Investment under Rational Inattention: Evidence from US Sectoral Data

Peter Zorn*

This Version: February 13, 2018.

Abstract

I document the effects of macroeconomic and sector-specific shocks on investment in disaggregate sectoral capital expenditure data. The response of sectoral investment to macroeconomic shocks is hump-shaped, just as in aggregate data, ruling out composition effects. By contrast, the effects of sector-specific innovations are monotonically decreasing. I build and calibrate a model of investment with convex capital adjustment costs and rational inattention to explain these features of the data. The model matches the empirical responses of sectoral investment to both shocks. The interaction of information frictions and physical adjustment costs is key to this result.

JEL Classification: E22, E32, D83, D92, C38.

Keywords: investment dynamics, dynamic factor model, rational inattention.

---

*University of Munich. Email: zorn@econ.lmu.de. I am very grateful to my advisors Rüdiger Bachmann and Mirko Wiederholt for their invaluable guidance and support. I would also like to thank Bartosz Maćkowiak and Matthias Doepke for helpful conversations, Jochen Güntner for comments on an earlier version, and seminar and conference audiences at Carlos III, the Chicago Fed, Cologne, the Computing in Finance and Economics Conference in Bordeaux, the Crash Course on Rational Inattention at CERGE-EI, Frankfurt, Ghent, HEC Paris, Northwestern, Notre Dame, Oxford, RWTH Aachen University, and the Verein für Socialpolitik in Augsburg for their comments. I gratefully acknowledge the hospitality of Northwestern University and the Federal Reserve Bank of Chicago during part of this research, and generous financial support from the Society for Computational Economics best student paper prize.

1 Introduction

The hump-shaped response of aggregate investment to macroeconomic shocks is a salient feature of the business cycle in the United States.[1] This paper establishes novel stylized facts that help to shed light on the propagation mechanism underlying this empirical regularity. I show that the response of investment to macroeconomic shocks in disaggregate sectoral data—and, hence, before aggregation—is hump-shaped, just like in aggregate data. In response to an aggregate shock that leads to a 1 percent increase on impact, sectoral investment spending in the median sector rises further to 1.2 percent at the 1-year horizon. At the 2-year horizon, sectoral investment then settles approximately at the long-run response. By contrast, the effects of sector-specific surprises on sectoral investment spending are monotonically decreasing.[2] In response to a sector-specific shock that leads to a 1 percent increase on impact, sectoral investment spending in the median sector falls to 0.7 percent at the 1-year horizon, which equals approximately the long-run response. Moreover, I find that sector-specific shocks account for 90 percent, aggregate shocks for 10 percent of sectoral investment volatility.

The second part of this paper seeks to understand the discrepancy in the empirical responses of sectoral investment to differential shocks. To this end, I build and calibrate a model of investment with convex capital adjustment costs and rational inattention following Sims (2003). My main qualitative result is that the model response of sectoral investment to aggregate shocks is hump-shaped, while the effects of sector-specific shocks are monotonically decreasing. The model matches this feature of the data because decision-makers in production units choose to obtain less than perfect information with costly information acquisition. The amount of information acquired about aggregate and sector-specific shocks is roughly the same. Given less than perfect information, the re-

[1] See, for example, Christiano et al. (2005) and Altig et al. (2011) for monetary policy shocks, Romer and Romer (2010) and Mertens and Ravn (2013) for tax policy shocks, Dedola and Neri (2007) for technology shocks, and Altig et al. (2011) for investment-specific technology shocks.

[2] A monotonically decreasing response peaks on impact and then decreases monotonically.

sponse of sectoral investment to both shocks is dampened in the impact period of the shock. At the 1-year horizon, more information becomes available and decision-makers learn that their optimal level of capital is actually higher and thus increase their investment spending. In the calibration that draws on empirical estimates from the literature, aggregate shocks are more persistent than sector-specific shocks. Hence, the optimal capital stock decays more slowly and investment under rational inattention increases more strongly at the 1-year horizon following these disturbances, hence the hump shape in the response of sectoral investment. On the other hand, by the time decision-makers in firms acquire more information about sector-specific shocks, the optimal level of capital has decreased more and there is less investment demand due to learning about the shock.

Without convex capital adjustment costs, the response to aggregate shocks becomes monotonically decreasing because decision-makers choose to adjust the level of capital immediately, given the information they acquire. At the 1-year horizon, as more information about the optimal capital stock becomes available, investment demand is positive but smaller than on impact. With capital adjustment costs, decision-makers smooth capital expenditures over time which leads to additional investment demand at the 1-year horizon. Thus, convex capital adjustment costs and rational inattention are essential for the model to explain the novel stylized facts documented in the first part of the paper.

Moreover, the form of the investment response to macroeconomic shocks is preserved under aggregation across all production units in the model. Hence, my results provide a new microfounded explanation for the hump-shaped response of aggregate investment to macroeconomic shocks and highlight rational inattention as a new propagation mechanism in the investment literature.

To establish my empirical results, I estimate a dynamic factor model using capital expenditure data from US manufacturing industries. The data set contains information about real investment spending for 462 industries at the 6-digit NAICS-level for the years from 1958 to 2009. The dynamic factor model represents sectoral investment as the sum of a common component, consisting of a common factor with industry-specific factor loading, and a sector-specific component. The common factor and the sector-specific component follow an autoregressive process each with reduced-form error terms that reflect a variety of macroeconomic and industry-specific shocks. Because the innovations to the sector-specific component are independent across industries, aggregate shocks lead to common dynamics in sectoral investment across all 462 industries while sector-specific shocks do not. I use Bayesian methods to estimate the model. Based on the joint posterior density, I study the effects of aggregate and sector-specific shocks and compute the variance shares of each shock in sectoral investment volatility.

The theoretical model has the following features. There is a representative production unit in each sector. Production units operate a production function that transforms capital services into output. Total factor productivity (TFP) consists of an aggregate and a sector-specific component, which are both affected by shocks. Decision-makers in production units maximize the expected discounted value of profits by choosing capital and, thus, investment spending, subject to convex capital adjustment costs. They must pay attention to learn about the realizations of TFP shocks. Paying attention reduces uncertainty about shock realizations, where uncertainty is measured by entropy following Sims (2003). Paying attention to aggregate and sector-specific shocks are independent activities.[3] Attention is costly and decision-makers optimally allocate their attention. I calibrate the model parameters using standard values from the literature.

In principle, other propagation mechanisms can also be consistent with the empirical findings presented in this paper. Following Christiano et al. (2005), many business cycle models feature investment adjustment costs so as to match the hump-shaped impulse response of aggregate investment to macroeconomic shocks.[4] In Appendix A, I solve an otherwise standard real business cycle model with investment adjustment costs, perfect

[3] Maćkowiak and Wiederholt (2009) also make this assumption.

[4] Following Christiano et al. (2005), investment adjustment costs are convex in the growth rate of investment while capital adjustment costs are convex in the growth rate of capital.

information, and aggregate and sector-specific TFP shocks. I calibrate the model parameters at the quarterly frequency using standard values from the existing literature and time-aggregate the model responses to the yearly frequency. This calibration strategy helps to rule out the case in which the response of sectoral investment following sector-specific innovations is hump-shaped at the quarterly frequency, but time aggregation to the yearly frequency obtains a monotonically decreasing response as observed in the data. My results show that in partial as well as in general equilibrium, the impulse responses of sectoral investment to aggregate and sector-specific shocks are hump-shaped at either frequency. Hence, under standard assumptions and using a standard calibration of the model parameters, a model with investment adjustment costs has difficulties to match my empirical findings.

Fiori (2012) explores another propagation mechanism that is consistent with the hump-shaped response of aggregate investment. He shows that if rapid output expansion in the investment good producing sector is costly, the relative price of investment increases in response to aggregate shocks. This general equilibrium price response initially depresses demand for investment goods in all other sectors of the economy. As the supply of investment goods increases over time, the relative price of investment falls and investment demand in the rest of the economy picks up. The impulse responses of sectoral investment to aggregate shocks are protracted in each sector, as in the data, but not hump-shaped in general. Only the consumption good producing sector displays a slowly building sectoral investment response. More importantly, in Appendix B, I provide evidence that the relative price of investment in the manufacturing sector does not move with the macroeconomic shock estimated in the statistical model of this paper.

