model naturally generates price dispersion. Figure 8 plots the price distribution, with the highest price in the support normalized to 1. The degree of price dispersion isn't as large as the retail data suggests (Kaplan & Menzio, 2015). Nonetheless, the implication that smaller shops have a higher variance in price (Figure 6) is reasonable. The limited variances in capital and price dispersion call for greater heterogeneity.

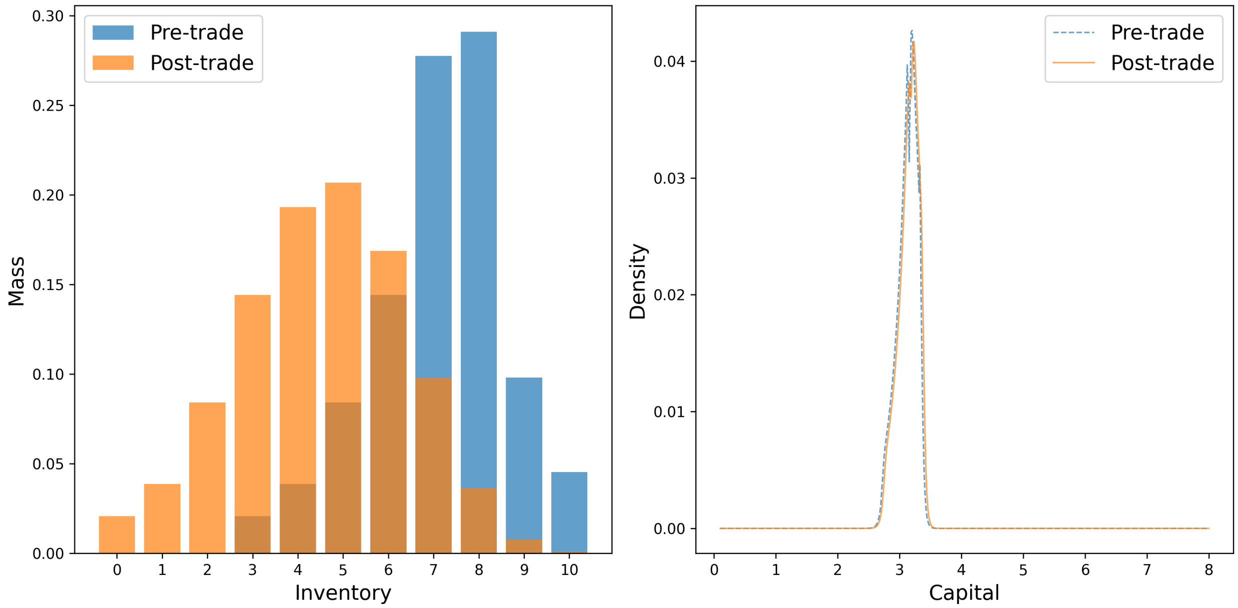

Figure 7: Marginal distributions

The distributions are from the stationary equilibrium. The left panel plots the marginal distribution for inventory. The right panel plots the marginal distribution for capital. The pre-trade state is after posting before trading. Appendix F reports the joint distributions.

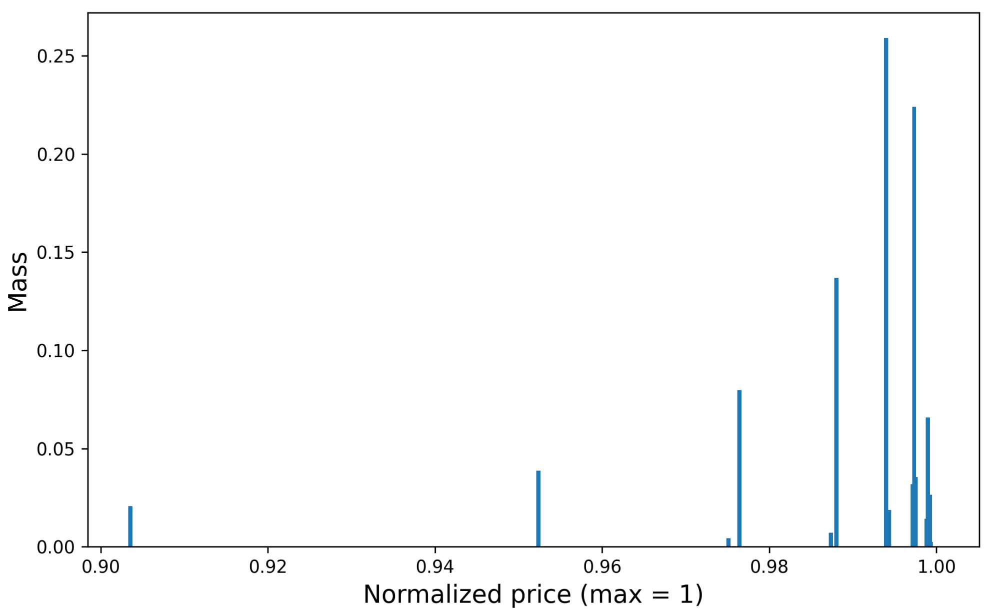

Figure 8: Stationary distribution: price

The graph depicts the price distribution. The normalization applies a constant multiplier to all prices, with the highest price being normalized to 1.

