Kahn, James A.

Working Paper
Durable goods inventories and the great moderation

Staff Report, No. 325

Provided in Cooperation with:
Federal Reserve Bank of New York

Suggested Citation: Kahn, James A. (2008) : Durable goods inventories and the great moderation, Staff Report, No. 325, Federal Reserve Bank of New York, New York, NY

This Version is available at:
http://hdl.handle.net/10419/60793

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Federal Reserve Bank of New York
Staff Reports

Durable Goods Inventories and the Great Moderation

James A. Kahn

Staff Report no. 325
May 2008

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in the paper are those of the author and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the author.
Durable Goods Inventories and the Great Moderation
James A. Kahn
*Federal Reserve Bank of New York Staff Reports*, no. 325
May 2008
JEL classification: D21, E20, N12

Abstract

This paper revisits the hypothesis that changes in inventory management were an important contributor to volatility reductions during the Great Moderation. It documents how changes in inventory behavior contributed to the stabilization of the U.S. economy within the durable goods sector, in particular, and develops a model of inventory behavior that is consistent with the key facts about volatility decline in that sector. The model is calibrated to evidence from survey data showing that lead times for materials orders in manufacturing shrank after the early 1980s. Simulations of the model show large reductions in the volatility of output growth and more modest reductions in the volatility of sales growth. In addition, the model addresses concerns raised by a number of researchers who criticize the inventory literature’s focus on finished goods inventories, given that stocks of works-in-process and materials are actually larger and more volatile than those of finished goods. The model adapts the stockout-avoidance concept to a production-to-order setting and shows that much of the intuition and results regarding production volatility still apply.

Key words: Great Moderation, inventories, volatility

Kahn: Federal Reserve Bank of New York and Stern School of Business, Economics Department, New York University (email: jkahn@stern.nyu.edu). The author thanks Meg McConnell and Steve Davis for beneficial comments, and Doug Elmendorf, Nick Bloom, and Valerie Ramey for comments on an earlier version of the paper. The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.
The large reduction in the volatility of GDP and other economic aggregates since the early 1980s, now commonly referred to as the “Great Moderation,” has spawned a number of hypotheses about its cause. But there have been surprisingly few truly structural explanations as an alternative to the hypothesis that volatility just exogenously fell. The primary alternative to the “good luck” view is that monetary policy changed discretely in the early 1980s and stabilized both output and inflation.¹ As discussed in Davis and Kahn (2008), this theory has a number of weaknesses: First, given the importance of credibility and transparency in determining monetary policy’s impact on the economy, it is implausible that the results of a policy change could have been felt so rapidly as to bring about an immediate drop in volatility. Second, some have argued that the policy adopted in the 1980s was not fundamentally different from that of the 1950s and early 1960s, when GDP volatility was considerably higher than it was post-1983.² In addition, structural econometric models of aggregate output and inflation have typically found that monetary policy has most of its impact on inflation volatility, and comparatively modest effects on the volatility of GDP growth.³ Hence there is little scope in these models for policy changes to account quantitatively for the much of the reduction in real volatility—especially if one compares the 1950s and 1960s to the post-1983 era.

This paper revisits the hypothesis, first conjectured in McConnell and Perez-Quiros (1999) and Blanchard and Simon (2001), that changes in inventory management were an important contributor to volatility reductions. The first part of the paper documents the disproportionate statistical contribution of changes in inventory behavior to the stabilization of the U.S. economy, within the durable goods sector in particular. The remainder of the paper develops a model of inventory behavior that is consistent with the key facts about the volatility decline in that sector. The model is calibrated to evidence from survey data showing that lead times for materials orders in manufacturing shrank after the early 1980s. Simulations of the model show that it implies large reductions in the volatility of output growth, and more modest reductions in the volatility of sales growth.

The model also addresses concerns raised by a number of researchers (e.g. Blinder and Maccini, 1991; Humphries et al., 2001) who criticize the inventory literature’s focus on finished goods inventories, given that stocks of work-in-process and materials are actually larger and more volatile than those of finished goods. In addition, durable goods manufacturing in particular consists primarily of industries best characterized as production-to-order industries rather than production-to-stock (see Belsley, 1969). One contribution of the model, aside from its contribution to understanding the Great Moderation, is to adapt the stockout-avoidance concept (as in Kahn, 1987) to a production-to-order setting, and to show that much of the intuition and results regarding production volatility still apply. In particular, the model shows that the need to make

¹See, for example, Clarida, Gali, and Gertler (2000).
²See Romer and Romer (2000).
³See, for example, Stock and Watson (2002); Ahmed et al. (2004).
production-related decisions (in this case materials orders and work-in-process) in advance of information about final product demand results in production being more volatile than shipments. This holds even though the model has no finished goods inventories at all.

Stock and Watson (2002) also express skepticism regarding the role of inventory management in the Great Moderation. But their analysis focuses on four-quarter growth rates of economic time series, a transformation that essentially filters out the higher frequency volatility that is more likely to be associated with inventory investment. Indeed, it should be no surprise that inventory behavior appears less important in explaining the (smaller) volatility declines of four-quarter growth rates. Still, looking at four-quarter growth rates might be a reasonable thing to do if changes in high frequency volatility were economically uninteresting or played a small role in the Great Moderation, but this paper will otherwise. Stock and Watson’s other grounds for skepticism—the fact that sales volatility declined, and that most inventories in manufacturing are raw materials or work-in-progress—are addressed in the next section.

It should be emphasized that the claim of this paper is not that better inventory management accounts for all of the Great Moderation. Rather it argues that it can explain a significant part of it, where “significant” means that absent the effects described in the paper, the Great Moderation might not have been so Great. But undoubtedly other factors came into play as well that can account for some of the volatility reductions.

1 The Reduced Volatility of the U.S. Economy: Revisiting the Facts

1.1 Overview

The seemingly sudden decline in the volatility of U.S. real GDP growth in the early 1980s provided the initial impetus for research on The Great Moderation. Early findings of a discrete break in volatility around 1983 (McConnell and Perez-Quiros, 2000) encouraged a focus on comparisons before and after 1983. This approach conceals the fact that many economic series did not undergo an abrupt volatility drop around 1983. Some did so much earlier, some later. The “sudden drop” view also directs attention away from factors that may play an important role in the long term decline of volatility. Structural shifts in the economy such as the rising share of services in aggregate output are unlikely to produce an abrupt drop in aggregate volatility but may contribute to gradual reductions in volatility over time..Blanchard and Simon (2001), in fact, argue, that the drop in volatility in the early 1980s was really just a return to a longer-term downward trend after an unusual period of turbulence in the 1970s.

The reality is that volatility does appear to have been trending downward for much of the postwar era;
yet there was also a relatively sudden decline in the early 1980s. Figure 1 depicts the volatilities, defined by rolling five-year standard deviations of annualized growth rates of various components of GDP, scaled by their nominal shares in GDP. Each chart includes the volatility of GDP itself for comparison. In order to omit from the sample the exceptionally high volatility of the Korean War years, these charts cover the period 1954-2004. The first four charts (Figure 1a) depict the volatilities of four “major product” categories: durable goods, nondurable goods, services, and structures. The second set (Figure 1b) does the same for six spending categories: Consumption spending (separately for nondurables plus services and durables), fixed investment (separately for nonresidential and residential), inventory investment, and government spending.

It is clear from Figure 1a that of the four product categories, only in durable goods did volatility change in much the same way as GDP volatility—both in terms of magnitude and timing—as GDP. Nondurables output volatility dropped, but it had also been lower in the 1960s before increasing in the 1970s, and in any case it was never anywhere nearly as volatile as durables. Service sector output was also never nearly as volatile as durable goods output, and moreover, its volatility dropped substantially in the early 1960s, and again in the 1970s, long before the break in GDP volatility. Structures output did experience a drop in volatility at the same time as overall GDP, but the size of the sector and the magnitude of the contribution is modest.4

On the expenditure side, inventory investment—despite a GDP share of less than one percent—stands out as the key component with changes in volatility similar to those of GDP. It is also noteworthy that consumption volatility appears to have trended modestly lower throughout the 50-year period depicted in the top panels of Figure 1b. The permanent income model predicts that consumption expenditure responds primarily only to real interest rates and to highly persistent “permanent” shocks to income. Thus, the downward trend in volatility, particularly in the nondurables and services component to which the permanent income model is most applicable, suggests either 1) there has been a steady decline in either real interest rate volatility or low frequency income volatility during the 50-year period depicted in the chart; or 2) changes in financial markets such as reduced credit market frictions have gradually increased consumers’ ability to smooth consumption expenditures. The fact that government spending volatility has also trended downward points toward the first explanation, since government spending is more likely to follow the permanent income model (in the sense of responding more to low frequency changes in revenues than to transitory ones) versus being constrained by financial market imperfections.