There are two empirical studies in the price setting literature to which this paper closely relates. Boivin et al. (2009) and Maćkowiak et al. (2009) examine the effects of macroeconomic and sector-specific shocks on sectoral price indices. This paper estimates the same impulse responses in the case of sectoral investment spending. While the statistical model and estimation methodology are similar to their work, there are differences that I will describe in more detail below. Interestingly, my empirical findings bear strong resemblance to those of Boivin et al. (2009) and Maćkowiak et al. (2009). Both studies find that aggregate shocks lead to gradual changes in sectoral price indices, whereas adjustment to sector-specific shocks is immediate. Also, they report that the bulk of sectoral inflation volatility is due to sector-specific shocks.

This article also adds to the literature on rational inattention following Sims (2003, 2006). To the best of my knowledge, this paper is the first to study the implications of investment under rational inattention.[5] Other applications include price setting decisions of firms (Woodford, 2009; Maćkowiak and Wiederholt, 2009; Matějka, forthcoming); the consumption-saving decision of households (Luo, 2008; Tutino, 2013); discrete choice behavior (Matějka and McKay, 2015); monetary policy (Paciello, 2012; Paciello and Wiederholt, 2014); and portfolio choice (Mondria, 2010; Van Nieuwerburgh and Veldkamp, 2009, 2010; Kacperczyk et al., 2016). Maćkowiak and Wiederholt (2015) formulate a dynamic stochastic general equilibrium model with rational inattention. However, their model abstracts from capital in production.

The remainder of this paper is organized as follows. Section 2 presents the statistical model for the sectoral data. Section 3 describes the data. Section 4 contains the main empirical results and several robustness checks. In Section 5, I lay out the model of investment with convex capital adjustment costs and rational inattention. Section 6 evaluates the model and contains the quantitative results. Section 7 concludes.

[5]In related work, Verona (2014) explores the implications of capital adjustment in a model with sticky information. Under this assumption, decision-makers must pay a fixed cost to acquire new information and, once they do so, have perfect information in the period of updating.

2 Statistical Model for Sectoral Capital Expenditure Data

I use the following dynamic factor model to study sectoral capital expenditure data:

$$y_{it} = H_i x_t + w_{it}, \tag{1}$$

where y_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$, denotes the period t log change of real investment in sector i, x_t is a single unobserved common factor, and the w_{it} are sector-specific error terms. The H_i are factor loadings that are possibly different across industries. In Equation (1), I omit a constant for ease of exposition and because I standardize the data in the next section.

The factor and the sector-specific terms each follow autoregressive (AR) processes:

$$x_t = F(\ell)x_{t-1} + v_t, \qquad v_t \sim i.i.d.\,\mathcal{N}(0, Q) \tag{2}$$

$$w_{it} = D_i(\ell)w_{it-1} + u_{it}, \qquad u_{it} \sim i.i.d.\,\mathcal{N}(0, R_i) \tag{3}$$

where $F(\ell)$ and $D_i(\ell)$ denote lag polynomials of order three, and v_t and the u_{it} are Gaussian white noise with variance Q and R_i, respectively. The u_{it} are pairwise independent and uncorrelated with v_t. Moreover, the u_{it} and v_t are uncorrelated with initial conditions, the w_{i0} and x_0. These assumptions imply that the w_{it} are pairwise independent and uncorrelated with x_t.

A few remarks are in order. First, it is worth pointing out that I do not attempt to identify structural innovations. Surprise movements in v_t and in the w_{it} are reduced-form and reflect a convolution of structural innovations. Second, given x_t, Equation (1) is a normal linear regression with serially correlated error term. Because the w_{it} are pairwise independent and uncorrelated with x_t, all comovement in sectoral investment comes from the factor x_t. It follows that, given x_t, Equation (1) can be estimated equation-by-equation for each sector. Note that sector-specific components are allowed to have different persistence and innovation variances across industries. Third, the dynamic response of sectoral investment to innovations in the factor, v_t, can be read off the coefficients of the infinite-order lag polynomial $H_i(1 - F(\ell)L)^{-1}$, where L denotes the lag operator. Hence, the statistical model imposes that the impulse responses of investment to aggregate shocks are proportional across industries.[6] It bears pointing out that the shape of the impulse responses itself is not pinned down by the model, but will be determined by the data. Furthermore, the model does not restrict the impulse responses of sectoral investment to sector-specific innovations to be proportional.

This paper uses Bayesian methods to estimate the model. In particular, I use Gibbs sampling with a Metropolis-Hastings step to sample from the joint posterior density of the factor and the model's parameters. Given a draw of the model's parameters, I sample from the conditional posterior density of the factor, x_t, using the Carter and Kohn (1994) simulation smoother. Given a draw of the factor, I sample from the conditional posterior densities of the parameters. Equation (2) is an AR process that can be estimated using a variant of Chib and Greenberg (1994). Equation (1) is a normal linear regression model with AR errors, which can be estimated using the method by Chib and Greenberg (1994).

The priors for the lag polynomials $F(\ell)$ and $D_i(\ell)$ are centered around zero at each lag. Like the Minnesota prior, the prior precision at more distant lags is higher. The factor loadings H_i also have zero prior mean and unit variance. For the sector-specific innovations R_i, I use the diffuse prior by Otrok and Whiteman (1998). More details on the estimation methodology and priors are available in Appendix C.

[6] Maćkowiak et al. (2009) point out this insight. In the spirit of Jordà (2005), their dynamic factor model estimates impulse responses at each horizon of interest without the restriction of proportionality. Like Ramey (2013), I found that this approach can lead to erratic impulse responses of sectoral investment that contradict economic intuition. For this reason, I use the specification in which impulse responses of sectoral investment to aggregate shocks are proportional.

3 Data

The disaggregate sectoral capital expenditure data comes from the NBER-CES Manufacturing Industry Database. This data set contains nominal investment spending and investment price deflators at the industry level for a representative sample of the US manufacturing sector. The sample starts in 1958 and the frequency of the data is annual. The level of aggregation is the 6-digit NAICS-level.[7] The data set contains a balanced panel of 462 sectors.[8] The median number of establishments per sector in the population is 342.[9] The data set ends in 2009.[10]

I compute sectoral real investment by dividing nominal capital expenditures in each year and sector by the corresponding investment price deflator. I convert each series into growth rates by taking log differences. Furthermore, I standardize each growth rate series to have zero mean and unit variance. The standardization helps to abstract from differences in the coefficients of the statistical model due to differences in sectoral volatility. This facilitates estimation and makes impulses responses easier to compare across sectors.

In terms of sectoral comovement, the first principal component of the standardized sectoral real investment growth rates explains roughly 14.5 percent of their total variance. The next four principal components add 5.46 percent, 4.15 percent, 3.82 percent, and 3.62 percent each to the total variance explained. The drop and leveling off in incremental explanatory power after the first principal component informally suggests the presence of one factor, which is why I assume a single factor in the statistical model described in the previous section. Also, the low portion of variation explained by the first princi-

[7] As an example, "Cookie and Cracker Manufacturing" is a 6-digit NAICS industry.

[8] In 1997, eleven industries were reclassified into manufacturing but capital expenditure data prior to 1997 is not available for them. Therefore, I do not consider them in the analysis.

[9] I obtain this number from the County Business Patterns as the median value for the years from 1998 to 2001. The industry classification used in the Country Business Patterns is different from the industry classification used in the NBER-CES Manufacturing Industry Database in other years.

[10] See Bartelsman and Gray (1996) and Becker et al. (2016) for a detailed description of the NBER-CES Manufacturing Industry Database.

pal component already suggests that investment dynamics at the sector-level are mostly driven by sector-specific shocks.

Aggregating over all sectors, the sample covers on average about 55 percent of US manufacturing nonresidential, private fixed investment spending. In real terms, the linear correlation between total investment expenditures in the sample and US manufacturing nonresidential, private fixed investment spending is 0.97.[11] These statistics suggest that the data is representative of the US manufacturing sector.

4 Empirical Results

The first part of this section presents the three main empirical findings of this paper: (i) the impulse response of sectoral investment to aggregate shocks is hump-shaped, (ii) the effects of sector-specific shocks on sectoral investment are not hump-shaped and decrease monotonically, and (iii) sector-specific shocks account for the bulk of sectoral investment volatility.

The second part assesses the robustness of my empirical findings by exploring whether (i) there are multiple common factors, (ii) the results change at the 4-digit and 3-digit NAICS industry-level, and (iii) the results are prone to the missing persistence bias pointed out by Berger et al. (2015). I find that the results are robust along these dimensions.