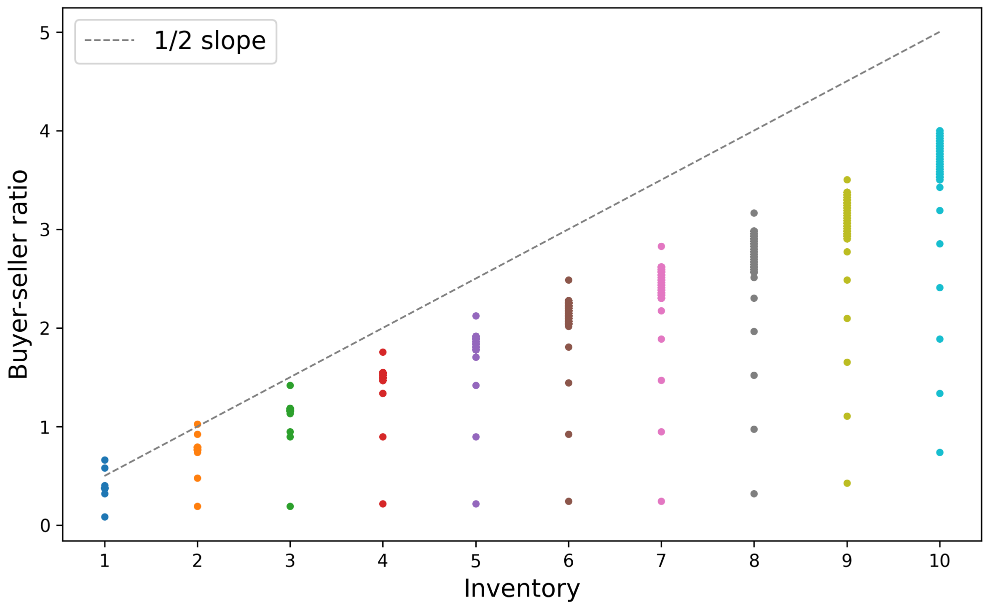

Figure 9: Overstock: tightness vs. inventories

The dots represent the equilibrium buyer-seller ratio across different inventory levels. For any given inventory level, the variation of buyer-seller ratio comes from the capital dispersion. More capital is associated with higher buyer-seller ratio. The dashed line has a slope of 0.5.

4.4 Efficiency

Figure 9 plots the buyer-seller ratio for each inventory level. The dashed line has a slope of 0.5, depicting the threshold where the inventory is sufficient to serve twice the expected number of customers. The first thing to note is that all sellers overstock—the chosen buyer-seller ratios are well below the inventory levels. Moreover, most sellers hold enough inventory to serve more than double the expected number of buyers. Are the inventory levels efficient?

Given the nature of search frictions, the constrained optimal inventory level can be determined by solving the following optimization problem. [11]

$$x^* = \max_x \ \mu\eta\alpha(x,\mu) - a^{-1}\sum_{s=0}^{x}\pi_s(x,\mu)\left[s^\kappa + \delta(x-s)\right] \tag{11}$$

Holding more inventory increases the consumption probability but incurs higher maintenance

[11]The urn-ball matching technology ensures that the optimal market structure consists of a single submarket, where the market tightness is determined by the aggregate buyer-seller ratio.

costs. The actual level of the constrained optimum depends on the specific parameters.

In the calibrated economy, the optimal inventory level is $x^* = 7$. However, as shown in the left panel of Figure 7, only 28% of sellers stock exactly 7 units. Among the remaining sellers, 44% stock 8 or more units, while 28% stock 6 or fewer units. Overall, the economy overstocks by 1.6%, which is equivalent to 0.2% of annual GDP.

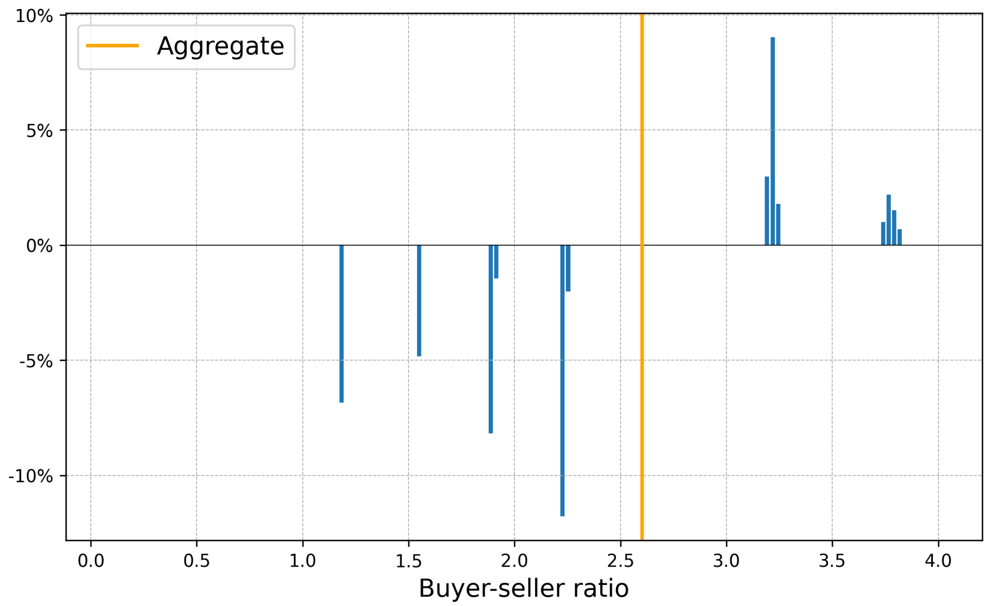

Figure 10: Over/under-stock heterogeneity

The graphs depicts the percentage of over/under-stock given the equilibrium tightness, weighted by the distribution density.

The actual welfare loss consists of two components, as illustrated in Figure 10. The first component is over- or under-stocking given market tightness, resulting in a 0.06% welfare loss relative to the constrained optimum. The second component is the dispersion in market tightness caused by capital heterogeneity, which accounts for a 0.49% welfare loss. Together, these contribute to a total welfare loss of 0.55%, equivalent to 0.27% of annual GDP.

Note that larger sellers tend to overstock, while smaller sellers understock. Therefore, to restore efficiency, the policy implication is to implement a progressive tax. A progressive tax can yield two positive effects: it improves inventory efficiency across different seller sizes and enhances the size distribution simultaneously. This paper focuses on understanding inventory dynamics, leaving policy design for future work.