Figure 1 is of course just accounting, and does not prove cause and effect. It is possible that the decline in GDP volatility caused the decline in durable goods output volatility, or in inventory investment volatility,

---

4 Note that the volatility contributions depicted in the charts are also affected by trends in sector shares over time, but the effect is very slight. The pictures would look virtually identical if sector shares were held constant.
or that all three had a common cause. Still, a challenge for any explanation of the overall decline in GDP volatility is to account for the patterns observed in Figure 1, as well as the more detailed facts regarding inventories and durable goods to be discussed below.

As another illustration of the evolution of volatility over this time period, we estimate a GARCH process for GDP growth, and for durables output growth, including time trends and other variables to explain changing volatility. The results are depicted in Table 1. The growth rates are assumed to be AR(1) processes apart from their time-varying volatilities. The variance equation includes the standard GARCH terms, plus a time trend, and the trend squared (given that the variance has to remain positive). Also included in some specifications was the (lagged) 10-year treasury bond rate, to proxy for inflation and the volatility of the 1970s, and a dummy variable that takes on the value of one for the observations beginning in 1984Q1. The results show a significant downward trend in the variance, with the 10-year rate also coming in significantly, but with the post-1983 dummy not significant when added to the equation. Thus once one accounts for the volatility trend, and the uptick in volatility in the 1970s and early 1980s, it would appear that the post-1983 decline in volatility is better represented as the continuation of a longer-term trend than as a one-time break.

To summarize: The large reduction in volatility that occurred in the early 1980s took place against the background of a long-term modest downward trend in volatility, though punctuated by cyclical increases in volatility associated with recessions, especially during the 1970s. Much of both the downward trend in volatility and the relatively sudden drop appears closely linked to changes in the volatility in the durable goods sector and in inventory investment.

1.2 Frequency Domain Analysis

The term “volatility” encompasses a variety of concepts: Conditional and unconditional, for example, or volatility at different frequencies (e.g. low, business cycle, and high), where the intermediate business cycle frequencies are conventionally defined as cycles of between 6 and 32 quarters. Looking at growth rates tends to emphasize higher frequency volatility, whereas detrended series (assuming they are stationary around the trend) includes more business cycle volatility. To see the difference, consider the GDP and durables sector volatility statistics in Table 2. The sample is divided into three periods: 1954-69, 1970-83, 1984-2007. Growth rate volatility is not very different across the first two subsamples, which is why they are often considered one long high-volatility era. But the detrended logarithmic levels data tell a somewhat different story. Volatility, especially for GDP, appears to have increased considerably in the 1970-83 period. In fact,

5The logarithms of the series were detrended using the HP filter with parameter 1600. Similar results were found using other detrending methods.
by this measure GDP volatility in the earlier 1954-69 period is closer to volatility in the period of the Great
Moderation than to that of 1970-83.

What does this mean? It suggests that higher frequency volatility was relatively stable in the 1954-83
period, whereas business cycle volatility increased in the second half of that period. We can formalize this
by looking at the data in the frequency domain. Following Ahmed, Levin, and Wilson (2004, hereafter
ALW) we divide the frequencies into three ranges as described above, but into three sample periods rather
than the standard two. Because of the relatively short duration of the subsamples, we also consider break
the frequency range into just two intervals, high and low, with the dividing point at $2\pi/12$, i.e. cycles of 12
quarters duration.

Because volatilities tended to be lower after 1983 across all frequency ranges, ALW emphasize the behavior
of the “normalized” spectrum, i.e. the values obtained by integrating over an interval but dividing by the
integral over the entire range $[0, \pi]$. Thus, if the entire spectrum just shifts down proportionally (ALW’s null
hypothesis) then the values of the normalized spectrum integrated over any interval do not change. ALW
analyze quarterly growth rates of a variety of series. To include more of the lower frequency variation in the
data, we analyze the series detrended as in Table 2. The results of this exercise are shown in Table 3. Note
that these are variances, so the magnitudes of the changes are exaggerated relative to Table 2. Although
most of the changes are not statistically significant, the general pattern of “medium, high, low” volatility at
the low and business cycle frequencies, and “high, high, low” at the higher frequencies. Focusing on the
last two columns for each series, the division of the frequency range above and below $2\pi/12$, we see that
only lower frequency volatility increased between the 1954-69 and 1970-83 periods. Comparing 1954-69
with 1984-2007, the declines in lower frequency volatility are considerably more modest than the declines
in high frequency volatility—significantly so. Although what we are calling “high” is a bit broader than
the conventional definition, at cycles of under three years duration it still omits the bulk of what one would
normally think of as business cycle variation.\footnote{2\pi/12 was not the highest cutoff at which significant results were obtained. For example, the same qualitative results obtained with a split at 2\pi/10.}

These results suggest that the middle period was exceptional in its high volatility at lower frequencies.
This fact contributes to ALW’s finding that for most of the series they examine, including GDP growth, they fail to reject their null hypothesis of a proportional reduction in volatility at all frequencies. Even
for the detrended log level data considered here, when the sample is split only between 1983 and 1984,
the data cannot reject that null hypothesis for GDP, and only marginally rejects it (the $p$-value is exactly
0.10) for durable goods output. The point is that when the 1970-83 period is lumped together with the
earlier subsample, it results in a larger reduction in lower frequency volatility than if that period is treated
separately.

Overall this analysis supports the idea that the volatility reductions associated with the Great Moderation are disproportionately attributable to changes at higher frequencies, so long as the 1970-1983 period is treated as exceptional. This provides additional impetus for looking at changes in inventory behavior, as inventory investment disproportionately adds volatility at higher frequencies.

1.3 Changing Inventory Behavior

Since the early 1980s there have been significant changes in the behavior of inventories in aggregate data. Here we focus on the durable goods sector. While the inventory literature has traditionally focused on more disaggregated data, and in particular on the 2-digit (SIC) level manufacturing data, for the questions examined in this paper it is more appropriate to look data that are “vertically” aggregated, i.e. that include inventories at all stages of production and final output. Disaggregated data can be misleading because it is impossible to tell whether changes in inventory behavior are genuine or just the result of a change in location or ownership of inventories. For example, if a manufacturer reduces its holdings of finished goods inventories, and instead speeds up its shipments to wholesalers or retailers, that would appear as a decline in manufacturing finished goods inventories, even though it could be largely offset by an increase in wholesale or retail inventories. Similarly, if manufacturers in one industry were to insist on “just-in-time” delivery of materials from suppliers, that could look like a dramatic decline in materials inventories for manufacturers, but it would likely be offset by increases in the inventories of suppliers, who could be from different industries or from outside of manufacturing entirely.

Apart from the issue of structural change, since durable goods manufacturing industries tend to be “production-to-order” rather than “production-to-stock,” focusing on manufacturing’s slice of the supply chain means overlooking a substantial component of inventory behavior in durable goods production. Thus, according to data from the Bureau of Economic Analysis, inventories in durable goods manufacturing industries represent only about 35 percent of inventories in the durable goods sector as a whole. (A related, and subtler, point is that the distinction between output and sales or shipments in production-to-order industries is not as clear cut.) And while it is possible that even the vertically aggregated data could be vulnerable to similar criticisms—for example, if materials inventories were held by foreign suppliers—research by Alessandria et al (2008) suggests this is unlikely to be a significant issue. They find evidence that firms that rely on foreign suppliers of materials actually hold larger inventories, due to the longer time and larger fixed costs of replenishing their stocks. Nonetheless, for the sake of completeness we will examine both types of data.

An important fact about the Great Moderation is that output volatility fell by substantially more than

---

7See, for example, Kahn et al (2002).
final sales volatility, particularly in the durable goods sector. Since the difference between output and final sales is the change in inventories, this fact implies a change in inventory behavior—either a reduction in the volatility of inventory investment, or a change in the covariance between inventory investment and sales. Note that by the conventions if the NIPAs, the service and structures sector do not carry inventories (in structures this is because final output includes construction in progress), so they do not contribute directly to these changes in inventory behavior. Figure 2 shows the behavior of output and sales volatility over time in the durable goods sector. In contrast to the behavior of output volatility, sales volatility shows only a modest decline.10

Given our focus on the volatility of real growth rates (as opposed to levels), we can examine the relationship between output, inventories, and sales in terms of growth contributions. Although inventory investment, because it can be negative, does not have a conventionally defined growth contribution, we can define it indirectly as the difference between the growth rate of output and the growth contribution of final sales (cf. Kahn et al, 2002). Following Whelan (2000) we can approximate the latter in terms of the real growth rate of sales and the nominal share of sales in output. Letting $\gamma_{xy}$ denote the growth contribution of $x$ to output $y$, where $x = s$ for sales and $x = i$ for inventories, we define the growth contribution of inventory investment as

$$\gamma_{iy} = \gamma_{yy} - \gamma_{sy}$$

where $\gamma_{sy} = \gamma_{ss} \theta_{sy}$, $\theta_{sy}$ is the nominal share of $s$ in $y$ (measured as the average of current and lagged shares).

In this notation, the growth contribution of a variable to itself is just its own real growth rate.