Before I present my main empirical findings, let me give two additional results. First, Figure 1 displays impulse responses of aggregate investment to a 1 percent innovation over a 5-year horizon. I estimate the following AR(3) process to obtain these impulse responses:

$$y_t = c + \sum_{j=1}^{3} \phi_j y_{t-j} + w_t, \tag{4}$$

[11] US manufacturing nonresidential, private fixed investment spending in nominal and real terms is available from the Bureau of Economic Analysis (BEA) Fixed Asset Accounts, Tables 4.7 and 4.8, respectively.

where y_t denotes the log change of aggregate investment in real terms and w_t is Gaussian white noise. The impulse response of the log-level of aggregate investment corresponds to the cumulative impulse response of y_t. Again, it is worth pointing out that this is a reduced-form impulse response and does not reflect the effects of a structural macroeconomic shock. I estimate Equation (4) using three different time series.[12] The blue line in Figure 1 shows the effects on US nonresidential, private fixed investment. In response to a 1 percent innovation, aggregate investment rises further to 1.6 percent at the 1-year horizon, giving rise to a hump-shape. The green line in Figure 1 is based on aggregate manufacturing investment data, while the red line is based on the aggregated micro data. The effects of an innovation on aggregate manufacturing investment are in both cases slightly less pronounced and more short-lived, but the hump shape is nevertheless preserved. Notice that the error bands do not contain 0.01 at the 1-year horizon.[13]

Second, in Figure 2, the solid blue line depicts the pointwise posterior median estimate of the common factor. The dashed black line depicts the growth rate of value added in the US manufacturing sector for comparison.[14] The gray-shaded regions correspond to NBER recessions. The figure suggests that the common factor is pro-cyclical. Indeed, the correlation with US manufacturing value added growth is 0.55. Moreover, the correlation between the factor and US manufacturing investment growth is 0.87.

In sum, these results show why the estimated statistical model for disaggregate sectoral capital expenditure data from manufacturing industries is useful. The impulse responses in the manufacturing sector are very similar to that of the total economy. Moreover, the statistical model provides a plausible estimate of common investment dynamics. We can now ask what are the effects of macroeconomic and sector-specific shocks on sectoral investment.

[12] See Footnote 11 for data sources of manufacturing and total economy data used in the following.

[13] These are 68 percent error bands obtained by direct Monte Carlo sampling from the posterior distribution of the AR parameters. I take 1,000 draws and use Jeffrey's noninformative prior in estimation.

[14] The data source for the US manufacturing value added series is the BEA Industry Economic Accounts.

4.1 Main Results

The first empirical main result is that the impulse response of sectoral investment to aggregate shocks is protracted and hump-shaped. To obtain this result, I first sample randomly 1,000 parameter draws from the joint posterior density. Second, for each sector and every draw, I compute the cumulative impulse response of investment growth in response to an aggregate shock that leads to a 1 percent increase on impact. The cumulative impulse response corresponds to the impulse response of the log-level of sectoral investment. Third, I define the median sector as the pointwise 50th percentile of the distribution of impulse responses obtained in the previous step. Recall that the impulse responses of investment to aggregate shocks are proportional across industries. Given a parameter draw, the pointwise cross-sectional median of impulse responses therefore corresponds to the same industry at all horizons. Moreover, the impulse responses are scaled to imply an increase of investment by 1 percent on impact in each sector. It follows that the impulse responses of investment to aggregate shocks are the same in all sectors for a given parameter draw. The form of impulse responses across draws varies, however. The median sector measures the central tendency of impulse responses at each horizon. Fourth, I also compute the pointwise 16th and 84th percentiles of the distribution of impulse responses obtained in the second step. I use these statistics to characterize posterior uncertainty about the impulse responses. From the above, it follows that posterior uncertainty reflects posterior parameter uncertainty only. Figure 3 shows the result of this procedure. In response to an aggregate shock that leads to a 1 percent increase on impact, sectoral investment spending in the median sector rises further to 1.2 percent at the 1-year horizon, giving rise to a hump-shape. Note that the posterior density at the 1-year horizon lies above 0.01. At the 2-year horizon, sectoral investment then settles approximately at the long-run response.

To shed light on posterior uncertainty from a different angle, I compute the percentage share of investment responses to aggregate shocks that have a hump-shaped form. I consider all investment responses obtained in the second step of the above procedure. About 83 percent of the investment responses peak between horizons 0 and 5. If, in addition, the requirement that the response is monotonically increasing to the left of the peak is imposed, approximately 76 percent of the impulse responses have a hump-shaped form.

The second empirical main result is that the effects of sector-specific shocks on sectoral investment are not hump-shaped but monotonically decreasing. I use the same procedure as above to conduct posterior inference on the impulse response to a sector-specific shock that leads to a 1 percent increase in sectoral investment. However, the median sector now measures the central tendency of impulses responses at each horizon both across sectors and draws. Similarly, the posterior uncertainty now reflects both posterior parameter uncertainty and cross-sectional variation. The reason for this difference with respect to the impulse responses to aggregate shocks is that the statistical model does not restrict the impulse responses of sectoral investment to sector-specific shocks to be proportional. Figure 4 depicts the result. In response to a sector-specific shock that leads to a 1 percent increase on impact, sectoral investment spending in the median sector falls to 0.7 percent at the 1-year horizon, which equals approximately the long-run response. In comparison to the impulse response to aggregate shocks, the effects of sector-specific shocks on sectoral investment are short-lived and monotonically decreasing.

In the case of sector-specific shocks, only about 14 percent of the investment responses drawn peak between horizons 0 and 5. This percentage share reduces further to 8 percent if only those responses that are monotonically increasing to the left of the peak are considered.

The third empirical main result is that sector-specific shocks explain the bulk of sectoral investment volatility. To obtain this result, recall that the assumptions of the econometric framework imply that the factor, x_t, and the sector-specific term, w_{it}, are uncorrelated. Hence, the variance of the sectoral investment growth rate, y_{it}, can be written as $\text{Var}[y_{it}] = H_i^2 \text{Var}[x_t] + \text{Var}[w_{it}]$. The first term captures the contribution of aggregate shocks, the second term the contribution of sector-specific shocks to sectoral investment volatility. First, I use the posterior median estimate of $F(\ell)$ to compute the unconditional variance of the process for x_t, $\text{Var}[x_t]$. Second, I compute the unconditional variance of the process for w_{it}, $\text{Var}[w_{it}]$, using the posterior median estimates of $D_i(\ell)$ and R_i for each sector. Third, I compute the variance shares of aggregate and sector-specific shocks in sectoral investment volatility for each sector. Fourth, I define the median industry as the 50th percentile of the cross-sectional distribution of variance shares. I find that sector-specific shocks account for about 90 percent, aggregate shocks for about 10 percent of sectoral investment volatility in the median sector.

4.2 Robustness

4.2.1 Number of Factors

The statistical model in Equation (1) assumes a single common factor. To test for the presence of additional common factors, I study the cross-sectional correlation of the sector-specific terms, w_{it}. Recall that the factors account for all the comovement in the observable data, whereas the sector-specific terms are assumed to be uncorrelated in the cross-section. If there are additional factors omitted from Equation (1), the comovement stemming from them has to be captured by the sector-specific terms. Therefore, I take a random draw from the posterior distribution of the factor, x_t, and the factor loading, H_i, to compute the w_{it}. Next, I compute the median of the absolute value of the cross-sectional correlation, $|\text{corr}[w_i, w_j]|, \forall i \neq j$. I repeat this procedure 1,000 times. Figure 5 displays the histogram of this statistic. The median of this distribution is low and equals 0.1091, which means that there is little cross-sectional correlation in the sectoral components. This exercise suggests that there are no additional factors relevant to explain the cross-sectional comovement in the sectoral investment.

4.2.2 Level of Aggregation

I re-estimate the model at the 4-digit and 3-digit NAICS industry level to test if the results depend on the level of aggregation.[15] Figure 6 contrasts the posterior median estimate of the common factor at different levels of aggregation. The solid blue line depicts the estimate based on 6-digit NAICS industry data shown in Figure 2. The red dash-dot line and the green dashed line show the estimates obtained from using 4-digit and 3-digit NAICS industry data, respectively. Figure 6 shows that the median estimates of the factor have virtually the same dynamics at different levels of aggregation. At higher levels of aggregation, the factor captures more comovement in sectoral investment, which is why the volatility of the estimates increases. Figures 7 and 8 show that the impulse responses to shocks also do not change with the level of aggregation. Figure 7 contrasts the impulse responses of sectoral investment to aggregate shocks at the 6-digit, the 4-digit, and the 3-digit NAICS industry level. The line styles and colors are the same as in Figure 6. The figures shows that the impulse responses to aggregate shocks are qualitatively and, to a large extent, quantitatively the same and do not depend on the level of aggregation. Similarly, Figure 8 depicts the effects of sector-specific shocks on sectoral investment at different levels of aggregation. The line styles and colors are again the same as above. In all three cases, the effects of sector-specific shocks are monotonically decreasing. As the sectors become more aggregate, the impulse responses become more gradual.