4.5 Impulse responses

The main exercise of this paper is to compare the impulse responses from the model to those from the data (Figure 3). To introduce a sales shock to the model, I inject a one-time shock to η, the buyers' utility for the SM goods. Specifically, the new utility $\tilde{\eta} = 0.42$ is set to match the magnitude of the initial sales jump in the VAR impulse responses. The timing of the shock occurs after the sellers' price posting and before the buyers' search. Since the posted inventories and prices are fixed, buyers alter their search strategies to yield a new market utility $\tilde{J}$ and a new tightness distribution $\tilde{G}$, as stated in Proposition 2.

To obtain the initial distribution of F_0 after the shock, I iterate $\tilde{J}$ until the average buyer-seller ratio equals the actual buyer-seller ratio μ. For the impulse responses, I first guess two sequences of interest rates r_t and market utility J_t for $t = 0, 1, ..., T$, where T is the final period. I then backward-compute all the optimal responses $(\hat{k}_t, \hat{x}_t, n_t, p_t)$, using these optimal responses to forward-compute the evolution of the joint distributions F_t. The distribution path yields two new sequences of r'_t and J'_t. I repeat the iteration and look for convergence in J_t and r_t, increasing T as needed.

Figure 11 compares the results. The impulse responses from the models match those from the VAR reasonably well. The directions of the initial movements are the same, and the responses are very persistent, although the magnitudes of the changes are not exact. Sales drop too quickly, while inventory and markup do not respond sufficiently. Nonetheless, the model predictions mostly remain within the 95% confidence interval of the VAR results.

The model performs well given the parsimonious parameterization. More importantly, all movements are driven by economic mechanisms, highlighting the significance of price adjustment. With this additional decision margin, agents' quantity responses align more closely with the data. The model allows agents to choose quantities and prices by replacing

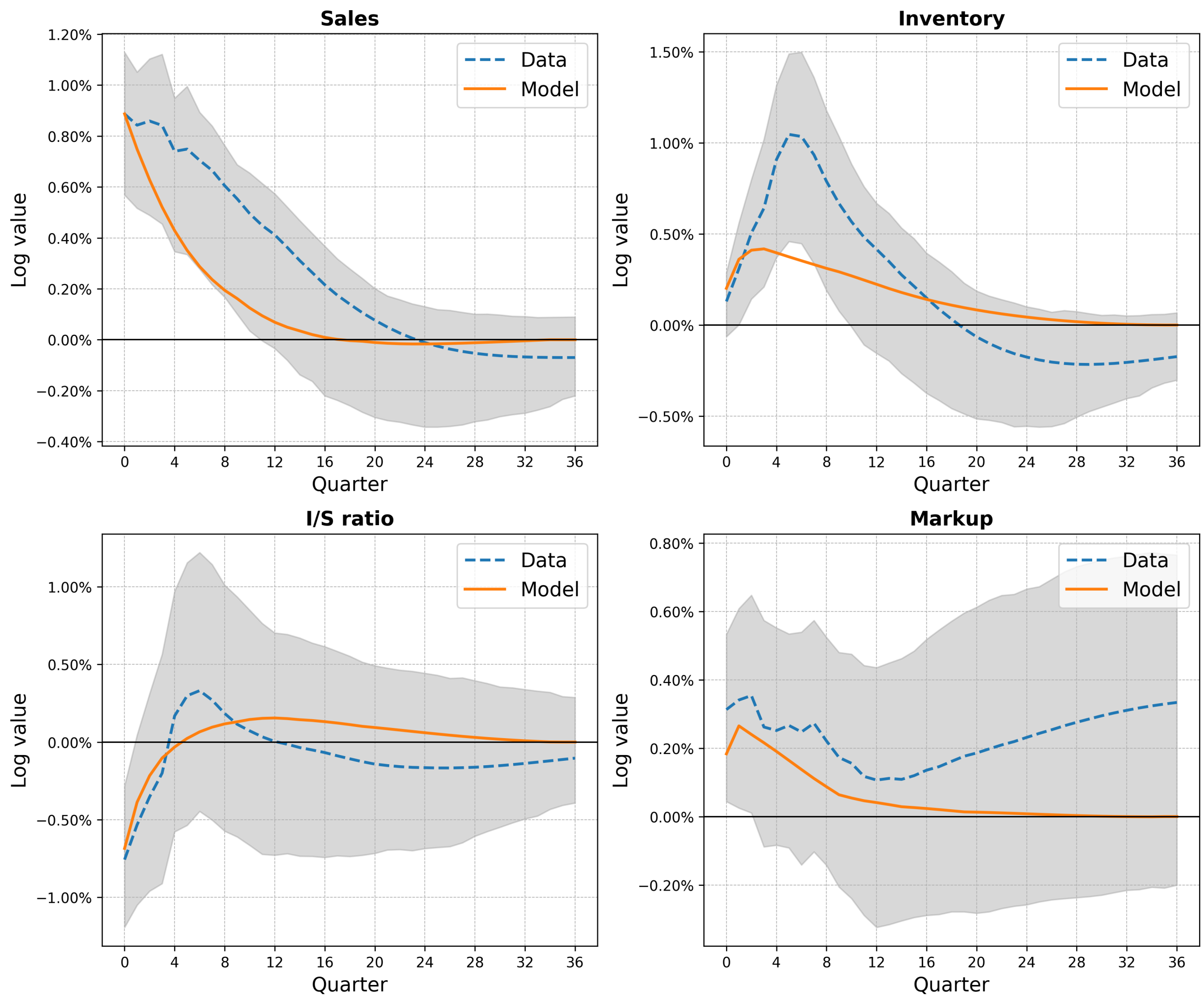

Figure 11: Model responses vs. VAR responses

A comparison between the equilibrium model and the VAR model. Both models construct a stationary relation among the variables. The impulse responses are induced by some one-time unexpected sales shocks. The VAR shock is a one standard deviation shock. The shock to the equilibrium is chosen to match the VAR shock. All variables are in log form.

the market-clearing condition with the market utility condition. Figure 14 in Appendix E reports the impulse responses for other major variables, with aggregate responses similar to those in a standard model.

