INVENTORY CYCLES

Katsuyuki Shibayama
University of Kent at Canterbury

February 2008

ABSTRACT:

This paper investigates a rational dynamic stochastic general equilibrium model with a stockout constraint and a production chain. Our model shows that both stockout avoidance and cost shock mechanisms replicate stylised inventory facts -- production is more volatile than sales and inventory investment is procyclical. In addition, production smoothing also works at very high frequencies. Note that the cost shock and production smoothing mechanisms are naturally embedded in our micro-founded general equilibrium framework. Moreover, as a by-product, the production chain causes the slow adjustment of inventories in aggregate. Consequently, our model generates (a) high labour volatility and (b) low correlation between labour productivity and output; the standard RBC cannot produce these two empirical findings. Finally, our model yields inventory cycles.

KEYWORDS: Inventories, Inventory cycles, Stockout constraint, Production Chain, Over-damped oscillations, Dynamic stochastic general equilibrium model

JEL CLASSIFICATION: E32, C68

Acknowledgements: I would like to thank my colleagues and seminar participants at the University of Kent at Canterbury and the London School of Economics and Political Science. This article is based on one chapter of my Ph.D. thesis at LSE.

Address for correspondence: Katsuyuki Shibayama, Department of Economics, University of Kent, Canterbury, Kent, CT2 7NP, U.K. Phone:+44(0)122082-4714. Email: k.shibayama@kent.ac.uk

1 Introduction

Inventories are important in understanding business cycles. Inventory investment accounts for a large share of GDP fluctuations, especially during recessions.[1] Despite this importance, most existing theoretical studies of inventories focus only on firm/industry level analyses; only a few general equilibrium analyses exist. The motivation of this article is to investigate a micro-founded rational dynamic stochastic general equilibrium (DSGE) model that satisfies two stylised inventory facts: (1) production is more volatile than sales and (2) inventory investment is procyclical. Specifically, we construct a DSGE model with a stockout constraint and a production chain; the stockout constraint means that no seller can sell more products than the inventories she holds, and the production chain means that one firm's output is used as a production input by other firms, and this repeats.

In a sense, this article is a general equilibrium extension of Kahn (1987, 1992), who first analysed the stockout constraint. The key trade-off under the stockout constraint is that having too much inventory is costly because unsold goods impose a carrying cost (Jorgenson's user cost), while having too little inventory is also costly because the risk of losing sales opportunity due to stockout is too high. Balancing carrying cost against stockout probability, firms choose the optimal level of inventories. As a result, the optimal level of inventories is an increasing function of expected demand; given the level of inventories, strong demand reduces the expected amount of unsold goods and raises the stockout probability.

Our research, however, is most closely related to Khan and Thomas' (2004b) fully rational DSGE for inventories. In comparing the (S,s) and stockout avoidance models, they conclude that the former is superior to the latter, partly because firms have almost no inventories in their stockout avoidance model.

However, we conjecture that the competitive goods market in their model is not compatible with the existence of unsold goods (inventories carried over to the next

[1] For example, Fitzgerald (1997) reports that "changes in inventory investment are, on average, more than one-third the size of quarterly changes in real GDP over the postwar period." See also Blinder and Maccini (1991).

period). Consider firms' decisions at different points in one period. Certainly, *when firms decide their production*, there is an incentive to hold inventories as buffers, because some factor inputs are decided before the realisation of aggregate shocks in their model. However, *when firms decide their sales*, there is little incentive to hold inventories,[2] because all aggregate shocks are already revealed. In their competitive goods market, the price of goods should rise if demand is strong and vice versa, until the market clears (i.e., no inventories exist). At the end of the day, no inventories are carried to the next period.

In contrast, in our non-Walrasian goods markets, price does not equate demand and supply; instead, we assume price posting. Indeed, we claim that neither instances of stockouts nor unsold goods take place under flexible price. In sum, the most important difference between Khan and Thomas' model and ours is that they assume a competitive goods market, while we assume non-Walrasian goods markets.

Simulating our model, we find several observations. First, our model quantitatively satisfies the two stylised inventory facts. The intuition is as follows. When a positive demand shock hits firms, their inventories are *initially* reduced, and thus firms want to replenish inventories. Moreover, the target level of inventories becomes higher than the normal level, because the demand is stronger than usual. Hence, *in subsequent periods*, firms have to produce more than they sell in order to accumulate inventories. Thus, inventory investment is positive when sales and production are high, while production is more volatile than sales. Although this mechanism was predicted by Kahn (1987) in his firm level analysis, one of our contributions is to quantitatively endorse his prediction in the dynamic stochastic general equilibrium framework.

Although our model only explicitly assumes the stockout constraint, it generates the mechanisms predicted by the cost shock and production smoothing models. Importantly, even though we do not intend to explicitly build these mechanisms in our model, they must, naturally and inevitably, appear in our fully rational, micro-founded

[2] It is still possible that, if the marginal cost is expected to increase very sharply, firms carry inventories to the next period (production smoothing motive). However, such a motive seems to be quantitatively too weak to generate a significant amount of steady-state inventories (See Section 2.2.2).

environment. On the one hand, with a positive productivity shock (i.e., a negative cost shock), production increases but sales do not increase very much; as a result, inventories increase when production increases, while production is more volatile than sales. On the other hand, inventories certainly decrease *right after* a positive demand shock, and production does not react quickly because of the convex cost function. More specifically, if a band-pass filter is applied to the simulated data series, our model finds that production is *less* volatile than sales and inventory investment is *countercyclical* at very high frequencies. In sum, in our model, the following three leading mechanisms are all working: cost shocks, production smoothing and stockout constraint. Or, equivalently, our model finds that these three mechanisms predicted by firm/industry level analyses are all alive even in our micro-founded DSGE framework.

Another important finding in our model simulation is the slow adjustment of inventories, which is found in several empirical studies.[3] The key mechanism behind this is the production chain[4]. When an intermediate goods producer (M-firm) wants to replenish its inventories of intermediate goods (M-goods),[5] it has to increase its own production and its use of M-goods provided by other M-firms. That is demands for other M-firms' goods and reduces their inventories. This process repeats. In other words, increasing inventories in one firm decreases inventories in other firms. Thus, *the adjustment of inventories (or intermediate goods) in aggregate is slow.*

This slow adjustment of inventories also generates two by-products: higher volatility of working hours, and lower correlation between labour productivity and output, than the standard real business cycle (RBC) model. For the former, different from the standard RBC model, there is one extra production factor in our model − M-goods. However, because the adjustment of M-goods is slow, firms are forced to use more labour input to compensate for the sluggish adjustment of M-goods during booms. Indeed, our model

[3]See Blinder and Maccini (1991), among others. Also, Ramey and West (1997) interpret the persistent inventory to sales ratio as one expression of the slow adjustment of inventories.

[4]However, the primary purpose of explicitly modelling the production chain is to generate a realistic sales volume, which is much larger than the volume of production due to the use of intermediate goods. Note that under representative firm models, production is (almost) equal to sales.

[5]Note that our model analyses the stockout constraint in M-goods markets. Thus, inventories in our model mean inventories of M-goods, unless otherwise mentioned.

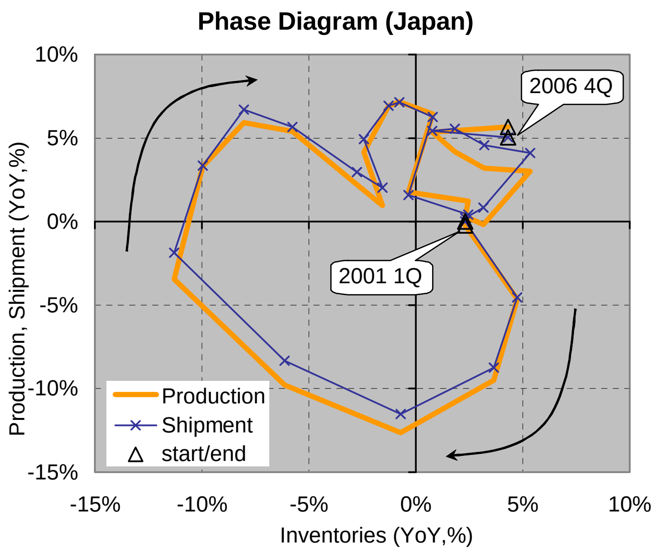


Figure 1: Inventory cycle in Japan. Source: MITI, Japan.

predicts that M-goods' price increases sharply after a positive demand shock, which encourages firms to substitute M-goods with labour. As a result, labour productivity (= output/hours) does not increase when output increases, because the increases in working hours are large enough to offset those in output; thus the correlation between labour productivity and output is low in our model. In sum, by adding stockout constraint and production chain, our model improves the standard RBC model in terms of labour.

Finally, our model can replicate so-called inventory cycles (Figure 1 and 2). In this respect, the model is successful to some extent. Our model generates simulated data, which exhibit cycles in the phase diagrams (See Figures 6 and 7 on page 33). However, although Shibayama (2007) finds sine curve impulse response functions (IRFs) by conducting VAR-based analyses, the theoretical model in this article only generates over-damped oscillations, which means that there is a mechanism that generates oscillation, but its effect is not strong enough to exhibit sine curve IRFs.

The plan of this paper is as follows. Section 2 reviews both theoretical and empirical literature, and summarises the stylised inventory facts. Our model satisfies not only the two famous stylised facts, but also additional detailed facts. Section 3 establishes the model environment. The key features of our model include: (i) in addition to the representative household, there are two types of firms: final goods producers (F-

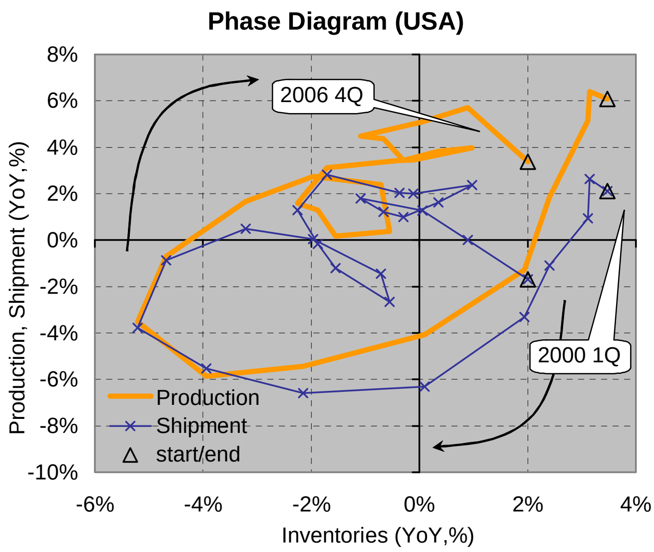


Figure 2: Inventory cycle in the U.S. Source: U.S. Bureau of Economic Analysis and Fed.

firms) and intermediate goods producers (M-firms), both of which use capital, labour and M-goods as inputs, while the former produce final goods (F-goods) which are used as consumption or investment goods, while the latter supply M-goods; (ii) individual M-goods are differentiated from each other, and hence an M-firm must use M-goods produced by other M-firms (production chain); and (iii) the sales of M-firms are subject to the stockout constraint. Section 4 presents numerical results. The final section concludes. The technical details are relegated to the Appendices.

In terms of terminology, note that this article uses "she" for a seller and "he" for a buyer. Also, the concept of inventories includes "goods on shelf" GoS_t and "unsold goods" U_{t+1}. Though this may sound ambiguous, we often need a word that represents both, because they are closely related to one another; indeed, $GoS_t = U_t$ under a simplified parameter setting. Inventory investment always means $U_{t+1} - U_t$.

2 Literature Review and Stylised Facts

This section reviews existing research. Despite inventory's importance in business cycle research, most existing theoretical inventory models focus only on firm/industry level analyses. There have been only a limited number of analyses of inventories in the setting of the DSGE model. In addition, key empirical research is also reviewed to reconsider stylised inventory facts.

2.1 Theories in Firm/Industry Level Analyses

Although we adopt more detailed facts to evaluate the model performance, the following two traditional stylised inventory facts have motivated the theoretical inventory research[6]:

(i) Production is more volatile than sales.

(ii) Inventory investment is procyclical.

2.1.1 Production Smoothing

The first attempt to understand inventories was the simple **production smoothing/buffer inventories** model, in which, analogous to consumption smoothing, firms want to avoid wild fluctuations in production because of a convex cost function (which should be present even with the CRS production function in general equilibrium), and inventories are used as buffers against demand shocks. However, it is obvious that smooth production is a concept opposite to volatile production, and it predicts that inventory investment is negative when there is a positive demand shock. In sum, its predictions contradict both of the above stylised facts.

2.1.2 Subsequent Models

Hence, subsequent researchers have made efforts to reconcile the production smoothing motive and the two stylised facts. In firm/industry level analyses, there are several strands of literature:[7]

[6] See Fitzgerald (1997) among others for a survey. Shibayama (2007) shows that these two facts essentially mention the same one fact.

[7] Of course, some researchers have contrived tricks to amend the problems pointed out here. The comments in the following list simply offer a glimpse of the models' basic features.

- **Serially correlated demand shocks** may explain to some extent why production is not very smooth, but it alone cannot explain why the volatility of production exceeds that of sales.[8]

- The **non-convex cost function** (or **bunching production**) has much empirical evidence from plant level studies, but it is uncertain as to whether the same mechanism works in aggregate.[9]

- The **cost shock model** successfully explains stylised fact (i), while its empirical evidence is mixed.[10] However, without any additional assumptions, it predicts that sales and inventories should be uncorrelated.

- **(S,s) ordering policies** successfully explains (i) under the assumption that production takes place no sooner than the order is placed; a fixed ordering cost induces **bunching orders**, and hence orders (production by suppliers) are more volatile than sales (of retailers).[11] However, it does not predict (ii). Moreover, the effect of bunching orders may disappear in aggregate, and it alone cannot explain why the stylised facts also hold within individual firms. In terms of empirical support, the (S,s) models have strong supports especially for trading firms.[12]

- **Inventories as production factors** can explain (ii) but not (i). In aggregate level analyses, where some simplification is inevitable, it may be difficult to discriminate inventory investment from capital investment in this model.[13]

It seems that the above lines of research have not yet reached successful results.

2.1.3 Micro-Founded Target Inventory Models

The following two models provide a micro-foundation for our general equilibrium model.[14]

[8] See Blinder (1986) for a more detailed discussion.
[9] See Ramey (1991) and Ramey and Vine (2004) for this line of research.
[10] For example, Kahn (1992) finds that "demand shocks are indeed more important" for U.S. automobile firms.
[11] See Caplin (1985) and Caballero and Engel (1991), among others.
[12] See Hall and Rust (2000) and Mosser (1991).
[13] See Ramey (1989).
[14] It is important to recognise the differences between the micro-founded target inventory models and the empirical models (LQ models) such as Blanchard (1983) and West (1986). LQ models assume that

- The **inventories as sales facilities** model is suggested from the standpoint of empirical studies.[15] The idea is that inventories, e.g., in showcases, are necessary to sell goods as samples or specimens. When sales are strong and serially correlated, a firm has to make up for the drop in inventories and, in addition, has to accumulate additional inventories to keep up with the new sales level, which is higher than before. Hence, in principle the model can explain both (i) and (ii).

- The **stockout avoidance motive** is probably the most natural setting, at least as a casual conjecture. Similar reasoning to that of the inventories as sales facility model shows that this can also explain (i) and (ii).[16]

Note that the inventories as sales facilities and stockout avoidance models are indeed special cases of a more general class of models. The generalised micro-founded target inventory model has the following sales function:

$$S_t = \left(D_t(P_t)^{\psi} + \phi GoS_t^{\psi} \right)^{\frac{1}{\psi}} \tag{1}$$

where $D_t(.)$ is demand as a function of price P_t, GoS_t is goods on shelf (inventories), and ψ and ϕ are parameters. The model reduces to the inventories as sales facilities model in Bils and Kahn (2000) if $\psi = 0$, while it reduces to the stockout avoidance model when $\psi = -\infty$. It is important to note that both models imply that the (target) *level of inventories*, rather than *inventory investment*, is an increasing function of demand.

The author personally believe that target inventory models for producers' final goods and (S,s) models for retailers' and wholesalers' inventories are the two most promissing approaches.[17]

there is a fixed target I/S ratio and any deviates from it incurs a cost. However, the micro-founded target inventory models emphasise that the target I/S ratio is endogenously determined.

[15] See Bils and Kahn (2000) and Pindyck (1994).

[16] See Kahn (1987, 1992). Abel (1985) provides early work on the stockout constraint. Wen(2002) also gives some empirical support for this idea.

[17] Blinder and Maccini (1991), for example, criticise the lack of research on producers' intermediate goods inventories.

2.2 General Equilibrium Analyses

As mentioned above, only a few general equilibrium analyses have been done to date. We list some of the theoretical works below.

2.2.1 (S,s) Models

Fisher and Hornstein (2000) and Kahn and Thomas (2004a, 2004b) focus on the (S,s) model in the settings of DSGE.