4.2.3 Missing Persistence Bias

Berger et al. (2015) prove that the estimated persistence of aggregate time series with lumpy behavior at the micro level is biased towards zero at low levels of aggregation. The reason for the bias is an identification problem: the econometrician cannot disentangle the adjustment in response to contemporaneous shocks from the adjustment to past shocks, and attributes all adjustment to the contemporaneous innovation. At higher

[15] I follow the approach by the BEA to aggregate chain-type quantity indices and aggregate the real investment quantity indices to the 4-digit and 3-digit NAICS industry level. There are 86 industries at the 4-digit and 21 industries at the 3-digit NAICS industry level in the US manufacturing sector.

levels of aggregation, the cross-sectional correlation of capital adjustments across sectors informs the econometrician and the bias vanishes. Indeed, Figure 8 suggests that the persistence of impulse responses of sectoral investment to sector-specific shocks increases with the level of aggregation. To account for this bias, Berger et al. (2015) propose to use proxy variables for the shocks.

To verify the robustness of my results, I follow Berger et al. (2015) and use proxy variables for the shocks to re-estimate impulse responses. More specifically, I calculate growth rates of Solow residuals for each sector from the NBER-CES data using a Cobb-Douglas production function for real value added in employment and real capital. Since the data set does not contain a deflator for value added, I use the GDP deflator. The employment share equals the average percentage share of payroll in value added in the ongoing and in the previous year. The capital share equals the residual factor share. Next, I decompose the sectoral Solow residual growth rates into common and sectoral components, denoted TFP_t^{Agg} and TFP_t^{Sect}, using principal components. Using these variables as proxies for aggregate and sector-specific shocks, I run a regression of the sectoral investment growth rate on the contemporaneous and lagged values of TFP_t^{Agg} and TFP_t^{Sect}:

$$y_{it} = \sum_{j=0}^{5} \alpha_{ij} TFP_{t-j}^{Agg} + \sum_{j=0}^{5} \beta_{ij} TFP_{t-j}^{Sect} + \varepsilon_{it}. \tag{5}$$

Using TFP_t^{Agg} and TFP_t^{Sect} as proxy variables for each shock, the impulse responses of sectoral investment to aggregate and sector-specific shocks after h years are just the sum of the coefficients on the contemporaneous value and the first h lags of aggregate and sector-specific TFP: $\sum_{j=0}^{h} \alpha_{ij}$ and $\sum_{j=0}^{h} \beta_{ij}$. To test if sectoral investment responds faster to sector-specific shocks than to aggregate shocks, I follow Maćkowiak et al. (2009) and measure the speed of adjustment for each sector i by the following statistic:

$$\tau_i^{Agg} = \frac{\sum_{h=0}^{1} |\sum_{j=0}^{h} \alpha_{ij}|}{\sum_{h=2}^{3} |\sum_{j=0}^{h} \alpha_{ij}|} \quad \text{and} \quad \tau_i^{Sect} = \frac{\sum_{h=0}^{1} |\sum_{j=0}^{h} \beta_{ij}|}{\sum_{h=2}^{3} |\sum_{j=0}^{h} \beta_{ij}|}. \tag{6}$$

For each shock, this statistic captures the short-run response of sectoral investment spending relative to the long-run response. I define the short-run response as the average absolute effect on sectoral investment in the impact period and at the 1-year horizon. Similarly, I take the long-run response as the average absolute effect at the 2-year and at the 3-year horizon.

Figure 9 plots the histogram of the cross-sectional distribution for the speed of adjustment. The upper panel shows the speed of adjustment to aggregate shocks, the lower panel the speed of adjustment to sector-specific shocks. The median of the distribution is 0.6241 in the top panel and 0.9113 in the bottom panel. This means that adjustment of the median sector to aggregate shocks in the short run is less than two-thirds of the adjustment in the long run, while the adjustment to sector-specific shocks in the short run is about as large as the adjustment in the long run. In other words, investment adjusts relatively faster to sector-specific TFP shocks than to aggregate TFP shocks. This exercise suggests that the main results of this paper are not prone to the missing persistence bias.

An interesting observation that emerges from this exercise regards the nature of the aggregate shock. In Figure 10, I contrast the pointwise posterior median estimate of the common factor with the aggregate component of sectoral TFP growth. The two shock measures are very similar, the correlation between both series is 0.63. This is at least suggestive that the estimated aggregate shock in the statistical model can be interpreted as innovations to TFP. In the theoretical model in the next section, I will assume that TFP shocks are the driving force of investment activity.

5 Investment under Rational Inattention

In this section, I build a model of investment with convex capital adjustment costs and rational inattention. In the next section, I calibrate and solve the model to investigate if it can account for the empirical findings presented in this paper.

5.1 Setup

The economy consists of a large number of sectors, which are each populated by a representative production unit indexed by i. Time is discrete. Production unit i operates the production function

$$Y_{it} = Z_t \mathcal{E}_{it} K_{it}^{\alpha}, \tag{7}$$

where K_{it} denotes the current stock of capital, Z_t and $\mathcal{E}_{it}$ are aggregate and sectoral total factor productivity (TFP), and α is a parameter.[16]

Production units own the capital stock, which is specific to their sector. The law of motion for capital is

$$K_{it+1} = (1-\delta)K_{it} + I_{it}, \tag{8}$$

where I_{it} is investment and δ denotes the rate of depreciation. Changing the level of capital is costly because of installation costs and results in a loss of profit. Capital adjustment cost are given by $\frac{\gamma}{2}\left(\frac{I_{it}}{K_{it}}\right)^2 K_{it}$. Period profits of production unit i thus read

$$Y_{it} - I_{it} - \frac{\gamma}{2}\left(\frac{I_{it}}{K_{it}}\right)^2 K_{it}. \tag{9}$$

The sectoral and aggregate components of TFP each follow stationary Gaussian first-order autoregressive processes in logs:

$$\ln Z_t = \rho_z \ln Z_{t-1} + e_t, \tag{10}$$

$$\ln \mathcal{E}_{it} = \rho_\varepsilon \ln \mathcal{E}_{it-1} + v_{it}, \tag{11}$$

where the error terms are Gaussian white noise with distributions $e_t \sim \mathcal{N}(0, \sigma_e^2)$ and $v_{it} \sim \mathcal{N}(0, \sigma_v^2)$, respectively. The sector-specific shocks, v_{it}, are pairwise independent in the cross-section. Moreover, the v_{it} are independent of aggregate shocks, e_t.

[16]Because each sector has a representative product unit, the term "sectoral" henceforth refers to the idiosyncratic variables of the production unit in that sector.

In each production unit, a decision-maker maximizes the expected net present value of current and future profits with discount factor β. In period -1, decision-makers decide how much attention to pay. Paying attention is costly and decision-makers will not attend to all available information. Given less than perfect information, decision-makers choose investment. I begin with the derivation of the objective function given the information they do acquire. The following section describes the attention problem of decision-makers.

Substituting the production function in Equation (7) and the law of motion for capital in Equation (8) into the expression for period profit in Equation (9) yields the period profit function

$$\pi\left(K_{it}, K_{it+1}, Z_t, \mathcal{E}_{it}\right) = Z_t \mathcal{E}_{it} K_{it}^{\alpha} - K_{it+1} + (1-\delta) K_{it} - \frac{\gamma}{2}\left(\frac{K_{it+1}}{K_{it}} - (1-\delta)\right)^2 K_{it}. \quad (12)$$

Rewriting Equation (12) in log-deviations from the non-stochastic steady state, multiplying with β^t, summing over all periods from 0 to ∞, and finally taking expectations conditional on information in period -1 yields the objective function for production unit i. I work with a log-quadratic approximation around the non-stochastic steady state. That is, I compute a second-order Taylor approximation to the objective function and derive the following expression for the expected discounted sum of losses in profit when the actual capital choice given less than perfect information, k_{it+1}, deviates from the profit-maximizing capital choice under perfect information, k_{it+1}^{*}:

$$\sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\frac{1}{2} H_0 \left(k_{it+1} - k_{it+1}^{*}\right)^2 + \left(k_{it+1} - k_{it+1}^{*}\right) H_1 \left(k_{it+2} - k_{it+2}^{*}\right)\right\}, \quad (13)$$

where $H_0 = \overline{K}\left[-\gamma + \beta\left(\alpha(\alpha-1)\overline{K}^{\alpha-1} - \gamma\right)\right]$ and $H_1 = \beta\gamma\overline{K}$. Here, $\overline{K}$ denotes the value of capital in the non-stochastic steady state and lower case letters denote log-deviations from the non-stochastic steady state, for example $k_{it+1} = \ln K_{it+1} - \ln \overline{K}$.

