Stylized (Arte)Facts on Sectoral Inflation*

Ferre De Graeve† and Karl Walentin‡

Sveriges Riksbank Working Paper Series No. 254

August 2011

Abstract

Research on disaggregate price indices has found that sectoral shocks generate the bulk of sectoral inflation variance, but no persistence. Aggregate shocks, by contrast, are the root of sectoral inflation persistence, but have negligible relative variance. We argue that these findings are largely an artefact of using overly simple factor models to characterize inflation. Sectoral inflation series are subject to particular features such as sales and item substitutions. In factor models, these blow up the variance of sectoral shocks, while reducing their persistence. Controlling for such effects, we find that inflation variance is driven by both aggregate and sectoral shocks. Sectoral shocks, too, generate substantial inflation persistence. Both findings contrast sharply with earlier evidence from factor models. However, these results align well with recent micro evidence. This has implications for the foundations of price stickiness, and provide quantitative inputs for calibrating models with sectoral heterogeneity.

Keywords: Inflation persistence, sticky prices, factor model, sectoral inflation
JEL Codes: E5, G01, G21

*We would like to thank Jean Boivin, Mikael Carlsson, Tim Cogley, Marco Del Negro, Roland Meeks, Andrea Tambalotti, Thijs van Rens and seminar participants at IIES, Federal Reserve Bank of New York, New York University Alumni Conference, Federal Reserve Bank of Philadelphia, Sveriges Riksbank and Uppsala University for useful comments and valuable suggestions. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank.

†Sveriges Riksbank, Research division. E-mail: ferre.de.graeve@riksbank.se.

‡Sveriges Riksbank, Research division. E-mail: karl.walentin@riksbank.se.

1 Introduction

The extent and nature of price rigidities are important inputs for many macroeconomic considerations. A recent body of research aims to shed light on this issue by identifying the sources of volatility and persistence in disaggregate (sectoral) inflation rates (Boivin, Giannoni and Mihov, 2009; Maćkowiak, Moench and Wiederholt, 2009). Based on a variety of estimated dynamic factor models for a number of different sectoral price data sets, two conclusions emerge: (i) Sectoral inflation *volatility* is mostly due to sector-specific disturbances, while aggregate shocks explain only a small fraction of movements in inflation. (ii) Sectoral inflation *persistence* is generated by aggregate shocks. The response to idiosyncratic or sector-specific shocks, by contrast, is close to instantaneous.

The empirical findings on the sources of inflation persistence and volatility are used to validate foundations of price stickiness. For instance, Maćkowiak and Wiederholt (2009, 2010) and Maćkowiak *et al.* (2009) argue for rational inattention as the root of price stickiness because it can replicate swift responses to sector-specific shocks and sluggish adjustment to aggregate shocks. In later work, however, Carvalho and Lee (2010) show that time-dependent nominal rigidities can generate similar impulse responses.

More generally, with the increased feasibility of quantitatively evaluating macro-models with heterogeneity comes the need for moments to calibrate them to. We argue that the empirical factor models used cannot serve this purpose. The basic reason is that factor models are not well suited for any *relative* assessment of common and idiosyncratic sources of fluctuations. Factor models yield identification of common factors and their effects reliably, under various forms of measurement error or misspecification (Stock and Watson, 1998). However, all other sources of fluctuations, including measurement error, are lumped together and labeled idiosyncratic. As a result, the residual treatment of idiosyncratic shocks invalidates almost any inference that involves them.

This issue particularly affects research on price indices, where measurement issues prevail. The sampling of prices across products, stores and cities is a huge endeavor, inherently subject to measurement error (Shoemaker, 2007). Moreover, price setting is naturally characterized by features such as sales and product substitutions, which have dramatic effects on evaluations of volatility and persistence (Bils and Klenow, 2004; Golosov and Lucas, 2007; Klenow and Kryvtsov, 2008; Nakamura and Steinsson, 2008, 2009; Kehoe and Midrigan, 2010; Klenow and Malin, 2010; Eichenbaum, Jaimovich and Rebelo, 2011).

Because factor models are not well-suited to assess the relative importance of sector-specific versus aggregate shocks, the facts (i) and (ii) are potentially misleading. In this sense, they may well be artefacts rather than stylized facts.

We reformulate the factor model such that it can cope with general concerns in price measurement. We then estimate it on U.S. personal consumption expenditures (PCE) price indices and compare the outcome with a simple factor model (essentially that of Boivin *et al.*, 2009). The conclusions differ drastically. First, sectoral inflation responds sluggishly not only to aggregate shocks but also in response to sector-specific shocks. Second, the volatility of idiosyncratic shocks is substantially smaller than previously argued.

The intuition of these results is as follows. Sales and item substitutions are non-persistent and largely idiosyncratic (unrelated to aggregate conditions) contributors to sectoral inflation. A simple factor model ignores that much of the volatility in prices is due to such measurement effects. Instead, it lumps them together with more persistent idiosyncratic shocks to inflation. Because the variance contribution of sales and substitutions is substantial, the resulting composite idiosyncratic process will be volatile and non-persistent.

The paper is organized as follows. We start by reproducing the so-called stylized facts using a simple factor model. Then, in Section 3, we lay out what can go wrong with factor models for inflation indices. Section 4 specifies our benchmark factor model that is able to cope with these concerns. Subsequently, in Section 5, we estimate the benchmark model for PCE data and compare it to the stylized facts. In Section 6 we discuss aggregation, validation and alternative interpretations. After assessing the robustness of our conclusions in Section 7, we conclude.

2 A simple factor model for sectoral inflation

Consider the following decomposition of sectoral inflation π_{it} into a common and a sector-specific component

$$\begin{aligned}\pi_{it} &= COM_{it} + SEC_{it} && (1)\\ &= \lambda_i' C_t + e_{it}. && (2)\end{aligned}$$

Here, $COM_{it} = \lambda_i' C_t$, and C_t is a $N \times 1$ vector of common factors. These factors are distilled from a large cross-section of macroeconomic and/or sectoral time series, X_t. The factor loadings λ_i measure the dependence of inflation in sector i on aggregate, or common, conditions. The remainder, e_{it}, is a purely sector-specific scalar process. The dynamics of sectoral inflation originate from both the common component and the sectoral component, through

$$
\begin{aligned}
C_t &= \Phi(L)C_{t-1} + v_t, &(3)\\
e_{it} &= \rho_i(L)e_{it-1} + u_{it}. &(4)
\end{aligned}
$$

With this kind of decomposition at hand, Boivin *et al.* (2009) and Maćkowiak *et al.* (2009) decompose the variance, $\sigma^2(\pi_{it})$, and persistence, $\rho(\pi_{it})$, of sectoral inflation into a common and a sector-specific part.[1]

As a quantitative reference for what follows, we use the data of Boivin *et al.* (2009) to estimate the model (1)-(4). The data for π_{it} are monthly PCE price indices for 190 sectors over the period 1976:1-2005:6. We extract 5 common factors C_t from a total of 653 monthly series. In particular, X_t consists of 111 macroeconomic indicators, 190 sectoral PCE and 154 Producer Price Index (PPI) inflation series as well as 190 sectoral PCE quantity series. In addition, X_t contains 4 PCE price aggregates and the corresponding quantity aggregates.[2] We set lag length to 13 for all lag polynomials, in analogy to Boivin *et al.* (2009), though results are very similar using standard lag selection criteria.

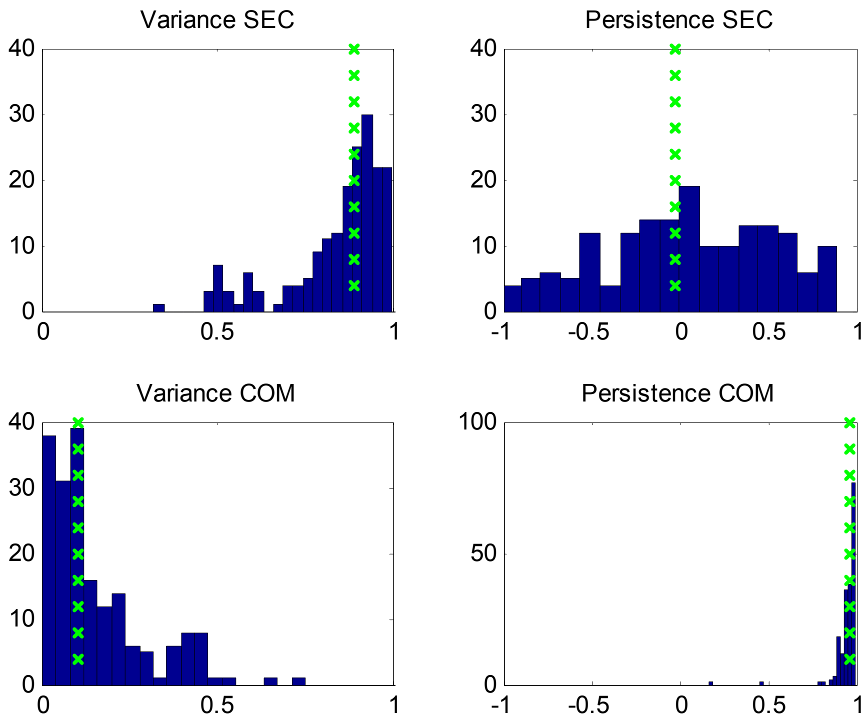
Figure 1 plots the breakdown of PCE inflation variance and persistence into a common and a sector-specific component across all sectors. Comparing the upper and lower left plots, it is clear that inflation variance is primarily induced by sector-specific shocks. The variance contribution of common shocks, by contrast, is concentrated toward zero. The right-hand plots of the figure show the decomposition of persistence across sectors. Sectoral shocks generally do not tend to cause much persistence. The distribution of persistence of the sectoral component is relatively flat, with the median sector having no persistence at all. The picture is dramatically different for the persistence of the aggregate component. Its distribution across sectors is strongly negatively skewed, with almost all sectors bunching up at very high levels of persistence.

These results are fully in line with those of Boivin *et al.* (2009) and Maćkowiak *et al.* (2009). In sum, from both the literature and our own simple factor model two seemingly

[1] There are different ways to estimate such a decomposition. Boivin *et al.* (2009) take a two step approach in which one first retrieves the common factors by principal components analysis, and subsequently estimates the observation equation (2) and the transition equations (3) and (4). Maćkowiak *et al.* (2009) opt for a Bayesian state-space model in which this is done jointly.