Although the (S,s) model has some difficulties in firm/industry level analyses, Fisher and Hornstein (2000) construct a DSGE model that satisfies the two stylised facts. In their model, general equilibrium feedback seems to be the key to understanding inventories.[18] By incorporating a matching scheme in the goods market,[19] they embed a mechanism by which a high level of inventories induces retailers to lower their sales prices so that consumers increase their search efforts (thus, sales are positively correlated with inventories).[20]

On the other hand Khan and Thomas (2004a, 2004b) also find that the (S,s) model can explain two stylised inventory facts. In Khan and Thomas (2004b), they compare (S,s) and stockout models and conclude that the former is better than the latter in terms of the two traditional stylised facts (see the next subsection).

2.2.2 Micro-Founded Target Inventory Models

Kahn, McConnell and Perez-Quiros (2002) constructed an **inventory in the utility model** as a proxy for the stockout avoidance motive with imperfect information. Their intuition is essentially the same as ours; when a positive shock hits a firm, its inventories decline, but the firm then has to replenish inventories and build up inventories to achieve the new, higher target level (because the sales shock is assumed to be persistent). They

[18] For the aggregation problem, they restrict the state space; the possible level of inventory holdings are limited to a few natural numbers.

[19] Note that in this sense their model also can be regarded as a non-Walrasian model. Their pricing mechanism is marginal (reservation) utility pricing, which is a special case of the Nash Bargain (sellers have all the bargaining power), and similar to ours.

[20] See Blinder (1982) and Bental and Eden (1993) for similar insights.

emphasise informational imperfection; firms cannot sell all of today's products in today's market due to an informational problem. However, inventory in the utility is not based on a micro-foundation, though it could be a useful short-cut.

Khan and Thomas (2004b) analyse the **stockout constraint** in a non-linear DSGE framework. In comparing the (S,s), which is very successful, and stockout avoidance models, they conclude that the former is superior to the latter, partly because firms have almost no inventories in the stockout avoidance model.

However, we conjecture that the competitive goods market in their model is not compatible with the existence of unsold goods (inventories carried over to the next period). Consider firms' decisions at different points in one period. Certainly, *when firms decide their production*, there is an incentive to hold *inventories as buffers* against imperfect information during one period.[21] This is because some factor inputs are decided before the realisation of aggregate shocks in their model. However, *when firms decide their sales*, there is little incentive to hold inventories,[22] because all aggregate shocks are already revealed. Having inventories just leads to a carry cost, but it no longer protects firms unless the marginal cost of the next period is expected to be very high (production smoothing motive). In their competitive goods market, the price of goods should rise if demand is strong and vice versa, until the market clears (i.e., no inventories exist), although Khan and Thomas (2004b) do not report the behaviour of the goods prices. At the end of the day, no inventories are carried to the next period. In a sense, their goods market is a Walrasian market with a vertical supply curve; unless the demand curve is unorthodox, the market finds a price to equate demand and supply.

In contrast, in our non-Walrasian goods markets, price does not adjust demand and supply; instead, we assume price posting. Indeed, we claim that neither instances of stockouts nor unsold goods take place under flexible price. In sum, our research is most closely related to Khan and Thomas' (2004b), but the most important difference

[21] Note that inventories in this sentence are goods on shelf in our terminology. However, because there is no unsold goods carried from the previous period in their model, goods on shelf are equal to today's production.

[22] Note that inventories in this sentence are unsold goods in our terminology. Note also that inventory investment means the time difference of unsold goods in general.

between their and our models is that they assume a competitive goods market, while we assume non-Walrasian goods markets.[23] Note that, because goods prices respond to all the aggregate shocks, though not to idiosyncratic shocks, our model falls into the class of flexible price models in aggregate.

2.2.3 Inventories with Sticky Price

Hornstein and Sarte (2001) and Boileau and Letendre (2004) incorporate inventories into a dynamic sticky price model.

The motivation to hold inventories used by Hornstein and Sarte is production smoothing. In their model, after a positive monetary shock, (i) for agents who have an opportunity to change prices, sales plummet down because their new prices become higher than other agents', but production does not move very much due to convex cost function, while (ii) for agents who do not change their price, sales and production increase. According to them, initial changes in sales are offset in aggregate, while changes in production are not. Thus, production is more volatile than sales.

Boileau and Letendre studied three types of models in the dynamic sticky price model. The most successful one is the model they call the shopping-cost model,[24] and it creates more persistence in output and inflation than the standard sticky price model. At first glance, their shopping-cost model seems to be similar to the micro-founded target inventory model such as ours, in the sense that both models share the feature that inventories help sales. However, it seems that their model should be regarded as an inventories as production factors model, at least, in aggregate. This is because, while inventories reduce the retailers' shopping cost, the authors impose the zero profit condition on the retailers at the same time. This means that, if retailers and producers can be regarded as one big sector, inventories work as a production factor in this big sector. Indeed, their final algebraic results look like those of the inventories as production factors model. In this sense, it is slightly questionable whether or not their model should

[23] In addition, while our model is solved by linearisation, they employ a non-linear solution method.

[24] The other two model investigated by Boileau and Letendre (2002) are a linear-quadratic model and inventories as factors of production.

be classified as the same class of the models as ours.

2.2.4 Other Important Research

Another important general equilibrium inventory paper is Diamond and Fudenberg (1989).[25] Although their model yields interesting results, including cyclical movements and multiple equilibria, their economy is highly stylised. They assume that each agent cannot have a (stochastic) production opportunity until she sells her products, and hence their "inventories" represent the number of people who had a production opportunity but have not yet sold their products.

2.3 Empirical Studies and Stylised Facts

This subsection briefly reviews empirical research and draws implications.

2.3.1 Stylised Inventory Facts

Although, as mentioned in the previous subsection, two stylised inventory facts are well known, we use more detailed facts in order to evaluate the model performance.

Most importantly, Wen (2002) reveals that the two traditional findings hold only at the business cycle frequencies (8 to 40 quarters); production is less volatile than sales and inventory investment is countercyclical at very high frequencies (2 to 3 quarters).[26] In addition, Ramey and West (1997) suggest that the inventory to sales ratio (I/S ratio) is persistent, which is perhaps essentially equivalent to the slow adjustment of inventories estimated by Blinder and Maccini (1991).[27] Finally, Bils and Kahn (2000) show that the I/S ratio is countercyclical.

In sum,

1.a Inventory investment is strongly countercyclical at very high frequencies (2 to 3 quarters).

[25] See also Diamond (1982).

[26] In this connection, Hornstein (1998) states that inventory investments are important for short-term output fluctuations (6 quarters or less), rather than business cycle fluctuations.

[27] Their model is often called an (empirical) target inventory model (though they are typically not micro-founded). See also Blanchard (1983) and West (1986).

1.b Inventory investment is procyclical at business cycle frequencies (8 to 40 quarters).

2.a Production is less volatile than sales at high frequencies.

2.b Production is more volatile than sales at business cycle frequencies.

3.a The I/S ratio is persistent and the adjustment of inventories is very slow.

3.b The I/S ratio is countercyclical.

There are a couple of supplementary comments. First, facts 1 and 2 (and hence traditional facts (i) and (ii)) are essentially equivalent to one another (see Shibayama (2007)). Second, while facts 1.a and 2.a support the production smoothing motive model, 1.b and 2.b are consistent with the target inventory models (see Sections 3 and 4 for details).

2.3.2 Inventory Cycles

Inventory cycles are cyclical movements in the phase plan, wherein typical year-on-year change (YoY) in inventories is on the x-axis, and YoY changes in production/shipment are on the y-axis (See Figures 1 and 2). This phenomenon is stable over time. The conjugate pair of complex roots in VAR coefficients is detected in Shibayama (2007), which is necessary for generating inventory cycles. Hence, in addition to the stylised facts listed above, the objective of this theoretical research is to construct a DSGE model that exhibits inventory cycles, as mentioned in the Introduction.

2.3.3 Other Empirical Issues

Negative Correlation Between I/S Ratio and Interest Rate: Bils and Kahn (2000) report that the correlation between the real interest rate and I/S is negative (see Table 2 in Bils and Kahn (2000)). They compute the correlation between expectations of real interest rate and I/S conditional on proper information sets. Then they argue that there must be some mechanism such as countercyclical markup to reconcile the FOC with respect to inventories to the data. Their finding is puzzling because the target inventory models suggest that the optimal inventories are decreasing in the interest rate (carrying cost). One possible way to understand this finding is that they essentially estimate the monetary policy rule, rather than the optimisation condition of inventories.[28] Nonetheless, we want to point out that a serious puzzle exists in the inventory literature.

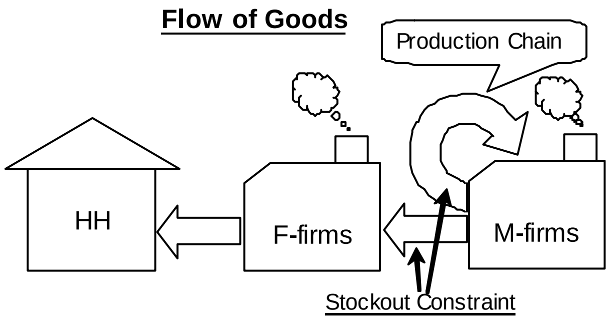
Inventories as Collateral: Related to the financial side of the economy, Kashyap, Lamont and Stein (1994) and Gertler and Gilchrist (1994) empirically show that small firms, whose access to financial markets is presumably limited, reduce their inventory holdings more than large firms during recessions. Thus, they both conclude that, for small firms, there is some form of interactions between inventories and financial/liquidity constraints.

Diminishing GDP Volatility and New Inventory Management Since mid 1980s, many industrialised countries have experienced a decline in the volatility of their GDP and prices (though some authors, such as Comin and Philippon (2005), find that the variability of output is increasing over time at the firm level). In this regard, Kahn et al. (2002) argue that improved inventory management (due to, say, new information technology) allows firms to protect themselves from shocks. They show that the decline in output volatility is salient more in the durable goods sector than in others. Their claim is also numerically evaluated by using our model.

3 Model Environment and Some Intuition

This section illustrates the key features of the model, but the full derivation of the most general model is relegated to Appendices A to C. This section motivates production chain (Section 3.1), discusses the implications of stockout constraint (Section 3.2) and shows price posting rule and other model assumptions (Section 3.3).

[28] Though controlling some information set looks like using the two-stage regression, their information set is presumably not independent of disturbances (i.e., the variables in the information sets do not work as IVs). Suppose, for example, that the monetary authority has a rule that it raises its policy interest rate when sales are strong and the inventory level is low. With no remedy, if the estimation of the FOC is less stable than that of the monetary policy rule, such computation essentially detects the monetary policy rule, rather than the FOC w.r.t. inventories.



Note: F- and M-firms mean final and intermediate goods producers, respectively. HH is household.

Among other assumptions, the stockout constraint and the production chain are essential – the model aims to analyse their implication in general equilibrium –, while idiosyncratic demand shock, price posting rule, etc. are rather technical assumptions. The latter are necessary devices for modelling the non-Walrasian goods markets; stockout implies that the goods markets do not clear.

3.1 Production Chain

There are three types of agents in the model: a representative household (HH), intermediate goods producers (M-firms) and final goods producers (F-firms), all of which optimise. HH works, consumes and invests. Production factors for both types of firms are labour, capital and intermediate goods (M-goods). Final goods (F-goods) are converted into consumption and investment goods (it is possible to interpret F-firms as retailers). A continuum of M-firms produce mutually differentiated M-goods (à la Dixit-Stiglitz monopolistic competition). A bundle of M-goods are necessary to produce not only F-goods, but also M-goods in the production chain.

Looking at the Leontief's input-output table, any two industries demand and supply M-goods from and to one another. Because the input of M-goods is subtracted from sales to compute value-added, sales are much larger than value-added in reality. On the other hand, the stockout constraint implies that the target level of inventories (or goods on shelf) is an increasing function of sales, not value-added. Hence, without modelling the production chain, we underestimate the volume of sales – and hence the volume of the target level of inventories.[29]

Note that if the M-goods markets are frictionless, then the model reduces to a single production sector model; the stockout constraint – a friction in M-goods markets – makes the production chain worth analysing.

3.1.1 Implications of Production Chain

Different from the standard RBC model, however, there is one additional production factor – M-goods. When a shock hits the model economy, capital cannot adjust quickly, as in the standard RBC model, because its evolution is governed by the capital accumulation equation. The adjustment of the additional production factor – M-goods – is also sluggish. This is because of the production chain; when one M-firm wants to increase its supply, it must use other firms' M-goods, which, in turn, implies that other firms want to increase their production by using other firms' M-goods.[30] In aggregate, to produce M-goods, M-firms must consume M-goods! In sum, due to the production chain, the adjustment of inventories is very sluggish in aggregate. In addition, this slow adjustment of M-goods inventories has several important implications for labour (see below for details).

3.2 Stockout Constraint

Our model explicitly analyses the effect of the stockout constraint, which was first examined by Kahn (1987), and our study is a general equilibrium extension of his market equilibrium analysis.

Our model considers the stockout constraint on the M-good markets, and defines

[29] In this connection, consider the Leontief production function where the elasticity of substitution between labour/capital and M-goods is zero $\eta_M = 0$ (see Section 4 for notations). Then the use of M-goods is proportional to the gross output: $Y_t^M = Z_t^{Mn} M_{t-1}^M/(1-\phi_M)(= Z_t^{Mn} V_t^M/\phi_M)$, where Z_t^{Mn} is the technology and ϕ_M is the share parameter of value-added component V_t^M. The Leontief's inverse matrix – the most important concept in the input-output table analysis – shows the increase in the output of one sector due to a unit increase in final demand. Noting that M-goods produced are used as inputs of F-firms and M-firms: $Y_t^M = M_t^F + M_t^M$, $\partial Y_{ss}^M/\partial M_{ss}^F = (1-(1-\phi_M))^{-1} = \phi_M^{-1} > 1$ in the symmetric steady state (in our model, the matrix is actually 1×1). Hence, gross output fluctuates more when the share of intermediate goods is larger. In industrial countries, $\phi_M \simeq 0.5$, which implies that the size of value-added is roughly one half of that of sales.

[30] To gain further intuition, see also the previous footnote.

goods on shelf GoS_t as the sum of unsold goods U_t and (a portion of) today's production Y_t^M. In terms of terminology, GoS_t and U_t are both (the level of) "inventories," but the former is measured before the opening of M-goods markets, while the latter is after the markets close.

The stockout constraint, the main friction in our model, means that no seller can sell more products than the stocks on shelf GoS_t. Hence,

$$S_t = \min \{GoS_t, M_t^p\} \tag{2}$$

where S_t is the sales and M_t^p is the potential demand for M-goods. The potential demand is "potential" simply because it may not be realised due to stockout.[31] In the simplest version of the model, the FOC with respect to inventories (unsold goods)[32] is

$$E_t \left[\beta \frac{\lambda_{t+1}^H}{\lambda_t^H} \left\{ Pr_{t+1} P_{t+1}^i + (1 - Pr_{t+1}) \lambda_{t+1}^M \right\} \right] = \lambda_t^M \tag{3}$$

where P_{t+1}^i is the price of the i-th M-good, $\beta\lambda_{t+1}^H/\lambda_t^H$ is the stochastic discount factor (SDF) between time t and $t+1$,[33] Pr_t is the probability that a seller faces stockout and λ_t^M is the marginal cost/shadow price of inventories (Lagrange multiplier on the law of motion of inventories). To understand the FOC, consider the additional one unit of inventory (marginal inventory). If it is sold (i.e., stockout takes place), it brings P_{t+1}^i units of revenue to the seller, while if it is not sold, then it remains on the seller's hand and its value is its shadow price λ_{t+1}^M. Thus, the euqation means that the shadow price of inventories today is equal to the PV of inventories in the next period. Or equivalently,

$$E_t \left[\beta \frac{\lambda_{t+1}^H}{\lambda_t^H} Pr_{t+1} \left(P_{t+1}^i - \lambda_{t+1}^M \right) \right] = \lambda_t^M - \left[\beta \frac{\lambda_{t+1}^H}{\lambda_t^H} \lambda_{t+1}^M \right] \tag{4}$$

This equation states that the user cost of inventories (RHS) is equal to the PV of the

[31]It may be possible to express the stockout constraint in the form of a non-negativity constraint on GoS_t, but adding the non-negativity constraint complicates the algebra.

[32]It is the same as (10e) in Appendix C with $\upsilon = 0$, where υ is the portion of today's output that can be sold in today's market.