With these definitions in hand, we can track the contributions of sales and inventory investment to the variance of output growth over time:

$$\sigma_y^2 = \sigma_s^2 + \sigma_i^2 + 2\sigma_{si}$$

where the variances and covariance on the right-hand side refer to the growth contributions defined above. Figure 3 plots the three components for the durable goods sector. We see that both the inventory term and especially the covariance term exhibit a substantial downward trend, with the covariance term in particular accounting for the big drop in the early 1980s. Thus, not only is the apparent break in 1984 associated with a change in inventory behavior, but the downward trend from the 1950s onward is as well.

In addition to this indirect evidence of changing inventory behavior, we can directly examine the inventory-sales ratio in the durable goods sector. Figure 4 shows that whether one looks at the ratio for durable goods

---

9 There is, however, evidence of a change in inventory behavior in residential construction, even though it is not treated as such in the NIPAs. See Kahn (2000).
output or durable goods manufacturing, and for the latter whether or not one includes materials stocks in the numerator (due to concerns about increased foreign trade affecting this number), the ratio began a downward trend in the early 1980s that continued until at least the mid-1990s. This is not by itself a proof of "progress;" it could just represent a shift along a fixed technological tradeoff in response to changing costs (e.g. changes in real interest rates, which were quite low in the 1970s), or a compositional change within the sector. It could even be the result of reduced volatility that is somehow exogenous to the firms, though such an explanation begs the question of why volatility declined. But the timing of the break in trend is striking, and fits with the broader pattern of changing inventory behavior.

In addition to declining, the inventory-sales ratio is clearly less volatile (relative to its varying trend), consistent with the idea that businesses either make smaller mistakes or are able to correct their inventories more quickly. Kahn et al. (2002) also describes results from a VAR with sales and inventories that indicates a change in the variance decomposition pre- and post-1983. Before 1983 sales accounted for much more of the variance of inventories than inventories did of sales (37.8 percent versus 5.4 percent); after 1983 they were almost even (18.2 versus 14.9), consistent with the idea that firms were better able to anticipate sales and adjust inventories in advance. Moreover, the residual variance of sales dropped precipitously, meaning that less of the variation in sales was unpredicted given prior sales and inventories. Again this is not definitive; it is possible that the shocks are smaller or that the industry composition has shifted.

Although we have focused on the durable goods sector as a whole, for reasons emphasized earlier, we can examine disaggregated manufacturing data as well. This helps to alleviate concerns that the patterns in the aggregate sector are somehow misleading, either because they stem from compositional change (e.g. relative growth of less volatile industries within the sector) or are unrepresentative of a broad range of subsectors. Table 4 shows the volatility of production and sales growth for 2-digit durable goods manufacturing industries over the periods 1967-83 and 1984-1997. (The data are not available prior to 1967, and after 1997 the industry classifications were changed, so the series are not continuous.) The table shows a similar pattern across all eleven industries: a large reduction in the volatility of both output and sales growth. The only qualitative difference with the NIPA aggregate durables data is that for many industries the decline in sales volatility is approximately as large as the decline in output volatility. This suggests that some of the change in inventory behavior may occur more downstream in the wholesale or retail sectors. But the reduced volatility is clearly not due to compositional change, nor is it confined to a small subset of industries.
2 A Model of Durable Goods Production and Inventory Behavior

One approach to assessing the role of improved inventory control is to be agnostic about the details, but look for changes in parameters and propagation in, for example, a structural VAR. This is the approached in McCarthy and Zakrajsek (2007). They do find evidence of structural change pre- and post-1983. They use conventional identifying restrictions in an effort sort out the role of, for example, monetary policy in altering the dynamics of the sales process.

A second approach is a more specific model of improved inventory control as in Kahn et al (2002), based on the approach in Kahn (1986) and Bils and Kahn (2000). Firms carry finished goods inventories to avoid stockouts in the face of uncertain demand, trading off the cost of foregone profits against the cost of carrying inventories. If demand is serially correlated, the mistakes will get magnified in production volatility, so that it will exceed the volatility of sales. If technology enables firms to have better information about demand disturbances, then they will make smaller errors in their production decisions, and the additional volatility induced by correcting those errors is reduced. Firms may also be able to hold fewer inventories.

This type of mechanism can account for reduced production volatility (relative to the volatility of sales), but has several drawbacks. First, depending on the timing of the arrival of information, either the volatility of sales actually increases substantially, or the covariance of sales with inventory investment increases. As we have seen, the opposite is the case in the data. The reason sales volatility increases in this model is that the improved information essentially allows firms to accommodate demand shocks as opposed to damping them via stockouts. The covariance of sales and inventory investment only becomes more negative if the firm gets the information in time to adjust production sufficiently in advance (due to a desire to smooth production if costs are convex) that inventory movements anticipate the demand shock. Then when the shock occurs, inventory investment moves in the opposite direction, as anticipated by the firm. But this tends to exacerbate sales volatility.

The second problem with this approach, as alluded to in the introduction, is that it does not apply so obviously or directly to the durable goods sector, much of which is best characterized as production-to-order rather than production-to-stock. And as pointed out by Humphries et al (2001) and many others, most inventories, particularly in durable goods, are of materials or works in process, not final goods. Third, while there is much anecdotal evidence of technology that might provide better information about future sales, there is no direct evidence to assist in specifying a model. And as this discussion suggests, the details matter.

Because the model involves at least two types of stocks (works-in-process or “intermediate” goods, and unfilled orders of final goods), and will distinguish between materials orders and deliveries, a lot of notation
is involved. Let $D_t$ denote deliveries of materials at date $t$, which get combined with labor $N_{Mt}$ to produce $Y_{Mt}$, which is the flow of intermediate goods that gets added to the stock $M_t$ of works-in-process inventories at the end of period $t$. $U_t$ is the stock of unfilled orders of final goods, $X_t$ the flow from $M_{t-1} + Y_{Mt}$ into final production (i.e. gross output at the intermediate stage), $S_t$, which corresponds to shipments of final goods, and $O_t$ the flow of new orders for final goods. Value added at the final stage is $V_{Ft} = S_t - X_t = A_{Ft}N_{Ft}$.

At the intermediate stage, value added is $V_{Mt} = Y_{Mt} - D_{t-1} = A_{Mt}N_{Mt}$. We assume a Leontief technology for non-labor inputs at each stage. Thus we have

$$D_t = b_M Y_{Mt} \quad (1)$$
$$X_t = b_F S_t \quad (2)$$

To simplify, we abstract from materials inventories and assume that all materials are immediately converted into works-in-process. Figure 5 provides a schematic diagram of the model.

A key element of the model is the delivery lag for materials orders, which we denote by $\tau$. A longer lag means that when the firm makes a decision about materials orders it has less information about what the state of the economy will be when the materials arrive. Consequently the decisions will be less accurate, and will (as the model will show) induce greater volatility in production. By the same token, if firms are, by whatever means, whether it be information technology or management resources, able to shorten the lead time, they can reduce this source of volatility, and also potentially reduce average inventory holding costs.

While much has been written about information technology and inventory control, and in particular the push toward "just-in-time" inventory management, it is difficult to find direct and tangible evidence of improved inventory management. There is time series evidence on $\tau$, however: the Institute for Supply Management (ISM) surveys manufacturers monthly on their lead time for materials orders, going back to 1955. It turns out that this lead time has varied substantially, and in particular has shortened since the early 1980s, at around the same time that output volatility declined. Figure 6 displays the time series. It is not ideal evidence, as it is not limited to durable goods producers (though volatility did decline in the nondurables sector as well). But it very clearly shows a lead time that is both distinctly lower and less volatile since the early 1980s. In particular, the average declines from 65.8 days in the 1955-83 period to 48 days in the period since 1984. Moreover, the series has stabilized since the early 1990s to a value of around 45 days, with much less volatility than in earlier years.

There are undoubtedly other ways in which inventory management may have progressed during this time. In addition to shorter lead times enabling the firm to have better information about the state of the economy at the time of delivery, the firm may just have better information at any point in time. There is considerable
anecdotal information about the use of information technology in inventory management to obtain improved information about product demand. For example, James Surowiecki writes about the retailer Zara in the September 18, 2000 *New Yorker* magazine:

"...instead of reacting quickly to what customers want now, most retailers must guess what they’ll want six or nine months hence. That’s hard enough if you’re selling televisions or bicycles. In the fashion business, it’s close to impossible.

Zara doesn’t have to worry about any of that.... It does not overstock, and unsuccessful designs are often whisked off shelves after just a week, so the company doesn’t have to slash prices. Equipped with handheld devices linked directly to the company’s design rooms in Spain, Zara’s store managers can report daily on what customers are buying, scorning, and asking for but not finding. Most important, the company takes just ten to fifteen days to go from designing a product—which, to be sure, often means knocking off a hot new look—to selling it.

This idea of better information at any point in time was incorporated into the Kahn et al (2002) model for a production-to-stock technology. But while it is undoubtedly part of the larger story of improved inventory management, it is difficult to quantify. The primary reason for focusing on shorter lead times is not because it is the only, or even the most important, aspect of improved inventory management. It is just that there is some quantitative evidence on it, however limited, that can be used as an input to the model described in more detail below.