[2] We closely follow Boivin *et al.* (2009), with two minor exceptions. First, we do not force the Fed Funds rate to be a separate factor. Second, we estimate the observation equation by maximum likelihood, which is useful for later reference. Neither difference is quantitatively important for what follows.

Figure 1: Benchmark model - variance and persistence



Note: Inflation is standardized, such that $\sigma^2(\pi_{it}) = 1, \forall i$. Following Boivin *et al.* (2009), persistence is measured as the sum of the polynomial coefficients estimated for COM_{it}, and SEC_{it}. There is no natural lower bound on this persistence measure. To maintain visibility in the figures, we limit the scale to [-1,1]. The medians -green x's- and histograms take into account all sectors.

robust conclusions emerge. Across sectors,

$$\begin{aligned} \text{Stylized fact 1} &: \quad \sigma^2(COM_{it}) < \sigma^2(SEC_{it}) \\ \text{Stylized fact 2} &: \quad \rho(COM_{it}) > \rho(SEC_{it}) \approx 0. \end{aligned}$$

In words, for almost all sectors, inflation volatility is predominantly driven by non-persistent sector-specific shocks, while inflation persistence is due to the common component.

3 Sales, substitutions and measurement error in factor models

3.1 Prices and measurement

The scope for measurement error in the collection of prices is widely recognized. Shoemaker (2007) provides variance estimates associated with sampling. Evaluations of persistence, too, almost invariably discuss the scope for measurement error. Bils and Klenow (2004) and Boivin *et al.* (2009) are but two examples. The micro price setting literature more generally is concerned with sales and forced item substitutions (Golosov and Lucas, 2007; Klenow and Kryvtsov, 2008; Nakamura and Steinsson, 2008, 2009; Kehoe and Midrigan, 2010; Eichenbaum *et al.*, 2011). The consensus view in the literature is that prior to evaluating volatility and persistence, one should filter sales and substitutions.

Both sales and substitutions will impart particular short-run dynamics on inflation. Sales are changes in the price level that are undone after a brief period of time. In this paper we use the most restrictive and unambiguous sales definition among the various measures used by Nakamura and Steinsson (2008). This definition is the one-period symmetric 'V-shaped' pattern of the price level illustrated in the top row of Figure 2. Sales generate negative autocorrelation in inflation. An item substitution implies a change in the measured price level that does not necessarily reflect an actual decision to change price, but nevertheless generate a one-off blip in observed inflation. This is shown in the bottom row of Figure 2. To summarize, both sales and substitutions affect volatility (positively) and persistence (negatively) in observed inflation. To correctly measure volatility and persistence of inflation, one should control for these two measurement issues.[3]

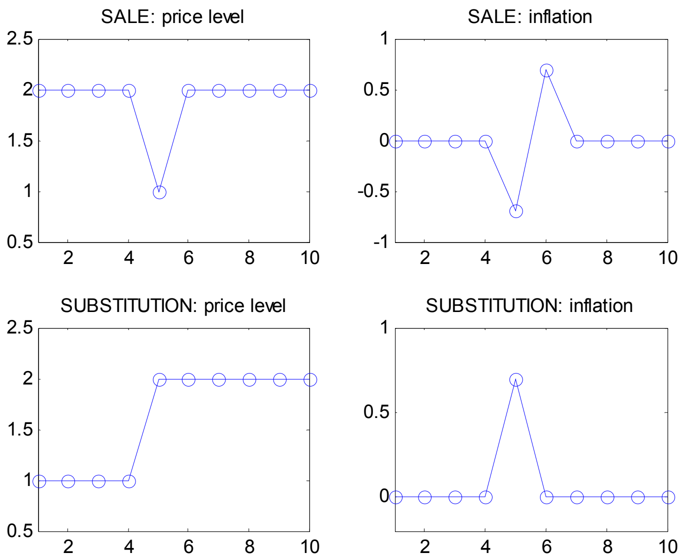
3.2 Factor models and measurement error

Factor models perform well in the presence of measurement error or misspecification, as shown in, among others, Stock and Watson (1998). This statement is, however, subject to an important qualification. The excellent performance of factor models relates to the identification of the common factors (C_t) and their loadings (λ_i). It does not pertain to inference on the residual.

This qualification is not always addressed in applied work. At times, this may well be innocuous. In research on prices, however, it is not. The reason is that measurement error in general, and sales and substitutions in particular, are important additional sources of

[3] A separate issue is to what degree these measurement issues are reduced by aggregating from the product level to the sectoral level. We address this issue quantitatively and in detail in Section 6.1.

Figure 2: Sales and Substitutions



sector-specific volatility.[4] Hence, in a factor model sales and substitutions will be subsumed in the residual e_{it}. But this points to a clear form of misspecification in the simple factor model (1)-(4): e_{it} is not a scalar process. Instead, it has multiple components.

3.3 Sales, substitutions and factor models

To convey the intuition of why the dimensionality of e_{it} matters for the study of inflation variance and persistence, consider the following example. Suppose inflation in sector i is driven by an aggregate component, COM_{it} as before, an AR(1) sector-specific shock P_{it}, with $\rho(P_{it}) > 0$, and an additional sector-specific component S_{it}, that captures sales and/or substitutions. Let S_{it} have positive variance, $\sigma^2(S_{it}) > 0$, and be orthogonal to P_{it}, $S_{it} \perp P_{it}$. Then

$$\begin{aligned}
\pi_{it} &= COM_{it} + SEC_{it} = \lambda_i' C_t + \underbrace{P_{it} + S_{it}}_{e_{it}} \\
\sigma^2(SEC_{it}) &= \sigma^2(e_{it}) = \sigma^2(P_{it}) + \sigma^2(S_{it}) \\
\rho(SEC_{it}) &= \rho(e_{it}) = \rho(P_{it} + S_{it}) = \frac{\sigma^2(P_{it})\rho(P_{it}) + \sigma^2(S_{it})\rho(S_{it})}{\sigma^2(P_{it}) + \sigma^2(S_{it})}.
\end{aligned}$$

[4]This is not to say that there cannot be an aggregate component to sales or substitutions. Rather, if there is one, a factor model is able to control for it, provided the number of common factors is sufficiently large.

Concerning persistence, it is apparent from Figure 2 that $\rho(S_{it}) \leqslant 0$. More specifically, substitutions result in uncorrelated spikes in inflation, while sales generate autocorrelation of -0.5. It is then immediate that

$$\begin{aligned} \sigma^2(SEC_{it}) &> \sigma^2(P_{it}) \\ \rho(SEC_{it}) &< \rho(P_{it}). \end{aligned}$$

In words, sales and substitutions make the sector-specific component of a factor model seem more volatile and less persistent. Interestingly, this works exactly in the direction of the stylized facts: simple factor models have invariably found sector-specific shocks to be very volatile and non-persistent.[5]

From the literature that analyses product-level prices, it is well known that the scope for sales and substitutions is huge. Cross-sectional heterogeneity aside, estimates for the monthly frequency of sales range from 7.4% (Nakamura and Steinsson, 2008) to over 20% (Klenow and Kryvtsov, 2008; Kehoe and Midrigan, 2010), and 3.4% (Bils and Klenow, 2004) to 5% (Nakamura and Steinsson, 2009) for item substitutions. The size of price changes induced by sales is also large - the median sale is 2.6 times the size of the median regular price change according to Nakamura and Steinsson (2008). Combined with the possibility of the biases in variance and persistence in the presence of sales and substitutions described above, this calls for a re-evaluation of the findings from simple factor models of sectoral inflation.

4 An extended factor model

To control for the effects of sales and substitutions we extend the simple factor model. We will refer to this extended model as the benchmark model. In eq. (2), as before, sectoral inflation π_{it} loads on a number of common factors C_t that evolve according to eq. (3). At the idiosyncratic level ($SEC_{it} = e_{it}$), inflation is still driven by a persistent process, P_{it}, but now also contains two additional components. On the one hand, we allow for an $iid-$component, I_{it}, that serves to absorb item substitutions. On the other hand, we introduce a moving average component M_{it} that serves to absorb the pattern implied by sales. Thus, the sector-specific component, previously eq. (4), now becomes

$$e_{it} = P_{it} + I_{it} + M_{it} \tag{5}$$

[5] Note that in the example, one would expect the identification of the factors and the estimation of factor loadings to be largely unaffected (Stock and Watson, 1998). The biases we study should therefore have negligible impact on studies that solely focus on aggregate components, e.g. Reis and Watson (2010).

where

$$
\begin{aligned}
P_{it} &= \rho_i(L)P_{it-1} + \varepsilon_{it} && (6)\\
I_{it} &= \epsilon_{it} && (7)\\
M_{it} &= \xi_{it} - \xi_{it-1} && (8)
\end{aligned}
$$

and

$$
(\varepsilon_{it}, \epsilon_{it}, \xi_{it})' \sim N(0_{3\times 1}, D),\ \ D^{1/2} = \begin{bmatrix} \sigma_i^{\varepsilon} & 0 & 0 \\ 0 & \sigma_i^{\epsilon} & 0 \\ 0 & 0 & \sigma_i^{\xi} \end{bmatrix}.
$$

The three (unobserved) components P_{it}, I_{it} and M_{it} have distinct persistence properties, and mutually orthogonal shocks ε_{it}, ϵ_{it} and ξ_{it}. We estimate the above factor model on the same data as Boivin *et al.* (2009). More precisely, we retain the factors from the simple model and estimate, for each sector, using maximum likelihood and the Kalman filter, the observation equation (2) accounting for (5)-(8).[6]

While the distinct persistence properties in the above specification ensure theoretical identification, this does not reveal much about the empirical performance of the estimator in finite samples. In Appendix A we document the favorable properties of the multi-component maximum likelihood procedure for various data-generating processes (DGP) of interest. In short, when the DGP has multiple components, the estimator identifies multiple components and estimates persistence close to that of the DGP. Not surprisingly, for lower underlying persistence, the estimator has lower precision. Importantly, estimating single component processes (ARs) on multi-component data generates estimates not even in the ballpark of the true persistence. On the other hand, when the DGP truly is a single component process, estimating a multi-component process does not imply substantial biases.