[33]Note that the SDF is the ratio of marginal utilities ($\beta\lambda_{t+1}^H = \partial U/\partial C_{t+1}$ and $\lambda_{t+1}^H = \partial U/\partial C_t$).

profit margin conditional that the marginal inventory is sold out. Note that the stockout probability Pr_t is decreasing in inventories U_t and increasing in potential demand M_t^p. In the simplest version of the model, its functional form[34] is

$$Pr_t = \frac{M_t^p - U_t}{\nu} + \frac{1}{2} \tag{5}$$

There is a fundamental trade-off; stockout is costly because it means the loss of a profitable sales opportunity, but unsold goods are also costly because they impose a carrying cost (or Jorgenson's user cost) of unsold goods. Note that the nature of the carrying cost is the cost of financing inventories.

Hence, the target level of inventories is an increasing function of the potential demand (which moves closely with sales), but is a decreasing function of the interest rate (financing cost) through SDF. When the potential demand is strong, for example, if U_t were kept unchanged, the stockout probability would be too high while the level of expected unsold goods would be too low; hence, firms have an incentive to accumulate inventories, and vice versa. Note that choosing optimal inventories is essentially equivalent to choosing optimal stockout probability.

3.2.1 Implications of Stockout Constraint

The stockout constraint can (at least potentially) explain the two inventory stylised facts (see also Kahn (1987)). One of the goals of this article is to quantitatively evaluate the effects of the stockout constraint in the DSGE framework.

The intuition is as follows. As mentioned above, under the stockout constraint, the target level of inventory is an increasing function of the potential demand, which shows movements quite similar to sales. Hence, inventory investment is naturally procyclical (fact 1b). Furthermore, production must increase more than sales because, otherwise, inventories decrease (fact 2b).

In addition, the I/S ratio is countercyclical because, during a recession, the interest

[34] It is essentially the same as (13b) in Appendix C. Note that $M_t^p = Q_t^{F\frac{-\theta}{\theta-1}} M_t^F + Q_t^{M\frac{-\theta}{\theta-1}} M_t^M$ and $U_t = GoS_t$ (υ is assumed to be zero in the simplest version).

rate is low and thus the carrying cost is low as well, which stimulates inventory holdings relative to sales.

3.2.2 Inventories as Buffers

It is important to note that the mechanism explained in the previous subsection is expected to materialise at business cycle frequencies.

At very high frequencies, on the other hand, production smoothing can be explained by the very basic convex cost function. Inventories work as buffers against demand shocks. Even if production technology ensures constant returns to scale (CRS), as long as the labour supply is convex (due to the concave utility function), this mechanism works. Because firms do not want to adjust their production quickly, inventories will decrease right after a positive demand shock, and vice versa.

Note that both mechanisms − buffer stocks and stockout constraint − do not contradict each other, and they indeed coexist in our model. Indeed, equation (3) also shows the production smoothing motive. If there is not stockout probability ($Pr_{t+1} = 0$), (3) simply shows the marginal cost smoothing. Moreover, the cost shock model, in which productivity shock directly affects λ_t^M, is also encompassed; when λ_t^M is low (i.e., positive productivity shock), (3) implies that the optimal Pr_{t+1} is also low (so sales is high), which means that M-firms produce more to accumulate U_{t+1} (see (5)).

In sum, the inventories' FOC (3) embraces the three mechanisms: the stockout avoidance, production smoothing and cost shock models. Our objective is not to pick up one single "true" mechanism out of the three models, but to compare them quantitatively.

3.2.3 Inventories as Options to Sell

Also, we can interpret (4) as a derivative pricing equation, in which inventories are interpreted as options to sell; having inventories, sellers have options to sell their goods. It can be shown that the (4) has almost one-to-one correspondences to the Black-Scholes option pricing formula (see Appendix C.2 for details).

Although sales (2) itself is not a smooth function, it is possible to find FOC, because *the expected sales* is a smooth function due to demand uncertainties. This technique is commonly used in the analyses of voting behaviours.

3.3 Structure of M-goods Markets

This subsection provides rather technical basis of the model. We recommend that interested readers consult Appendix A. Here, only the key assumptions are listed:

- Due to **idiosyncratic shocks**, individual sellers face different levels of demand. Both stockout and unsold goods exist, implying that the M-market is **non-Walrasian**.
- Hence, we cannot use market clearing conditions as a pricing mechanism. Instead, we assume **price posting** by sellers, wherein buyers decide on the trading quantities. Buyers' FOCs are regarded as demand curves.
- M-goods are differentiated from each other (Dixit-Stiglitz' monopolistic competition). Two-stage budgeting is modified by the **cost effect of losing variety**.

However, note the following two model features. First, our model is a **flexible price model** in aggregate. In price posting, the M-goods prices can respond to all aggregate shocks, but not to idiosyncratic shocks (hence, all sellers post the same price because their sales prices are posted before observing idiosyncratic shocks). Such price rigidity disappears in aggregate. Second, due to the CRS production function and price posting, our model falls in the class of **representative agent models in aggregate**, despite the heterogeneity caused by stockout.

Note also that it is possible to linearise the stockout constraint (2), intuitively because the *numbers* of sellers and buyers with binding stockout constraint (2) are smooth functions *in aggregate*, even though (2) itself is not a smooth function *from the individual sellers' viewpoint* (See Appendix A.2 for details). In the simplest version of the model, aggregate sales and the law of motion of inventories are

$$
\begin{aligned}
S_t &= U_t - \frac{\nu}{2}\left\{\frac{M_t^b - GoS_t}{\nu} - \frac{1}{2}\right\}^2 \qquad (= E\,[\text{sales}] \ \text{ for individual sellers}) \\
U_{t+1} &= \frac{\nu}{2}\left\{\frac{M_t^b - GoS_t}{\nu} - \frac{1}{2}\right\}^2 + Y_t
\end{aligned}
$$

where Y_t is today's output[35] and M_t^b is the baseline demand which is the average of potential demand M_t^p; $M_t^{pi} = M_t^b + e_t^{pi}$ where e_t^{pi} is the idiosyncratic shock to seller i.

4 Numerical Experiments

This section shows the calibration results. The analytical solution is linearised around the non-stochastic steady state, and is simulated to obtain the second moments and impulse response functions (IRFs).

4.1 Parameter Selection

To select parameters, we do not employ any optimal selection criteria. Rather, for the sake of comparability, we follow the convention in the RBC literature. For the parameters that are specific to our model, we select values to match some steady state values to the data. The difficulty, however, is that two parameters ν and υ govern the steady state I/S ratio, which means that there are many possible combinations of ν and υ that match the observed I/S ratio. In addition, there are six coefficients for the adjustment costs, which are not pinned down by the first moments. Hence, perhaps one possible criticism is that our model has too many degrees of freedom in choosing parameters.

4.1.1 RBC Parameters

For exact values of the RBC parameters, see Table 1. For the elasticity of substitution among varieties, we borrow the number that is commonly used in the sticky price models

[35] Under the assumption that Y_t cannot be sold in today's market (i.e., $\upsilon = 0$), the first equation is equivalent to (6b) and (13a), while the second one is equivalent to (6c) and (13c). Also, $\frac{\nu}{2}\{\frac{M_t^b - GoS_t}{\nu} - \frac{1}{2}\}^2$ is the amount that is not sold due to stockout.

($\theta = 10$). We select values for AR(1) coefficients for technology shocks to match the autocorrelation function of GDP (i.e., $Corr\{GDP_t, GDP_{t-5}\} \simeq 0$). Though these values are smaller than in the standard RBC model, perhaps this is merely due to the existence of adjustment costs and does not signify endogenous persistence.

Table 1: Benchmark Parameters for Model Simulations

Symbol	Meaning	In Benchmark
β	Subjective discount factor (4% annual interest rate)	$1.04^{-1/4}$
γ	Reciprocal of elasticity of intertemporal substitution of consumption	1.00
γ_L	Reciprocal of elasticity of intertemporal substitution of labour	0.00
ψ	Weight on leisure in period utility (Working hours = 1/3)	0.68
θ	Elasticity of substitution among M-goods	10.0
ν	Range parameter of idiosyncratic shock ($U_{ss}/S_{ss} = 2$ months)	0.40
υ	Share of today's output that can be sold in today's market	0.50
α_M, α_F	Capital share in value added	0.35
η_M, η_F	Elasticity of substitution btw M-goods and value-added compo.	0.30
ϕ_M	Weight on value-added compo. of M-firms	0.50
ϕ_F	Weight on value-added compo. of F-firms	0.05
δ_M, δ_F	Depreciation rate of capital (Capital/GDP = 10)	0.015
χ_{MK}, χ_{FK}	Coefficient on quadratic adjustment cost of investment	0.10
χ_{MH}, χ_{FH}	Coefficient on quadratic adjustment cost of labour	1.50
χ_{MM}, χ_{FM}	Coefficient on quadratic adjustment cost of M-goods use	1.00
ρ_{Mn}	AR(1) coefficient of Hicks-neutral technology shock to M-firms	0.75
ρ_{Fn}	AR(1) coefficient of Hicks-neutral technology shock to F-firms	0.85

4.1.2 Parameters Specific to the Model

Share Parameter of Value-Added: For the share parameter of the (notional) value-added in production functions, we set $\phi_M = 0.5$ so that the share of M-goods M_t^M / Y_t^M in the M-firms is roughly 45%; the value-added is roughly 55% of sales. This number is taken from the Japanese and U.S. Leontief's input-output tables.[36] Also, we set $\phi_F = 0.05$ so that F-firms act as if they were the retailers who simply convert M-goods into F-goods.

Note that the notional value-added V_t^M and V_t^F, which appear in the definitions of our production functions, are not consistent with the statistical concept of GDP. This

[36] Ministry of Internal Affairs and Communications, Government of Japan (2004) and Bureau of Economic Analysis, U.S. Department of Commerce (2007).

article uses terminology "GDP" to mean gross output minus the use of M-goods; e.g., $GDP_t^M \equiv Y_t^M - P_{ss}^M M_{t-1}^M$ for M-firms. Also, we assume the Laspeyres price index so that goods are evaluated by the price of the steady state as the base year.

Table 2: Endogenous Variables in the Steady State

Symbol	Meaning	In Steady State
SDF_t	Stochastic discount factor (= real interest rate)	β
W_t	Wage rate	1.76
P_t^M	M-goods price	0.9996
Q_t	Pr[cannot buy] (= number of available varieties)	0.999
Pr_t	Pr[stockout]	0.074
λ_t^M	Marginal cost of M-goods production (shadow price of M-goods)	0.89
C_t	Consumption	0.83
H_t^H	Labour supply (= 1 − leisure = $H_t^M + H_t^F$)	0.28
S_t	Sales of M-goods	1.79
M_t^M, M_t^F	Use of M-goods as production factors	0.86, 0.93
Y_t^M	Gross output of M-goods (= $M_{t+1}^M + M_{t+1}^F$)	1.79
Y_t^F	Gross output of F-goods (= $C_t + I_t^M + I_t^F$)	0.98
V_t^M	(Notional) value-added in M-firms (= $K_t^{M\ \alpha_M} H_t^{M\ (1-\alpha_M)}$)	0.93
V_t^F	(Notional) value-added in F-firms (= $K_t^{F\ \alpha_F} H_t^{F\ (1-\alpha_F)}$)	0.053
$H_t^{M_p}$, $H_t^{F_p}$	Labour input for production	0.27, 0.015
I_t^M, I_t^F	Investment	0.14, 0.008
K_t^M, K_t^F	Capital at the beginning of period t	9.37, 0.53
U_t	Unsold M-goods at the beginning of period t	1.25
Z_t^H	Preference shock	1.00
Z_t^{Mn}, Z_t^{Fn}	Hicks-neutral technology shock in production function	1.00, 1.00

Elasticity of Substitution Between Value-Added and Input M-goods: For the elasticity of substitution between the notional value-added component and intermediate goods η_K $(K = F, M)$, we do not have much guidance. Rotemberg and Woodford (1996) used a value of 0.7, while Bruno (1984) suggested 0.3 to 0.4.[37] Because, presumably, the substitution should be low, we use 0.3.

Magnitude of Idiosyncratic Shock and Proportion of Output that Can be Sold in Today's Markets: There are two parameters that affect the steady state

[37] Basu (1996) regards Bruno's survey as an upper bound.

I/S ratio: the upper and lower supports of the uniform idiosyncratic shock $\nu/2$, and the portion of today's products that can be sold in today's market υ. In the data, the I/S ratio is roughly 2 months (0.67 quarter).[38]

On one hand, if we set $\upsilon = 1$, as in most firm/industry level analyses, aggregate inventories do not move very much. This is because we assume that production is decided after observing all of the aggregate shock. Hence, if M-goods firms can sell all of their products in the current period market, they, as a collective agent, can respond to aggregate shocks almost fully. Certainly, inventories still vary over time as the interest rate changes over time, and so does the carrying cost. However, in a sense, inventories merely follow other key variables in this case; hence, the model behaves very similarly to the standard RBC model. On the other hand, if we set $\upsilon = 0$ (i.e., $GoS_t = U_t$), (2) implies that $U_{ss} > S_{ss}$, which clearly contradicts the data. If we could know how well firms responded to contemporary aggregate shocks in the real world, we could pin down the value of υ.

Our strategy is as follows. We first naively set $\upsilon = 1/2$, as simply the midpoint between the two extremes, and then choose $\nu = 0.4$ so that the I/S in the model economy is 2 months.

Convenience Yield on Inventories: Stockout probability, which is roughly 5% to 9% in the data according to Bils (2004), is mainly affected by the subjective discount factor β elasticity of substitution among varieties θ and convenience yield c_1. Essentially, any parameters that determine the opportunity cost of holding inventories affect the steady state stockout probability. If the opportunity cost of lost sales is high, the optimal stockout probability is lower. Given $\theta = 10$, we select $c_1 = 0.00$ (we assume no convenience yield), so that $Pr_{ss} = 7.4\%$.

Adjustment Costs: We assume quadratic adjustment costs, which are rather standard in DSGE research. Specifically, we set $\chi_{MK} = \chi_{FK} = 0.1$, $\chi_{MH} = \chi_{FH} = 1.5$ and $\chi_{MM} = \chi_{LM} = 1.0$.

[38] See Ramey and West (1997), for example.

4.2 Numerical Results

A shock to F-firms' production function (F-shock) can be regarded as a *pure* demand shock for M-firms, while a shock to M-firms' production function (M-shock) works as a demand shock and a supply shock from the viewpoint of *individual M-firms*.

In this subsection, all the simulated data are HP-filtered, unless otherwise mentioned. Also, "relative volatilities" are standard deviations relative to that of total GDP or M-firms' GDP. Similarly, "correlations" are correlations with total GDP or M-firms' GDP.

4.2.1 Second Moments

Table 3 summarises the second moments generated by the model simulations. The main results are (i) compared to the RBC model, our model considerably decreases the correlation between labour productivity and hours worked and (ii) it satisfies the two stylised inventory facts.

Correlation of Inventory Investment with GDP: Inventory investment is positively correlated with M-firms' GDP for both shocks. With M-shocks, it is not surprising to observe this positive correlation (0.65); this is exactly what the cost shock model expects. However, it is more important to find a positive correlation (0.31) even with a pure demand shock. In data, the correlation is 0.66.

The near-zero correlation between inventory investment and total GDP (M-firms' GDP plus F-firms' GDP) with F-shocks is the artefact of the model assumptions because the F-shock directly increases the F-firms' value-added, but it decreases M-firms' inventories. Indeed, if we use preference shocks instead of the F-shocks, the correlation is even higher. However, preference shocks deteriorate other dimensions of the model performance, so we do not choose this option.

Relative Volatility of Sales: Sales are less volatile than output for both types of shocks; the standard deviation of sales relative to that of M-firms' GDP is 0.77 for both F- and M- shocks in our model, while this value is 0.71 in the data. With M-shocks, this is not surprising, because the source of the shock lies on the production side, as the cost shock models predict. However, it is important to note that, even when the source of the shock lies on the demand side, production is more volatile than sales.

Table 3: Simulation results (comparison to the standard RBC model).

Cited from Cooley and Prescott (1995)

	Output	Sales	Hours	Consumption	Investment	d(inventories)	Output/Hours	Corr{Productivity, Hours}
Standard RBC Model								
relative s.d.	*1.35*	-	**0.57**	0.24	4.41	-	0.45	
corr	1.00	-	0.99	0.84	0.99	-	0.98	**almost 1**
Data								
relative s.d.	*1.72*	0.71^	**0.92**	0.50	4.79	0.27^	0.52	
corr	1.00	0.94^	0.86	0.83	0.91	**0.66^**	0.41	**-0.26***

Notes: "relative s.d." means s.d. relative to s.d. of output. Italics are s.d., not relative s.d.
"corr" means correlation with GDP.
^ indicates that numbers are taken from Khan & Thomas (2004)
* indicates that numbers are taken from Gali (1999).