### 2.1 One-Period Lead Time

Given the discrete time nature of the model, for the sake of simplicity we will assume that \( \tau \) is an integer, representing the number of periods ahead (“lead time”) the firm must order materials before they will arrive. To start with we will assume a one-period lead time for materials. So the timing is as follows: An materials order \( Z_t \) results in a delivery \( D_{t+1} \) and intermediate production \( Y_{M_{t+1}} = b_M^{-1} D_{t+1} \), with final production and shipment \( S_{t+1} \) at \( t+1 \). Materials costs are incurred upon delivery.

Assuming a constant wage \( w \) and a price \( q \) of materials, the cost of producing \( Y_M \) is \( c_M Y_M = (w/A_M + q b_M) Y_M \), and total the cost of producing \( S_t \) is \( c S_t \), where

\[
c = (w/A_F + c_M b_F) = w/A_F + (w/A_M + q b_M) b_F
\]

The firm incurs additional costs from carrying inventories of works-in-process.

We assume that prices are fixed and final goods orders \( O_t \) follow a stochastic process, which for concreteness we assume is a simple AR(1):

\[
O_t = (1 - \rho) \bar{O} + \rho O_{t-1} + \eta_t
\]

(3)

We also assume that there is no “spec” final production, so \( X_t \) is chosen only to fill known (i.e. unfilled) orders as of the beginning of period \( t \), which do not include new orders \( O_t \). The idea is that final production involves
customization that can only be done for a specific order. Intermediate production is more generic, and can be done speculatively. \( X_t \) may also be constrained by the availability of works-in-process \( M_{t-1} + b_M^{-1} D_t \).

We assume the firm maximizes profits subject the various technological constraints:

\[
E_0 \left\{ \sum_{\tau=t}^{\infty} \beta^{\tau-t} \left[ pS_{\tau} - qD_{\tau} - w (N_{M\tau} + N_{F\tau}) \right] \right\}
\]

subject to

\[
\begin{align*}
M_t & = M_{t-1} + Y_{Mt} - X_t \quad (4) \\
U_t & = U_{t-1} + O_t - S_t \quad (5) \\
M_t & \geq 0 \quad (6) \\
U_t & \geq O_t \quad (7) \\
V_{it} & = A_iN_{it}, \quad i = M, F \quad (8) \\
Z_{t-1} & = D_t \quad (9) \\
M_{t-1} + Y_{Mt} & \geq X_t \quad (10) \\
U_{t-1} & \geq S_t \quad (11)
\end{align*}
\]

where \( V_i \) denotes value added, \( Y_i \) gross output at each stage \( (i = M, F) \), and \( \beta < 1 \) is a discount factor. For \( p \) sufficiently large (that is, for a positive markup), the firm will always try to fill all unfilled orders \( U_{t-1} \) at date \( t \). This implies

\[
\begin{align*}
X_t & = \min \left\{ M_{t-1} + b_M^{-1} D_t, b_F U_t \right\} \quad (12) \\
U_{t+1} & = U_t + O_{t+1} - b_F^{-1} X_t \quad (13) \\
& = O_{t+1} + \max \left\{ 0, U_t - b_F^{-1} (M_{t-1} + D_t) \right\} \quad (14)
\end{align*}
\]

That is, subject to availability of works-in-process, \( X_t \) is chosen to fill all unfilled orders \( U_t \).

Note the timing assumptions here: All new orders in period \( t \) are unfilled as of the end of \( t \); shipments during \( t \) are for orders placed at \( t - 1 \) or earlier. Whether or not new orders \( O_t \) are filled by the end of period \( t + 1 \) depends on the adequacy of materials orders \( Z_t \), which were made before \( O_t \) was known. Also note that under certainty, with constant final goods orders \( \bar{O} \), the above setup implies that \( X = Y_M = b_F \bar{O}, \ S = \bar{O}, \ Z = D = b_M b_F \bar{O}, \ N_M = b_F \bar{O}/A_M, \ N_F = \bar{O}/A_F, \) and \( M_t = 0, U_t = \bar{O} \ldots \) There is no reason to hold inventories if both deliveries and orders are known in advance. Unfilled orders are held only because of the
assumption that “spec” production is infeasible due to customization requirements.

We can show that the optimal ordering rule decision rule is of the form

\[ Z_t = b_M [b_FE_t \{U_{t+1}\} + \kappa_1 - M_t] = D_{t+1} \]

where \( \kappa_1 \) is a constant to be determined (see below). It is then straightforward to show that

\[ U_{t+1} = \eta_t + O_{t+1} - \min \{b_F^{-1}\kappa_1, \eta_t\} \]

\[ X_t = b_F (\eta_{t-1} + E_{t-1} \{O_t\}) - \min \{\kappa_1, b_F \eta_{t-1}\} + \min \{\kappa_1, b_F \eta_t\} \]

\[ M_t = \max \{\kappa_1 - b_F \eta_t, 0\} \]

\[ D_t = b_M b_F (\eta_{t-1} + E_{t-1} \{O_t\}) \]

\[ Z_t = b_M b_F (\eta_t + E_t \{O_{t+1}\}) \]

Whether or not the constraint (10) is binding is reflected in terms like \( \min \{\kappa_1, b_F \eta_{t-1}\} \) and \( \max \{\kappa_1 - b_F \eta_t, 0\} \). But note that \( Z_t \) is not affected by past constraints, nor does it end up depending on \( \kappa_1 \). If \( \kappa_1 < b_F \eta_t \), \( X_t \) is constrained by materials. This should add to \( E_t \{U_{t+1}\} \), which it does. But the effect of \( \eta_t \) on \( Z_t \) occurs regardless of the outcome of \( \min \{\kappa_1, b_F \eta_{t-1}\} \), because if there is no stockout (so no increase in unfilled orders) there is the same impact on \( M_t \), which also adds to \( Z_t \). In other words, as \( \eta_t \) increases, either \( M_t \) falls or \( E_t \{U_{t+1}\} \) rises, and both have the same impact on \( Z_t \). \( X_t \) is affected by the outcome of \( \min \{\kappa_1, b_F \eta_t\} \) because higher \( \eta_t \) can mean that \( X_t \) is constrained by the stock of works-in-process. Note that if \( \kappa_1 < b_F \eta_t \), there will be unfilled orders carried over into \( t+1 \) (i.e. \( U_{t+1} > O_{t+1} \)), as the firm had insufficient works-in-process to fill \( U_t \), so some of those unfilled orders get carried over into \( t+1 \). If \( \kappa_1 > b_F \eta_t \) then \( U_{t+1} = O_{t+1} \).

In the National Income and Product Accounts (NIPA), “sales” are really total expenditures on final goods. This corresponds to shipments \( S_t \), i.e.

\[ S_t = b_F^{-1} X_t = \eta_{t-1} + E_{t-1} \{O_t\} - \min \{b_F^{-1}\kappa_1, \eta_{t-1}\} + \min \{b_F^{-1}\kappa_1, \eta_t\} \]

\[ = \Delta \eta_t + \Delta \min \{b_F^{-1}\kappa_1, \eta_t\} \]

“Production” \( Y_t \) is shipments plus the change in inventories, i.e.

\[ Y_t = \eta_{t-1} + E_{t-1} \{O_t\} - \min \{b_F^{-1}\kappa_1, \eta_{t-1}\} + \min \{b_F^{-1}\kappa_1, \eta_t\} + b_F \max \{b_F^{-1}\kappa_1 - \eta_t, 0\} - b_F \max \{b_F^{-1}\kappa_1 - \eta_{t-1}, 0\} \]
It is easy to see that production can be more volatile than sales, in particular when the $O_t$ process exhibits positive serial correlation. Letting $\nu_t \equiv \min \left\{ b_F^{-1} \kappa_1, \eta_t \right\}$, we can simplify the above expressions (using the fact that $\max \left\{ b_F^{-1} \kappa_1 - \eta_t, 0 \right\} = b_F^{-1} \kappa_1 - \nu_t$):

\[
S_t = \eta_{t-1} + E_{t-1} \{ O_t \} + \Delta \nu_t
\]

\[
Y_t = \eta_{t-1} + E_{t-1} \{ O_t \} + (1 - b_F) \Delta \nu_t
\]

Note that $\Delta \nu_t$ is negatively correlated with $\eta_{t-1} + E_{t-1} \{ O_t \}$ (at least for $\rho \geq 0$), which helps to explain why the variance of production can exceed the variance of sales: Inventory investment covaries positively with shipments. The focus in this paper, however, is not on this, but on how production and shipments volatility vary with $\tau$.