[6] Note that when $\rho_i(.)$ has zero coefficients at all lags, there is an identification issue, as the likelihood is then flat in σ_i^{ϵ} and σ_i^{ε}. If this occurs, we set $\sigma_i^{\varepsilon} = 0$, such that $P_{it} = 0, \forall t$ and I_{it} absorbs all the variance allocated to P_{it}. Related, at $\sigma_i^{\varepsilon} = 0$ the likelihood is flat in $\rho_i(L)$, $\forall L$. Similarly, we then set zero coefficients at all lags. However, these cases hardly occur in practice. In other words, these ridges are typically located away from the likelihood's maximum. We have also estimated Bayesian versions of the model. While these make it easier to achieve identification through the prior, they also tend to attribute non-zero prior variance to each component, which we prefer to not impose.

5 Re-evaluating the stylized facts

5.1 Model selection

Observe that the benchmark factor model, through eq. (5), nests the simple factor model, via eq. (4). Therefore, standard model selection criteria are available to choose between the simple model and the benchmark factor model. If the new components I_{it} and M_{it} are of no importance, the increase in the likelihood of the benchmark factor model relative to the simple model will be marginal. Selection criteria penalizing for the additional number of parameters (i.e. σ_i^{ϵ}, σ_i^{ξ}) will then favor the more parsimonious simple model.

Table 1 shows that in almost 90% of the sectors the data is better described by the benchmark factor model than by the simple model. In only 12% of all sectors is there no notable improvement in terms of fit by allowing multiple components at the sectoral level.

Table 1: Model selection criteria

	Simple	Benchmark
AIC	12%	88%
$SBIC$	12%	88%

Table 2 provides an alternative view on the estimated benchmark factor model. It characterizes sectors by the relevance of their idiosyncratic components.[7] A number of features stand out. First, all sectors have a persistent component. Second, for more than half of the sectors both I and M play a role. Third, only 10% of the sectors are well captured by a single component process.[8] Thus, from this perspective too, the scope for additional components is substantial.

Table 2: Sectors and idiosyncratic components

Components	% sectors
P	10%
I	0%
M	0%
$P+I$	24%
$P+M$	14%
$I+M$	0%
$P+I+M$	52%

[7] For the purpose of this table, we consider a component irrelevant for a particular sector if it accounts for less than 1% of the variance in the sectoral component.

[8] Not surprisingly, these are also the sectors for which the information criteria select the simple model over the extended model.

5.2 Variance

The additional components are also quantitatively important. Figure 3 decomposes the variance of the sectoral component into P, I and M for all sectors. A point at the origin implies that all the sectoral variance is attributed to the I component. A sector located at the top corner signifies 100% of its sectoral variance stems from the P component, and analogously the right bottom corner signifies $\sigma^2(SEC_{it}) = \sigma^2(M_{it})$. If a sector is located on, say, the $I - P$ axis, this implies it has no M component. The key message from Figure 3 is the enormous degree of heterogeneity across sectors. Further details about the variance decomposition are also documented in Table 3. First, in half of the sectors, most of the variance in SEC is due to P. Conversely, the other half of the sectors have most of their sectoral variance coming from sales and substitutions. Second, substitutions appear to be quantitatively more important than sales at the sectoral level.

At face value, are these numbers reasonable? Micro evidence conveys similar magnitudes. As discussed in Section 3.3, product-level data indicate that sales are not only very frequent, but also tend to be large in magnitude. Substitutions are far less frequent, but can have important effects on measured dynamics nonetheless (see Nakamura and Steinsson, 2009).[9] In addition, heterogeneity between sectors prevails in micro data, too.

Table 3: Variance decomposition - SEC

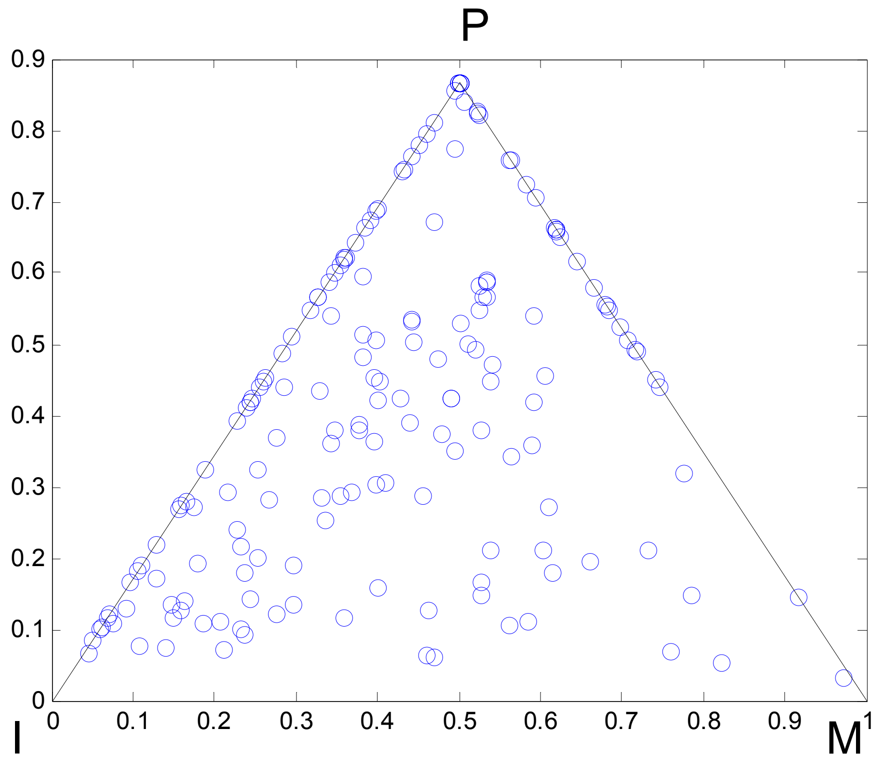
	Median	Mean
P	0.51	0.51
I	0.28	0.32
M	0.12	0.16

Table 4 shows, for each component, the median and mean variance contribution to π_{it} across sectors. As expected, the variance contribution of the common component is around 10-15%, consistent with the evidence in the literature. The remaining 85-90% inflation variance is driven by sector-specific shocks. But as the next three rows in the table (and Figure 3) indicate, a non-negligible part of the sectoral variance is due to the I and M component. The median contribution of the persistent sectoral component P to total sectoral inflation is 43%.

Factor models in the literature have the sharp result that for the median sector, sector-specific shocks are almost an order of magnitude more important than aggregate shocks. This large difference dominates any cross-sectional heterogeneity. Taking the ratio of common to

[9] High frequency of sales (or substitutions) in the underlying micro data of the indexes does not necessarily translate into a high variance contribution of the M (I) component. We elaborate on this issue in Section 6.2.

Figure 3: Variance contributions - SEC



sectoral variance contributions in the simple model, it appears that only 5 out of 190 sectors (3%) are more affected by aggregate shocks than by sectoral shocks. The first row of Figure 4 shows that result, with almost no mass below 1.

However, simple factor models ignore that much of the variance of the sectoral component is driven by sales and substitutions. Filtering those out, the benchmark model estimates sectoral shocks to be three to four times as volatile as aggregate shocks for the median sector, as is apparent in the second row of Figure 4. Importantly, aggregate shocks are more important than sector-specific shocks for one sector in four. Thus, while sectoral shocks tend to dominate, this is certainly not true for all sectors.

5.3 Persistence

Because sales and substitutions generate non-negligible sector-specific variance, they are likely to influence evaluations of persistence. In Section 3.3 we showed how multiple components could lead to underestimating persistence for the simple example of an AR(1) data generating process. For more elaborate processes (e.g. with longer lags) and persistence measures (e.g. sum of polynomial coefficients) the direction and size of the bias induced by

Table 4: Variance decomposition - inflation

	Median	Mean
COM	0.10	0.17
SEC	0.89	0.85
P	0.43	0.44
I	0.25	0.27
M	0.10	0.14

sales and substitutions is less clear cut *a priori*. Whether persistence in the simple factor model is substantially biased is thus ultimately an empirical question.

Figure 5 therefore compares persistence in the simple model (on the x-axis) to persistence in the benchmark factor model (y-axis). The result is overwhelmingly clear: 89% of all sectors lie above the 45°-line. In other words, the simple factor model substantially underestimates the persistence of sectoral shocks. The two right-hand quadrants contain sectors that exhibit positive persistence in the simple factor model (about 50 % of all sectors). For these, the median bias in persistence is 45%. In the upper left quadrant, the benchmark factor model finds positive persistence, where the simple model fails to detect any. This quadrant contains 16% of all sectors. For the remaining sectors, in the bottom left quadrant, neither of the factor models find any positive persistence.

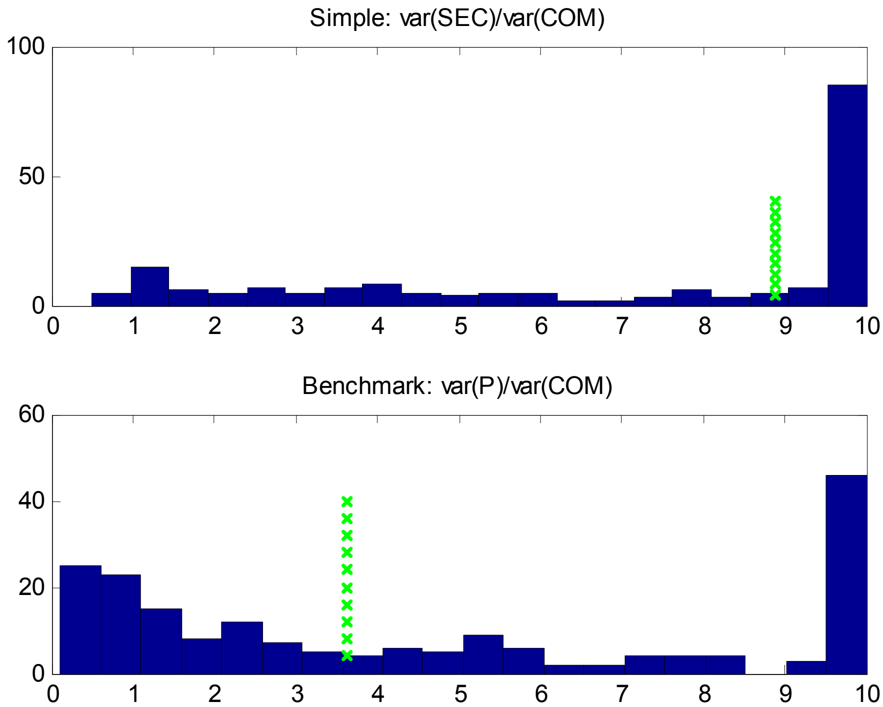
These biases substantially alter the view on the persistence of sectoral shocks. The top row of Figure 6 first reprints the cross-section of persistence measures in the simple model. It is a rather flat distribution, with the median sector having zero persistence. This is the second stylized fact. The benchmark factor model (bottom row) shows that, actually, sectoral persistence is strongly negatively skewed. A lot of sectors cluster at very high levels of persistence. For the median sector, persistence is estimated at just above 0.4.