After the log-quadratic approximation, the profit-maximizing capital choice under perfect information is given by

$$k^*_{it+1} = \frac{\gamma k^*_{it} + \beta E_t \left\{ \gamma k^*_{it+2} + \alpha \overline{K}^{\alpha-1}(z_{t+1} + \varepsilon_{it+1}) \right\}}{\gamma + \beta\gamma - \beta\alpha(\alpha-1)\overline{K}^{\alpha-1}}. \tag{14}$$

Here, E_t denotes the expectation operator conditioned on the history of the economy up to and including period t.[17] Equation (14) is the usual log-linearized optimality condition for capital in a partial equilibrium model with capital adjustment costs, which can be expressed as a linear function of current and past shocks:

$$k^*_{it+1} = \underbrace{A_1(\ell)e_t}_{k^{*z}_{it+1}} + \underbrace{A_2(\ell)v_{it}}_{k^{*\varepsilon}_{it+1}}, \tag{15}$$

where $A_1(\ell)$ and $A_2(\ell)$ are infinite-order lag polynomials.[18]

The actual capital choice by decision-makers given less than perfect information follows the stochastic process

$$k_{it+1} = \underbrace{B_1(\ell)e_t + C_1(\ell)u^e_{it}}_{k^z_{it+1}} + \underbrace{B_2(\ell)v_{it} + C_2(\ell)u^v_{it}}_{k^\varepsilon_{it+1}}, \tag{16}$$

where $B_s(\ell)$ and $C_s(\ell)$ with $s = 1, 2$ are infinite-order lag polynomials. Moreover, u^e_{it} and u^v_{it} are Gaussian white noise with unit variance, independent of e_t and v_{it}, independent of each other, and independent across production units.

Given less than perfect information, the actual capital choice by decision-makers differs from the profit-maximizing capital choice under perfect information along two dimensions. First, capital may respond with dampening and delay to aggregate and sector-

[17] Appendix E contains the derivation of Equations (13) and (14).

[18] Let λ_1 and λ_2 denote the roots of the characteristic equation of the linear difference equation appearing in Equation (14). Without loss of generality, suppose that $\lambda_1 < \lambda_2$. The coefficient corresponding to lag j in $A_1(\ell)$ equals $\frac{\rho_z^j}{\gamma} \frac{\alpha\overline{K}^{\alpha-1} - \rho_z}{\lambda_2 - \rho_z} \frac{1-(\lambda_1/\rho_z)^{j+1}}{1-\lambda_1/\rho_z}$, and the coefficient corresponding to lag j in $A_2(\ell)$ equals $\frac{\rho_\varepsilon^j}{\gamma} \frac{\alpha\overline{K}^{\alpha-1} - \rho_\varepsilon}{\lambda_2 - \rho_\varepsilon} \frac{1-(\lambda_1/\rho_\varepsilon)^{j+1}}{1-\lambda_1/\rho_\varepsilon}$.

specific shocks, i.e., $B_s(\ell) \neq A_s(\ell)$ for some s. Second, the actual capital choice may be noisy, i.e., $C_s(\ell) \neq 0$ for some s.[19] Clearly, if decision-makers know the history of the economy up to and including period t, they will choose $B_s(\ell) = A_s(\ell)$ and $C_s(\ell) = 0$ for $s = 1, 2$ and the actual capital choice coincides with that under perfect information.

5.2 Information Structure

All information is freely available in the economy. Paying attention is costly, however. It takes time and mental capacity to process information about shocks and translate it into decisions. Following Sims (2003), I assume that paying attention is modelled as uncertainty reduction, where uncertainty is measured by entropy. The amount of information that the actual capital choice, k_{it+1}, contains about the profit-maximizing capital choice under perfect information, k^*_{it+1}, cannot be greater than $\kappa \geq 0$. Formally,

$$\mathcal{I}\left(\left\{k^{*z}_{it+1}, k^{*\varepsilon}_{it+1}\right\}, \left\{k^{z}_{it+1}, k^{\varepsilon}_{it+1}\right\}\right) \leq \kappa, \tag{17}$$

where the operator $\mathcal{I}$ is defined in Appendix D.

Decision-makers choose how much attention to pay. Paying attention is costly and results in loss of profit. The per-period marginal cost of paying attention equals λ.

5.3 Attention Problem

In period -1, the decision-maker in production unit i chooses the allocation of attention and hence a stochastic process for k_{it+1} to minimize the expected discounted value of current and future profit losses:

$$\max_{\kappa, B(\ell), C(\ell)} \left\{ \sum_{t=0}^{\infty} \beta^t E_{i,-1} \left\{ \frac{1}{2} H_0 \left(k_{it+1} - k^*_{it+1}\right)^2 + \left(k_{it+1} - k^*_{it+1}\right) H_1 \left(k_{it+2} - k^*_{it+2}\right) \right\} - \frac{\lambda}{1-\beta}\kappa \right\} \tag{18}$$

[19] Maćkowiak and Wiederholt (2015) also make these assumptions.

subject to the law of motion for the profit-maximizing capital choice under perfect information

$$k^*_{it+1} = A_1(\ell)e_t + A_2(\ell)v_{it}, \tag{19}$$

the law of motion for the actual capital choice

$$k_{it+1} = B_1(\ell)e_t + C_1(\ell)u^e_{it} + B_2(\ell)v_{it} + C_2(\ell)u^v_{it}, \tag{20}$$

and the information flow constraint

$$\mathcal{I}\left(\left\{k^{*z}_{it+1}, k^{*\varepsilon}_{it+1}\right\}, \left\{k^z_{it+1}, k^\varepsilon_{it+1}\right\}\right) \leq \kappa. \tag{21}$$

Decision-makers weigh the benefit of paying more attention so that their actual capital choices follow more closely the profit-maximizing capital choices under perfect information against the cost of paying attention. Note that the decision to pay more attention to one shock does not have an effect on the information acquisition about the other shock, given a constant marginal cost of attention.

6 Model Results

This section calibrates and solves the model. I find that the model is able to explain the discrepancy in the empirical responses of sectoral investment to differential shocks.

6.1 Calibration

I calibrate the model parameters to standard values from the investment literature to evaluate the model. A period in the model corresponds to a year. The parameters for β and δ are chosen to match empirical moments reported by Khan and Thomas (2008). The discount factor β is set to imply discounting of future profits by decision makers at an annual real interest rate of 4 percent, which gives $\beta = 0.9615$. The depreciation rate is $\delta = 0.10$, which implies that the steady-state investment-to-capital-ratio equals 10 percent. Bachmann et al. (2013) estimate the value-added-weighted average persistence and value-added-weighted average standard deviation of sectoral TFP from Solow residuals measured using the same data source as this paper, which leads to the values $\rho_\varepsilon = 0.55$ and $\sigma_v = 0.0501$. Khan and Thomas (2008) estimate the persistence and volatility of aggregate TFP from Solow residuals and find $\rho_z = 0.8590$ and $\sigma_e = 0.0140$. Because the production function of production units implicitly reflect the output of a whole sector, I assume that the arguments invoked to justify decreasing returns to scale such as span-of-control do not apply. Indeed, averaging over the returns-to-scale estimates by Basu et al. (2006) for 2-digit manufacturing industries gives 0.94. However, for the steady state level of capital to be uniquely defined, some curvature in production is required. For this reason, the parameter α is set to 0.99. The capital adjustment costs parameter γ equals 0.5, a value at the lower end of estimates in the literature. Finally, the parameter λ is set to imply a per-period marginal cost of attention equal to 0.06% of steady state profits. This value corresponds to the value for the marginal cost of attention estimated by Maćkowiak and Wiederholt (2015) in the case of the price setting decisions. Given that rational inattention is a friction that sits on the level of decision-makers, the marginal costs of paying attention should be the same order of magnitude for any profit-relevant decision.