5 Conclusive remarks

Despite the known importance of inventory to the business cycle, the literature typically overlooks it in its analyses. Following earlier studies, I document three persistent empirical regularities of inventory: it is pro-cyclical, the inventory-sales ratio is counter-cyclical, and markup is pro-cyclical.

Two questions remain: Why do sellers hold inventory, and what motives can explain their inventory behavior? To answer these questions, this paper rationalizes inventory holding through search friction and explains the empirical regularities of inventory via market liquidity, the trade-off between markup and the speed of sales. In the calibrated equilibrium, sellers hold enough inventory for more than twice the expected number of buyers, as holding more inventory improves buyers' shopping experience, allowing sellers to charge higher prices. This new motive for holding inventory helps explain the inventory cycle. When sales increase, sellers not only stock more inventory but also charge higher prices. Since price adjustments are permitted, the quantity responses are not as large as in the Walrasian case, where sellers cannot adjust prices. Incorporating the price margin achieves two goals: the inventory-sales ratio becomes counter-cyclical, and markup becomes pro-cyclical. Obtaining the correct signs for quantity and price simultaneously is not trivial. For example, the New Keynesian framework typically relies on nominal rigidity as a key transmission mechanism, which often implies a counter-cyclical markup.

In addition to the directions of impulse responses, the model also shows potential regarding the persistence of shock effects. This persistence does not rely on the shock process; instead, it is driven by the model mechanism. High inventories sustain high prices, while high prices incentivize high inventories. This complementarity prolongs the effect of a one-time shock. Given the improved predictions in terms of direction and persistence, it is promising to extend the model to examine business cycles in general by incorporating aggregate shocks.

However, some shortcomings of the model call for further study. For example, the predicted wealth inequality and price dispersion are insignificant compared to the data. The model requires more heterogeneity. One way to achieve this is by adding a credit market. In the current model, sellers must accumulate capital themselves, leading the wealth distribution to move away from zero. In reality, wealth distribution spikes at zero as individuals rely on credit for economic activities. To match the wealth inequality observed in the data, incorporating a credit market is necessary. Additionally, greater wealth dispersion will generate greater price dispersion in the model.

Lastly, this paper provides a framework to handle many-to-one matching, which may enhance our understanding of other markets. A close example is financial assets, where market liquidity affects prices and volumes. Another example is the labor market. In the labor search literature, matching is mostly one-to-one, with vacancies being independent. In other words, we lack a comprehensive definition of a firm. Modeling many-to-one matching is a starting point for conceptualizing a firm. However, the inventory model has clear limitations. Investors commonly hold a portfolio of financial assets, firms often post multiple complementary vacancies, and stores typically sell varieties of substitutes. The inventory model in this paper considers only a special case of perfect substitutes. Understanding allocations and prices together remains a challenge that requires further research.

References

Aiyagari, S Rao. 1994. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, **109**(3), 659–684.

Arrow, Kenneth J, Harris, Theodore, & Marschak, Jacob. 1951. Optimal inventory policy. *Econometrica: Journal of the Econometric Society*, 250–272.

Auernheimer, Leonardo, & Trupkin, Danilo R. 2014. The role of inventories and capacity utilization as shock absorbers. *Review of Economic Dynamics*, **17**(1), 70–85.

Berry, Steven, Levinsohn, James, & Pakes, Ariel. 1995. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Bils, Mark, & Kahn, James A. 2000. What inventory behavior tells us about business cycles. *American Economic Review*, **90**(3), 458–481.

Blinder, Alan S. 1986. *More on the speed of adjustment in inventory models.*

Blinder, Alan S, & Maccini, Louis J. 1991. Taking stock: a critical assessment of recent research on inventories. *Journal of Economic perspectives*, **5**(1), 73–96.

Burdett, Kenneth, Shi, Shouyong, & Wright, Randall. 2001. Pricing and matching with frictions. *Journal of Political Economy*, **109**(5), 1060–1085.

Cavallo, Alberto, & Kryvtsov, Oleksiy. 2023. What can stockouts tell us about inflation? Evidence from online micro data. *Journal of International Economics*, **146**, 103769.

Chang, Briana, & Zhang, Shengxing. 2018. Endogenous market making and network formation. *Available at SSRN 2600242*.

Christiano, Lawrence J. 1988. Why does inventory investment fluctuate so much? *Journal of Monetary Economics*, **21**(2-3), 247–280.

Coen-Pirani, Daniele. 2004. Markups, aggregation, and inventory adjustment. *American Economic Review*, **94**(5), 1328–1353.

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

Duffie, Darrell, Gârleanu, Nicolae, & Pedersen, Lasse Heje. 2007. Valuation in over-the-counter markets. *The Review of Financial Studies*, **20**(6), 1865–1900.

Gale, Douglas M, & Kariv, Shachar. 2007. Financial networks. *American Economic Review*, **97**(2), 99–103.

Galenianos, Manolis, & Kircher, Philipp. 2012. On The Game-Theoretic Foundations Of Competitive Search Equilibrium. *International economic review*, **53**(1), 1–21.

Geromichalos, Athanasios. 2012. Directed search and optimal production. *Journal of Economic Theory*, **147**(6), 2303–2331.

Geromichalos, Athanasios. 2014. Directed search and the Bertrand paradox. *International Economic Review*, **55**(4), 1043–1065.

Haltiwanger, John C, & Maccini, Louis J. 1988. A model of inventory and layoff behaviour under uncertainty. *The Economic Journal*, **98**(392), 731–745.

Huggett, Mark. 1993. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of economic Dynamics and Control*, **17**(5-6), 953–969.

Kahn, James A. 1987. Inventories and the volatility of production. *The American Economic Review*, 667–679.

Kaplan, Greg, & Menzio, Guido. 2015. The morphology of price dispersion. *International Economic Review*, **56**(4), 1165–1206.

Khan, Aubhik, & Thomas, Julia K. 2007. Inventories and the business cycle: An equilibrium analysis of (S, s) policies. *American Economic Review*, **97**(4), 1165–1188.