5.4 Facts or artefacts?

In this section we have re-evaluated the two stylized facts. In our view, they appear to be artefacts. The first stylized fact in the literature is that the variance of sector-specific shocks is almost an order of magnitude higher than aggregate shocks across sectors. However, this ignores that much of the variance of the sectoral component is driven by sales and substitutions. Filtering those out, we estimate sectoral shocks to be three to four times as volatile as aggregate shocks for the median sector. Importantly, heterogeneity across sectors is large and we find that aggregate shocks are more important than sector-specific shocks for one sector in four.

The second stylized fact is that median persistence of sector-specific shocks is zero and,

Figure 4: Variance ratios

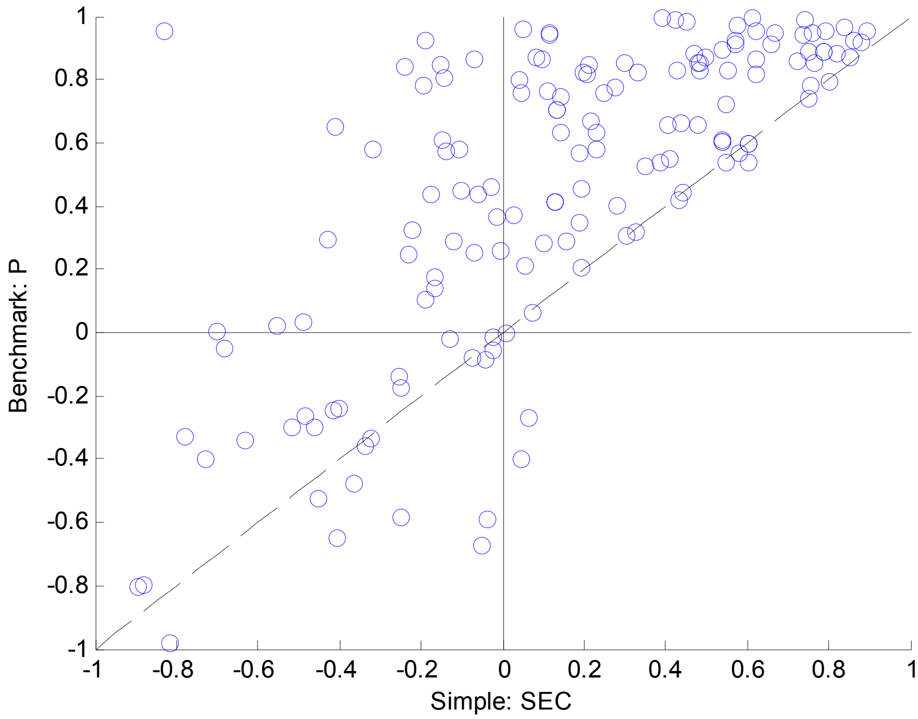


Note: Due to the presence of sectors with virtually no variance in the common component, values above 10 are truncated at 10.

accordingly, persistence in sectoral inflation almost exclusively driven by aggregate shocks. Filtering out sales and substitutions, we eliminate a bias present in previous estimates and obtain a median persistence of the sectoral component around 0.4. The mode of this persistence is around 0.8.

This establishes the main result of the paper - the two so-called stylized facts are everything but facts. Rather, they are an artefact of failing to appropriately account for sales and substitutions in simple factor models. In the remainder of the paper we discuss alternative interpretations, quantify aggregation properties from product to sector-level indexes, validate the components in the benchmark model with evidence from micro studies and provide numerous robustness exercises.

Figure 5: Persistence - Bias



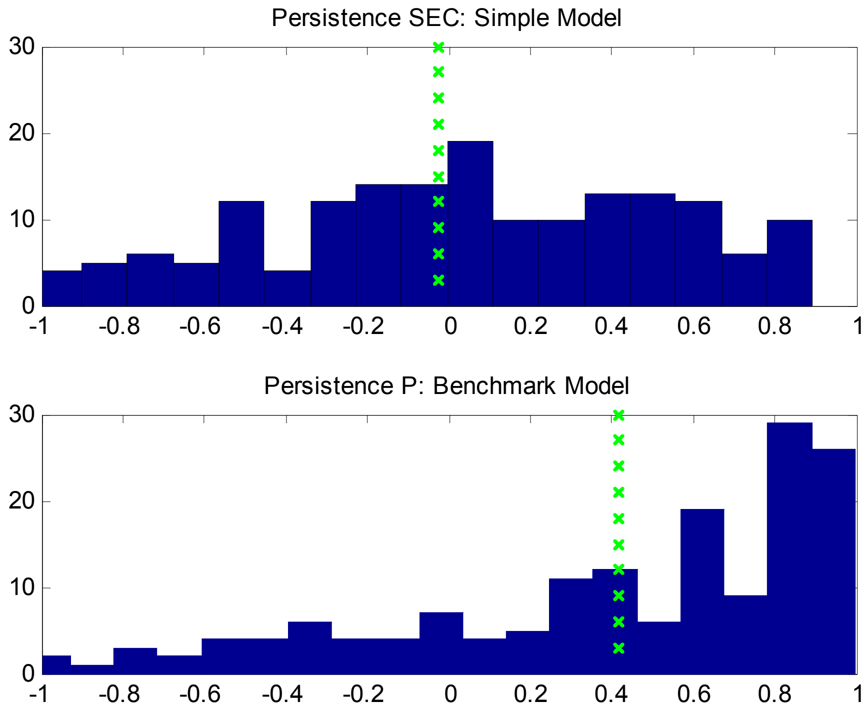
6 Discussion

6.1 Aggregation

Since sectoral price indices are combining price quotes across multiple cities, stores and products, one might expect sales, substitutions and general measurement error to average out at the sectoral level. While there definitely is scope for aggregation to reduce the need for our additional components, there are a number of elements that reduce the tendency of these components to be aggregated away at the sector level and at the sampled (monthly) frequency. In what follows, we first discuss aggregation under ideal conditions - uncorrelated homogenous-size price changes. We then discuss and quantify two aspects that decrease the power of aggregation: correlated sales or substitutions and heterogeneity in the size of price changes. Throughout we make the simplifying assumption that all products receive equal weights in the sector-level indices.

The discussion below concerns what fraction of the volatility of product-level sales and substitutions remains at the sector level. But let us start by stating that the dynamics,

Figure 6: Persistence - Simple vs. benchmark factor model



in particular the persistence properties, induced by these phenomena remain unchanged by aggregation: An *iid* movement induced by substitution at the product level induces an *iid* movement in the corresponding sector index. Similarly for the MA component induced by sales.[10]

[10] Recall eq. (8), which at the sector level yields

$$M_{it} = \sum_j (\xi_{jit} - \xi_{jit-1})$$

where j indexes products within a sector and ξ_{jit} is uncorrelated across t. Then $Var(M_{it}) = 2Var(\xi_{ji})$ and autocorrelation at the sector level is

$$\begin{aligned}
\rho(M_{it}, M_{it-1}) &= \frac{1}{Var(M_{it})} Cov\left(\sum_j (\xi_{jit} - \xi_{jit-1}), \sum_j (\xi_{jit-1} - \xi_{jit-2})\right) \\
&= \frac{1}{Var(M_{it})} Cov\left(\sum_j (-\xi_{jit-1}), \sum_j (\xi_{jit-1})\right) = \\
&= -\frac{Var\left(\sum_j \xi_{jit-1}\right)}{Var(M_{it})} = -0.5
\end{aligned}$$

The first reason product level measurement errors do not completely cancel out at the sectoral level is that the number of product prices sampled per month is limited. The consumer price index (CPI), which is the main source of the sectoral PCE price indices we use, is based on 70.000-80.000 prices across 388 entry-level items (ELIs) roughly corresponding to the PCE sectors we study, yielding a mean number of observations slightly above 200 product prices per ELI/PCE sector and month. Theoretically, in absence of any aggregation problems, the ratio of the standard deviation of the index, σ_{index}, to the standard deviation of the product price, $\sigma_{product}$, is $\frac{1}{\sqrt{N}}$. This implies that for the sector with the mean number of observations $1/\sqrt{200} = 7\%$ of the variation induced by sales and substitutions at the product level would remain at the sector level.[11] The first column in Table 5 present the corresponding numbers for the empirically relevant range of sample sizes.

Correlated sales or product substitutions could occur due to sector-specific shocks: low demand can build up inventory and induce larger sales, technical progress can generate product turnover and induce product substitutions, etc.[12] To illustrate the impact of correlated sales or substitutions we perform the following exercise. For a sample length equal to ours (T=353) we randomly generate sequences of sales (the outcomes are indistinguishable for the case of substitutions). At any point in time, an individual product is on sale with a particular frequency. If there is no sale, the price remains constant. When there is a sale, the price change is a sum of two random components from the normal distribution: A common component generates correlated variation across products within an index and an idiosyncratic component generates uncorrelated variation. We generate many product level price series, and construct inflation indices from them, for a variety of numbers of goods in the index, N. In this exercise the only reason that the theoretical prediction of the effect of aggregation, $1/\sqrt{N}$, does not obtain is that the size of sales contain a common component that makes them correlated. We let the correlation equal 0.25. In Table 5 we present the results for a range of frequencies, recalling from Section 3.3 that micro evidence indicates that the median monthly frequency of sales are in the range from 7.4% to over 20%, and 3.4% to 5% for item substitutions. The first, and least surprising, result to note is that correlated sales do not aggregate away very well. Secondly, aggregation actually works better the lower the frequency is. The intuition is that for low frequencies the realized correlation

which coincides with the product-level autocorrelation of M_{jit}.

[11] Whether that 7% represents a large fraction of the index's variance, which also contains regular price changes, is a different question. It depends on the relative volatility of sales and substititions vs. regular price changes at the product level. Micro level data suggest that sales and to a smaller degree, substitutions, may well cause substantially more volatility than regular price changes (see Section 3.3 for details). This makes effectively controlling for them at the index level all the more needed.