Stockout Model (elasticity btw Value-add & M-goods = 0.3)

	Output	Sales	Hours	Consumption	Investment	d(inventories)	Output/Hours	Corr{Productivity, Hours}
Technology shock to M-firms: rho = 0.75, sigma = 0.7%								
relative s.d.	*2.83*	0.77	0.99	0.18	4.64	0.29	0.36	
corr	1.00	0.81	0.93	0.44	0.53	0.62	0.23	-0.13
of which M-firms								
relative s.d.	1.04	**0.77**	**0.96**		4.25	0.28	0.36	
corr	1.00	0.90	0.93		0.51	**0.65**	0.30	**-0.06**
Technology shock to F-firms: rho = 0.85, sigma = 0.7%								
relative s.d.	*1.57*	0.55	0.87	0.20	4.52	0.15	0.26	
corr	1.00	0.96	0.97	0.73	0.99	0.01	0.63	0.42
of which M-firms								
relative s.d.	0.53	**0.77**	**1.63**		8.46	0.28	0.67	
corr	0.94	0.99	0.98		0.96	**0.31**	-0.88	**-0.95**

Notes: For "of which M-firms," "relative s.d." and "corr" show s.d. relative to that of M-firms' output and correlation with M-firms' output, respectively.
Relative s.d. of M-firms' output shows s.d. of M-firms' output relative to that of total output. See also notes above.

Intuition: For F-shocks, the target inventory models explain the mechanism behind two observations: (i) procyclical inventory investment and (ii) output more volatile than sales, as follows. When a positive demand shock hits M-firms, of course, their inventories initially decline, simply because buyers take away M-goods from the shelf of M-firms.

However, keeping such a low level of inventories is costly, because it leads to a too high stockout probability (in the stockout model) and because of an inefficient sales activity without enough samples in showcases (in the inventories as sales facility model). The common prediction among the target inventory models is that the target level of inventories is an increasing function of demand/sales. Hence, with a positive demand shock, the target level of inventories is higher than usual and, as a result, M-firms have an incentive not only to replenish their declined inventories but also to accumulate more inventories to meet the higher demand. However, as the law of motion of inventories (10j) shows,

$$U_{t+1} - U_t = Y_t^M - S_t$$

the output of M-goods Y_t^M must increase more than the sales of M-goods S_t to build up inventories U_{t+1}, suggesting that (i) Y_t^M increase more volatile than S_t and (ii) $U_{t+1} - U_t$ is positive when Y_t^M and S_t increase. Indeed, this article confirms this mechanism quantitatively in the DSGE setting.

Relative Volatilities of Consumption and Investment: For both shocks, our model inherits the basic nature of the standard RBC model. That is, the relative volatility of consumption is too low, while that of investment roughly matches the data. This is not surprising since our model is an extension of the standard RBC model. The correlation of investment and value-added is too low for the M-shock (0.53), though. The reason for this is that the M-shock, opposed to F-shock, raises the price of investment goods (F-goods), relative to M-goods price.

Persistence of I/S Ratio: According to Ramey and West (1997), the first and second autocorrelations of the inventory-sales relationship (akin to I/S ratio) range from 0.88 to 0.97 and 0.80 to 0.91, respectively. This persistency is regarded as another expression of the slow adjustment of inventories. In our model, the first and second autocorrelations of the I/S ratio are 0.88 and 0.61 for F-shocks and 0.71 and 0.25 for M-shocks, respectively.[39]

[39] These values are defined as U_t/S_t, where S_t is the M-firms' sales. The results are almost the same if we define the I/S ratio as unsold goods divided by total sales.

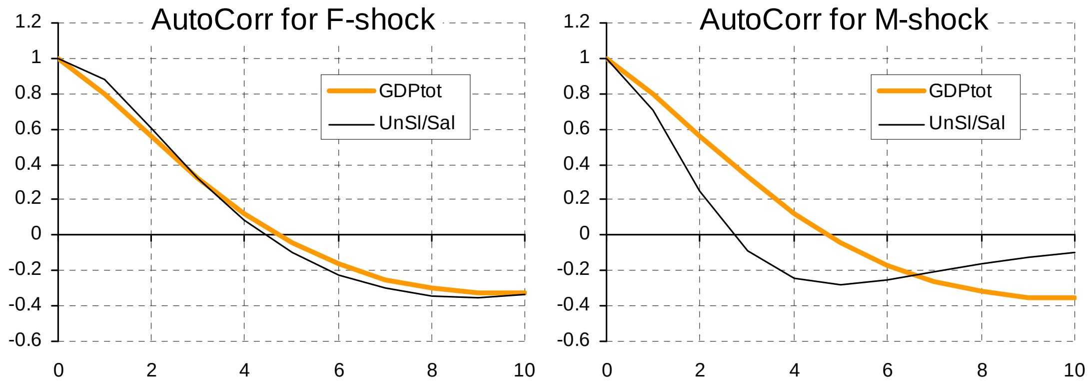


Figure 3: Autocorrelation functions. "GDPtot" and "Unsl/Sal" mean gross output minus the use of M-goods, and unsold goods divided by sales (I/S ratio), respectively.

The I/S ratios in our model are considerably persistent (see Figure 3), though they are somewhat less persistent than the data. Moreover, in our model the I/S ratio is countercyclical because of the procyclical interest rate.

The key mechanism behind this is the production chain. Suppose a positive demand shock hits an M-firm. This firm faces a decrease in its inventories and expects strong future sales, so it wants to replenish its inventories; much more, it raises its inventory level to catch up with the new higher level of sales. As a consequence, it has to increase its production and, hence, the use of production factors, including M-goods. However, this, in turn, implies that the demands (and hence the sales) of other M-firms increase, and that their inventories are reduced. In other words, the production chain implies that one firm's replenishment of inventories reduces other firms' inventories. Therefore, the adjustment of inventories is slow in aggregate. It is important to note that M-goods price increase sharply after a positive F-shock, while M-goods price does not decrease very much after a positive M-shock. Note that unit labour cost (wage/labour productivity) decreases after a positive M-shock (= a negative cost shock), implying that M-goods becomes expensive in relative term.

In this regard, our model can suggest a very simple reason that reduced form target inventory models estimate an *implausibly* slow adjustment speed; it is indeed slow! Certainly, Blinder and Maccini (1991) persuasively argue that "One major difficulty with stock-adjustment models is that adjustment speeds generally turn out to be extremely low; the estimated λ is often less than 10 percent per month. This is implausible when even the widest swings in inventory stocks amount to no more than a few days of production."[40] Reiterating our finding, the inventories' adjustment is slow in aggregate due to production chain, although it seems to be implausible from the viewpoint of individual firms. Partial equilibrium analyses may miss the general equilibrium feedback through volatile M-prices; during a boom, high M-prices discourage M-firms from replenishing their inventories quickly by producing more.

Working Hours: In our model, working hours are more volatile than in the standard RBC model. As a result, the correlation between hours and labour productivity is lower than the standard RBC model. If we focus on M-firms, this correlation is -0.06 and -0.95 with M- and F-shocks, respectively.

One of the major drawbacks of the standard RBC model is that it counterfactually exhibits an almost perfect correlation between labour productivity and working hours. Although one way to overcome this caveat is to add demand shocks (see Christiano and Eichenbaum (1992) for government expenditure, and Bencivenga (1992) for preference shocks), such demand shock models are criticized by Gali (1999), in which a structural VAR shows that the correlation between labour productivity and hours is negative for technology shocks, but positive for other shocks. Gali (1999) suggested that a dynamic sticky price model with a labour effort model can, at least potentially, generate a negative correlation. However, our model improves the model performance in this respect even without price rigidity.

The mechanism that generates volatile working hours in our model is the slow adjustment of inventories; due to the production chain, one firm's replenishment of inventories reduces other firms' inventories in aggregate. The right panels of Figures 4 and 5 show the IRFs of production factors. It is clear that, for both types of shocks, the increase in M-goods use is less volatile than M-goods production and labour input compensates

[40] See Blinder and Maccini (1991, p.81).

Table 4: Model behaviour at different frequency domains.

	High Frequencies (2-3quaters)		Business Cycle Frequencies (8-40quaters)	
Data	Var(sales)/Var(output)	Cor(d(inventory),	Var(sales)/Var(output)	Cor(d(inventory), sales)
	1.10	**-0.43**	**0.72**	**0.58**
Model (**M-firms and F-firms**)	Var(sales)/Var(output)	Cor(d(inventory),	Var(sales)/Var(output)	Cor(d(inventory), sales)
tech shock to M-firms: rho = 0.75, sigma = 0.7%	0.18	0.21	0.83	0.60
tech shock to F-firms: rho = 0.85, sigma = 0.7%	0.36	-0.97	0.56	0.62
Model **(M-firms only)**	Var(sales)/Var(output)	Cor(d(inventory),	Var(sales)/Var(output)	Cor(d(inventory), sales)
tech shock to M-firms: rho = 0.75, sigma = 0.7%	**0.32**	**0.36**	**0.82**	**0.29**
tech shock to F-firms: rho = 0.85, sigma = 0.7%	**1.02**	**-0.92**	**0.76**	**0.63**

Note: Data is OECD average (cited from Wen (2003)).

such sluggish adjustment of M-goods. Note that, because an increase in technology directly contributes to the increase in output, the increase in labour is roughly 50 to 60% of that in output in the standard RBC model (see Table 1).

The overly low volatility of working hours predicted by the standard RBC model is closely related to the overly high correlation between labour productivity and output. For example, in the standard RBC model, the increase in working hours during a boom is not large relative to the increase in output, and hence output/hours increases during a boom. However, in our model, hours increase enough to decrease output/hours, and hence corr{output/labour, output} becomes negative.

4.2.2 Frequency Analysis

This subsection exploits the band-pass filter developed by Baxter and King (1999) to the simulated data. For the summary, see Table 4. At business cycle frequencies (8-40 quarters), both shocks perform quantitatively well.

At high frequencies (2-3 quarters) the results with M-shocks fail to mimic the data;

the sales volatility relative to output volatility (0.32) is too low and inventory investment is positively correlated to sales (0.32). This finding supports the view that the main driving force of the economy is demand shocks not supply shocks.

On the other hand, F-shocks generate results qualitatively similar to the data, especially for M-firms; inventory investment is negatively correlated to sales (-0.92) and sales is more volatile than output (the relative volatility is 1.02).Intuitively, as the pro-

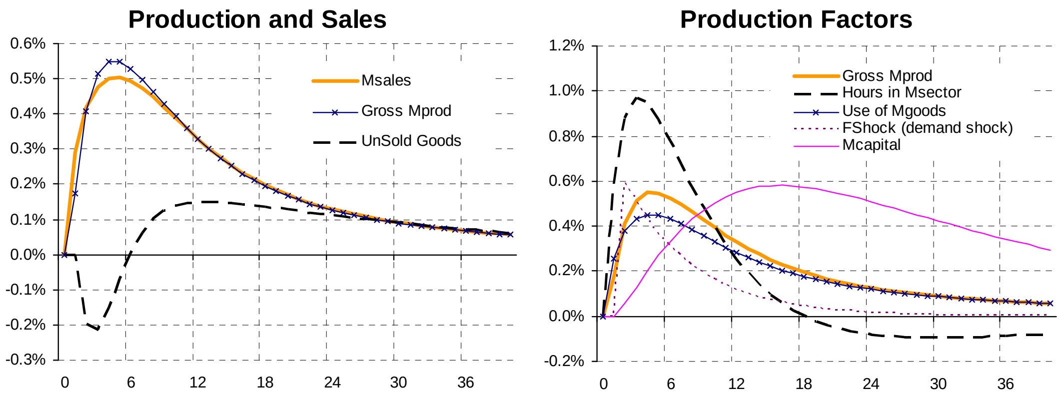


Figure 4: Selected impulse response functions to a positive demand shock (a shock to F-firms' production).

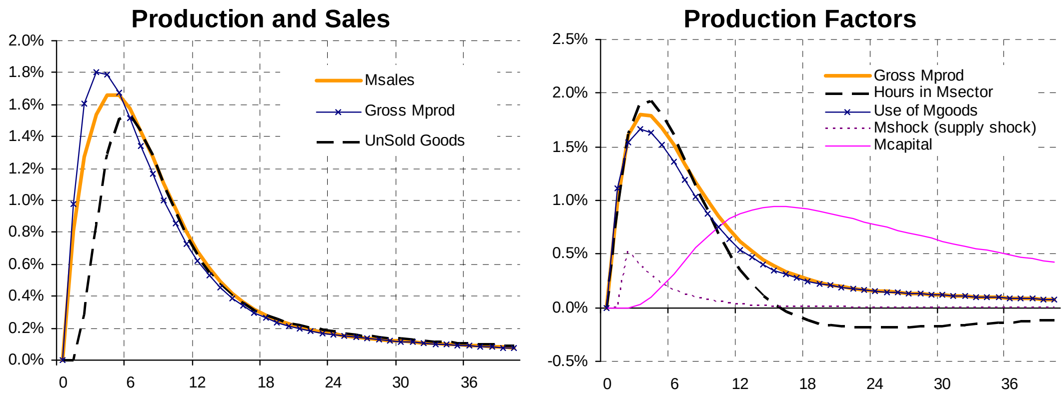


Figure 5: Selected impulse response functions to a positive supply shock (a shock to M-firms' production).

duction smoothing model predicts, inventories work as buffers at high frequencies. Due to the convex cost function, it is costly to change the production level very frequently;

hence firms use inventories as buffers to prevent their production from wildly varying over time.

4.2.3 Impulse Response Functions and Inventory Cycles

Our model has two (or one, depending on parameters) pairs of conjugate complex roots whose absolute values are less than one. Because no impulse response functions exhibit clear oscillations (see Figures 4 and 5), we can say that our model shows over-damped oscillations. Roughly speaking, in our model, there exist a potential mechanism to yield cycles, but it is not strong enough to generate sine waves IRFs.

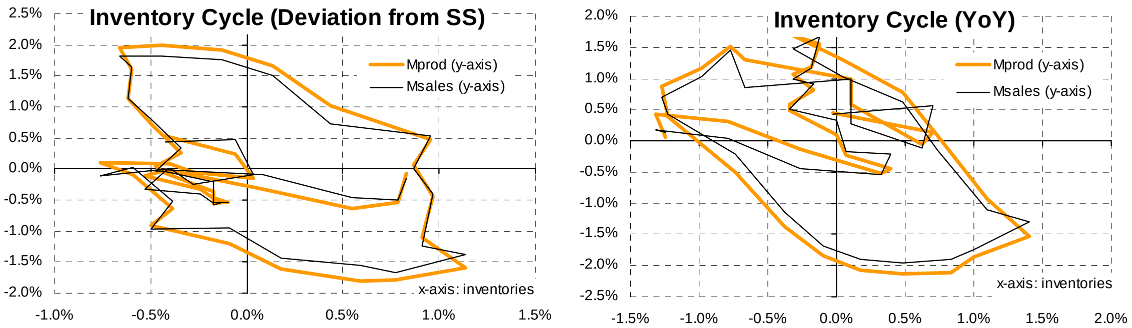


Figure 6: A sample path in phase diagrams generated by shocks to F-firms' production. Simulated data are converted to the year-on-year (YoY) growth rate in the right panel.

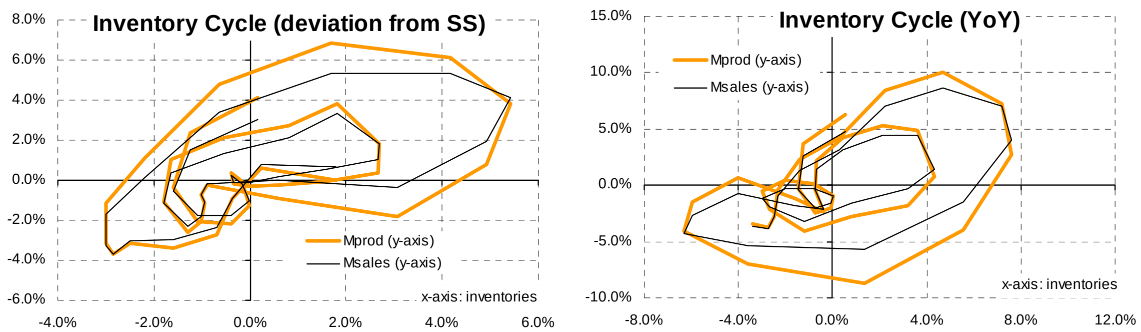


Figure 7: A sample path in phase diagrams generated by shocks to M-firms' production. Simulated data are converted to the year-on-year (YoY) growth rate in the right panel.