6.2 Numerical Solution

I use numerical methods to solve the firm's attention problem. Following Maćkowiak and Wiederholt (2015), I parametrize the infinite-order lag polynomials $B_s(\ell)$ and $C_s(\ell)$ with $s = 1, 2$ as lag polynomials of ARMA(2,2) and AR(1) processes, respectively. To make the problem finite-dimensional, I truncate the lag polynomials to degree 250. Similarly, I evaluate the information flow constraint in Equation (17) for 250 periods. I use the nonlinear optimization routine by Kuntsevich and Kappel (1997) to solve for the coefficients in the lag polynomials and the allocation of attention. To concentrate the numerical search on regions of the parameter space that imply invertibility of the AR parts in the lag polynomials $B_s(\ell)$ and $C_s(\ell)$ with $s = 1, 2$, I reparameterize the problem by adapting the method of Monahan (1984).

6.3 Model Investment Responses

The main quantitative result from the model with capital adjustment costs and rational inattention, depicted in Figure 11, is that the response of sectoral investment to aggregate shocks displays a hump-shaped form. By contrast, the response of sectoral investment to sector-specific shocks is monotonically decreasing.

Figure 11 shows the model responses of sectoral investment to aggregate and sector-specific shocks over a 5 year horizon in the top and bottom panel, respectively. The solid black lines in both panels show the case of investment with capital adjustment costs under perfect information. The dashed blue lines in both panels show the case of investment with capital adjustment costs under rational inattention. The size of each shock is scaled to imply a 1 percent increase of sectoral investment under perfect information.

It is well-known that capital adjustment costs under perfect information do not give rise to hump-shaped investment responses; in Figure 11 the peak response of sectoral investment to both aggregate and sector-specific shocks occurs in the impact period in this case. Due to increasing marginal costs of capital adjustment, however, decision-makers delay some of their investment spending to future periods, which explains the persistence in sectoral investment responses. Note that the effects in the top panel are longer-lasting than those in the bottom panel. In the calibration, aggregate shocks are more persistent than sector-specific shocks. Hence, the optimal level of capital decays more slowly in response to these shocks.

Now consider the case with the information flow constraint binding. The response of sectoral investment to aggregate shocks becomes hump-shaped. An aggregate shock that increases sectoral investment spending by 1 percent under perfect information leads to a 0.84 increase on impact under rational inattention. At the 1-year horizon, sectoral investment rises further to 0.90 percent. On the other hand, the response of sectoral investment to sector-specific shocks is still monotonically decreasing. A sector-specific shock that increases sectoral investment spending by 1 percent under perfect information leads to a 0.78 increase on impact under rational inattention. At the 1-year horizon, sectoral investment falls to 0.59 percent.

Under rational inattention, the effects of both shocks on sectoral investment are dampened in the impact of period of the shock. The reason for this dampening is that decision-makers have less than perfect information about the current values of aggregate and sector-specific shocks. Note that the dampening in both responses is about equal. Indeed, decision-makers on average attend to information about aggregate shocks equal to 1.2943 bits per period and information about sector-specific shocks equal to 1.0867 bits per period.

Decision-makers pay about as much attention to aggregate and sector-specific shocks even though the unconditional variance of the latter is greater by a factor of about five. To understand this perhaps surprising result, consider the expression for loss of profit due sub-optimal investment decisions in Equation (13). The first term in expectation captures the variance of errors when the actual capital choice given less than perfect information deviates from the profit-maximizing capital choice under perfect information. The second term in expectation captures the first-order autocovariance of errors. The goal of decision-makers is to minimize the variance of errors and to make only those mistakes that do not persistent extensively over time. Notice that these two objectives do not necessarily coincide. On the one hand, because the unconditional variance of sector-specific shocks is larger, decision-makers wish to pay more attention to these shocks. On the other hand, because aggregate shocks are more persistent, the mistakes from not paying attention to these shocks last longer over time. In the calibrated version of the model, these two effects together are about the same for both shocks, which is why decision-makers roughly pay the same amount of attention.

At the 1-year horizon, there is further uncertainty reduction. Decision-makers learn that their optimal capital stock is larger and increase investment spending. This effect is absent under perfect information. At the 1-year horizon, the optimal level of capital is higher in response to aggregate shocks than in response to sector-specific shocks because the former are more persistent than the latter. As a result, decision-makers expand their capital expenditures more strongly and the response of sectoral investment to aggregate shocks becomes hump-shaped.

In the model, the response of sectoral investment to aggregate shocks is the same in every sector. Therefore, aggregation across all production units preserve the form of the investment response to aggregate shocks. My results therefore provide a new microfounded explanation for the hump-shaped response of aggregate investment which is a salient feature of aggregate data.

Crucially, both capital adjustment costs and rational inattention are necessary to obtain these results. The solid black lines in Figure 11 illustrated that capital adjustment costs alone do not give rise to hump-shaped investment responses. Next, I will consider a model without capital adjustment costs and rational inattention. In this model, the response of investment to aggregate shocks is also not hump-shaped.

6.4 Model without Capital Adjustment Costs

The attention problem of decision-makers simplifies when adjusting the capital stock is not costly. Setting $\gamma = 0$ in Equation (13), we have $H_0 = \beta\alpha(\alpha - 1)\overline{K}^{\alpha}$ and $H_1 = 0$. The profit-maximizing capital choice under perfect information in Equation (14) becomes

$$k^*_{it+1} = \frac{E_t \left\{ z_{t+1} + \varepsilon_{it+1} \right\}}{1 - \alpha}, \tag{22}$$

and the expression for loss of profit in Equation (13) reads:

$$\sum_{t=0}^{\infty} \beta^t E_{i,-1} \left\{ \frac{1}{2} \left(\beta\alpha(1-\alpha)\overline{K}^{\alpha} \right) \left(k_{it+1} - k^*_{it+1} \right)^2 \right\}. \tag{23}$$

Notice that the per-period loss becomes static and does not depend on past or future values of capital, even though choosing capital is an intertemporal decision. The reason for this result is the fact that the capital choice for the next period is independent of the current level of capital without capital adjustment costs.

I use the same calibration and the same numerical solution method to solve the decision maker's attention problem in the model without capital adjustment costs. In order to make the two models comparable, however, I fix the amount of attention, κ, at the same level as in the model with capital adjustment costs. The main results from this exercise are that (i) the effects of aggregate shocks on sectoral investment are protracted, but not hump-shaped and (ii) the effects of sector-specific shocks on sectoral investment are short-lived and monotonically decreasing.

Figure 12 displays the response of sectoral investment to aggregate and sector-specific shocks in the model without capital adjustment costs over a 5 year horizon. The solid black lines in both panels show the case of investment under perfect information. The dashed blue lines in both panels show the case of investment under rational inattention. The size of each shock is scaled to imply a 1 percent increase of sectoral investment under perfect information.

Without capital adjustment costs and the constraint on information flow, a decision-maker optimally chooses instantaneous adjustment of capital to the optimal level. Sectoral investment consequently spikes on impact. The effects of shocks to TFP dissipate over time and the optimal level of capital reverts to the non-stochastic steady state. In response to aggregate shocks, the amount of depreciation per period roughly corresponds to the decrease in the optimal capital level, which is why sectoral investment is essentially zero at the 1-year horizon and thereafter. Because sector-specific shocks are less persistent, the optimal level of capital decays faster, which is why sectoral investment turns negative at the 1-year horizon and thereafter.

Now consider the case with the information flow constraint binding. Relative to the perfect information case, the response of sectoral investment to aggregate shocks is dampened. Moreover, the effects of aggregate shocks are protracted; there is still some positive investment at the 1-year horizon, but the response is not hump-shaped. On the other hand, the response of sectoral investment to sector-specific shocks is almost identical to the perfect information case. The reason for this result is that decision-makers now allocate a larger share of attention to sector-specific shocks, about 2/3. Decision-makers choose a different allocation of attention without capital adjustment costs because their errors do not persist over time in this case. The information flow about sectoral TFP thus closely resembles that under perfect information. The information about aggregate shocks is more noisy. On impact the decision-maker dampens the response of sectoral investment because of higher uncertainty. At the 1-year horizon, uncertainty declines, decision-makers learn that the optimal capital stock is larger, and choose to invest. However, the bulk of capital adjustment occurs in the impact period of the shock in the absence of capital adjustment costs and the response does not display a hump-shaped form in this case.