[12] Note that the price data we work with is seasonally adjusted, so correlation in sales that follow a seasonal pattern are filtered out.

tends towards zero as most prices are unchanged. To specifically address the question of how well aggregation works for the median sector, we read from the table that for $N = 200$, the ratio of the standard deviation of the index relative to the standard deviation of its underlying products $\frac{\hat{\sigma}_{index}}{\sigma_{product}}$ is roughly 0.2 at the empirical frequency of sales and roughly 0.1 at the empirical frequency of substitutions. Interestingly, results at the empirical frequency of sales are approximately unchanged for $N = 500$ and $N = 1000$. In other words, roughly 20% (10%) of the product level volatility from sales (substitutions) remains at the sector level if correlation is 0.25. This is substantially more than for uncorrelated price changes.

Table 5: Aggregation and sales/substitutions - correlation

		Frequency			
Number of products in index: N	$1/\sqrt{N}$	0.25	0.1	0.05	0.01
50	0.1414	0.2849	0.2110	0.1796	0.1495
100	0.1000	0.2685	0.1865	0.1497	0.1113
200	0.0707	0.2595	0.1728	0.1319	0.0864
500	0.0447	0.2536	0.1640	0.1205	0.0670
1000	0.0316	0.2519	0.1611	0.1162	0.0591

Note: The table reports the ratio of the standard deviation of an index, $\hat{\sigma}_{index}$, relative to the (homogenous) standard deviation of its underlying products, $\sigma_{product}$, for various N and frequencies, but for a fixed correlation of 0.25. The first column is the theoretical relation without correlation and the four subsequent columns the small-sample (T=353) results across 5000 replications.

It is plausible that not all products within a sector exhibit the same unconditional size of sales or substitutions. Heterogeneity in size of sales or substitutions within a sector weakens aggregation. Intuitively, the degree to which various sales or substitutions cancel out at the sector level decreases with size heterogeneity.

To quantify the effect of heterogeneity we perform a similar exercise to the one above. We let the size of the sale or substitution be a random draw from a normal distribution whose standard deviation is drawn from a uniform distribution to induce heterogeneity in size. As a rough reference for the within-sector size heterogeneity we use heterogeneity between major groups from Nakamura and Steinsson (2008). It shows that the standard deviation of the sales size, σ_{size}, is one third of the mean sales size, μ_{size}, for both of the sample periods they report.

We report the results for a range of heterogeneity in Table 6. We note that the quantitative impact of heterogeneity in size is limited for this range of heterogeneity. Results are indistinguishable for sales and substitutions, and independent of frequency.

In this section we have quantified how much of product-level variation in prices due to sales and substitutions remains at the sector-level. We first noted that the empirical sample size in the mean sector is limited. This makes it likely that sales and substitutions generate

Table 6: Aggregation and sales/substitutions - heterogeneity

		σ_{size}/μ_{size}				
Number of products in index: N	$1/\sqrt{N}$	0.95	0.75	0.5	0.25	0.05
50	0.1414	0.1952	0.1761	0.1577	0.1456	0.1416
100	0.1000	0.1376	0.1247	0.1118	0.1031	0.1000
200	0.0707	0.0973	0.0882	0.0790	0.0728	0.0707
500	0.0447	0.0616	0.0558	0.0500	0.0460	0.0447
1000	0.0316	0.0436	0.0395	0.0353	0.0325	0.0317

Note: The table reports the ratio of the standard deviation of an index, $\hat{\sigma}_{index}$, relative to the mean of the heterogenous standard deviation of its underlying products, $\sigma_{product}$, for various ratios of the within sector standard deviation of the size of sales, σ_{size}, to the mean size of sales, μ_{size}. The first column is the theoretical relation without heterogeneity, the four subsequent columns the small-sample (T=353) results for lower frequencies of price change across 5000 replications.

significant variance at the sectoral index level. We then separately quantified the impact of two factors that further weaken aggregation: correlation and heterogeneity. Empirically, across sectors, there are different numbers of products per sector, varying degrees of heterogeneity across products within each sector, and varying degrees of correlation between those products. Each of these factors, and possible interactions between them affect how well aggregation works.

6.2 Validation across sectors - sales and substitutions

The relationship between the variance of our sales (substitutions) component and the fraction of price changes that are sales (substitutions) is not predicted to be one-to-one. Several factors, including heterogeneity across sectors in the relative size of sales price changes and in aggregation properties, weaken this link.[13] Keeping this in mind, we nevertheless provide an informal validation of our results by examining to what degree the presence of the I and M components in our benchmark factor model coincide with Nakamura and Steinsson's (2008, henceforth NS) product-level CPI data evidence. Given the tenuous theoretical relationship we focus on the extreme results: we compare whether a sector has a sales or substitution component at all in our results to the prevalence of sales and substitutions in that 'major

[13]For an additional reason why aggregation need not preserve the relation between our components and the micro data, consider the following extreme example. Two sectors A and B each have 100 products sampled. In sector A all products have sales, while in sector B only 1 product is ever on sale. Sales in sector B have no hope of averaging out across products, and will thus generate an M component in the index of sector B. The index for sector A, by contrast, may well not be affected much by product-level sales, as they have the scope to average out across products. Thus, despite being a sales-intensive sector, sector A may not require a M component. The opposite is true for sector B, despite having very few sales at the micro level. A similar logic applies to substitutions.

group' according to NS.[14] As documented above in Table 2, sales and product substitution components, M and I respectively, are only present in a subset of the PCE sectors we study. In particular, Table 2 documents that 24% of sectors have no M component while 34% of sectors have no I component.

NS document that Utilities, Vehicle fuel, Services (excl. travel) and Travel have virtually no sales, and at the opposite end of the spectrum that Apparel, Household Furnishing and Food (processed and unprocessed) have the highest prevalence of sales.

Comparing our results for which sectors lack a sales component we note that they coincide to a reasonable degree with NS. Key utilities sectors (Electricity and Gas) have no sales component. Gasoline, on the other hand, does have a sale component contrary do what NS results indicate.[15] In line with NS most travel sectors (Taxicab, Bus and Other) have no sales component. Services (excl. travel) is a very diverse group. We note that an above average fraction (31%) of the PCE service sectors lacks a sales component, in line with NS results.

Switching to sectors which have lots of sales according to NS, we confirm that sectors within Apparel (clothes for men, women and children, respectively) have a sales component. Four of the five Household Furnishing sectors have a sales component. For food sectors a non-negligible fraction of them lack a sales component, contrary to the evidence in NS.

The analogous exercise for product substitution validates our method by lining up very well with NS. Their product-level data indicates that product substitution is most common in Apparel and Transportation goods (mainly cars), and least common in Vehicle fuel and Utilities. We find no substitution component in Gasoline or the utilities sectors Electricity and Water. Furthermore, and also in line with NS, we find a substitution component in all three clothes sectors and in all of the nine transportation good sectors.

To summarize, we find that our results on which sectors have sales and substitutions coincide roughly with what NS find. But recall that this is only indicative in terms of validation of our method, as several factors may distort the relation between product level sales and substitutions and the corresponding sector level component.

6.3 An alternative rationale for I and M

Note that classical uncorrelated measurement error has very similar effects to those of sales and substitutions. In particular, measurement error in prices will result in negative auto-

[14] An additional factor that complicates comparisons is the imperfect mapping between PCE sectors and the CPI 'major groups' and ELIs that NS reports.

[15] The contradiction is with NS's benchmark results which are based on the BLS flag for sales. But, NS explain why the 'V-shaped' filter finds substantial amounts of sales for gasoline, also on product-level data. The issue is caused by high volatility in the price in combination with a tendency for discrete price changes.

correlation in inflation and can thus generate a M component. Analogously, measurement error in inflation will result in an *iid*-component, similar to I. As such, measurement error is observationally equivalent to sales and substitutions.

First, whether it is sales and substitutions or other measurement issues is not the primary concern. Irrespective of which it is, it seems important to filter out such non-fundamental variation prior to evaluating variance and persistence of sectoral shocks. Our benchmark factor model does just that.

Second, for some purposes, it may actually be useful to quantify how much of the non-persistent sector-specific fluctuations is due to measurement error, rather than due to sales or substitutions. For instance, many studies make conjectures about plausible degrees of measurement error, in order to verify whether it could drive their results (e.g. Bils and Klenow, 2004). To inform such questions, we here adapt our factor model to shed light on the importance of measurement error, relative to sales and substitutions.

One way to overcome the observational equivalence between sales and substitutions on the one hand, and measurement error in prices and inflation on the other, is to use quantities. *A priori*, there is no apparent reason to expect measurement error in prices to affect quantities. Sales and substitutions, by contrast, can be expected to influence quantities. In Appendix B, we lay out an extension to the factor model that separates measurement error from sales and substitutions. We here summarize the results of that model specification briefly, while the appendix contains the results on variance and persistence across sectors.

Table 7: Variance decomposition - measurement error

	Benchmark model		Accounting for measurement error	
	Median	Mean	Median	Mean
COM	0.10	0.17	0.11	0.16
SEC	0.89	0.85	0.89	0.85
P	0.43	0.44	0.47	0.46
I	0.25	0.27	0.18	0.23
M	0.10	0.14	0.07	0.15
η	–	–	0.11	0.16

Table 7 indicates that for the median sector, 11% of inflation variance is due to measurement error (η). In the benchmark model (without quantities that isolates measurement error), the I and M components seem to soak up that variance, as expected. Nevertheless, even in the model that accounts separately for pure measurement error, the I and M components still appear very relevant. Importantly, the conclusions for the relative variance and persistence of common and sectoral shocks remain unchanged from our benchmark model.