However, *in sample paths*, our model yields cycles that are quite similar to the observed inventory cycles (see Figures 6 and 7), although the shape of cycle is not clear with F-shocks. The typical length of cycles (if they exist) seems to be around 15 to 19 quarters, which is somewhat longer than Kitchin cycles (13 quarters), is close to the Japanese post-war average (16.8 quarters), and is shorter than the U.S. post-war average (21 quarters).[41] Importantly, the sample paths with M-shocks (right in Figure 8) show a time lag between peaks and bottoms of production/sales and unsold goods (inventories). Such a time lag, perhaps caused by the slow adjustment of inventories, is called a phase shift. The phase shift between production (or sales) and inventories is important to generate inventory cycles.[42]

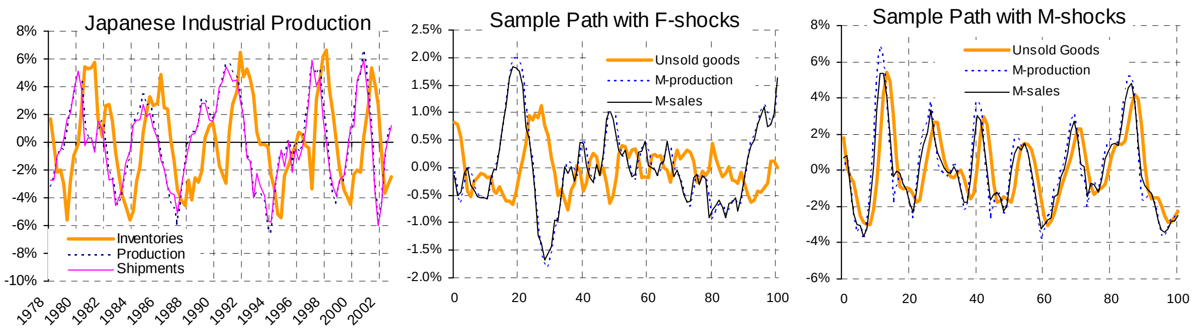


Figure 8: Sample paths of selected variables. Left panel shows the actual data (Japanese industrial production); middle and right panels show samples paths generated by F- and M-shocks, respectively.

4.2.4 Changing Magnitude of Friction

Kahn et al. (2002) and McConnell and Perez-Quiros (2000) argue that the decline in GDP volatility is due to an improvement in inventory management technology. To test this idea, we simulate the model for various values of ν and υ. We interpret an improvement in inventory management as a lower value of ν (smaller magnitude of idiosyncratic shock) or a higher value of υ (a larger portion of today's output that can be sold in today's market). The results are summarised in Figures 9 and 10.

The effects of changing the variation in idiosyncratic shock. The source of aggregate shock is shocks to M-firms. A lower ν (x-axis) implies lower goods market frictions.

[41] For Japanese business cycles, the number is the average of all business cycles See Economic and Social Research Insutitute, Cabinet Office, Government of Japan (2004) and NBER (n.d.).

[42] See Shibayama (2007).

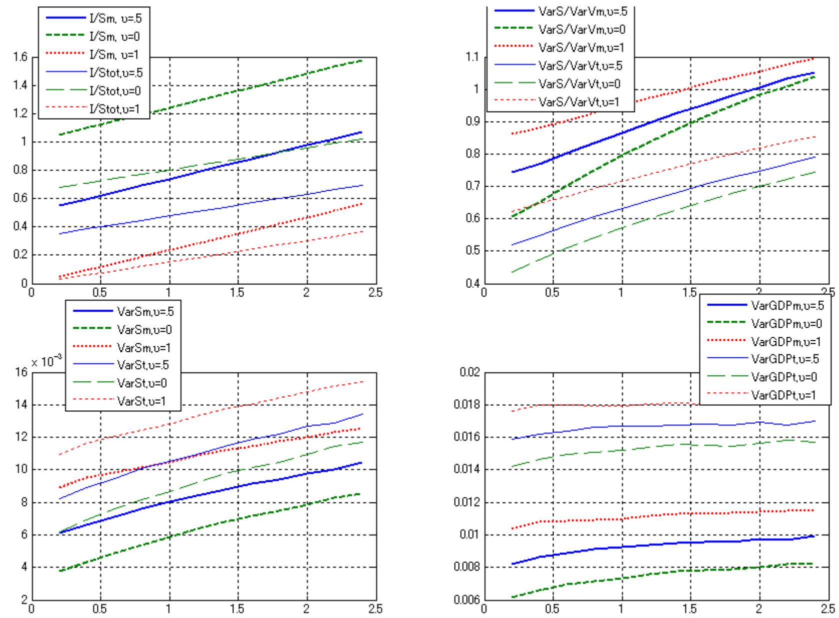


Figure 9: The effects of changing the variation in idiosyncratic shock with demand shocks (shocks to F-firms). A lower ν (x-axis) implies lower goods market frictions. "I/S_M" is unsold goods divided by M-goods sales, and "I/S_{tot}" is unsold goods divided by total sales (M- and F-goods). Results are shown for $\upsilon = 0$, 0.5 and 1.

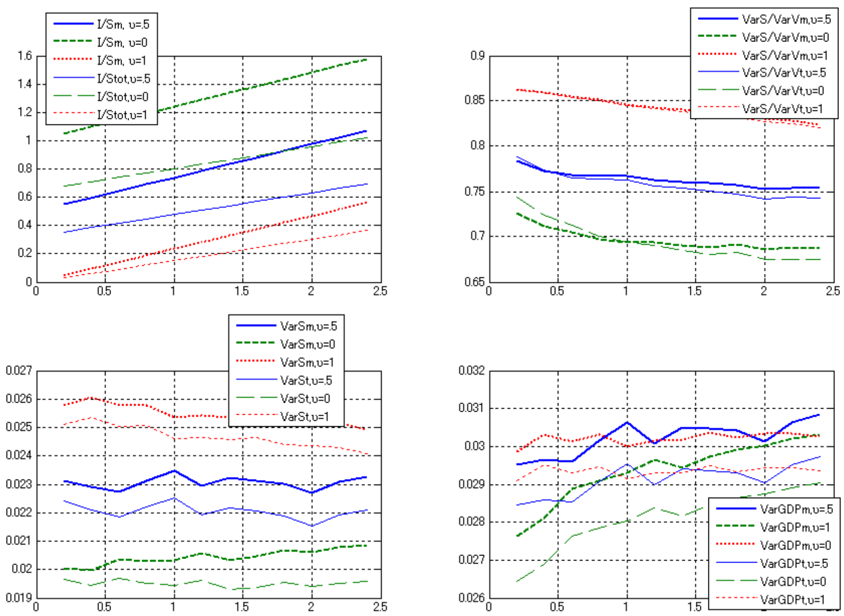


Figure 10: The effects of changing the variation in idiosyncratic shock with supply shocks (shocks to M-firms). A lower ν (x-axis) implies lower goods market frictions. "I/S_M" is unsold goods divided by M-goods sales, and "I/S_{tot}" is unsold goods divided by total sales (M- and F-goods). Results are shown for $\upsilon = 0$, 0.5 and 1.

"I/S_M" is unsold goods divided by M-goods sales, and "I/S_{tot}" is unsold goods divided by total sales (M-goods + F-goods). Results are shown for $\upsilon = 0$, 0.5 and 1.

Changing the magnitude of idiosyncratic shock ν does not significantly change the volatility of GDP in either case (see the lower-right panels). Interestingly, an increase in the portion of today's products that can be sold in today's market υ increases, rather than decreases, GDP volatility for F-shocks, as opposed to their conjecture. This is perhaps because inventories are a stabilising factor at very high frequencies, as shown above. The more quickly M-firms can react to today's demand shocks, the more quickly those shocks are transmitted to M-firms' production.

The I/S ratio decreases when either ν goes down or υ goes up in our experiments. This supports Kahn et al. (2002), in the sense that they regard a declining I/S ratio as evidence for their hypothesis. However, judging from the results of other experiments, it seems that the observed decline in the durable goods sector's I/S ratio is not the cause but the result of the decline in GDP volatility; the less volatile an economy is, the weaker is firms' incentive to hold inventories to hedge their loss of sales opportunities.

Overall, our model shows a negative implication for the hypothesis that an improvement in inventory management is the main reason for the decline in GDP volatility. The key intuition is that inventories are destabilising factors at business cycle frequencies but stabilising factors at very high frequencies; hence, it is uncertain whether holding lower inventories implies a more stable economy.

5 Conclusion

This article investigates a fully rational dynamic stochastic general equilibrium model with a stockout constraint and a production chain. Here, the stockout constraint simply means that no seller can sell goods more than what she holds on the shelf (i.e., inventories), even if she faces a strong demand. The key trade-off in this market friction is that a stockout is costly because it means the loss of a profitable sales opportunity, while having excess inventories is also costly because it imposes a too high carrying cost (financing cost). The production chain means that a firm's product is used as an input by other firms. Our model has two types of firms: final goods producers and intermediate goods producers, both of which take a basket of intermediate goods as production factors. The model constructed in this article is in the class of representative agent models without any price rigidity; however, the intermediate goods market is non-Walrasian.

The model quantitatively satisfies stylised inventory facts. On the one hand, if the source of the shock lies on the supply side, a positive technology shock pushes up production, and such an increase in production leads to an increase in inventories, while sales do not increase very much. This is exactly what the cost shock approach predicts. One the other hand, if the source of the shock lies on the demand side, a positive demand shock increases sales, and inventories *initially* decrease; inventories work as buffers at very high frequencies. However, due to stronger demand, the target level of inventories also increases. *In subsequent periods*, production must increase more than sales, because firms must not only replenish decreased inventories but also accumulate inventories to meet the stronger demand. Because inventories increase as demand increases, inventory investment is procyclical. These results are consistent with the buffer stock view and the micro-founded target inventory models. In this sense, our model supports three leading firm/industry level analyses: cost shock, production smoothing and target inventory models. Note that, while our model explicitly assume the stockout constraint, it does not assume anything for cost shock and buffer stock mechanisms; they both naturally appear in our micro-founded environment.

In addition, due to the production chain, adjustment of aggregate inventories is quite slow. When one firm want to replenish its inventories, it must increase its production. However, such an increase in production must use other firms' inventories as production factors. Hence, the adjustment of inventories is slow *in aggregate*. Note that such a tight intermediate goods market leads to an increase in intermediate goods price; if the change in intermediate goods price is ignored (i.e., the general equilibrium feedback through price is ignored), it may seem easy to adjust inventory level quickly; but a sharp increase in the intermediate goods price discourages firms from using them.

The most important finding in this article is that the stockout constraint and production chain generate a low correlation between labour productivity and output. The key intuition behind this is the slow adjustment of inventories. When a positive shock hits the model economy, firms cannot increase their use of intermediate goods because inventories of intermediate goods cannot adjust swiftly in aggregate; as a result, intermediate goods price increases. Thus, firms are encouraged to substitute their intermediate goods input with more labour input (capital cannot adjust as in the standard RBC model). Although the standard RBC model predicts the low volatility of working hours, our model yields working hours volatile enough to match the data. When output increases, because working hours increase considerably, labour productivity (i.e., output/hours) does not increase very much. Compared to the standard RBC model, the stockout constraint and production chain improve the behaviour of labour without deteriorating other properties of the model.

Appendix

The full derivation is shown in the following. The equation numbers indicated in the MATLAB codes correspond exactly to the equation numbers in Appendix.[43] We use the word "number" instead of "measure" unless there is a risk of confusion.

A Structure of M-goods Markets

This subsection provides the details of technical assumptions.

A.1 Agents Distribute over $[0, 1] \times [0, 1]$ ($\subset \mathbb{R}^2$)

Unlike the standard monopolistic competition models, we assume that agents distribute over a rectangle rather than over a line segment. Specifically, there is a continuum of markets over $[0, 1]$, and there is a continuum of sellers distributed over $[0, 1]$ in each market. In different markets, different varieties (types) of goods are traded; in each market, all sellers sell the same variety of goods (there are one-to-one correspondences between markets and varieties of goods).

In a discrete example, there are, say, 1,000 markets and 1,000 sellers in each market, yielding a total of 1,000,000 sellers. If all sellers behave as buyers at the same time (production chain), then there are 1,000,000 buyers as well. If each buyer visits all markets, then 1,000,000 buyers appear in every market.[44] Thus, each seller in a market meets (on average) 1,000 buyers.[45] Note that, though the discrete example is often used in the sequel, the formal derivation is based on the continuum of agents.

A.2 Idiosyncratic Shock

Next, we assume that buyers do not distribute evenly in each market. That is, some sellers meet many buyers while others meet only a few in every market. The uncertainty

[43] The MATLAB codes used in this paper are available upon request.

[44] Note that this exposition ignores F-firms. If F-firms are taken into account as buyers, then there is a total of 2,000,000 buyers in each market. In the continuous model, the measure of sellers (M-firms) is 1 (in $\mathbb{R}^2$), and the measure of buyers (M-firms plus F-firms) is 2 (in $\mathbb{R}^2$).

[45] Note that in a continuous setting, this means that each seller meets a positive measure of buyers.

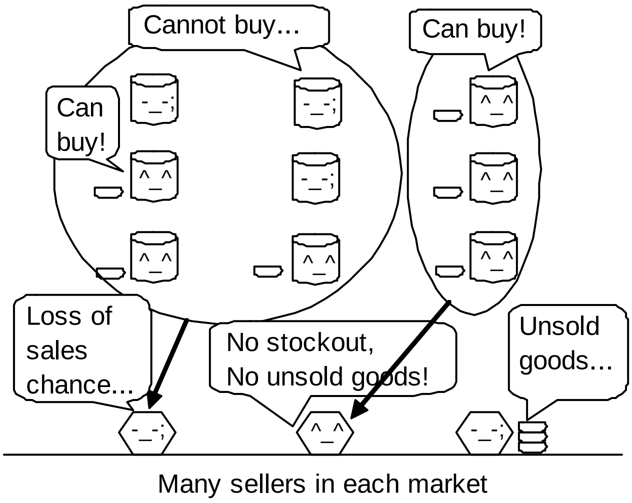


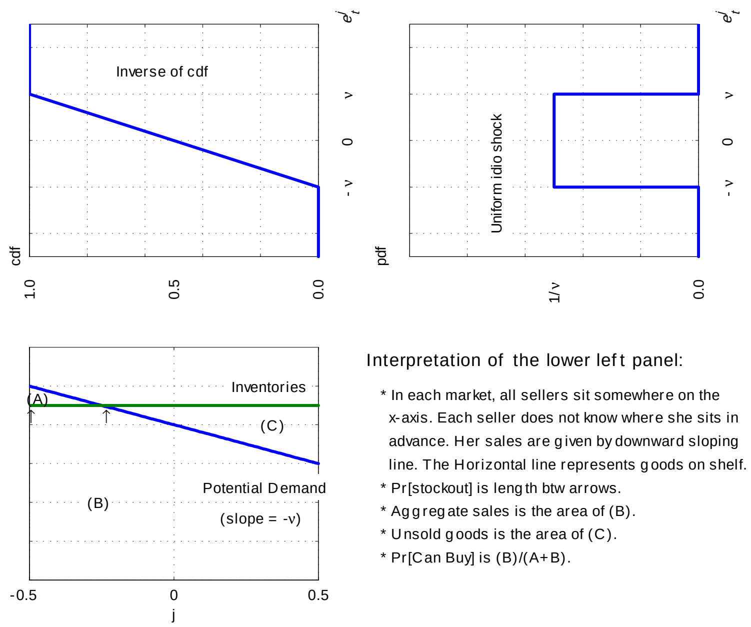
Figure 11: An example with nine buyers and three sellers. Buyers do not distribute evenly over sellers.

in the number of buyers is called an idiosyncratic shock. A simple example is illustrated in Figure 11. It should be clear that the idiosyncratic shock causes the mismatch between buyers and sellers in every market.

From the sellers' viewpoint, if a seller meets more buyers than GoS_t/M_t^b, where M_t^b be the baseline demand (demand per buyer), she faces a stockout; she sells all of goods on her shelf but she loses some of her customers due to the stockout. Otherwise, she has unsold goods U_{t+1} which she carries to the next period. There is a key trade-off between stockout and unsold goods. Having too low GoS_t leads to too high a stockout probability (loss of sales opportunity), but having too high GoS_t leads to too high a carrying cost of U_{t+1}.

In each market, one specific type (variety) of goods are traded. Thus, from the buyers' viewpoint, some buyers, who visit a busy seller in a market, cannot buy that specific type of goods; because we assume imperfect substitution among varieties, these buyers experience a utility cost.[46] Buyers determine M_t^p taking into account such losses in variety (see Appendix A.4 for details).

[46]We assume that, once buyers visit a shop, they cannot visit other shops in the same market. This assumption is necessary to make the idiosyncratic shock meaningful; otherwise, all buyers will buy each variety of goods in the end, reducing our M-goods markets to Walrasian markets.



A.2.1 Uniform Distribution

We assume that the idiosyncratic demand shock follows a uniform distribution.[47] More specifically, we assume that the potential demand for a seller M_t^p is the sum of the baseline demand M_t^b (= demand per buyer) and the idiosyncratic shock e_t^p,[48] where M stands for M-goods.

$$M_t^p = M_t^b + e_t^p, \qquad e_t^p \sim U\left[-\frac{\nu}{2}, \frac{\nu}{2}\right]$$

where ν is the parameter that governs the support, and the variance $(\nu^2/12)$ of the distribution of e_t^p.

A.2.2 Derivation of Key Equations

The easiest way to understand the following results is by examining the graph above. The two panels in the upper half show how to derive the lower right panel; the downward sloping lines in the two left panels are identical, and represents potential demand M_t^p.