Finally, what is $\kappa_1$? It is straightforward to show that it follows from the first-order condition:

\[
\beta p \Pr (b_F U_{t+1} > M_t + D_{t+1}/b_M) - \beta c + \beta^2 c [1 - \Pr (b_F U_{t+1} > M_t + D_{t+1}/b_M)] = 0.
\]

where $c$ is as defined earlier, the total unit cost of producing the final good. Given

\[
U_{t+1} = \eta_t + O_{t+1} - v_t
\]

we have

\[
U_{t+1} = E_t \{ U_{t+1} \} + \eta_{t+1}.
\]

The probability is of a materials stockout at date $t + 1$, and also represents $dE_t (X_{t+1}) / dZ_t$, the expected impact on “completions” at $t + 1$ from an additional order of materials at $t$. The intuition is that by ordering an additional unit at date $t$ at cost $\beta c$, with some probability the firm gains an additional sale at date $t + 1$ (the event of a work-in-process stockout at date $t + 1$), and with one minus that probability it results in surplus stocks and an offsetting decrease in materials orders at date $t + 1$.

Suppose $\eta_t$ has a c.d.f. of $G_1$. We then get

\[
\Pr (b_F [E_t \{ U_{t+1} \} + \eta_{t+1}] > M_t + Z_t/b_M) = \frac{c (1 - \beta)}{p - \beta c}
\]

which implies

\[
\kappa_1 = b_F G_1^{-1} \left( \frac{p - c}{p - \beta c} \right).
\]
which, as one would expect, is increasing in the markup and decreasing in the discount rate $1/\beta - 1$. 

### 2.2 Shorter or Longer Delivery Lags

#### 2.2.1 Zero Lead Time

Now consider one extreme: No delivery lag, so $D_t = Z_t$. With a little algebra, we can show that

\begin{align*}
X_t &= b_F O_t \quad (30) \\
M_t &= 0 \quad (31) \\
Z_t &= D_t = b_M b_F O_t \quad (32) \\
S_t &= O_{t-1} = Y_t \quad (33)
\end{align*}

Clearly going from $\tau = 1$ to $\tau = 0$ changes the relationship between production and sales volatility. As we shall see, for the realistic case in which $O_t$ exhibits positive serial correlation, shrinking the delivery lag from one to zero obviously reduces the gap between the variance of production and the variance of sales. This case is relatively trivial, but nonetheless it provides some insight into why in the more general of shrinking the length of the delivery lag has the same qualitative impact of reducing the volatility of both production and shipments, but with the former declining more than the latter.

#### 2.2.2 $\tau$-Period Lead Time ($\tau > 1$)

We now consider a delivery lag of $\tau$ periods. We hypothesize a decision rule

\[ D_t = Z_{t-\tau} = b_M \left[ b_F E_{t-\tau} \{ U_t \} + \kappa_{\tau} - E_{t-\tau} \{ M_{t-1} \} \right] \quad (34) \]

where, again, $\kappa_\tau \neq \kappa_1$ is to be determined. We then have

\begin{align*}
X_t &= \min \left\{ M_{t-1} + b_F E_{t-\tau} \{ U_t \} + \kappa_{\tau} - E_{t-\tau} \{ M_{t-1} \} , b_F U_t \right\} \quad (35) \\
U_{t+1} &= O_{t+1} + \max \left\{ U_t - E_{t-\tau} \{ U_t \} - b_F^{-1} \kappa_{\tau} - b_F^{-1} \left[ M_{t-1} - E_{t-\tau} \{ M_{t-1} \} \right] , 0 \right\} \quad (36) \\
M_t &= \max \{ \kappa_{\tau} - \left[ b_F ( U_t - E_{t-\tau} \{ U_t \} ) - ( M_{t-1} - E_{t-\tau} \{ M_{t-1} \} ) \right] , 0 \} \quad (37)
\end{align*}

Let

\[ r_{t,\tau} = \sum_{s=0}^{\tau-1} \left( \sum_{j=0}^{s} \rho^j \right) \eta_{t-s} \]
for \( \tau \geq 1 \). Also, let \( u_t = \min \{ b^{-1}_F \kappa, r_{t, \tau} \} \). With considerably more algebra than in the \( \tau = 1 \) case, we can show (see Appendix) that

\[
U_t = O_t + \max \{ r_{t-1, \tau} - b^{-1}_F \kappa, 0 \} \\
M_t = \max \{ \kappa - b_F r_{t, \tau}, 0 \} = b_F \max \{ b^{-1}_F \kappa - r_{t, \tau}, 0 \} \\
\text{(38)} \\
= \kappa - b_F \min \{ b^{-1}_F \kappa, r_{t, \tau} \} \\
\text{(39)} \\
S_t = O_t + \max \{ r_{t-1, \tau} - b^{-1}_F \kappa, 0 \} - \max \{ r_{t, \tau} - b^{-1}_F \kappa, 0 \} \\
= O_t - \Delta r_{t, \tau} + \Delta \min \{ b^{-1}_F \kappa, r_{t, \tau} \} \\
\text{(40)} \\
Y_t = O_t - \Delta r_{t, \tau} + (1 - b_F) \Delta \min \{ b^{-1}_F \kappa, r_{t, \tau} \} \\
\text{(41)} \\
\text{Determining } \kappa_t \text{ is analogous to the } \tau = 1 \text{ case. It is straightforward to show that it follows from the first-order condition}

\[
p \Pr (b_F U_{t+\tau} > M_{t+\tau-1} + D_{t+\tau}/b_M) - c + \beta^T c [1 - p \Pr (b_F U_{t+\tau} > M_{t+\tau-1} + D_{t+\tau}/b_M)] = 0. \\
\text{(44)}
\]

The probability is now of a works-in-process stockout at data \( t + \tau \), and also represents \( dE_t (X_{t+\tau})/dZ_t \). The intuition is that by ordering an additional unit at date \( t \) (to be delivered at date \( t + \tau \)) at cost \( \beta^T c \), with some probability the firm gains an additional sale at date \( t + \tau \) (the event of a works-in-process stockout), and with one minus that probability it results in surplus stocks and an offsetting decrease in materials orders at date \( t + \tau \). Suppose \( r_{t, \tau} \) has a c.d.f. of \( G_{\tau} \). We then get

\[
\Pr \left( r_{t+\tau, \tau} > \frac{D_{t+\tau}}{b_M b_F} - \frac{E_t \{ U_{t+\tau} \} - b^{-1}_F E_t \{ M_{t+\tau-1} \}}{b_M b_F} \right) = \frac{c (1 - \beta^T)}{p - \beta^T c} \\
\text{(45)} \\
G_{\tau} \left( \frac{D_{t+\tau}}{b_M b_F} - \frac{E_t \{ U_{t+\tau} \} - b^{-1}_F E_t \{ M_{t+\tau-1} \}}{b_M b_F} \right) = \frac{p - c}{p - \beta^T c}. \\
\text{(46)}
\]

So we have

\[
Z_t = b_M \{ \kappa_t + b_F E_t \{ U_{t+\tau} \} - E_t \{ M_{t+\tau-1} \} \} \\
\text{(47)}
\]

where

\[
\kappa_t = b_F G_{\tau}^{-1} \left( \frac{p - c}{p - \beta^T c} \right). \\
\text{(48)}
\]

Any difference between \( \kappa_1 \) and \( \kappa_\tau \) would stem from the difference in the relevant distribution function (\( G_1 \) vs. \( G_{\tau} \)) and from the fact that excess materials orders are more costly because they take \( \tau \) periods to offset. In addition, the firm’s markup could differ in the two cases.
Finally, what is $b_F E_t \{U_{t+\tau}\} - E_t \{M_{t+\tau-1}\}$? We have

$$
\begin{align*}
&b_F E_t \{U_{t+\tau}\} - E_t \{M_{t+\tau-1}\} \\
&= b_F E_t \left[ O_{t+\tau} \max \left\{ r_{t+\tau-1,\tau} - b_F^{-1} \kappa_\tau, 0 \right\} - \max \left\{ b_F^{-1} \kappa_\tau - r_{t+\tau-1,\tau}, 0 \right\} \right] \\
&= b_F E_t \left[ O_{t+\tau} + r_{t+\tau-1,\tau} \right] - \kappa_\tau. 
\end{align*}
$$

So

$$Z_t = b_M b_F E_t \{O_{t+\tau} + r_{t+\tau-1,\tau}\} \quad (51)$$

If we compare (51) and (20) we see that now in addition to having to order based on two-period-ahead expected orders, the date $t$ innovation has a magnified impact (to the extent $\rho > 0$).

Regarding implications for inventory-sales ratios, it turns out that the model is incomplete: Absent a theory of the markup $p/c$ as a function of $\tau$, the relative size of inventory-sales ratios is ambiguous. But under the reasonable intermediate assumption that $(p(\tau) - c)/(p(\tau) - \beta^\tau c)$ is invariant to $\tau$, then inventory-sales ratios will be larger under $\tau = 2$, essentially because of the greater uncertainty at the time of ordering.

### 2.3 Aggregation

The solutions for the time series behavior output, shipments, and inventories are not realistic characterizations of any data that are likely to be observed. In practice, we observe aggregates, even when we look at relatively disaggregated data. The data are typically aggregates of different goods, different locations, and different firms. Consequently we never see zeros of any stocks, whether of unfilled orders or inventories.