7 Robustness

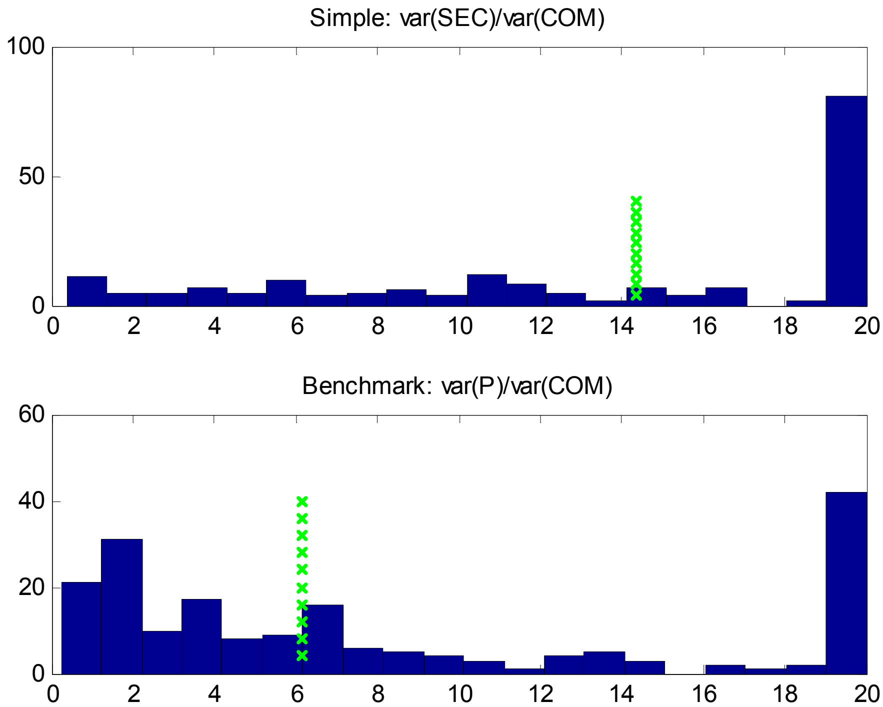
The main results of our benchmark model go through for other data sets and for variations in the model specification considered. First, we consider the effect of shortening the sample period to 1984-2005. This subsample is also considered in Boivin *et al.* (2009), and serves to isolate the results from the very different behavior of macroeconomic aggregates prior to and during the early eighties disinflation and the start of the so-called Great Moderation. Figures 7 and 8 document the variance and persistence of the various components for this period. Compared to the full sample results documented in Figure 4 the relative variance of aggregate shocks is substantially smaller already in the simple model. This is not unexpected, since decreased variance of aggregate conditions is exactly what the Great Moderation represents. Comparing the relative importance of aggregate shocks in the simple factor model with that of the benchmark model, which accounts for sales and substitutions, again shows how the former model substantially overestimates the relative importance of the sector-specific component. While the traditional approach suggests that in the median sector idiosyncratic shocks are roughly 14 times more important than aggregate shocks, the benchmark model finds this to be only 6 times as large. One could argue that this high relative variance of idiosyncratic shocks was particular to the Great Moderation era and might well disappear when considering more recent data.[16] Turning to persistence in Figure 8, the results for the subsample are very similar to those for the full sample. A simple factor model reveals no persistence due to sectoral shocks for the median sector, while substantial persistence is visible in the model that accounts for sales and substitutions. Again, one observes the strong concentration of sectors at very high levels of persistence.

Second, to assess the generality of their results, Boivin *et al.* (2009) also consider sectoral PPI series, and document that the stylized facts continue to hold. As an additional robustness check, we therefore re-estimate the simple model and the benchmark factor model for the PPI data. Here too, the results are very similar: The simple model confirms the first stylized fact and estimates sectoral shocks to be 9 times more volatile than aggregate shocks for the median sector (Figure 9). The benchmark model reduces this ratio to below 4. In terms of persistence, too, a similar bias appears to be present. As is clear from Figure 10, the standard, simple approach finds no persistence (stylized fact 2), while the benchmark approach indicates substantial persistence.

Third, we now switch from documenting robustness in terms of data to robustness in terms of model specification. Recall that our sales definition, operationalized by eq. (8), is

[16] Unfortunately, a change in the PCE definition makes extending the sample and verifying this conjecture infeasible.

Figure 7: Variance - subsample



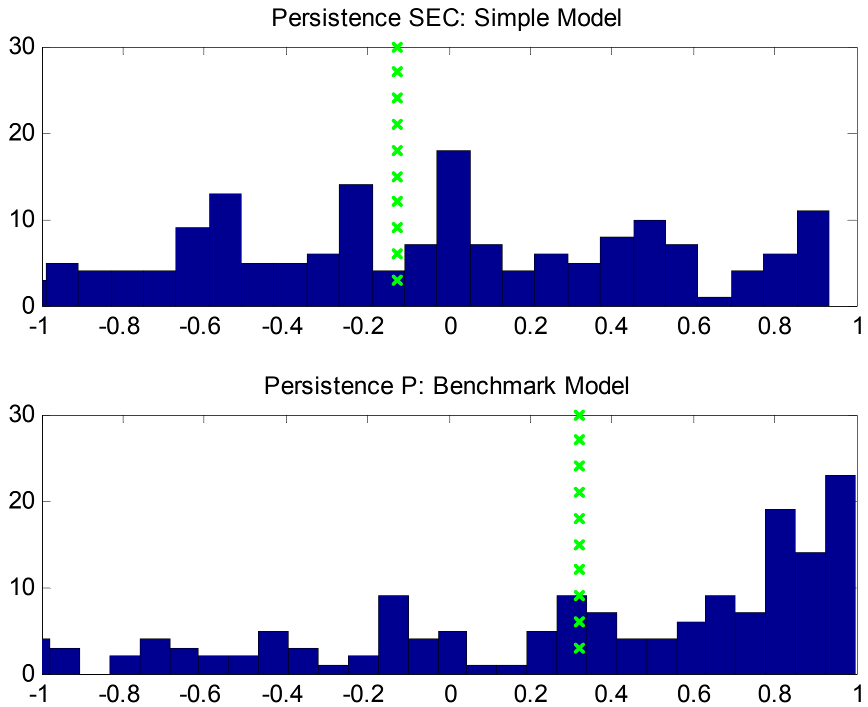
the most restrictive among the alternatives in the literature. We also explore a less restrictive sales definition that replaces eq. (8) by

$$M_{it} = \rho_{m,i}(L)M_{it-1} + \xi_{it}$$

and where identification is achieved by restricting the sum of the lags to be negative, $\rho_{m,i}(1) < 0$, while for the persistent component, P_{it}, we require $\rho_i(1) > 0$. Also this alternative specification yields very similar results to our benchmark model, both in terms of volatility of each component and persistence of P_{it}.

Finally, we perform a robustness exercise where we reduce the lag length of the persistent component, P_{it}. The reason for this exercise is that 13 lags might appear to over-parameterize the model, in particular in the presence of the two additional components. The results are very similar to our benchmark specification when either imposing 3 lags or using standard lag selection criteria.

Figure 8: Persistence - subsample

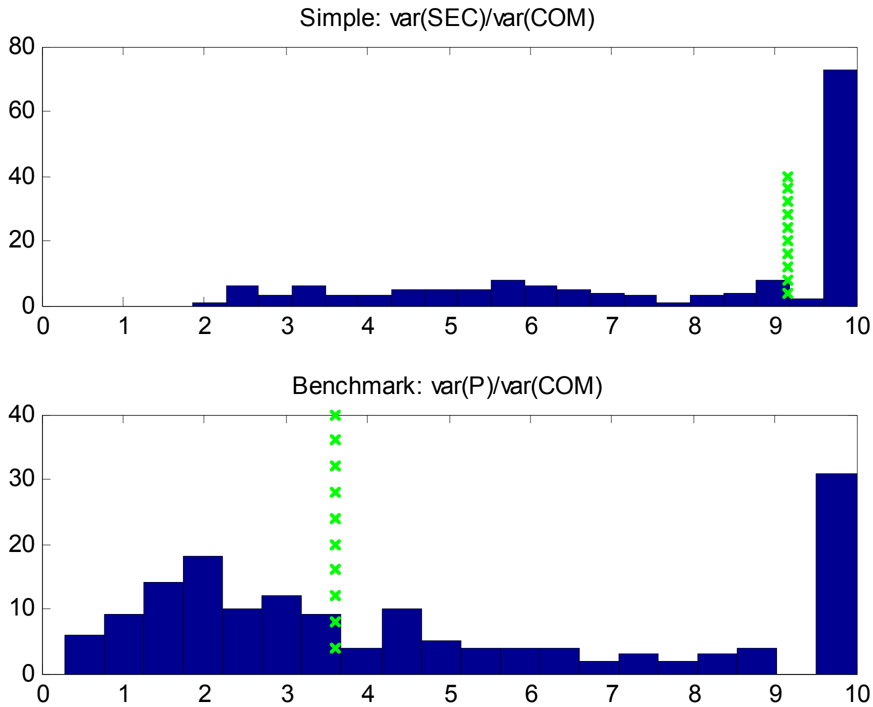


8 Conclusion

The variance contribution of idiosyncratic shocks to sectoral inflation may be a lot smaller than previous evidence suggests. While earlier factor models indicate that sectoral shocks are almost ten times as volatile, our estimates point to a ratio of sector-specific to aggregate volatility of three to four for the median sector. Moreover, heterogeneity prevails: for a quarter of the sectors in our data, aggregate shocks appear to be a more important source of fluctuations than sector-specific shocks. Persistence in inflation arises from both aggregate and sector-specific shocks. Our results show that the absence of persistence in the response to sectoral shocks in earlier empirical analysis is method-driven; an artefact.