[47] Unfortunately, a simple urn-ball analysis implies a degenerate distribution; if buyers visit sellers randomly, all sellers meet an equal number (measure) of buyers.

[48] It could be more natural to assume that e_t^i is the shock on the number of buyers, so that $M_t^p = M_t^i(N_b + e_t^i)$ where N_b is the average number of buyers. However, it turns out that the following computation becomes extremely messy with this specification.

If the number of buyers is normalised to one, the area under this line (i.e., (A) and (B)) is equal to baseline demand M_t^b.

In the lower left panel, the downward sloping line M_t^p shows how buyers distribute over sellers. Each point on the x-axis represents a seller, and the height of the downward sloping line at each point on the x-axis shows the number of buyers who meet that seller. Note that our assumptions about CRS and price posting (see below) guarantee that all sellers hold the same level of GoS_t, which is, thus, represented by the horizontal line in the lower left panel. Hence, area (A) implies that potential demand M_t^p exceeds GoS_t, and thus the area shows unsatisfied (potential) demand. From areas (A) and (B), we can compute the probability that, in the market for each type of good, a buyer can buy that type of good: $\Pr[\text{a buyer can buy a good}] = Q_t = (B)/((A)+(B))$.

From the viewpoint of each seller, she does not know in advance where her position is on the x-axis in the lower left panel before the realisation of the idiosyncratic shock. Hence, the probability that a seller faces a stockout is represented by the line segment between the two arrows in the lower left panel.

Area (C) implies that GoS_t that exceeds M_t^p; they are carried to the next period as unsold goods U_{t+1}. Also, a portion of today's production $(1-\upsilon)\,Y_t^M$ is not placed in today's market. Thus, U_{t+1} equals area (C) plus $(1-\upsilon)\,Y_t^M$. Area (B) shows the aggregate sales S_t, which equals $E[\text{sales of each seller}]$ for each seller.

A.2.3 Key Equations

Therefore, primary school arithmetic yields the following results:

$$\begin{aligned}
\Pr[\text{a seller faces stockout}] &\equiv Pr_t = \frac{M_t^i - GoS_t}{\nu} + \frac{1}{2} && \text{(6a)}\\
\text{aggregate sales of market} &\equiv S_t = GoS_t - \frac{\nu}{2}\left\{\frac{M_t^i - GoS_t}{\nu} - \frac{1}{2}\right\}^2 &&\\
&= E[\text{sales of a seller}] && \text{(6b)}\\
\text{unsold goods} &\equiv U_{t+1} = \frac{\nu}{2}\left\{\frac{M_t^i - GoS_t}{\nu} - \frac{1}{2}\right\}^2 + (1-\upsilon)\,Y_t^M && \text{(6c)}\\
\Pr[\text{a buyer can buy a good}] &= Q_t = \frac{1}{M_t^i}\left\{GoS_t - \frac{\nu}{2}\left\{\frac{M_t^i - GoS_t}{\nu} - \frac{1}{2}\right\}^2\right\} && \text{(6d)}
\end{aligned}$$

Several comments are in order. First, neither M_t^p nor e_t^p appears in these expressions, which implies that the idiosyncratic shocks in all markets average out. Second, because there is a continuum of markets with a unit measure, Pr[a buyer can buy a good] is equal to Q_t, the measure (number) of the available varieties for each buyer. If goods are considered collectively, a low Q_t deteriorates the quality of goods due to imperfect substitution among varieties (see A.4.1 for an intuitive example). Third, because there is a continuum of sellers in each market with a unit measure, and because the measure of market is unity, E [sales of a seller] is equal to the aggregate sales S_t. Fourth, regardless of the distribution assumption, the following relationship must hold:

$$\begin{aligned} U_{t+1} &= GoS_t + (1-\upsilon)\,Y_t^M - S_t \quad \text{where} \qquad GoS_t = U_t + \upsilon Y_T^M \\ Q_t &= S_t / M_t^i \end{aligned}$$

where υ is the portion of today's output that can be placed in today's market. Remember that we exogenously assume that only a portion of today's output can be placed in today's market. Finally, the first term of (6c) represents the unsold goods that cannot be sold due to the idiosyncratic shock (area (C)) and the second term represents goods that are not on sale in today's market.

A.3 Miscellaneous Comments for Assumptions

The idiosyncratic shock is necessary to deal with a kinked constraint; the stockout constraint $S_t = \min\{GoS_t, M_t^p\}$ is not smooth and non-differentiable. However, $E[S_t]$ becomes smooth by adding idiosyncratic shock from the viewpoint of each agent. This technique to smooth non-smooth constraints by adding shocks is not new; it is commonly used in analyses of voting behaviour, and was first used for inventory analysis by Kahn (1987). However, this article shows a nice interpretation: inventories as options to sell (see C.2 for details).

The large number of agents is necessary for aggregation. In terms of sellers, due to the law of large numbers (LLN), aggregate sales equal the expected sales ($S_t = E[S_t]$), which is a smooth function. Hence, we can linearise aggregate S_t. In terms of buyers, Q_t (the number of available varieties = probability of facing stockout) is also a smooth function, because there are infinitely many varieties (LLN).

It is also important to note that we need to confine our focus to the constant returns to scale (CRS) for aggregation. Individual M-firms (sellers) have different levels of U_t carried from the previous period, while the target level of goods on shelf $GoS_t(= U_t + \upsilon Y_T^M)$ is the same for all M-firms, meaning that Y_t varies among M-firms. Hence, if production technology is not CRS, it is not possible to aggregate individual productions.

A.3.1 Timing Assumption

There is another assumption; firms cannot use M-goods they purchase today for today's production. This assumption is logically necessary, especially for M-firms, because M-firms must produce before M-markets open, while they can use M-goods only after M-markets close.[49]

A.4 Monopolistic Competition and Cost of Losing Varieties

A.4.1 Imperfect Competition

In addition, we assume monopolistic competition à la Dixit and Stiglitz (1977). There are two reasons not to assume perfect substitution among varieties. First, if goods were perfect substitutes for each other, buyers would not need to visit all markets. Second, because perfect substitution implies zero profit, no seller wants to hold inventories; sellers earn zero profit from their sales if they can sell their inventories, while they suffer from a carrying cost of unsold goods if they cannot.

In our environment, two-stage budgeting with quantity and price indices still holds. However, as mentioned above, because the number of available varieties fluctuates over time, we need to consider the cost effect of losing varieties.[50]

[49] Certainly, it is possible to assume that F-firms (but not M-firms) produce, say, in the second half of each period, while M-firms produce in the first half. However, it is a bit cumbersome if the timing assumptions differ between F- and M-firms.

[50] Interestingly, one of the main motivations of Dixit and Stiglitz Dixit and Stiglitz (1977) is to analyse

The intuition of the utility cost is as follows. Let us consider a familiar example, say, ice cream. Suppose a consumer prefers vanilla and chocolate ice creams equally, but vanilla and chocolate ice creams are not perfect substitutes for one another. Also suppose that their costs are the same. Then, one vanilla and one chocolate give higher utility than two vanillas, because they are differentiated from one another. However, the costs of vanilla + vanilla and vanilla + chocolate are the same. Thus, given the level of expenditures, having fewer varieties provides lower utility, and vice versa. Or, equivalently, with fewer varieties available, the pecuniary cost of achieving a certain level of utility is higher.

A.4.2 Number of Available Varieties

The cost effect of losing varieties is not, in itself, of interest, and quantitatively its effect seems very weak under the plausible parameter range. It is a logical consequence of the combination of Dixit-Stiglitz monopolistic competition and stockout. Thus, we show only the key results without derivations. Note that they are defined and discussed *from the viewpoint of a buyer*.

First, Q_t^j is defined as an indicator function which is 1 if a buyer can buy the j-th good, and 0 otherwise. Then, the measure of the available varieties Q_t is:

$$
\begin{aligned}
Q_t &= \int_0^1 Q_t^j dj \\
Q_t^j &= \begin{cases} 1 & \text{if } j\text{-th variety is available} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Due to LLN, Q_t has two meanings: the number (measure) of available varieties and the probability that a *buyer* can buy a variety without encountering a stockout. Note that Q_t is a distinct concept from $1 - Pr_t$, the probability that a *seller* does not face a stockout.

firms' entry and exit, explicitly addressing the effect of a changing number of firms (or varieties, in our language).

A.4.3 Price Index

Next, we define the price index of intermediate goods as:

$$P_t^M \equiv \left[\int_0^1 \frac{Q_t^j}{Q_t}\left(P_t^{j\ 1-\theta}\right)dj\right]^{\frac{1}{1-\theta}} = \left[\int_0^1 \frac{Q_t^j}{Q_t}dj \int_0^1 \left(P_t^{j\ 1-\theta}\right)dj\right]^{\frac{1}{1-\theta}} = \left[\int_0^1 P_t^{j\ 1-\theta}dj\right]^{\frac{1}{1-\theta}}$$

where θ is the elasticity of intratemporal substitution among varieties. Several comments are in order. First, (a) multiplying by Q_t^j means that unavailable goods are not taken into account,[51] and (b) dividing by Q_t means that the index is the "average" of individual prices. Second, the integral is factorised as shown by the second equality because Q_t^j and P_t^j are, in a sense, not correlated; P_t^j is assumed to be fixed before the realisation of the idiosyncratic shock (see below), while Q_t^j is not the choice of an agent (determined exogenously by the idiosyncratic shock). Third, at optimum all sellers set the same price (i.e. $P_t^i = P_t^j$ for $\forall i, j \in [0,1]$) due to the price posting and CRS production technology. As a result, $P_t^j = P_t^M$ for $\forall j \in [0,1]$. Indeed, many combinations of definitions of price and quantity indices are logically consistent. We have chosen our definitions so that $P_t^j = P_t^M$ at optimum.

A.4.4 Quality-Adjusted Quantity Index

In this regard, the definition of the quantity index of M-goods that is consistent with our price index is:

$$M_t^K \equiv \left[\int_0^1 Q_t^j \left(M_t^{j\ \frac{\theta-1}{\theta}}\right)dj\right]^{\frac{\theta}{\theta-1}}$$

where $K = F, M$; i.e., M_t^F is the index of M-goods purchased by F-firms, and M_t^M is that of M-firms. Again, there are several comments parallel to the price index. First, multiplying by Q_t^j means that unavailable goods are not taken into account, and (b) *not* dividing by Q_t means that the index is the "sum" of individual quantities. Second, at optimum $M_t^i = M_t^j$ for $\forall i, j \in [0,1]$, because all prices are equal due to symmetricity. Third, it is shown that the baseline demand M_t^j in equations (6) is not an index, but

[51] In general, the price index could be different among buyers, because they have different baskets of goods. However, in our model, the price index is common to all buyers because of LLN.

instead is measured in terms of a physical unit. Thus,

$$M_t^j = Q_t^{\frac{-\theta}{\theta-1}} M_t^F + Q_t^{\frac{-\theta}{\theta-1}} M_t^M \tag{7}$$

since both F- and M-firms use M-goods for their production. Since $Q_t < 1$ and $\theta > 1$, $M_t^j > M_t^F + M_t^M$. In other words, physical demand is larger than the index. This difference becomes larger as Q_t becomes smaller. In this connection, M_t^K can be interpreted as a quality adjusted quantity index – with fewer varieties, the quality of the M-goods index becomes lower. Finally, Q_t and hence M_t^K have the same value for any buyer due to LLN.[52]

A.4.5 Two Stage Budgeting

From these two indices, the expenditure for M-goods of a buyer in sector K can be written as

$$\int_0^1 Q_t^i P_t^i M_t^j dj = Q_t^{\frac{-1}{\theta-1}} P_t^M M_t^K \quad \text{for} \quad K = F, M \tag{8}$$

where the LHS is the direct definition of expenditures on M-goods, and the RHS means that we can restate this definition with price and quantity indices. The first multiplicative term $Q_t^{\frac{-1}{\theta-1}} = Q_t Q_t^{\frac{-\theta}{\theta-1}}$ in (8) represents the cost of losing varieties. This is because, under non-perfect substitution, to achieve a certain level of quantity index, an increase in quantity in each variety must compensate for a loss of varieties (see (7)).

A.5 Price Posting

An important consequence of non-Walrasian intermediate goods markets is that we cannot use the market clearing conditions as a pricing mechanism. Hence, we assume the following price posting rule as an alternative. The rule follows a simple extensive game, in which first sellers set their price, then buyers are distributed among sellers *unevenly* (idiosyncratic shock), and finally buyers choose optimal quantity if they are

[52] Although the exact components of available varieties may differ among buyers (say, some can buy vanilla+strawberry, while others mint+chocolate), the number (measure) of available varieties is the same (2 varieties in this ice cream example).

not subject to a stockout. This extensive game is played in each M-market in every period. We assume that (i) in each market, only one identical variety of goods are traded (varieties and markets are one-to-one correspondences to each other), (ii) in each period, each buyer visits only one seller for each variety (i.e., only one visit in each market), and (iii) even if he fails to buy a variety due to a stockout, he cannot visit other shops in that market.

0. All the aggregate shocks are revealed.

1. Anticipating the buyers' action, sellers set their sales price before the realisation of the idiosyncratic shock. Once a seller decides her price, she cannot change it until the next period (price posting).

2. The idiosyncratic shock is revealed; buyers are distributed among sellers unevenly. As a result, some sellers meet many buyers while others meet only a few.

3. At each shop, all buyers stand in a queue, and then buyers, in order, choose an optimum amount to buy until goods on shelf run out. The order in the queue is stochastic for buyers; a buyer cannot buy the good if goods on shelf run out before his turn. In this case, he simply loses one variety.

A few remarks are in order here. First, due to the assumption that sellers set their sales price before observing the idiosyncratic shock, and the assumption of constant returns to scale, all sellers choose the same sales price. Second, the measure of available goods varies over time but, in each period, the LLN guarantees that all buyers enjoy the same measure of available varieties, although the varieties' components differ among agents. Third, analytically this price posting rule implies that sellers take buyers' demand function as a given, while the buyers take the M-price as a given. Algebraically, we first obtain the FOC w.r.t the use of M-goods for *each* M-price, and then we obtain the FOC w.r.t. M-price subject to the demand function. Note that (i), individual sellers cannot deprive other sellers' customers in our market structure (ii) sellers exploit the slope of the demand curve as monopolists, and (iii) the quantity traded is not socially optimum.[53]

Finally, the resulting pricing is a slightly generalised version of the markup formula in the standard Dixit-Stiglitz monopolistic competition model. Namely, there exists $\tilde{\theta}$ such that $P_t^M = \tilde{\theta}/\left(\tilde{\theta}-1\right)\lambda_t^M$, where λ_t^M is the marginal cost of producing M-goods and $\tilde{\theta} \gtrless \theta$ is the elasticity of substitution that is adjusted by Q_t and Pr_t.

B Optimisation Problems of Individual Agents

This section defines the optimisation problems of individual agents.

B.1 Household

The infinitely-lived representative household (HH) maximises its expected lifetime utility.

$$\max_{\{C_s, H_s^H\}_{s=0}^{\infty}} E_0\left[\sum_{t=0}^{\infty} \beta^t U\left[C_t, 1-H_t^H\right]\right]$$

s.t.

$$C_t + B_{t,t+1} = R_{t-1,t}B_{t-1,t} + W_t H_t^H + D_t^{iv}$$

The period utility $U[.,.]$ is time additive, is discounted by the subjective discount factor β^t, and takes consumption C_t and leisure $1-H_t^H$ as arguments, where the total time endowment is normalized to one and H_t^H is the labour supply.

The period budget constraint has cash outflow in the LHS and inflow in the RHS. The LHS means that HH spends its resources on consumption and one-period bonds $B_{t,t+1}$, while the RHS implies that cash inflows are the sum of bond redemption $R_{t-1,t}B_{t-1,t}$, wage income $W_t H_t^H$ and dividends D_t^{iv}.[54]

HH takes the real interest rate $R_{t-1,t}$, wage rate W_t and D_t^{iv} as givens. All the first order conditions (FOCs) are quite standard.

[53] This is not only because of the price posting, but also because of externalities (see C.3).

[54] Alternatively, we can assume that there are infinitely many HHs which own both F- and M-firms. In that case, dividends are assumed to be state contingent, and thus all household enjoy the same level of cash inflow; as a result, whole HHs reduce to one sector in aggregate.

B.1.1 Functional Form

Throughout this article, we assume the following functional form for the period utility.

$$U\left[\ C_t, 1-H_t^H\ \right] = (1-\psi)\frac{\left(\ C_t\ \right)^{1-\gamma}}{1-\gamma} + \psi\frac{\left(\ 1-H_t^H\ \right)^{1-\gamma_L}}{1-\gamma_L}$$

where ψ is the weight for leisure, and γ and γ_L are the elasticities of intertemporal substitutions of consumption and leisure, respectively. When $\gamma_L = 0$, our utility function reduces to Hansen's indivisible labour model.