Fortunately the model is amenable to aggregation as follows. We now suppose a continuum of symmetric firms, each of which faces stochastic orders as above, but with an idiosyncratic shock $v_{it}$. That is, for firm $i$,

$$O_{it} = (1 - \rho) \hat{O} + \rho O_{i,t-1} + \eta_t + v_{it}. \quad (52)$$

where for concreteness we can assume that $v_{it}$ is normally distributed with mean zero and variance $\sigma_v^2$, and that $\int_{-\infty}^{\infty} \phi(v) = 0$. Of course, the derivations of $\kappa_\tau$ must now be revised to reflect idiosyncratic risk. For example, $G_1$ in (29) should be the distribution function for $\eta_t + v_{it}$.
First consider the $\tau = 1$ case. Let

$$
\zeta_{1t} = \int_{-\infty}^{\infty} \min \left\{ b_F^{-1} \kappa_1, \eta_t + \sigma_v v \right\} \phi(v) dv
= \int_{-\infty}^{b_F^{-1} \kappa_1 - \eta_t} (\eta_t + \sigma_v v) \phi(v) dv + b_F^{-1} \kappa_1 \left[ 1 - \Phi \left( \frac{b_F^{-1} \kappa_1 - \eta_t}{\sigma_v} \right) \right]
$$

(53)

The idea here is that each firm’s outcome varies depending on its idiosyncratic shock, so we integrate to get the aggregate. But because of the linear technology, the firms adjust and go into the next period looking identical. So for aggregate inventories we have

$$
M_t = \max \{ \kappa_1 - b_F \eta_t, 0 \} = \kappa_1 - b_F \min \{ \eta_t, b_F^{-1} \kappa_1 \}
$$

(54)

$$
= \kappa_1 - b_F \zeta_{1t}
$$

(55)

Similarly,

$$
S_t = O_t - \Delta \eta_t + \Delta \zeta_{1t}
$$

(57)

$$
S_t + \Delta M_t = O_t - \Delta \eta_t + (1 - b_F) \Delta \zeta_{1t}
$$

We can simplify further by linearizing $\zeta_{1t}$ around $\eta_t = 0$:

$$
\zeta_{1t} \approx \kappa_1 \left[ 1 - \Phi \left( \frac{\kappa_1}{\sigma_v} \right) \right] + \Phi \left( \frac{\kappa_1}{\sigma_v} \right) \eta_t = \theta_1 \eta_t
$$

(58)

where

$$
\theta_1 = \Phi \left( \frac{\kappa_1}{\sigma_v} \right)
$$

(59)

and we ignore the constant term, which is not relevant for the exercise of computing volatility. Note that for finite $\kappa_1$, $\theta_1 \in (0, 1)$. So we have

$$
S_t = (1 + \rho - \theta_1) \eta_{t-1} + \rho^2 O_{t-2} + \theta_1 \eta_t
$$

(60)

$$
Y_t = S_t + \Delta M_t = (1 + \rho - \theta_1 (1 - b_F)) \eta_{t-1} + \rho^2 O_{t-2} + \theta_1 (1 - b_F) \eta_t
$$
\[
\text{Var}(S_t) = (1 + \rho - \theta_1)^2 \sigma_y^2 + \rho^4 \sigma_y^2 / (1 - \rho^2) + \theta_1^2 \sigma_y^2
\]

\[
\text{Var}(Y_t) = (1 + \rho - \theta_1 (1 - b_F))^2 \sigma_y^2 + \rho^4 \sigma_y^2 / (1 - \rho^2) + \theta_1^2 (1 - b_F)^2 \sigma_y^2
\]

So

\[
\frac{\text{Var}(Y_t) - \text{Var}(S_t)}{\sigma_y^2} = \frac{(1 + \rho - \theta_1 (1 - b_F))^2 - (1 + \rho - \theta_1)^2 + \theta_1^2 (1 - b_F)^2 - \theta_1^2}{\sigma_y^2}
\]

(61)

\[
= 2\theta_1 b_F (\rho - 2\theta_1 + \theta_1 b_F + 1)
\]

(62)

\[
> 0 \quad \text{if } 1 + \rho > (2 - b_F) \theta_1
\]

(63)

So for sufficiently large \(\rho\) or \(b_F\), or small \(\theta_1\) (that is, small \(\kappa_1\)), production is more volatile than sales. Note that \(\rho\) need not even be positive for this to be true.

With general \(\tau\) we define (assuming for simplicity that idiosyncratic risk is i.i.d.):

\[
\zeta_{\tau t} = \int_{-\infty}^{\infty} \min \{ b_F^{-1} \kappa_{\tau t}, r_{t\tau} + \sigma_v v \} \phi(v) dv
\]

(64)

\[
= \int_{-\infty}^{\kappa_{\tau t} - r_{t\tau} / \sigma_v} \phi(v) dv + \kappa_{\tau t} \left[ 1 - \Phi \left( \frac{\kappa_{\tau t} - r_{t\tau}}{\sigma_v} \right) \right]
\]

(65)

\[
\approx \Phi \left( \frac{\kappa_{\tau t}}{\sigma_v} \right) r_{t\tau} = \theta_\tau r_{t\tau}
\]

(66)

We then get

\[
M_t = \max \{ \kappa_{\tau t} - b_F r_{t\tau}, 0 \}
\]

(67)

\[
= \kappa_{\tau t} - b_F \min \{ r_{t\tau}, b_F^{-1} \kappa_{\tau 1} \}
\]

(68)

\[
= \kappa_{\tau t} - b_F \theta_\tau r_{t\tau}
\]

(69)

and

\[
S_t = O_t + \max \{ r_{t-1\tau} - b_F^{-1} \kappa_{\tau t}, 0 \} - \max \{ r_{t\tau} - b_F^{-1} \kappa_{\tau t}, 0 \}
\]

\[
= O_t - \Delta r_{t\tau} + \Delta \min \{ b_F^{-1} \kappa_{\tau t}, r_{t\tau} \}
\]

\[
= O_t - \Delta r_{t\tau} + \Delta \zeta_{\tau t}
\]

\[
\approx O_t - (1 - \theta_\tau) \Delta r_{t\tau}
\]
\[ Y_t = O_t - \Delta r_{t, \tau} + (1 - b_F) \Delta \zeta_{t, \tau} \]  
\[ \approx O_t - [1 - (1 - b_F) \theta_{\tau}] \Delta r_{t, \tau}. \]

With \( \tau = 2 \), for example, we have

\[ \text{Var}(S) / \sigma_n^2 = \rho^6 / (1 - \rho^2) + \theta_2^2 (1 + \rho^2) + [(1 - \theta_2) (1 + \rho) + \rho^2]^2 \]  
(72)

\[ \text{Var}(Y) / \sigma_n^2 = \rho^6 / (1 - \rho^2) + (1 - b_F)^2 \theta_2^2 (1 + \rho^2) + [(1 - (1 - b_F) \theta_2) (1 + \rho) + \rho^2]^2 \]  
(73)

and

\[ [\text{Var}(Y) - \text{Var}(S)] / \sigma_n^2 = 2 \theta_2 b_F (\rho^2 + \rho + 1) (\rho - 2 \theta_2 + \theta_2 b_F + 1) \]  
(74)

Now compare \( 2 \theta b_F (\rho - 2 \theta + \theta b_F + 1) \) versus \( 2 \theta_2 b_F (\rho^2 + \rho + 1) (\rho - 2 \theta_2 + \theta_2 b_F + 1) \). Clearly if the first is positive, the second is larger for \( \rho > 0 \). So holding fixed the shock variance \( \sigma_n^2 \), for parameters in the empirically relevant range, a reduction in the delivery lag (i.e. the lead time for materials orders) results in a reduction in both output and sales volatility, but a greater reduction in output volatility.

### 2.4 Simulations

We can get some feel for the capability of this approach to account for changes in volatility. We choose parameters that roughly match the relevant characteristics of the data in the early part of the sample (1954-1983) under the assumption that \( \tau = 3 \), and then compare that with \( \tau = 2 \). Given the reduction in lead times indicated in Figure 6, this means a period corresponds to about three weeks; that is, \( \tau = 3 \) corresponds to 63 days, \( \tau = 2 \) to 42 days. Of course, there was not an abrupt change; the simulation exercise should be thought of as comparing across steady states.

The benchmark simulation makes the following assumptions about the parameters: \( \sigma_n^2 = 0.07 \), \( \rho = 0.99 \), \( \sigma_\nu^2 = 1 \), \( b_F = 0.75 \), \( \beta = 0.99 \). Table 5 displays results for the benchmark assumptions and for alternative values of \( \rho \) and \( p/c \). The results are not at all sensitive, even quantitatively, to \( \sigma_n^2 \), and qualitatively similar across a range of the other parameters. The choice for \( \rho \) is intended to capture the strong persistence of sales and output measures in the data. The value for \( \sigma_n^2 \) was chosen so that the simulation would roughly match the pre-1984 standard deviation of output with \( \tau = 3 \).