Our results bring the micro and macro evidence on sluggishness closer together. Initially, high frequency volatility in sectoral price series seemed puzzling from the perspective of inflation inertia at the macro level. Boivin *et al.* (2009) reconcile this (non-filtered) fast-micro and slow-macro evidence by invoking conditionality: it matters whether a shock is aggregate or sector-specific. Our results, by contrast, reveal that there is no conflict between the micro and macro evidence: Applying filters similar to those used in research on micro

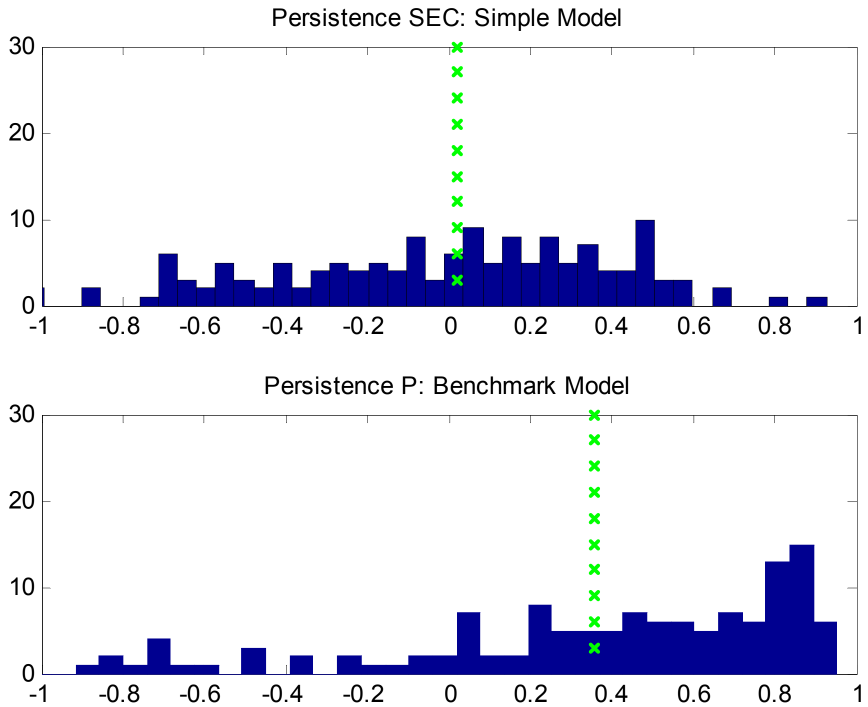
Figure 9: Variance - PPI



(product-level) price data, thereby taking account of sales and substitutions, one obtains very similar results. Lower volatility and higher persistence are obtained when sales and substitutions are accounted for, which is apparent from micro studies such as Nakamura and Steinsson (2008), Kehoe and Midrigan (2010) and Eichenbaum *et al.* (2011) as well as from our benchmark factor model. Thus, our results align well with the micro evidence. Furthermore, these results contrast starkly with those obtained at both micro and macro level for non-filtered data. In particular, prices then appear very volatile, and have low persistence. This is evident from the simple factor model (Boivin et al., 2009) and micro studies that do not control for sales (e.g. Bils and Klenow, 2004).

Our results have important implications for model calibration and validation. As discussed in Maćkowiak and Smets (2009), models of rational inattention (Maćkowiak and Wiederholt, 2010) and menu costs (Golosov and Lucas, 2007), for instance, often rely on sector-specific shocks that are an order of magnitude larger than aggregate shocks. Our analysis suggests that this is not necessarily what sectoral price data convey. Rather, we find that in one quarter of all sectors aggregate shocks are a more important source of fluctuations than sector-specific ones.

Figure 10: Persistence - PPI



In light of the above evidence, models of price rigidities should not be rejected because they fail to generate a sluggish response to aggregate shocks and a fast response to idiosyncratic disturbances. Persistence occurs irrespective of the source of the shock. Finally, there is a tremendous amount of heterogeneity between sectors in these findings, again consistent with the micro-evidence (Nakamura and Steinsson, 2008).

The results of the present paper also have implications for the appropriate design of core inflation indices. The fact that sector-specific dynamics are best characterized as multi-component processes means that sectors should not be excluded from a core index based on simple statistics such as unfiltered persistence or volatility. Such exclusion-based core measures are commonly used by central banks, most explicitly by Bank of Canada. The Federal Reserve's motivation for focusing on PCE excluding food and energy is a related short-cut in that direction.

References

[1] Bils, M., Klenow, P.J., 2004, "Some evidence on the importance of sticky prices", Journal of Political Economy 112, 947-985.

[2] Boivin, J., Giannoni, M.P., Mihov, I., 2009. "Sticky prices and monetary policy: evidence from disaggregated US data", American Economic Review 99, 350-384.

[3] Carvalho, C., Lee, J.W., 2010. "Sectoral price facts in a sticky-price model", Federal Reserve Bank of New York, mimeo.

[4] Eichenbaum, M., Jaimovich, N., Rebelo, S., 2011. "Reference prices, costs, and nominal rigidities", American Economic Review 101, 234-262.

[5] Golosov, M., Lucas, R.E. Jr., 2007. "Menu costs and Phillips curves", Journal of Political Economy 115, 171-199.

[6] Kehoe, P., Midrigan, V., 2010. "Prices are sticky after all", NYU and Princeton University, mimeo.

[7] Klenow, P.J., Malin, B.A., 2010. "Microeconomic evidence on price-setting", in Handbook of Monetary Economics Vol 3a, by Friedman, B., Woodford, M. (eds.), Elsevier.

[8] Klenow, P.J., Kryvtsov, O., 2008. "State-dependent or time-dependent pricing: Does it matter for recent U.S. inflation?", Quarterly Journal of Economics 123, 863-904.

[9] Maćkowiak, B., Moench, E., Wiederholt, M., 2009. "Sectoral price data and models of price setting", Journal of Monetary Economics 56, 78-99.

[10] Maćkowiak, B., Smets, F.R., 2009. "Implications of microeconomic price data for macroeconomic models", in Understanding Inflation and the Implications for Monetary Policy: A Phillips Curve Retrospective by Fuhrer, J., Sneddon Little, J., Kodrzycki, Y.K, Olivei, G.P. (eds.), MIT Press.

[11] Maćkowiak, B., Wiederholt, M., 2009. "Optimal sticky prices under rational inattention", American Economic Review 99, 769-803.

[12] Maćkowiak, B., Wiederholt, M., 2010. "Business cycle dynamics under rational inattention", ECB and Northwestern University, mimeo.

[13] Nakamura, E., Steinsson, J., 2008. "Five facts about prices: A reevaluation of menu cost models", Quarterly Journal of Economics 123, 1415-1464.

[14] Nakamura, E., Steinsson, J., 2009. "Lost in transit: product replacement bias and pricing to market", Columbia University, mimeo.

[15] Reis, R., Watson, M.W., 2010. "Relative goods' prices, pure inflation, and the Phillips correlation", American Economic Journal: Macroeconomics 2, 128-57.

[16] Shoemaker, O.J., 2007. "Variance estimates for price changes in the consumer price index: January-December 2006", Bureau of Labor Statistics, mimeo.

[17] Stock, J.H., Watson, M.W., 1998. "Diffusion indexes", NBER WP No. 6702.

Appendix A: Estimator properties in finite samples of simulated data

This appendix documents empirical properties of the maximum likelihood estimator used in the paper. In particular, we simulate data from various one- and multicomponent processes for sample lengths equal to our data ($T = 353$). For each of these, we estimate single component (P, as in eq. (4), henceforth AR) and multicomponent processes ($P + I + M$, as in eq. (5)-(8), henceforth PIM). For each process we use one lag for the AR (P) component. The Monte Carlo results are based on 100 time series per data-generating process. The data is generated from

$$e_t = P_t + I_t + M_t$$

with

$$\begin{aligned} P_t &= \rho P_{t-1} + \varepsilon_t \\ I_t &= \epsilon_t \\ M_t &= \xi_t - \xi_{t-1} \end{aligned}$$

for the parameter values in Table 8.

Table 8: Data generating processes for artificial data

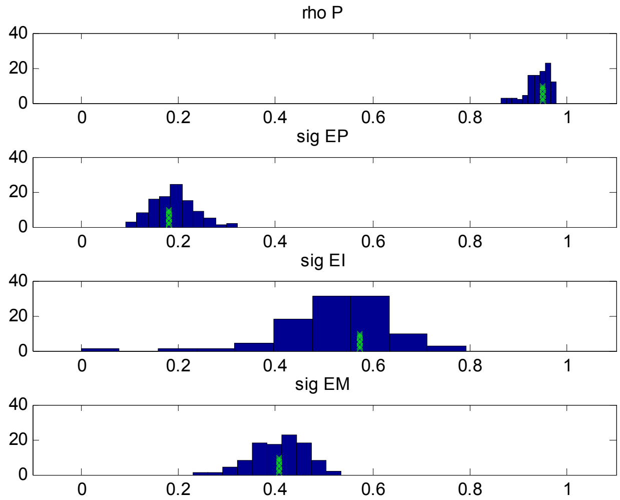
	IID	AR low	AR high	PIM low	PIM high
ρ	0	0.5	0.95	0.5	0.95
σ_P^2	1	1	1	.33	.33
σ_I^2	0	0	0	.33	.33
σ_M^2	0	0	0	.33	.33

Note: To facilitate evaluation of the relative importance of the various components, the table specifies volatility of the components rather than the innovations. Thus, $\sigma_P^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$, $\sigma_I^2 = \sigma_\epsilon^2$, $\sigma_M^2 = 2\sigma_\xi^2$ and the three shocks are orthogonal and follow $(\varepsilon_t, \epsilon_t, \xi_t)' \sim N(0_{3\times 1}, D)$.

Consider the last column of Table 8, PIM high. Here all three components are equally important, and the persistent component is very persistent. Figure 11 shows how, even for data with a limited time dimension, the estimator has no problem disentangling the various components.

It is plausible that high persistence makes identification easier. Therefore, now consider a PIM process with intermediate persistence, PIM low in Table 8. In this case, as apparent from Figure 12, there is more dispersion in point estimates. Persistence tends to be slightly underestimated (and, accordingly, the volatility of the persistent shock slightly overestimated). The M component is still consistently identified, while the I component is not always easily detected.

Figure 11: Estimation on simulated data: PIM on PIM high

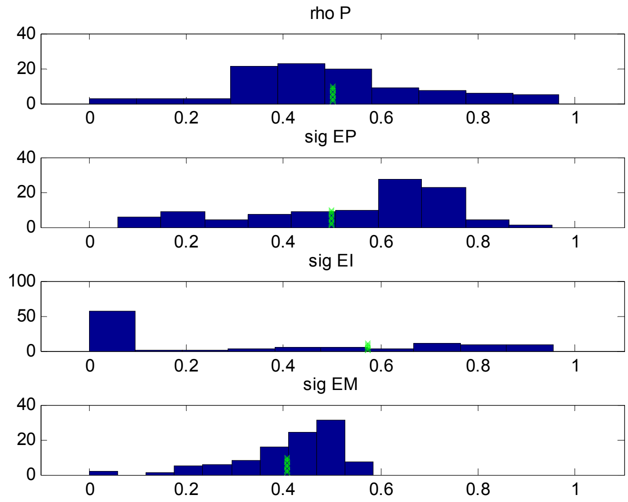


Note: Green x's mark data-generating parameters

Now consider the alternative; estimating an AR specification on these data. Irrespective of the persistence of the underlying process, estimating an AR fails to detect any significant amount of persistence, as illustrated in Figure 13 and Figure 14. We interpret these simulations as follows. While for low-persistence multicomponent processes, PIM-specifications may imply substantial imprecision regarding the variances of the components, they allow a fairly adequate evaluation of persistence. When persistence is high, they are both unbiased and precise across repeated samples, for the empirically relevant sample lengths. For the same DGP's, AR-specifications are clearly inadequate. These simulations establish one type of risk: if the DGP is a multicomponent process, AR estimation will fail to detect persistence.