B.2 Firms

We assume that quadratic adjustment costs apply to changing labour demand and input of M-goods, as well as investment.

B.2.1 M-Firms' Optimization Problem

As shown in Appendix A, we can exploit the slightly modified two-stage budgeting ($\int_0^1 Q_t^j P_t^j M_t^j dj = Q_t^{\frac{-1}{\theta-1}} P_t^M M_t^M$).

$$\max E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{\lambda_t^H}{\lambda_0^H}\left\{\begin{array}{c} P_t^i S_t - W_t H_t^{Mp} - I_t^M - Q_t^{\frac{-1}{\theta-1}} P_t^M M_t^M \\ -\chi_{MH}\left(\ H_t^{Mp} - H_{t-1}^{Mp}\ \right)^2 / H_{t-1}^{Mp} - \chi_{MM}\left(\ M_t^M - M_{t-1}^M\ \right)^2 / M_{t-1}^M \end{array}\right\}\right]$$

s.t.

$$\begin{aligned}
U_{t+1} &= U_t - S_t + Y_t^M \\
S_t &= \min\left\{U_t + \upsilon Y_t^M, M_t^p\right\} \\
Y_t^M &= Y^M\left[K_t^M, H_t^{Mp}, M_{t-1}^M; \mathbf{Z}_t^M\right] \\
K_{t+1}^M &= (1-\delta_M) K_t^M + I_t^M - \chi_{MK}(I_t^M - \delta_M K_t^M)^2 / K_t^M
\end{aligned}$$

The objective function says that M-firms maximise the present value (PV) of their net cash inflows, which are discounted by the stochastic discount factor $SDF_t = \beta^t \lambda_t^H / \lambda_0^H$ $= \beta^t \left(\partial U_t / \partial C_t\right) / \left(\partial U_0 / \partial C_0\right)$. The cash inflow is only the sales revenue $P_t^i S_t$, where P_t^i is the sales price of producer i. While sales price is a choice variable, the purchase price P_t^M is given for all agents, though $P_t^i = P_t^M$ for $\forall i$ in equilibrium. On the other hand, cash outflow is composed of the wage payment $W_t H_t^{Mp}$, which is wage rate W_t times labour hours H_t^{Mp}, the expenditure on investment goods I_t^M (the price of F-goods is normalized to 1) and the expenditure on M-goods $Q_t{}^{\frac{-1}{\theta-1}} P_t^M M_t^M$, where Q_t is the number of available varieties, and P_t^M and M_t^M are the price and quantity indices of M-goods, respectively. In addition, the adjustment costs of labour and M-goods inputs $\chi_{MH} \left\{ H_t^{Mp} - H_{t-1}^{Mp} \right\}^2$ $/ H_{t-1}^{Mp}$ and $\chi_{MM} \left(M_t^M - M_{t-1}^M \right)^2 / M_{t-1}^M$ also constitute M-firms' cash outflow. χ_{MH} and χ_{MM} are both given parameters. These costs are evaluated in terms of F-goods. In sum, the net cash inflow is the sales revenue minus expenditure on labour, investment goods and M-goods, as well as the adjustment costs.

The first constraint is the evolution of unsold goods. The second represents the stockout constraint; sales S_t is the minimum of GoS_t or potential demand M_t^p. Note that M_t^p is the sum of the baseline demand and idiosyncratic shock in our notation. The third constraint shows the production function, in which $\mathbf{Z}_t^M$ is exogenous shocks. The production function takes capital K_t^M, labour H_t^{Mp} and M-goods M_{t-1}^M as production factors. The fourth constraint is the evolution of capital, in which we assume a quadratic adjustment cost, where δ_M and χ_{MK} are given parameters.

B.2.2 F-firms' Optimization Problem

The optimization problem of F-firms is as follows.

$$\max E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{\lambda_t^H}{\lambda_0^H} \left\{ \begin{array}{c} Y^F \left[K_t^F, H_t^{Fp}, M_{t-1}^F; Y_t^{Ftot}, \mathbf{Z}_t^F \right] - W_t H_t^{Fp} \\ -Q_t^F{}^{\frac{-1}{\theta-1}} P_t^M M_t^F - I_t^F \\ -\chi_{FH} \left(H_t^{Fp} - H_{t-1}^{Fp} \right)^2 / H_{t-1}^{Fp} \\ -\chi_{FM} \left(M_t^F - M_{t-1}^F \right)^2 / M_{t-1}^F \end{array} \right\} \right]$$

s.t.

$$K_{t+1}^{F}=(1-\delta_{F})\,K_{t}^{F}+I_{t}^{F}-\chi_{FK}(I_{t}^{F}-\delta_{F}K_{t}^{F})^{2}/K_{t}^{F}$$

The objective function again says that firms maximize the PV of their net cash inflows. The modified version of two-stage budgeting holds, as in the case of M-firms. $W_t H_t^{Fp}$ and I_t^F refer to labour costs and expense on investment, respectively. The production of final goods Y_t^F takes capital K_t^F, labour H_t^{Fp} and M-goods M_{t-1}^F as production factors, where the superscript F implies F-firms. $\chi_{FH}\left(H_t^{Fp}-H_{t-1}^{Fp}\right)^2/H_{t-1}^{Fp}$ and $\chi_{FM}\left(M_t^F-M_{t-1}^F\right)^2/M_{t-1}^F$ denote the adjustment costs of labour and M-goods, respectively, in which χ_{FH} and χ_{FM} are given parameters.

The constraint represents the evolution of capital with the quadratic adjustment cost. Note that, in this formulation, the level of capital in the steady state is not affected by the parameter χ_{FK}, which governs the adjustment cost of investment.

B.2.3 Functional Form

We assume a CES production function with a Hicks-neutral technology shock $\mathbf{Z}_t^K = Z_t^{Kn}$. For $K = F, M$,

$$\begin{aligned}
Y_t^K &= Y^K\left[K_t^K, H_t^{Kp}, M_{t-1}^K; \mathbf{Z}_t^K\right] \\
&= Z_t^{Kn}\left[\phi_K\left(\frac{V_t^K}{\phi_K}\right)^{\frac{\eta_K-1}{\eta_K}}+(1-\phi_K)\left(\frac{M_{t-1}^K}{1-\phi_K}\right)^{\frac{\eta_K-1}{\eta_K}}\right]^{\frac{\eta_K}{\eta_K-1}} \\
V_t^K &= {K_t^K}^{\alpha_K}\,{H_t^{Kp}}^{1-\alpha_K}
\end{aligned}$$

where ϕ_K is the share parameter of the value-added component and η_K is the elasticity of substitution between the value-added component and M-goods as inputs. The value-added component V_t^K is assumed to be a Cobb-Douglas function, in which the share of capital is α_K.Parameters ϕ_K, η_K and α_K are exogenous.

C Analytical Results

This section summarises the analytical results (see also Section 3 in the main text).

C.1 Equilibrium

There are 26 endogenous variables and 26 equations (excluding the law of motions of exogenous shocks), of which four variables and four equations (H^{Mp}_{t-1}, H^{Fp}_{t-1}, M^{M}_{t-1} and M^{F}_{t-1}) are merely lagged variables and their definitions due to the adjustment costs. With proper initial and terminal conditions, these equations define the equilibrium.

Omitting lagged variables and their definition equations, this subsection summarises the 22 equations. See Table 2 on page 24 for the list of variables used.

Two equations are derived from the FOCs of the representative household's optimisation.[55]

$$\beta \frac{\partial U_t}{\partial C_t} = \beta \lambda^H_t = SDF_t \tag{9a}$$

$$W_t = \frac{\partial U_t / \partial L_t}{\partial U_t / \partial C_t} \tag{9b}$$

Nine equations come from the FOCs and constraints of M-firms' optimisation.

$$E_t \left[\beta \frac{\lambda^H_{t+1}}{\lambda^H_t} \left\{ \lambda^M_{t+1} \frac{\partial Y^M_{t+1}}{\partial K^M_{t+1}} + \frac{(1-\delta_M) + \chi_{MK} \left(\left(I^M_{t+1}/K^M_{t+1} \right)^2 - \delta^2_M \right)}{1 - 2\chi_{MK} \left(I^M_{t+1}/K^M_{t+1} - \delta_M \right)} \right\} \right] = \frac{1}{1 - 2\chi_{MK} \left(I^M_t / K^M_t - \delta_M \right)} \tag{10a}$$

[55] We omit the equation for the real interest rate $R_{t-1,t}$

$$\beta E_t \left[\frac{\partial U_{t+1} / \partial C_{t+1}}{\partial U_t / \partial C_t} R_{t,t+1} \right] = 1$$

because SDF_t and $R_{t-1,t}$ move in exactly the same way in the linearised model.

$$\chi_{MH}\left\{E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\left\{\left(\frac{H_{t+1}^{Mp}}{H_t^{Mp}}\right)^2-1\right\}\right]-2\left\{\frac{H_t^{Mp}}{H_{t-1}^{Mp}}-1\right\}\right\}+\lambda_t^M\frac{\partial Y_t^M}{\partial H_t^{Mp}}=W_t \tag{10b}$$

$$\chi_{MM}\left\{\begin{array}{c} E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\left\{\left(\frac{M_{t+1}^M}{M_t^M}\right)^2-1\right\}\right] \\ -2\left\{\frac{M_t^M}{M_{t-1}^M}-1\right\}\end{array}\right\}+E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\lambda_{t+1}^M\frac{\partial Y_{t+1}^M}{\partial M_t^M}\right]=Q_t^{M\,\frac{-1}{\theta-1}}P_t^M \tag{10c}$$

$$S_t=\theta Pr_t^-\left(1-\frac{\lambda_t^M}{P_t^M}\right)\left(Q_t^{F\,\frac{-\theta}{\theta-1}}M_t^F+Q_t^{M\,\frac{-\theta}{\theta-1}}M_t^M\right) \tag{10d}$$

$$\begin{aligned}\left(Pr_t^++Pr_t^-\right)\lambda_t^M &= \upsilon P_t^iPr_t^++E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\left\{(1-\upsilon)\,Pr_{t+1}^+P_{t+1}^i+Pr_{t+1}^-\lambda_{t+1}^M+c_1\right\}\right] && (10e)\\ \text{where } Pr_t^+ &\equiv Pr_t/\left(1-\upsilon Pr_t\right) \text{ and } Pr_t^-\equiv\left(1-Pr_t\right)/\left(1-\upsilon Pr_t\right) && (10f)\end{aligned}$$

$$\begin{aligned} Y_t^M &= Z_t^{Mn}\left[\phi_M\left(\frac{V_t^M}{\phi_M}\right)^{\frac{\eta_M-1}{\eta_M}}+(1-\phi_M)\left(\frac{Z_t^{Mm}M_{t-1}^M}{1-\phi_M}\right)^{\frac{\eta_M-1}{\eta_M}}\right]^{\frac{\eta_M}{\eta_M-1}} && (10g)\\ V_t^M &= Z_t^{Mv}\,{K_t^M}^{\alpha_M}\,{H_t^{Mp}}^{1-\alpha_M} && (10h)\\ K_{t+1}^M &= \left(1-\delta_M\right)K_t^M+I_t^M-\chi_{MK}(I_t^M-\delta_MK_t^M)^2/K_t^M && (10i)\\ U_{t+1} &= U_t-S_t+Y_t^M && (10j)\end{aligned}$$

Six equations come from the FOCs and the constraints of F-firms.

$$\begin{aligned} & E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\left\{\frac{\partial Y_{t+1}^F}{\partial K_{t+1}^F}+\frac{(1-\delta_F)+\chi_{FK}\left(\left(I_{t+1}^F/K_{t+1}^F\right)^2-\delta_F^2\right)}{1-2\chi_{FK}\left(I_{t+1}^F/K_{t+1}^F-\delta_F\right)}\right\}\right]\\ = & \frac{1}{1-2\chi_{FK}\left(I_t^F/K_t^F-\delta_F\right)} \end{aligned} \tag{11a}$$

$$\chi_{FH}\left\{E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\left\{\left(\frac{H_{t+1}^{Fp}}{H_t^{Fp}}\right)^2-1\right\}\right]-2\left\{\frac{H_t^{Fp}}{H_{t-1}^{Fp}}-1\right\}\right\}+\frac{\partial Y_t^F}{\partial H_t^{Fp}}=W_t \tag{11b}$$

$$\chi_{FM}\left\{\begin{array}{c} E_t\left[\beta\frac{\lambda^H_{t+1}}{\lambda^H_t}\left\{\left(\frac{M^F_{t+1}}{M^F_t}\right)^2-1\right\}\right] \\ -2\left\{\frac{M^F_t}{M^F_{t-1}}-1\right\}\end{array}\right\}+E_t\left[\beta\frac{\lambda^H_{t+1}}{\lambda^H_t}\frac{\partial Y^F_{t+1}}{\partial M^F_t}\right]=Q^{F\,\frac{-1}{\theta-1}}_t P^M_t \tag{11c}$$

$$Y^F_t = Z^{Fn}_t\left[\phi_F\left(\frac{V^F_t}{\phi_F}\right)^{\frac{\eta_F-1}{\eta_F}}+(1-\phi_F)\left(\frac{Z^{Fm}_t M^F_{t-1}}{1-\phi_F}\right)^{\frac{\eta_F-1}{\eta_F}}\right]^{\frac{\eta_F}{\eta_F-1}} \tag{11d}$$

$$V^F_t = Z^{Fv}_t\ K^F_t{}^{\alpha_F}\ H^{Fp}_t{}^{1-\alpha_F} \tag{11e}$$

$$K^F_{t+1} = (1-\delta_F)K^F_t+I^F_t-\chi_{FK}(I^F_t-\delta_F K^F_t)^2/K^F_t \tag{11f}$$

Two equations are the market clearing conditions for labour and F-goods. Because all adjustment costs other than investments are measured in terms of F-goods, they are deducted from the market clearing condition for the final goods.

$$H^H_t = H^{Mp}_t+H^{Fp}_t \tag{12a}$$

$$Y^F_t = C_t+I^M_t+I^F_t-AdjC_t \tag{12b}$$

$$\begin{aligned} AdjC_t &= \chi_{FH}\frac{(H^{Fp}_t-H^{Fp}_{t-1})^2}{H^{Fp}_{t-1}}+\chi_{MM}\frac{\left(M^M_t-M^M_{t-1}\right)^2}{M^M_{t-1}} \\ &\quad +\chi_{MH}\frac{(H^{Mp}_t-H^{Mp}_{t-1})^2}{H^M_{t-1}}+\chi_{FM}\frac{\left(M^M_t-M^M_{t-1}\right)^2}{M^M_{t-1}}+c_1U_t \end{aligned} \tag{12c}$$

Three equations are derived from the specification of the idiosyncratic shock (6).

$$S_t = GoS_t-\frac{\nu}{2}\left\{\frac{Q^{F\,\frac{-\theta}{\theta-1}}_t M^F_t+Q^{M\,\frac{-\theta}{\theta-1}}_t M^M_t-GoS_t}{\nu}-\frac{1}{2}\right\}^2 \tag{13a}$$

$$(S_t = \min\left\{U_t+\upsilon Y^M_t, M^p_t\right\}\ \text{for individual sellers})$$

$$\begin{aligned} Pr_t &= \Pr\left[\text{a seller faces stockout}\right] \\ &= \frac{Q^{F\,\frac{-\theta}{\theta-1}}_t M^F_t+Q^{M\,\frac{-\theta}{\theta-1}}_t M^M_t-GoS_t}{\nu}+\frac{1}{2} \end{aligned} \tag{13b}$$

$$U_{t+1} = \frac{\nu}{2}\left\{\frac{Q^{F\,\frac{-\theta}{\theta-1}}_t M^F_t+Q^{M\,\frac{-\theta}{\theta-1}}_t M^M_t-GoS_t}{\nu}-\frac{1}{2}\right\}^2+(1-\upsilon)Y^M_t \tag{13c}$$

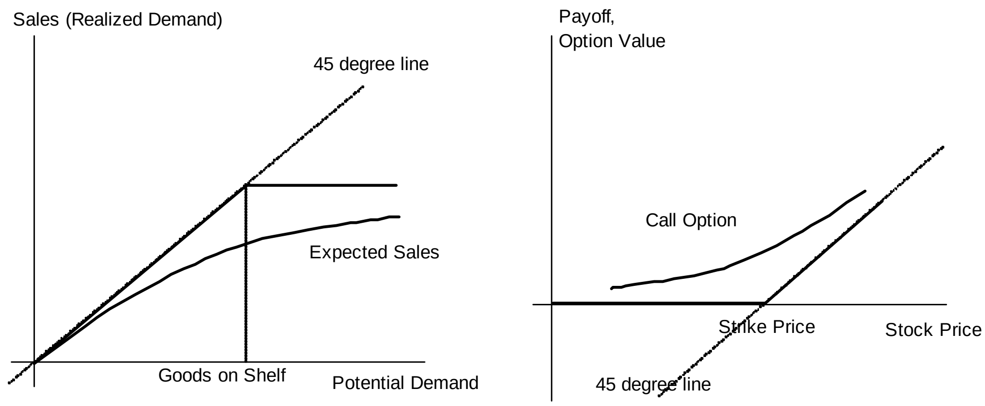


Figure 12: Comparison between a financial option and inventories.