Table 5 shows that the model yields sizeable reductions in the volatility of quarterly growth rates of output, along with very small reductions in sales volatility. Neither is as large as in the data—a good thing, as the intent is not to explain the entire Great Moderation. It also, as expected, does better at explaining the high frequency facts than the lower frequency facts as represented by 4-quarter growth rates, though it
does replicate the feature of the data that the decline of output volatility relative to that of sales volatility is more modest with 4-quarter growth rates.

It should be emphasized that the model, while realistic in many dimensions, makes important simplifying assumptions. In particular, the model lacks a production smoothing motive for inventories, which presumably explains why the gap between the volatility of output growth and sales growth is much larger than in the data. Nonetheless, the fact that it at least qualitatively matches the facts suggests that it may be a reasonable basis for pursuing richer and more realistic general equilibrium extensions.

3 Discussion and Conclusions

In the decade since economists took notice of the decline in aggregate volatility, there have been surprisingly few structural models geared to explaining the phenomenon. The most influential empirical efforts to distinguish between different explanations (e.g. Stock and Watson, 2002; Justiniano and Primiceri, forthcoming) have relied on the ability of small-scale models—either structural vector autoregressions (SVARS) or dynamic stochastic general equilibrium models (DSGEs) to distinguish between shocks and propagation. Such efforts seem inevitably to result in a finding that reduced volatility of observable variables is due to the reduced volatility of unobservable shocks. Although Justiniano and Primiceri do supplement their finding of an important decline in the volatility of “shocks specific to the equilibrium condition of investment” with data on the volatility of the relative price of investment to consumption, it is worth noting as well that their “investment” variable includes inventory investment. Thus, their findings are potentially consistent with an important role for changing inventory behavior that they simply do not model.

Ramey and Vine (2004) offer another sort of explanation relying on an exogenous change in a stochastic process. They provide a detailed analysis of automobile industry production decisions, and argue that a reduction in the persistence of the sales process can explain the reduction in output volatility relative to sales volatility. The change in persistence itself is left unexplained, nor is evidence provided to suggest that their explanation applies to the rest of the durable goods sector. In particular it seems to fly in the face of broader evidence (see Kahn and Davis, 2008, for example) that transitory shocks have diminished in volatility.

One reason for the lack of detailed structural explanations for the Great Moderation is simply that models that can explain volatility endogenously, without resort to exogenous changes in the parameters of shocks, are inevitably complex, highly specific, and therefore unlikely to be able to explain the entire Great Moderation. This would seem to be an unfair standard, as the Great Moderation is unlikely to have a single “magic bullet” explanation. Giannone et al (2008) make a similar point when they argue that “typical
macroeconomic models, looking only at a handful of variables, overstate the role of good luck.” They find that when they look at models involving a larger set of explanatory variables find a greater role for changes in propagation.

The model in this paper, while relatively simple in terms of the number of variables, exhibits a different type of complexity in which the magnitudes of the “shocks”—at least what would appear to be shocks—are themselves functions of structural parameters that could change over time. Propagation also changes endogenously. In order to be able to solve the model analytically, and to isolate the impact of the proposed structural, the model is simplified in other ways: It is partial equilibrium, with only one source of shocks. There is no production-smoothing motive for holding inventories, and the entire production technology is simplified. Future work will put the model in a more general framework, as a number of papers have done with the model of finished goods inventories in Kahn (1987).¹¹ That work suggests that the qualitative results in this paper—that empirically plausible improvements in supply chain management can account for the patterns of volatility reductions during the Great Moderation—paper will survive. Of course there may be other models that can account for the same pattern of facts, but explanations relying on purely exogenous (or at least unexplained) changes in volatility or propagation arguably fall short of being true theories: They offer no insight into the economic forces underlying the changes, little ability to predict out of sample, and few if any testable implications.

¹¹See, for example, Yi (2008), Bils and Kahn (2000), Kahn et al (2002).
Appendix

This appendix provides a derivation of the result for a lead time of $\tau$ periods, materials orders take the form (34)

$$D_t = Z_{t-\tau} = b_M [b_F E_{t-\tau} \{U_t\} + \kappa_t - E_{t-\tau} \{M_{t-1}\}],$$

and that, consequently,

$$M_t = \max \{\kappa_t - b_F r_{t,\tau}, 0\}$$

$$S_t = O_t + \max \{r_{t-1,\tau} - b_F^{-1} \kappa_t, 0\} - \max \{r_{t,\tau} - b_F^{-1} \kappa_t, 0\}$$

The strategy is to assume the form, derive the implications for the evolution of the endogenous variables, and then show that (34) satisfies the first-order condition. Let

$$x_t(\tau) \equiv b_F (U_t - E_{t-\tau} \{U_t\})$$

$$z_t(\tau) \equiv M_t - E_{t-\tau + 1} \{M_t\}$$

Then

$$U_t = O_t + b_F^{-1} \max \{x_{t-1}(\tau) - z_{t-2}(\tau) - \kappa_t, 0\}$$

$$M_t = \max \{\kappa_t - [x_t(\tau) - z_{t-1}(\tau)], 0\}.$$ 

Therefore (suppressing the $\tau$ argument in $x$ and $z$)

$$U_t - b_F^{-1} M_{t-1} = O_t + b_F^{-1} (x_{t-1} - z_{t-2} - \kappa_t)$$

$$E_{t-\tau} \{U_t - b_F^{-1} M_{t-1}\} = E_{t-\tau} \{O_t + b_F^{-1} (x_{t-1} - z_{t-2} - \kappa_t)\}$$

$$U_t - b_F^{-1} M_{t-1} - E_{t-\tau} \{U_t - b_F^{-1} M_{t-1}\} = O_t - E_{t-\tau} \{O_t\} + b_F^{-1} [x_{t-1} - E_{t-\tau} \{x_{t-1}\} - (z_{t-2} - E_{t-\tau} \{z_{t-2}\})]$$

$$= O_t - E_{t-\tau} \{O_t\} + (U_{t-1} - E_{t-\tau - 1} \{U_{t-1}\}) -$$

$$E_{t-\tau} \{(U_{t-1} - E_{t-\tau - 1} \{U_{t-1}\})\} - b_F^{-1} (M_{t-2} - E_{t-\tau - 1} \{M_{t-2}\}) - E_{t-\tau} \{M_{t-2} - E_{t-\tau - 1} \{M_{t-2}\}\})$$

$$= O_t - E_{t-\tau} \{O_t\} + U_{t-1} - E_{t-\tau} \{U_{t-1}\} - b_F^{-1} (M_{t-2} - E_{t-\tau} \{M_{t-2}\})$$

$$U_{t-1} - b_F^{-1} M_{t-2} - E_{t-\tau - 1} \{U_{t-1} - b_F^{-1} M_{t-2}\} = O_{t-1} - E_{t-\tau - 1} \{O_{t-1}\} + U_{t-2} -$$

$$E_{t-\tau - 1} \{U_{t-2}\} - b_F^{-1} (M_{t-3} - E_{t-\tau - 1} \{M_{t-3}\})$$

23
From our initial assumption about $D_t$, we then have

\[ D_t/b_M = b_F E_{t-\tau} \{ O_t \} + \kappa_\tau - E_{t-\tau} \{ M_{t-1} \} \]

\[ = E_{t-\tau} \{ b_F O_t + \max \{ b_F r_{t-1,\tau} - \kappa_\tau, 0 \} - \max \{ \kappa_\tau - b_F r_{t-1,\tau}, 0 \} \} + \kappa_\tau \]

\[ = b_F E_{t-\tau} \{ O_t + r_{t-1,\tau} \} \]
For example, if $\tau = 2$, $r_{t-1,\tau} = \eta_{t-1} + (1 + \rho) \eta_{t-2}$, and $E_{t-\tau} \{ r_{t-1,\tau} \} = (1 + \rho) \eta_{t-2}$, so

$$D_t/b_M = b_F (1 + \rho + \rho^2) \eta_{t-2}.$$  

Finally

$$S_t = b_F^{-1} X_t$$

$$= b_F^{-1} \min \{ M_{t-1} + b_F E_{t-\tau} \{ U_t \} + \kappa_{t\tau} - E_{t-\tau} \{ M_{t-1} \}, b_F U_t \}$$

$$= U_t - \max \{ r_{t,\tau} - b_F^{-1} \kappa_{t\tau}, 0 \}$$

$$= O_t + \max \{ r_{t-1,\tau} - b_F^{-1} \kappa_{t\tau}, 0 \} - \max \{ r_{t,\tau} - b_F^{-1} \kappa_{t\tau}, 0 \}$$