The question remains as to how PIM-specifications perform in the case of AR-DGPs. It is possible that the cure is worse than the disease. Figure 15 shows that this type of risk is limited. In particular, for an AR-DGP with high underlying persistence estimating a PIM-specification comes at little cost. As persistence decreases, see Figure 16, PIM-estimation attributes some variation to the I component, which entails a minor overestimation of persistence. Taken to the limit, estimating PIM-specifications on *iid* data, as in Figure 17, identification of separate components is cumbersome: there is a lot of dispersion in all the estimates. Firstly, however, note that the modes of the distributions are typically located at the truth. Secondly, for persistence close to zero, the likelihood is flat in certain dimensions. This occurs as P and I become equivalent. This is further discussed in footnote 6 in the paper.

Figure 12: Estimation on simulated data. PIM on PIM low



Appendix B: Isolating measurement error using quantities

The observation equation for sector i becomes

$$\pi_{it} = \lambda_i^{\pi\prime} C_t + P_{it} + I_{it} + M_{it} + \eta_{it} \tag{9}$$

$$q_{it} = \lambda_i^{q\prime} C_t + \alpha_i^P P_{it} + \alpha_i^I I_{it} + \alpha_i^M M_{it} + \varsigma_{it}. \tag{10}$$

or

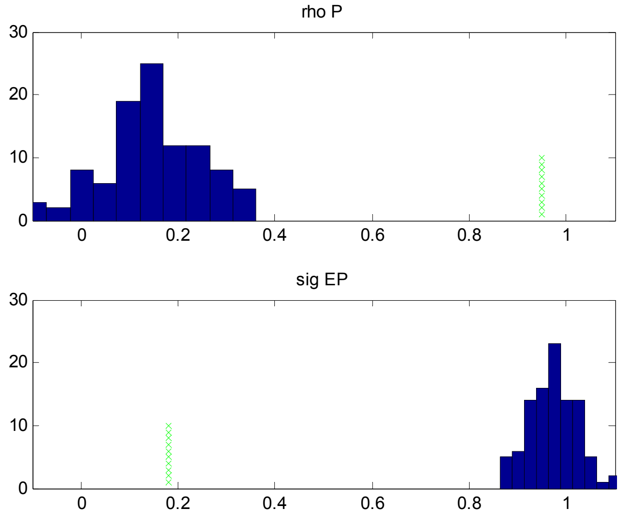
$$\begin{bmatrix} \pi_{it} \\ q_{it} \end{bmatrix} = \begin{bmatrix} \lambda_i^{\pi\prime} \\ \lambda_i^{q\prime} \end{bmatrix} C_t + \begin{bmatrix} 1 & 1 & 1 \\ \alpha_i^P & \alpha_i^I & \alpha_i^M \end{bmatrix} \begin{bmatrix} P_{it} \\ I_{it} \\ M_{it} \end{bmatrix} + \begin{bmatrix} \eta_{it} \\ \varsigma_{it} \end{bmatrix}$$

Here q denotes quantity growth. In addition to the requirement that the three components P, I and M affect quantities, their persistence properties continue to hold, as in eqs. (6)-(8). Measurement error in inflation and quantity growth are denoted by η_{it} and ς_{it} respectively. They are identified because they affect price or quantity respectively, but not both.

In the PCE data used by Boivin *et al.* (2009) real quantities are available, as part of X_t. However, real quantities are not measured independently, but calculated as nominal quantity deflated by the price index. To ensure that measurement error does not affect the quantity variable we therefore use nominal quantities.

In eq. (9), as before, the I and M components absorb substitutions and sales, respectively. The importance of measurement error is now captured separately by the sector-specific com-

Figure 13: Estimation on simulated data. AR on PIM high



ponent η_{it}. Note that substitutions related to sampling (a product not being available at the surveyed retailer) will not be captured by the I component in this setting, but instead by the measurement error component for inflation, η_{it}.

We allow both the idiosyncratic inflation and quantity components η_{it} and ς_{it} to exhibit unrestricted autoregressive dynamics. The reason for this flexible specification is that, for the inflation equation, for instance, measurement error in prices would generate negative autocorrelation.

Note that the identification assumption that the P, I and M components affect quantities does not hold at $\alpha_i^{\cdot} = 0$. This case does not turn out to be practically important. We have also estimated Bayesian versions where the sector-specific loadings are identified through the prior, with very similar results.

Table 7 in the main text summarizes the results of estimating (9)-(10), subject to (6)-(8). The following figures show the results for the relative variance (Figure 18) and persistence (Figure 19). They are very similar to the results of the benchmark factor model presented in the main text.

Figure 14: Estimation on simulated data. AR on PIM low

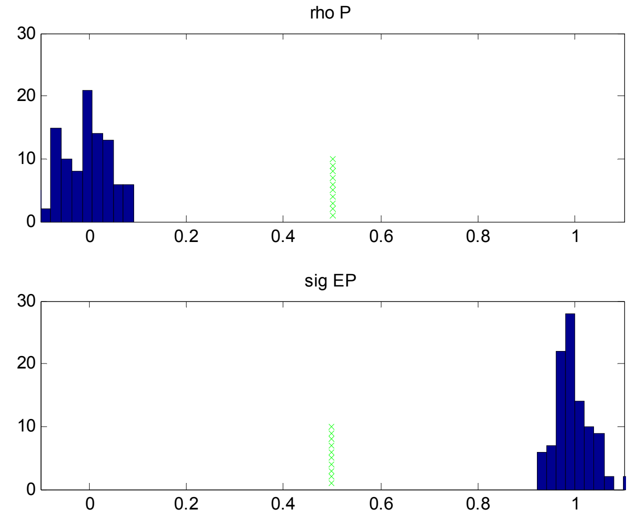


Figure 15: Estimation on simulated data. PIM on AR high

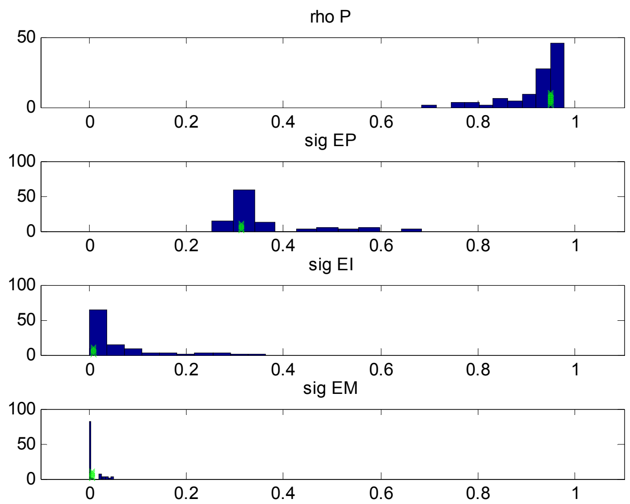


Figure 16: Estimation on simulated data. PIM on AR low

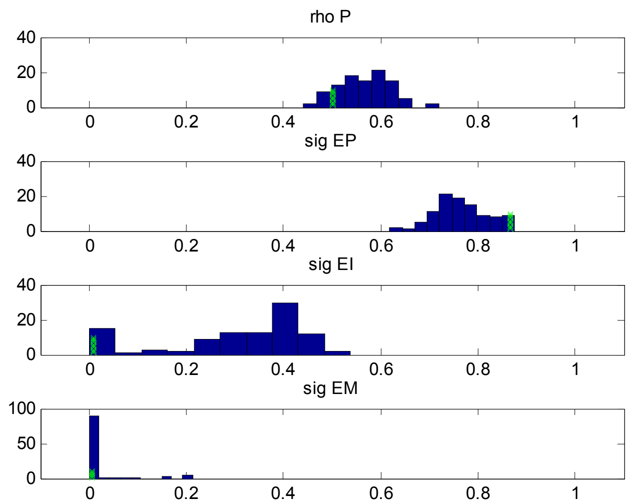


Figure 17: Estimation on simulated data. PIM on iid

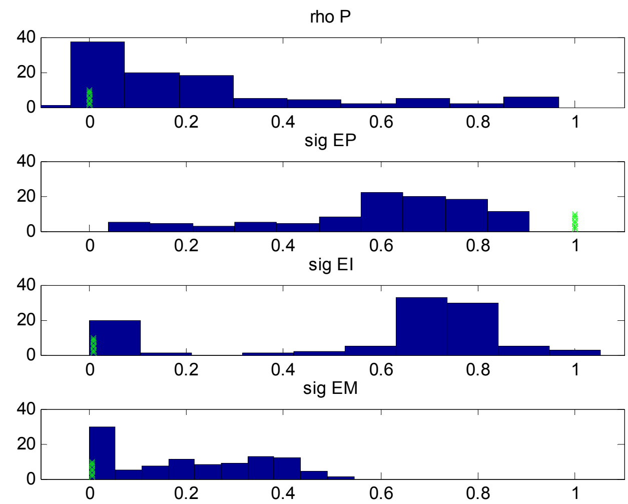


Figure 18: Identification using quantities - variance

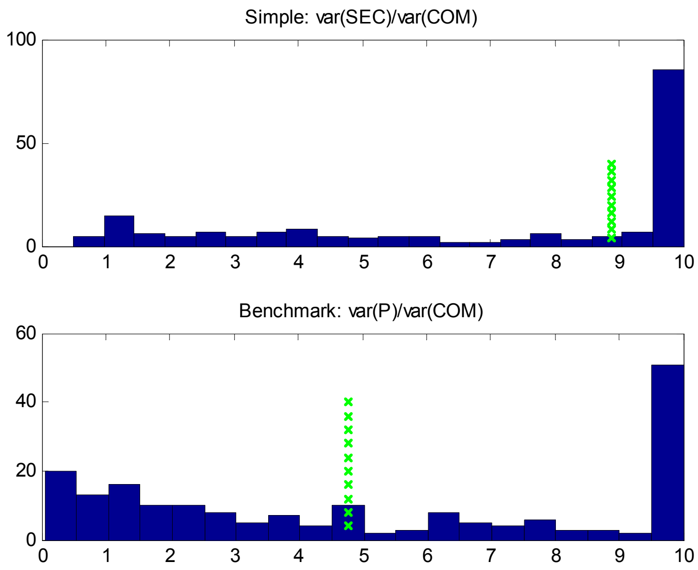


Figure 19: Identification using quantities - persistence