In a sense, (13) is the alternative to the market clearing condition of (the index of) intermediate goods. In the limit $\nu \to 0$ (i.e., no idiosyncratic shock), if $\upsilon = 1$ (all products today can be sold in today's market), (13a) and (13c) show that $U_{t+1} = 0$ (no unsold goods) and $M_t^i = GoS_t$ (M-markets clear), where $M_t^i = Q_t^{F\frac{-\theta}{\theta-1}} M_t^F + Q_t^{M\frac{-\theta}{\theta-1}} M_t^M$.

The last two equations show the law of motions of exogenous shocks. In the basic version, we use only AR(1) Hicks-neutral technology shocks in intermediate and final goods productions.

$$\begin{aligned} \ln Z_t^{Mn} &= \ln Z_{t-1}^{Mn} + \xi_t^{Mn} \\ \ln Z_t^{Fn} &= \ln Z_{t-1}^{Fn} + \xi_t^{Fn} \end{aligned}$$

where ξ_t^{Mn} and ξ_t^{Fn} are *iid* innovations that follow proper normal distributions.

C.2 Inventories as Options to Sell

This subsection discusses the key trade-off in the stockout model: the FOC with respect to unsold goods (10e). Assume, for simplicity, that $\nu = c_1 = 0$. Then, (10e) reduces to

$$E_{t-1}\left[SDF_t\left\{ \left(P_t^i - \lambda_t^M\right) \Pr\left[GoS_t < M_t^p\right] \right\}\right] = \lambda_{t-1}^M - E_{t-1}\left[SDF_t \lambda_t^M \right] \tag{14}$$

where λ_t^M is the marginal cost of producing M-goods (Lagrange multiplier for the law of motion of unsold goods), $SDF_t = \beta\lambda_t^H/\lambda_{t-1}^H$ is the stochastic discount factor, $P_t^i - \lambda_t^M$ is the marginal profit margin (P_t^i is the sales price of seller i), and $\Pr[GoS_t < M_t^p] = \partial E[S_t]/\partial U_{t-1}$ is the stockout probability from the viewpoint of individual sellers. This equation states that the carrying cost of one additional unit of inventory (RHS) is equal to the expected value of the marginal cost of the lost sales opportunity (LHS).

Equivalently, we can treat inventories as financial assets in the asset pricing equation,

$$E_{t-1}\left[SDF_t\left\{\frac{P_t^i \Pr[GoS_t < M_t^p] + \lambda_t^M \Pr[GoS_t > M_t^p]}{\lambda_{t-1}^M}\right\}\right] = 1 \tag{15}$$

Note that the inside of the curly bracket shows the gross return on having one more unit of unsold goods.

It is important to note that the expression $\Pr[GoS_t < M_t^p]$ is essentially equivalent to an "option delta" in finance;[56] having one more unit of inventory means having an option to sell one more unit (see Figure 12). In this sense, inventories have a feature similar to options on financial assets. While an option delta is defined as the sensitivity of the option price to a change in the underlying stock price in finance, $(P_t^i - \lambda_t^M)\Pr[GoS_t < D_t^p]$

[56] Remember that the delta of a call option is

$$\Delta_c = \Phi\left[\frac{(s + r\tau - k)}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}\right]$$

where s is the log of the underlying stock price today, k is the strike price of the option, r is the (constant) risk-free rate, τ is the time to maturity, σ is the volatility and Φ is a (standard normal) distribution function. (One way to understand the term $\sigma\sqrt{\tau}/2$ is Jensen's inequality. The Black-Scholes model assumes a log-normal, rather than normal, distribution for stock price.)

We can see the following correspondences: value of holding inventories (value of option), derivative of the expected profit w.r.t. inventories (option delta), and demand change (price change of underlying stock) relative to the inventory holdings (strike price). The correspondence of $(P_t^M - \lambda_t^M)$ is always 1 in the case of a call option, because a 1-pound increase in stock price trivially leads to a 1-pound increase in payoff, if the stock price at the exercise date is higher than the strike price. Remember that, if the potential demand is less than goods on shelf, 1 unit of increase in the potential demand leads to an increase in profit by $(P_t^M - \lambda_t^M)$.

Related to the importance of the CRS assumption, note that, ignoring the effect of Jensen's inequality, $s + r\tau$ represents the expected stock price at the exercise date under the equivalent martingale measure (in the risk neutral world, the stock price must grow at the same rate as the risk free rate); hence, the option delta can be regarded as the probability that the stock price exceeds the strike price under the risk neutral measure. The real world probability measure should be changed to the equivalent martingale measure because investors are risk averse. However, such a change of measure is not necessary in our model, because, roughly speaking, our CRS assumption (with some other technical assumptions) implies that sellers are risk neutral. So we can use *risk neutral pricing* without changing the measure.

is the sensitivity of profit to a change of GoS_t.[57]

C.3 Search Externalities

There are search externalities in M-markets.

On the buyers' side, each buyer ignores the negative effect of *congestion*. Intuitively, if buyers buy more, then available varieties ($Q_t = \Pr[\text{can buy}]$) become fewer because stockouts arise more often, but infinitesimal buyers ignore such an effect. In our model, FOC w.r.t M-goods input is

$$E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\frac{\partial Y_{t+1}^M}{\partial M_t^M}\right] = Q_t^{\frac{-1}{\theta-1}} P_t^M \tag{18}$$

However, if there were, say, a strong union of purchasing managers, which coordinated buyers' decisions, the FOC w.r.t M_t^M would be

$$E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\frac{\partial Y_{t+1}^M}{\partial M_t^M}\right] = Q_t^{\frac{-1}{\theta-1}} P_t^M\left(1+\frac{-1}{\theta-1}\frac{\partial Q_t/Q_t}{\partial M_t^M/M_t^M}\right) \tag{19}$$

which implies that the social cost (RHS of (19)) is larger than the private cost (RHS

[57] Certainly, it is potential demand rather than inventories that is stochastic, but we can show the following result:

$$\frac{\partial}{\partial U_t}E\left[\min\{GoS_t, M_t^p\}|\tilde{\Omega}_t\right] = 1-\frac{\partial}{\partial M_t^i}E\left[\min\{GoS_t, M_t^p\}|\tilde{\Omega}_t\right] \tag{16}$$

Since, as mentioned above,

$$\begin{aligned}\frac{\partial}{\partial U_t}E\left[\min\{GoS_t, M_t^p\}|\tilde{\Omega}_t\right] &= \Pr[GoS_t < M_t^p]\\ \frac{\partial}{\partial M_t^p}E\left[\min\{GoS_t, M_t^p\}|\tilde{\Omega}_t\right] &= \Pr[GoS_t > M_t^p]\end{aligned}$$

it is clear that (16) is equivalent to

$$\Pr[GoS_t < M_t^p] = 1-\Pr[GoS_t > M_t^p] \tag{17}$$

Thus, the first derivative of the expected sales w.r.t. unsold goods means one minus a *decrease* in the expected sales due to an *increase* in the underlying demand. Here, in the equation (17) "one" means that, without the stockout constraint, one unit of increase in demand would trivially lead to one unit of increase in sales, but, due to the second term ($\partial E[\min\{GoS_t, M_t^p\}|\tilde{\Omega}_t]/\partial M_t^i =$ effect of the probability of stockout), the incremental expected sales must be smaller than they would be without the constraint. Therefore, we can restate our claim more precisely; the first derivative of sales with respect to inventories means *a reduction in the loss of sales opportunity* by holding one more unit of inventories.

of (18)). The additional term shows the effect of congestion, which infinitesimal buyers ignore.

On the sellers' side, if there were a powerful union of sellers which coordinated sellers, the FOC w.r.t. unsold goods of intermediate goods producers would be

$$E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\left\{\begin{array}{c}\left(P_{t+1}^i-\lambda_{t+1}^{M2}\right)Pr_{t+1}\\ +\frac{M_{t+1}^M}{U_{t+1}}\left(\frac{\partial M_{t+1}^M/M_{t+1}^M}{\partial Q_{t+1}/Q_{t+1}}-\frac{\theta}{\theta-1}Q^{\frac{-\theta}{\theta-1}}\right)\frac{\partial Q_{t+1}/Q_{t+1}}{\partial U_{t+1}/U_{t+1}}\left(1-Pr_{t+1}\right)\end{array}\right\}\right]$$
$$=\lambda_t^{M2}-E_t\left[\beta\frac{\lambda_{t+1}^H}{\lambda_t^H}\lambda_{t+1}^{M2}\right]$$

but infinitesimal sellers ignore two effects. The first is *the cost of losing varieties* ($\frac{M_{t+1}^M}{U_{t+1}}\frac{\partial M_{t+1}^M/M_{t+1}^M}{\partial Q_{t+1}/Q_{t+1}}\frac{\partial Q_{t+1}/Q_{t+1}}{\partial U_{t+1}/U_{t+1}}\geq 0$). When inventories are higher, the measure of varieties that a buyer can enjoy is larger; hence the effective cost is lower, which, in turn, stimulates the demand for M-goods. However, such a mechanism is ignored. The second is the *squeezing effect* due to fewer varieties ($-\frac{M_{t+1}^M}{U_{t+1}}\frac{\theta}{\theta-1}Q^{\frac{-\theta}{\theta-1}}\frac{\partial Q_{t+1}/Q_{t+1}}{\partial U_{t+1}/U_{t+1}}\leq 0$). As mentioned in the previous subsection, when fewer varieties are available, the (physical unit of) potential demand of one buyer becomes larger to achieve a certain level of the quantity index. These two effects offset one another; the net effect may be positive or negative.

Nonetheless, some numerical experiments suggest that the overall effect of the search externalities seems to be very small.

References

Abel, Andrew B., "Inventories, Stock-Outs and Production Smoothing," *Review of Economic Studies*, April 1985, *52* (2), 283–93.

Basu, Susanto, "Procyclical Productivity: Increasing Returns or CyclicalUtilization?," *Quarterly Journal of Economics*, August 1996, *111* (3), 719–752.

Baxter, Marianne and Robert G. King, "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, November 1999, *81* (4), 575–93.

Bencivenga, Valerie, "An Econometric Study of Hours and Output Variation with Preference Shocks," *International Economic Review*, May 1992, *33* (2), 449–71.

Bental, Benjamin and Benjamin Eden, "Inventories in a Competitive Environment," *Journal of Political Economy*, October 1993, *101* (5), 863–86.

Bils, Mark, "Material for "Studying Price Markups from Stockout Behavior" (Very Preliminary and Very Incomplete)," *Working paper*, December 2004.

_ **and James A. Kahn**, "What Inventory Behavior Tells Us about Business Cycles," *American Economic Review*, June 2000, *90* (3), 458–81.

Blanchard, Olivier J., "The Production and Inventory Behavior of the American Automobile Industry," *Journal of Political Economy*, June 1983, *91* (3), 365–400.

Blinder, Alan S., "Inventories and Sticky Prices: More on the Microfoundations of Macroeconomics," *American Economic Review*, June 1982, *72* (3), 334–48.

_ , "Can the Production Smoothing Model of Inventory Behavior Be Saved?," *Quarterly Journal of Economics*, August 1986, *101* (3), 431–53.

_ **and Louis J. Maccini**, "Taking Stock: A Critical Assessment of Recent Research on Inventories," *Journal of Economic Perspectives*, 1991, *5* (1), 73–96.

Boileau, Martin and Marc-Andre Letendre, "Inventories, Sticky Prices and the Propagation of Nominal Shocks," *Working paper*, 2004.

Bruno, Michael, "Raw Materials, Profits, and the Productivity Slowdown," *Quarterly Journal of Economics*, February 1984, *99* (1), 1–30.

Bureau of Economic Analysis, U.S. Department of Commerce, "Benchmark Input-Output Accounts," 2007. http://www.bea.gov/industry/index.htm.

Caballero, Ricardo J. and Eduardo M. R. A. Engel, "Dynamic (S, s) Economies," *Econometrica*, November 1991, *59* (6), 1659–1686.

Caplin, Andrew S., "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *Econometrica*, November 1985, *53* (6), 1395–1410.

Christiano, Lawrence J. and Martin Eichenbaum, "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review*, June 1992, *82* (3), 430–450.

Diamond, Peter, "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, October 1982, *90* (5), 881–94.

_ **and Drew Fudenberg**, "Rational Expectations Business Cycles in Search Equilibrium," *Journal of Political Economy*, June 1989, *97* (3), 606–19.

Dixit, Avinash K. and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, June 1977, *67* (3), 297–308.

Economic and Social Research Insutitute, Cabinet Office, Government of Japan, "The Determination of Business Cycle Peak and Trough," November 2004. http://www.esri.cao.go.jp/en/stat/di/041112rdates.html.

Fisher, Jonas D. M. and Andreas Hornstein, "(S, s) Inventory Policies in General Equilibrium," *Review of Economic Studies*, January 2000, *67* (1), 117–45.

Fitzgerald, Terry J., "Inventories and the Business Cycle: An Overview," *Federal Reserve Bank of Cleveland Economic Review*, 1997, *33* (3).

Gali, Jordi, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, March 1999, *89* (1).

Gertler, Mark and Simon Gilchrist, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *Quarterly Journal of Economics*, May 1994, *109* (2).

Hall, George and John Rust, "An empirical model of inventory investment by durable commodity intermediaries," *Carnegie-Rochester Conference Series on Public Policy*, June 2000, *52.*

Hornstein, Andreas, "Inventory Investment and the Business Cycle," *Federal Reserve Bank of Richmond Economic Quarterly*, Spring 1998, *84* (2), 49–71.

_ **and Pierre-Daniel G. Sarte**, "Sticky Prices and Inventories : Production Smoothing Reconsidered," *Federal Reserve Bank of Richmond Working Paper*, 2001, (01-6).

Kahn, James A., Margaret M. McConnell, and Gabriel Perez-Quiros, "On the Causes of the Increased Stability of the U.S. Economy," *FRBNY Economic Policy Review*, May 2002, pp. 183–206.

Kahn, James J., "Inventories and the Volatility of Production," *American Economic Review*, September 1987, *77* (4), 667–79.

_ , "Why Is Production More Volatile Than Sales? Theory and Evidence on the Stockout-Avoidance Motive for Inventory-Holding," *Quarterly Journal of Economics*, May 1992, *107* (2), 481–510.

Kashyap, Anil K., Owen A. Lamont, and Jeremy C. Stein, "Credit Conditions and the Cyclical Behavior of Inventories," *Quarterly Journal of Economics*, August 1994, *109* (3).

Khan, Aubhik and Julia K. Thomas, "Inventories and the business cycle: an equilibrium analysis of (S,s) policies," *Federal Reserve Bank of Philadelphia Working Papers*, 2004a, (11).

_ **and** _ , "Modeling inventories over the business cycle," *Federal Reserve Bank of Philadelphia, Working Papers*, 2004b, (04-13).

McConnell, Margaret M. and Gabriel Perez-Quiros, "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?," *American Economic Review*, December 2000, *90* (5), 1464–1476.

Ministry of Internal Affairs and Communications, Government of Japan, "Input-Output Tables for Japan," 2004. http://www.stat.go.jp/english/data/io/index.htm.

Mosser, Patricia C., "Trade Inventories and (S,s)," *Quarterly Journal of Economics*, November 1991, *106* (4), 1267–86.

NBER, "US Business Cycle Expansions and Contractions." http://www.nber.org/cycles.html.

Pindyck, Robert S., "Inventories and the Short-Run Dynamics of Commodity Prices," *RAND Journal of Economics*, Spring 1994, *25* (1), 141–59.

Ramey, Valerie A., "Inventories as Factors of Production and Economic Fluctuations," *American Economic Review*, June 1989, *79* (3), 338–54.

_ , "Nonconvex Costs and the Behavior of Inventories," *Journal of Political Economy*, April 1991, *99* (2), 306–334.

_ **and Daniel J. Vine**, "Tracking the Source of the Decline in GDP Volatility: An Analysis of the Automobile Industry," *NBER*, 2004, (10384).

_ **and Kenneth D. West**, "Inventories," *NBER*, 1997, (6315).

Rotemberg, Julio J. and Michael Woodford, "Imperfect Competition and the Effects of Energy Price Increases on Economic Activity," *NBER working paper*, June 1996, (5634).

Shibayama, Katsuyuki, "Periodicity of Inventories," 2007.

Wen, Yi, "Understanding the Inventory Cycle," *CAE Working Paper*, 2002, (No.02-04).

West, Kenneth D., "A Variance Bounds Test of the Linear Quadratic Inventory Model," *Journal of Political Economy*, April 1986, *94* (2).