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

Kryvtsov, Oleksiy, & Midrigan, Virgiliu. 2012. Inventories, markups, and real rigidities in menu cost models. *Review of Economic Studies*, **80**(1), 249–276.

Kryvtsov, Oleksiy, & Vincent, Nicolas. 2021. The cyclicality of sales and aggregate price flexibility. *The Review of Economic Studies*, **88**(1), 334–377.

Kydland, Finn E, & Prescott, Edward C. 1982. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 1345–1370.

Lagos, Ricardo, Rocheteau, Guillaume, & Wright, Randall. 2017. Liquidity: A new monetarist perspective. *Journal of Economic Literature*, **55**(2), 371–440.

Metzler, Lloyd A. 1941. The nature and stability of inventory cycles. *The Review of Economics and Statistics*, **23**(3), 113–129.

Nekarda, Christopher J, & Ramey, Valerie A. 2013. *The cyclical behavior of the price-cost markup*. Tech. rept. National Bureau of Economic Research.

Ramey, Valerie A, & West, Kenneth D. 1999. Inventories. *Handbook of macroeconomics*, **1**, 863–923.

Rogerson, Richard, Shimer, Robert, & Wright, Randall. 2005. Search-theoretic models of the labor market: A survey. *Journal of economic literature*, **43**(4), 959–988.

Shevchenko, Andrei. 2004. Middlemen. *International Economic Review*, **45**(1), 1–24.

Wen, Yi. 2011. Input and output inventory dynamics. *American Economic Journal: Macroeconomics*, **3**(4), 181–212.

West, Kenneth D. 1990. The sources of fluctuations in aggregate inventories and GNP. *The Quarterly Journal of Economics*, **105**(4), 939–971.

Appendix A VAR using HP-filtered variables

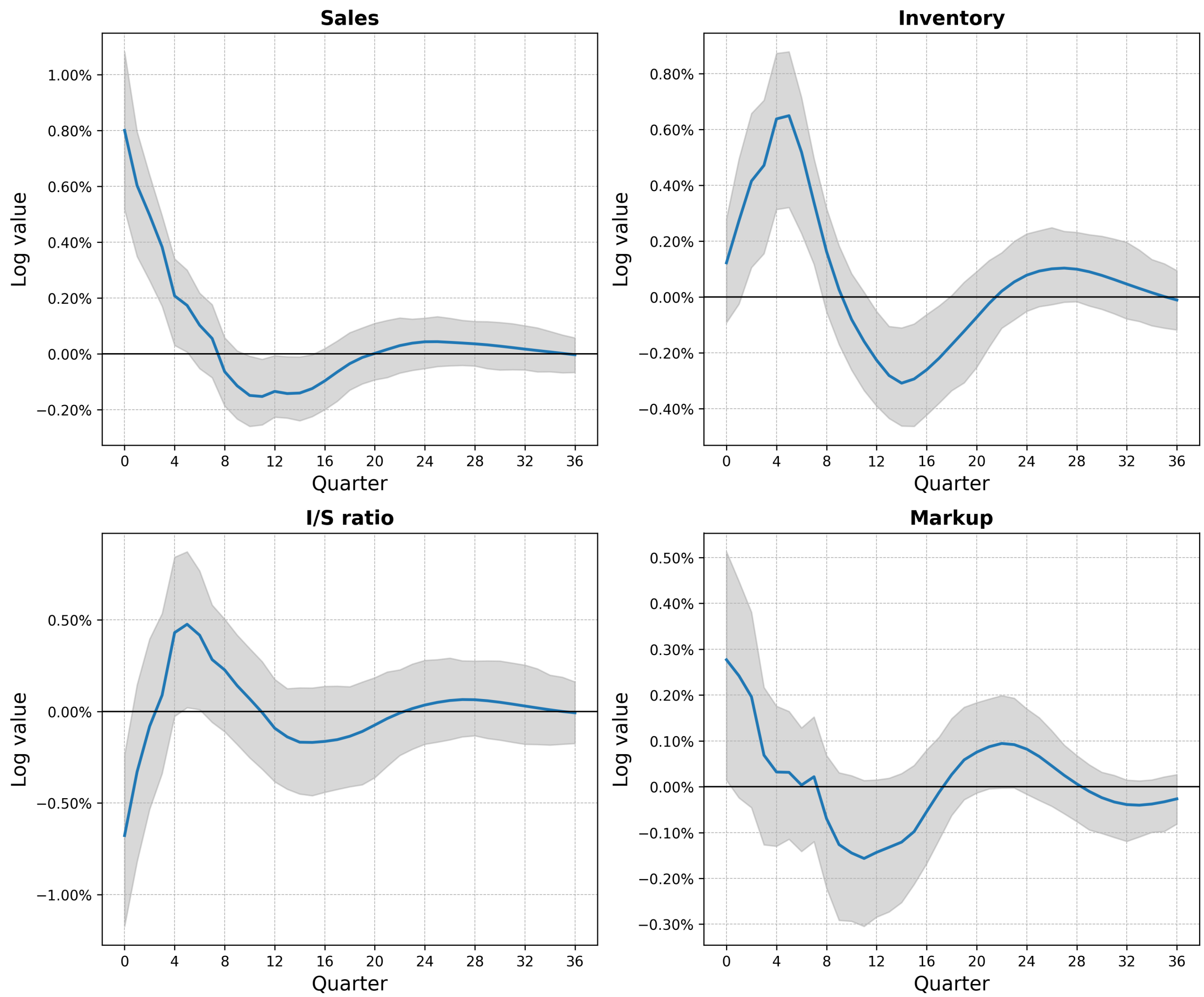

Figure 12: Impulse responses to sales shock, 1964 – 2023

Structural VAR using Cholesky decomposition. Vector order: (sales, inventory, markup). VAR specification: 8 lags with both constant and trend. All variables are in log form and HP-filtered.

This section reports the VAR results from the HP-filtered variables. Figure 12 plots the results where the shaded areas represent the 95% confidence interval. The directions of movements are same as using the selected sample. Inventory is pro-cyclical, I/S ratio is counter cyclical and aggregate markup is pro-cyclical. The confidence intervals cover origin in long-run as using the selected sample.