7 Conclusion

This paper shows that, in the median US manufacturing sector, the impulse response of sectoral investment to aggregate shocks is hump-shaped, just as in aggregate data. By contrast, the effects of sector-specific shocks are monotonically decreasing. I solve a model of investment with convex capital adjustment costs and rational inattention. The model predicts that the response of sectoral investment to aggregate shocks is hump-shaped, and monotonically decreasing in response to sector-specific shocks, hence matching the empirical findings of this paper.

There are two different ways in which I will explore the model further in future research. First, I will introduce a household sector to examine feedback effects of the real interest rate on investment activity in general equilibrium. Second, I will formally estimate the model by matching impulse responses of the theoretical model with impulse responses from the statistical model.

References

Altig, D., L. J. Christiano, M. Eichenbaum, and J. Linde (2011): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," *Review of Economic Dynamics*, 14, 225–247.

Bachmann, R., R. J. Caballero, and E. M. Engel (2013): "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model," *American Economic Journal: Macroeconomics*, 5, 29–67.

Bai, J. and P. Wang (2015): "Identification and Bayesian Estimation of Dynamic Factor Models," *Journal of Business & Economic Statistics*, 33, 221–240.

Bartelsman, E. J. and W. Gray (1996): "The NBER Manufacturing Productivity Database," Technical Working Paper 205, National Bureau of Economic Research.

Basu, S., J. G. Fernald, and M. S. Kimball (2006): "Are Technology Improvements Contractionary?" *American Economic Review*, 96, 1418–1448.

Becker, R., W. Gray, and J. Marvakov (2016): "NBER-CES Manufacturing Industry Database: Technical Notes," *mimeo*.

Berger, D., R. Caballero, and E. Engel (2015): "Missing Aggregate Dynamics: On the Slow Convergence of Lumpy Adjustment Models," *mimeo*.

BOIVIN, J., M. P. GIANNONI, AND I. MIHOV (2009): "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data," *American Economic Review*, 99, 350–84.

CARTER, C. K. AND R. KOHN (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

CHIB, S. AND E. GREENBERG (1994): "Bayes Inference in Regression Models with ARMA (p, q) Errors," *Journal of Econometrics*, 64, 183–206.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

DEDOLA, L. AND S. NERI (2007): "What Does a Technology Shock Do? A VAR Analysis with Model-Based Sign Restrictions," *Journal of Monetary Economics*, 54, 512–549.

DEL NEGRO, M. AND C. OTROK (2008): "Dynamic Factor Models with Time-Varying Parameters: Measuring Changes in International Business Cycles," *FRB of New York Staff Report*, 326.

DEL NEGRO, M. AND F. SCHORFHEIDE (2011): "Bayesian Macroeconometrics," in *The Oxford Handbook of Bayesian Econometrics*, Oxford University Press.

FIORI, G. (2012): "Lumpiness, Capital Adjustment Costs and Investment Dynamics," *Journal of Monetary Economics*, 59, 381–392.

JORDÀ, Ò. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 161–182.

KACPERCZYK, M. T., S. VAN NIEUWERBURGH, AND L. VELDKAMP (2016): "A Rational Theory of Mutual Funds' Attention Allocation," *Econometrica*, 84, 571–626.

KHAN, A. AND J. K. THOMAS (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76, 395–436.

Kim, C. and C. Nelson (1999): *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, The MIT Press.

Kuntsevich, A. and F. Kappel (1997): "SolvOpt: The Solver For Local Nonlinear Optimization Problems," *mimeo*.

Luo, Y. (2008): "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics*, 11, 366–385.

Maćkowiak, B., E. Moench, and M. Wiederholt (2009): "Sectoral Price Data and Models of Price Setting," *Journal of Monetary Economics*, 56, S78–S99.

Maćkowiak, B. and M. Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769–803.

——— (2015): "Business Cycle Dynamics under Rational Inattention," *Review of Economic Studies*, 82, 1502–1532.

Matějka, F. (forthcoming): "Rationally Inattentive Seller: Sales and Discrete Pricing," *Review of Economic Studies*.

Matějka, F. and A. McKay (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–298.

Mertens, K. and M. O. Ravn (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, 103, 1212–1247.

Monahan, J. F. (1984): "A Note on Enforcing Stationarity in Autoregressive-Moving Average Models," *Biometrika*, 71, 403–404.

Mondria, J. (2010): "Portfolio Choice, Attention Allocation, and Price Comovement," *Journal of Economic Theory*, 145, 1837–1864.

Otrok, C. and C. H. Whiteman (1998): "Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa," *International Economic Review*, 39, 997–1014.

Paciello, L. (2012): "Monetary Policy and Price Responsiveness to Aggregate Shocks under Rational Inattention," *Journal of Money, Credit and Banking*, 44, 1375–1399.

Paciello, L. and M. Wiederholt (2014): "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," *Review of Economic Studies*, 81, 356–388.

Ramey, V. (2013): "Comment," *NBER Macroeconomics Annual*, 27, 147–153.

Robertson, J. C. and E. W. Tallman (1999): "Vector Autoregressions: Forecasting and Reality," *Economic Review*, First Quarter, 4–18.

Romer, C. D. and D. H. Romer (2010): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100, 763–801.

Sims, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

——— (2006): "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review Papers and Proceedings*, 96, 158–163.

Tutino, A. (2013): "Rationally Inattentive Consumption Choices," *Review of Economic Dynamics*, 16, 421–439.

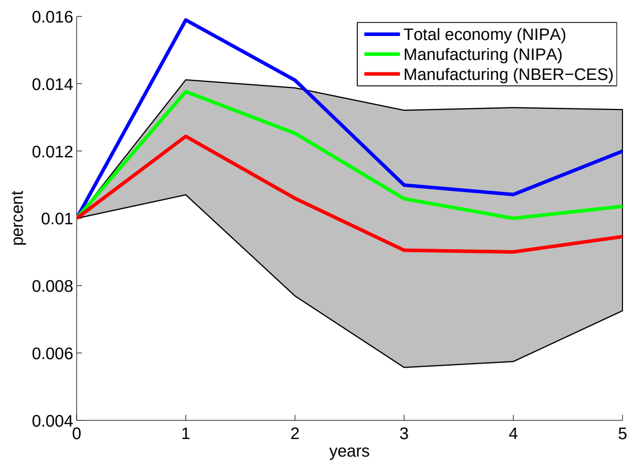
Van Nieuwerburgh, S. and L. Veldkamp (2009): "Information Immobility and the Home Bias Puzzle," *The Journal of Finance*, 64, 1187–1215.

——— (2010): "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77, 779–805.

Verona, F. (2014): "Investment Dynamics with Information Costs," *Journal of Money, Credit and Banking*, 46, 1627–1656.

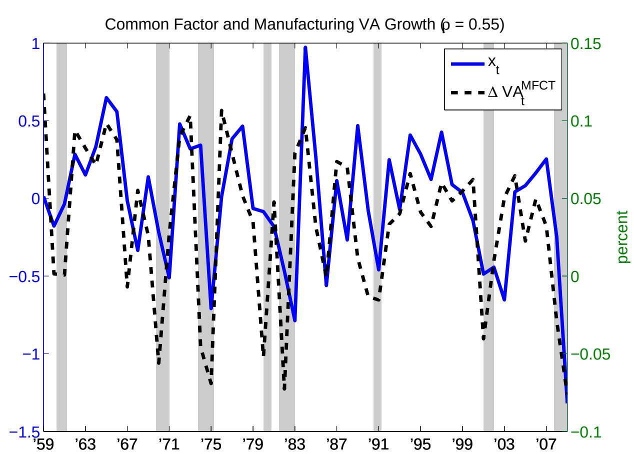
Woodford, M. (2009): "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics*, 56, S100–S124.

Figure 1 – Estimated Response of Aggregate Investment to 1 Percent Innovation.