Again if $\tau = 2$,

$$S_t = O_t + \max \{ [\eta_{t-1} + (1 + \rho) \eta_{t-2}] - b_F^{-1} \kappa_2, 0 \}$$

$$- \max \{ [\eta_t + (1 + \rho) \eta_{t-1}] - b_F^{-1} \kappa_2, 0 \}$$

$$= O_t + \max \{ r_{t-1,2} - b_F^{-1} \kappa_2, 0 \} - \max \{ r_{t,2} - b_F^{-1} \kappa_2, 0 \}$$

$$= O_t + (r_{t-1,2} - u_{t-1}) - (r_{t,2} - u_t)$$

$$= E_{t-2} \{ O_t \} + \eta_t + \rho \eta_{t-1} + \Delta u_t - \Delta r_{t,2}$$

$$= E_{t-2} \{ O_t \} + \Delta u_t + (1 + \rho) \eta_{t-2}$$
References


Table 1: GARCH results

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>GDP growth</th>
<th>Durables output growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Const</strong></td>
<td>2.250</td>
<td>2.249</td>
</tr>
<tr>
<td></td>
<td>(0.272)</td>
<td>(0.310)</td>
</tr>
<tr>
<td><strong>Lagged Dep Var.</strong></td>
<td>0.317</td>
<td>0.309</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.074)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance equation</th>
<th>GDP growth</th>
<th>Durables output growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Const</strong></td>
<td>6.515</td>
<td>11.118</td>
</tr>
<tr>
<td></td>
<td>(0.781)</td>
<td>(2.139)</td>
</tr>
<tr>
<td><strong>Resid$^2$</strong></td>
<td>0.096</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.063)</td>
</tr>
<tr>
<td><strong>GARCH(-1)</strong></td>
<td>0.713</td>
<td>0.658</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.149)</td>
</tr>
<tr>
<td><strong>Trend</strong></td>
<td>-0.045</td>
<td>-0.121</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.012)</td>
</tr>
<tr>
<td><strong>Trend$^2$</strong></td>
<td>8.03E-05</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>3.95E-05</td>
<td>(4.11E-06)</td>
</tr>
<tr>
<td><strong>10-year T-rate(-1)</strong></td>
<td>—</td>
<td>0.294</td>
</tr>
<tr>
<td></td>
<td>(0.202)</td>
<td>(0.181)</td>
</tr>
<tr>
<td><strong>Dummy(≥1984.1)</strong></td>
<td>—</td>
<td>-7.001</td>
</tr>
<tr>
<td></td>
<td>(5.192)</td>
<td>(66.05)</td>
</tr>
</tbody>
</table>

Note: Estimated on the sample 1947Q1-2007Q3
Table 2: Statistics from Detrended Data

<table>
<thead>
<tr>
<th>Durables</th>
<th>Durables</th>
<th>Durables</th>
<th>Durables</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Output</td>
<td>Sales</td>
<td>GDP</td>
</tr>
<tr>
<td>54.1-69.4</td>
<td>1.54</td>
<td>5.96</td>
<td>3.63</td>
</tr>
<tr>
<td>70.1-83.4</td>
<td>2.25</td>
<td>6.92</td>
<td>4.99</td>
</tr>
<tr>
<td>84.1-07.4</td>
<td>0.90</td>
<td>3.25</td>
<td>2.72</td>
</tr>
</tbody>
</table>

Note: The logarithmic series were detrended using an HP-filter with parameter 1600.

Table 3: Frequency Domain Analysis

<table>
<thead>
<tr>
<th>Durables</th>
<th>Durables</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Output</td>
</tr>
<tr>
<td>54.1-69.4</td>
<td>3.58</td>
</tr>
<tr>
<td>70.1-83.4</td>
<td>5.69</td>
</tr>
<tr>
<td>84.1-07.4</td>
<td>3.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Durables</th>
<th>Durables</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Output</td>
</tr>
<tr>
<td>54.1-69.4</td>
<td>2.67</td>
</tr>
<tr>
<td>70.1-83.4</td>
<td>3.43</td>
</tr>
<tr>
<td>84.1-07.4</td>
<td>2.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Durables</th>
<th>Durables</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td>Output</td>
</tr>
<tr>
<td>54.1-69.4</td>
<td>18.16</td>
</tr>
<tr>
<td>70.1-83.4</td>
<td>33.89</td>
</tr>
<tr>
<td>84.1-07.4</td>
<td>8.47†</td>
</tr>
</tbody>
</table>

*Reduction in normalized spectrum relative to 1954-69 period significant at 10% level

**Reduction in normalized spectrum relative to 1954-69 period significant at 5% level

†Increase in normalized spectrum relative to 1954-69 period significant at 10% level
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lumber (24)</td>
<td>23.15</td>
<td>17.18</td>
<td>20.36</td>
<td>13.70</td>
</tr>
<tr>
<td>Furniture/Fixtures (25)</td>
<td>15.38</td>
<td>8.77</td>
<td>14.57</td>
<td>7.41</td>
</tr>
<tr>
<td>Stone, Clay, Glass (32)</td>
<td>14.60</td>
<td>10.30</td>
<td>13.93</td>
<td>9.89</td>
</tr>
<tr>
<td>Primary Metals (33)</td>
<td>27.88</td>
<td>9.96</td>
<td>27.78</td>
<td>8.94</td>
</tr>
<tr>
<td>Fabricated Metals (34)</td>
<td>16.35</td>
<td>8.59</td>
<td>14.29</td>
<td>7.14</td>
</tr>
<tr>
<td>Industrial Machinery (35)</td>
<td>14.63</td>
<td>9.76</td>
<td>12.24</td>
<td>8.97</td>
</tr>
<tr>
<td>Electronic Machinery (36)</td>
<td>14.94</td>
<td>7.64</td>
<td>13.57</td>
<td>7.47</td>
</tr>
<tr>
<td>Motor Vehicles (371)</td>
<td>46.04</td>
<td>22.98</td>
<td>45.59</td>
<td>21.59</td>
</tr>
<tr>
<td>Other Transportation (37x)</td>
<td>23.37</td>
<td>17.09</td>
<td>17.14</td>
<td>16.19</td>
</tr>
<tr>
<td>Instruments (38)</td>
<td>11.42</td>
<td>7.22</td>
<td>10.84</td>
<td>5.79</td>
</tr>
<tr>
<td>Miscellaneous (39)</td>
<td>20.79</td>
<td>11.51</td>
<td>18.62</td>
<td>10.25</td>
</tr>
<tr>
<td>Durable Manufacturing</td>
<td>14.66</td>
<td>7.09</td>
<td>13.72</td>
<td>6.39</td>
</tr>
<tr>
<td>Durable Goods (aggregate)</td>
<td>17.34</td>
<td>8.02</td>
<td>10.92</td>
<td>8.60</td>
</tr>
</tbody>
</table>
Table 5: Simulation Results

<table>
<thead>
<tr>
<th>ρ</th>
<th>p/c</th>
<th>θ</th>
<th>τ</th>
<th>Output growth</th>
<th>Sales growth</th>
<th>4-q output growth</th>
<th>4-q sales growth</th>
<th>Period growth</th>
<th>Output growth</th>
<th>Sales growth</th>
<th>4-q output growth</th>
<th>4-q sales growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.975</td>
<td>3</td>
<td>20.16</td>
<td>5.30</td>
<td>5.60</td>
<td>2.68</td>
<td>54.1-83.4</td>
<td>18.32</td>
<td>11.04</td>
<td>6.84</td>
<td>4.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.99</td>
<td>1.1</td>
<td>0.911</td>
<td>2</td>
<td>14.93</td>
<td>5.20</td>
<td>4.46</td>
<td>2.70</td>
<td>84.1-2007.4</td>
<td>7.80</td>
<td>8.26</td>
<td>2.78</td>
<td>2.98</td>
</tr>
<tr>
<td>0.992</td>
<td>3</td>
<td>18.85</td>
<td>5.33</td>
<td>5.24</td>
<td>2.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>1.2</td>
<td>0.953</td>
<td>2</td>
<td>14.14</td>
<td>5.18</td>
<td>4.23</td>
<td>2.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Units are annualized logarithmic annual rates—e.g. "output growth" is \(400 \Delta \ln(Y)\). Simulations were based on one million random draws.*
Volatility is measured as a 5-year rolling standard deviations of growth rates scaled by GDP share.
In each chart, the dotted line is the volatility of GDP. Shaded periods represent NBER-designated recessions.
Figure 1b: Scaled Volatility by Expenditure Category, Relative to GDP Volatility

Note: Volatility is measured as a 5-year rolling standard deviations of growth rates scaled by GDP share. In each chart, the dotted line is the volatility of GDP. Shaded periods represent NBER-designated recessions.
Figure 2: Durable Goods Sector Volatility

Note: Volatility is measured as a 5-year rolling standard deviations of growth rates. Shaded periods represent NBER-designated recessions.
Figure 3: Contributions to the Volatility of Durable Goods Output

Note: Volatility is measured as a 5-year rolling variances of growth contributions.
Figure 4: Inventory Sales Ratios in Durable Goods*

Note: The discontinuities in the late 1990s are due to the change in industry classifications from the SIC system to the NAICS.
Figure 5: A Schematic Diagram of the Model
Figure 6: Average Lead Time for Production Materials Orders

Source: Institute for Supply Management