Appendix B Cointegration test

For the concern of cointegration, I consider a stationary relation between log inventory and log sales $i_t - \theta s_t$. To estimate θ and test the cointegration, I use Johansen procedure with eigenvalue method and long-run VECM for error correction. Table 3 reports the results.

Table 3: Johansen test

	Unfiltered		HP-filtered		Critical value		
Lag	$\hat{\theta}$	Test stat.	$\hat{\theta}$	Test stat.	0.10	0.05	0.01
2	0.89	19.49	1.45	72.16	12.91	14.90	19.19
4	0.90	15.95	1.41	80.52	12.91	14.90	19.19
8	0.92	14.46	2.67	40.55	12.91	14.90	19.19

[a] All variables are in log form. [b] Test with eigenvalue
[c] Error correction: long-run VECM

$\hat{\theta}$ is around 0.9 for the unfiltered cycle, robust to the lag lengths from 2 to 8. The HP-filtered cycle has $\hat{\theta}$ ranges from 1.4 to 2.7. In either setting, we reject the null of no cointegration at the 0.10 level. Most cases also reject the null at the 0.05 level, except the case with log difference and lag length 8. The test statistic 14.46 is very close to the critical value 14.90. Overall, these results suggest that inventory and sales are cointegrated.

To deal with the cointegration, I construct the inventory-sales (I/S) relations using the estimated $\hat{\theta}$ for the two cycle measurements. Figure 13 plots the corresponding impulse responses. The inventory-sales relation is also counter-cyclical. Replacing inventory-sales ratio by the cointegrated relation doesn't change the empirical regulation.

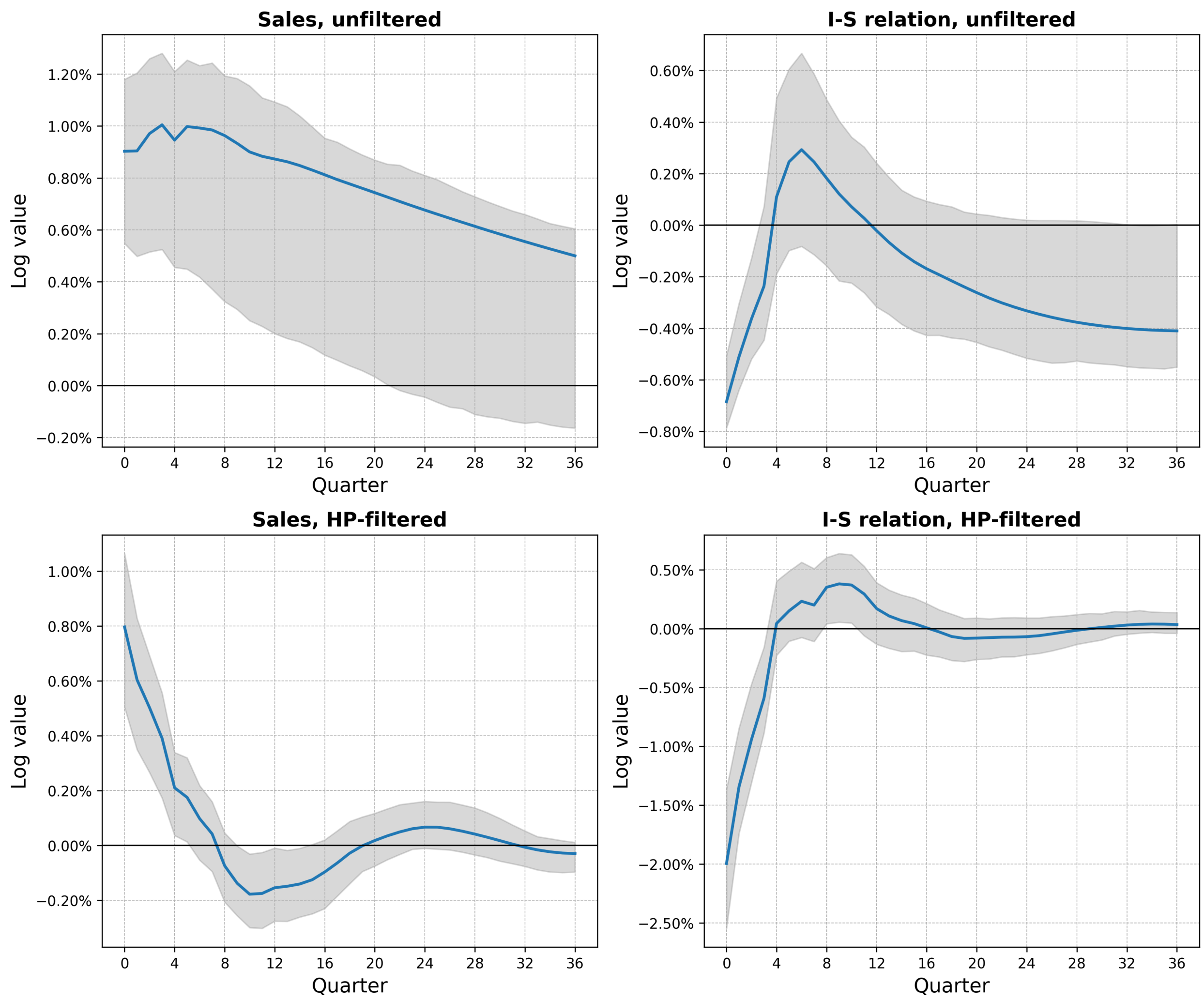

Figure 13: Impulse responses to sales shock, 1964 – 2023

Structural VAR using Cholesky decomposition. Vector order: (sales, inventory-sales relation). VAR specification: 8 lags with both constant and trend. All variables are in log form. The variables in the top panels are unfiltered. The variables in the bottom panels are HP-filtered.

Appendix C Proof of Proposition 1

The calculation below shows how inventory holding affects the elasticity of the consumption probability. From

$$\alpha(\hat{x},n) = \sum_{i=0}^{\hat{x}-1} \frac{n^i e^{-n}}{i!} + \sum_{i=\hat{x}}^{\infty} \frac{n^i e^{-n}}{i!} \frac{\hat{x}}{i+1} \tag{12}$$

we can calculate the derivative

$$\begin{aligned} \alpha_n(\hat{x},n) &= \sum_{i=0}^{\hat{x}-1} \frac{in^{i-1}e^{-n} - n^i e^{-n}}{i!} + \sum_{i=\hat{x}}^{\infty} \frac{in^{i-1}e^{-n} - n^i e^{-n}}{i!} \frac{\hat{x}}{i+1} \\ &= \sum_{i=0}^{\hat{x}-1} \frac{in^{i-1}e^{-n}}{i!} + \sum_{i=\hat{x}}^{\infty} \frac{in^{i-1}e^{-n}}{i!} \frac{\hat{x}}{i+1} - \alpha(\hat{x},n) \\ &\equiv g(\hat{x},n) - \alpha(\hat{x},n) < 0 \end{aligned} \tag{13}$$

Note that

$$\alpha(\hat{x}+1,n) - \alpha(\hat{x},n) = \frac{n^{\hat{x}}e^{-n}}{\hat{x}!} - \frac{n^{\hat{x}}e^{-n}}{\hat{x}!}\frac{\hat{x}}{\hat{x}+1} > 0 \tag{14}$$

$$g(\hat{x}+1,n) - g(\hat{x},n) = \frac{\hat{x}}{n}\left[\alpha(\hat{x}+1,n) - \alpha(\hat{x},n)\right] \tag{15}$$

It follows that

$$|\alpha_n(\hat{x}+1,n)| - |\alpha_n(\hat{x},n)| = [\alpha(\hat{x}+1,n) - \alpha(\hat{x},n)]\left(1 - \frac{\hat{x}}{n}\right) \begin{cases} > 0 & \text{if } \hat{x} < n \\ = 0 & \text{if } \hat{x} = n \\ < 0 & \text{if } \hat{x} > n \end{cases}$$

Therefore, when $\hat{x} \geq n$, the case of overstock

$$\left|\frac{\alpha_n(\hat{x}+1,n)n}{\alpha(\hat{x}+1,n)}\right| < \left|\frac{\alpha_n(\hat{x},n)n}{\alpha(\hat{x},n)}\right| \tag{16}$$

Appendix D Proof of Proposition 2

The steady state market utility satisifes

$$J = \alpha(\hat{x}, n)(\eta - p) + wl^* \tag{17}$$

for all $\hat{x}$ and p in the equilibrium distribution $G(x, p)$. Note that

$$\frac{\partial J}{\partial \eta} = \alpha(\hat{x}, n) > 0 \tag{18}$$

for all sellers. By the Envelope theorem, a change in buyer utility $\tilde{\eta} > \eta$ leads to a higher market utility $\tilde{J} > J$. Now evaluate the new market equilibrium. Since $\hat{x}$ and p are pre-committed, a change from η to $\tilde{eta}$ only affects n. Consider two sellers (x_1, p_1) and (x_2, p_2). In the old and new equilibrium, we have

$$\alpha(x_1, n_1)(\eta - p_1) = \alpha(x_2, n_2)(\eta - p_2) \tag{19}$$

$$\alpha(x_1, \tilde{n}_1)(\tilde{\eta} - p_1) = \alpha(x_2, \tilde{n}_2)(\tilde{\eta} - p_2) \tag{20}$$

Take a ratio of the two equations.

$$\frac{\alpha(x_1, \tilde{n}_1)}{\alpha(x_1, n_1)} \cdot \frac{\tilde{\eta} - p_1}{\eta - p_1} = \frac{\alpha(x_2, \tilde{n}_2)}{\alpha(x_2, n_2)} \cdot \frac{\tilde{\eta} - p_2}{\eta - p_2} \tag{21}$$

If $p_1 > p_2$, $\frac{\alpha(x_1, \tilde{n}_1)}{\alpha(x_1, n_1)} < \frac{\alpha(x_2, \tilde{n}_2)}{\alpha(x_2, n_2)}$ because

$$\frac{\tilde{\eta} - p_1}{\eta - p_1} = 1 + \frac{\tilde{\eta} - \eta}{\eta - p_1} > 1 + \frac{\tilde{\eta} - \eta}{\eta - p_2} = \frac{\tilde{\eta} - p_2}{\eta - p_2} \tag{22}$$

It follows that the expected sales increases with the posted price.

$$\frac{\sum_{s=0}^{x_1} s\pi(x_1, \tilde{n}_1)}{\sum_{s=0}^{x_1} s\pi(x_1, n_1)} > \frac{\sum_{s=0}^{x_2} s\pi(x_2, \tilde{n}_2)}{\sum_{s=0}^{x_2} s\pi(x_2, n_2)} \tag{23}$$

Since this is true for any (p_1, x_1) and (p_2, x_2) with $p_1 > p_2$, the total expected sales increases.

$$\int p \sum_{s=0}^{x} s\pi_s(x, \tilde{n}_{x,p}) \, \mathrm{d}G(x, p) > \int p \sum_{s=0}^{x} s\pi_s(x, n_{x,p}) \, \mathrm{d}G(x, p) \tag{24}$$

Appendix E Impulse responses

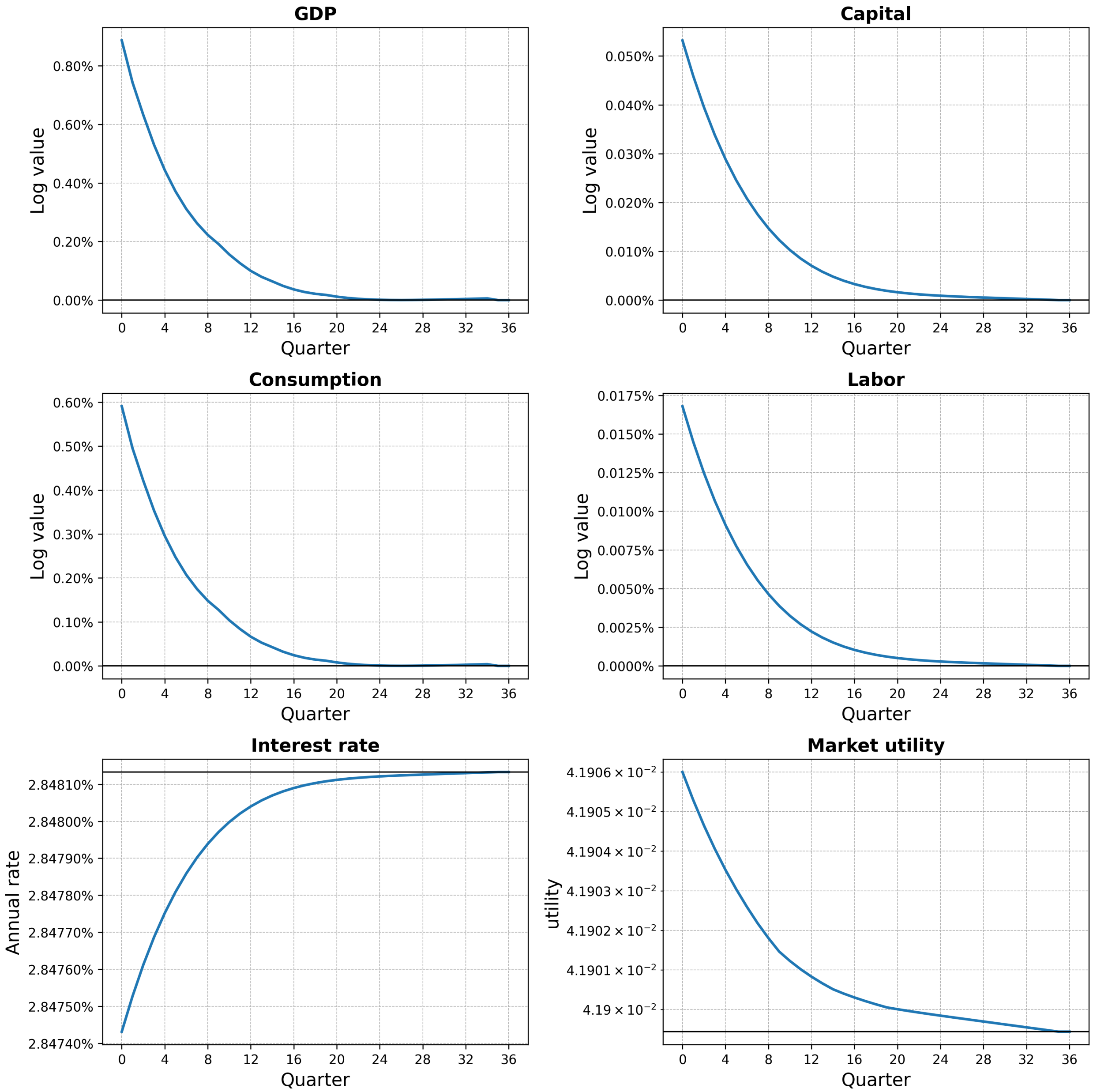

Figure 14: Model responses, other variables

The model responses to the sales shock. In the last graph, the market utility J is buyers' expected utility in the search market.

Appendix F Joint distributions

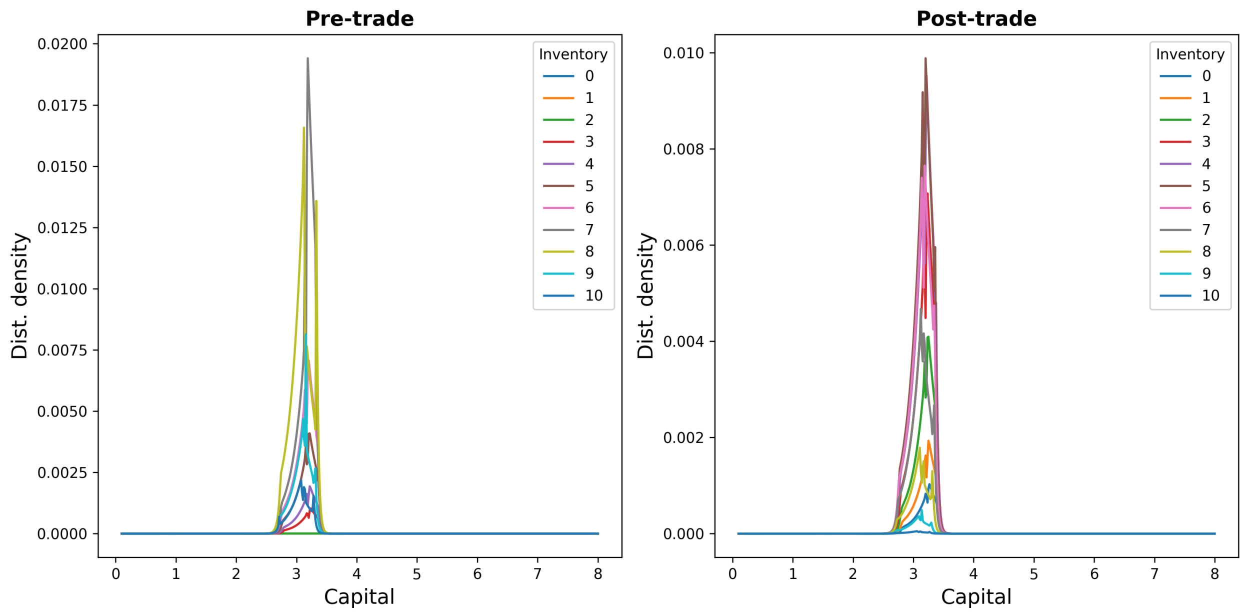

Figure 15: Joint distributions: steady state