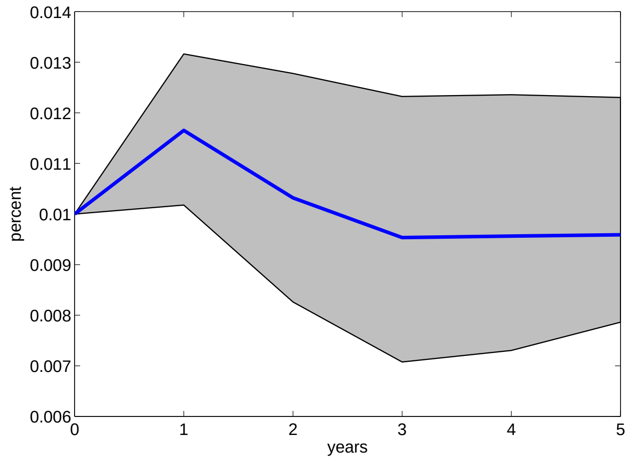
Notes: This figure depicts three impulse responses of aggregate investment to a one percent innovation in year zero. *Total economy (NIPA)* is the response of nonresidential, private fixed investment in the total economy using data from the Bureau of Economic Analysis (BEA) Fixed Asset Accounts, Table 4.8. *Manufacturing (NIPA)* is the response of nonresidential, private fixed investment in the manufacturing sector using data from the same source. *Manufacturing (NBER-CES)* is the response of the aggregated sectoral real capital expenditure data from the NBER-CES Manufacturing Industry Database. Each impulse response is obtained by estimating Equation (4) and computing the cumulative effects of an innovation in w_t that leads to a one percent increase on impact. The gray-shaded area corresponds to the 68 percent error bands for the response of *Manufacturing (NBER-CES)* generated by taking 1,000 draws from the joint posterior density as described in the text.

Figure 2 – Estimated Common Factor.



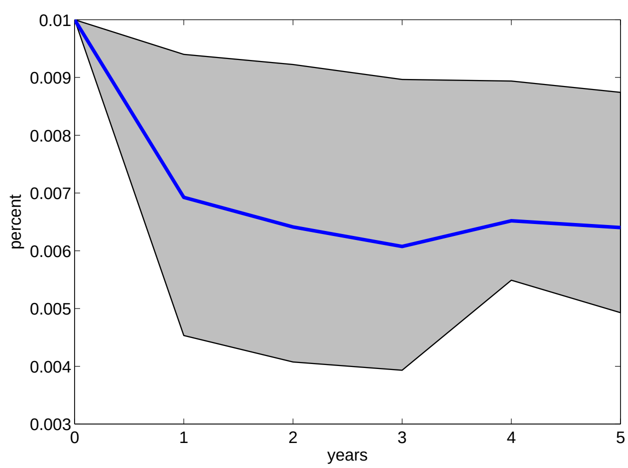
Notes: This figure shows the pointwise posterior median estimate of the common factor, x_t (left axis), in the dynamic factor model given by Equations (1)-(3). The model is estimated using Gibbs-sampling with a Metropolis step as described in the text. ΔVA_t^{MFCT} (right axis) is the growth rate of real value added in the manufacturing industry using GDP-by-industry data from the BEA Annual Industry Account. The correlation coefficient between x_t and ΔVA_t^{MFCT}, ρ, is 0.55. The gray-shaded regions show NBER recessions.

Figure 3 – Estimated Response of Sectoral Investment to 1 Percent Aggregate Shock.



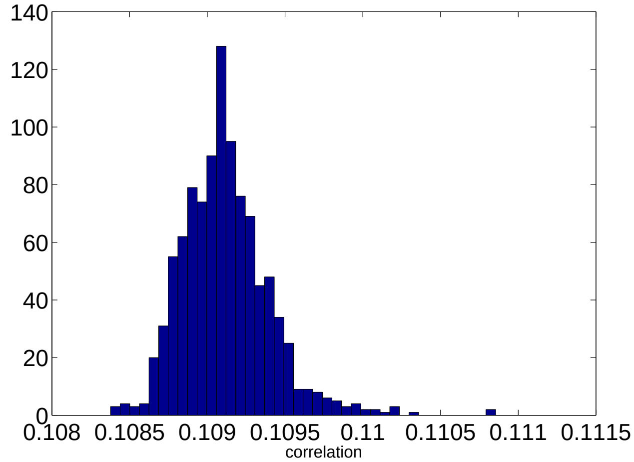
Notes: This figure plots the impulse response of sectoral investment in the median industry to a one percent aggregate shock in year zero. The impulse response is obtained by estimating the dynamic factor model in Equations (1)-(3) and computing the cumulative effects of an innovation in v_t that leads to a one percent increase on impact. The gray-shaded area corresponds to the 68 percent error bands generated by taking 1,000 draws from the joint posterior density as described in the text. The median industry is defined as the pointwise median impulse response across all draws and sectors.

Figure 4 – Estimated Response of Sectoral Investment to 1 Percent Sector-Specific Shock.



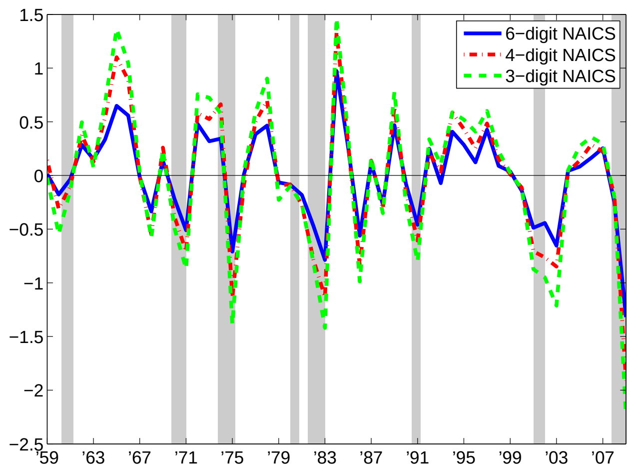
Notes: This figure plots the impulse response of sectoral investment in the median industry to a one percent sector-specific shock in year zero. The impulse response is obtained by estimating the dynamic factor model in Equations (1)-(3) and computing the cumulative effects of an innovation in u_{it} that leads to a one percent increase on impact. The gray-shaded area corresponds to the 68 percent error bands generated by taking 1,000 draws from the joint posterior density as described in the text. The median industry is defined as the pointwise median impulse response across all draws and sectors.

Figure 5 – Testing for the Number of Common Factors.



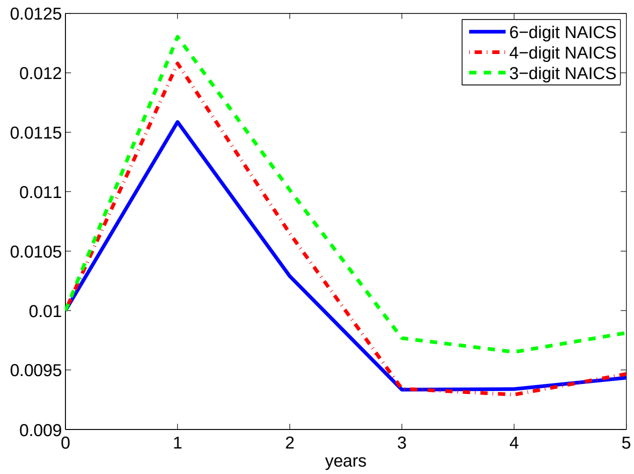
Notes: The histogram in this figure depicts the posterior density of the statistic defined in the text to test for the number of common factors. The test statistic is the median absolute value of cross-sectional correlations between the sector-specific components of the dynamic factor model in Equations (1)-(3), obtained by taking a draw from the joint posterior density of the common factor and the model's parameters, computing the pairwise cross-sectional correlations between the w_{it} in Equation (1), corr $\left[w_i, w_j\right], \forall i \neq j$, taking absolute values, and retaining the median across sectors. The posterior density of this statistic is simulated for 1,000 draws.

Figure 6 – Estimated Common Factor by Level of Aggregation.



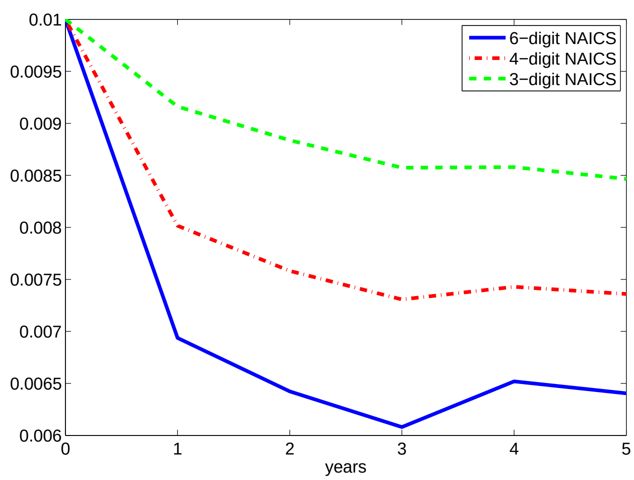
Notes: This figure shows three estimates of the common factor, x_t, obtained by estimating the dynamic factor model in Equations (1)-(3) using sectoral real capital expenditure data at different levels of aggregation. *6-digit NAICS* is the pointwise posterior median estimate shown in Figure 2 using data at the 6-digit North American Industry Classification System (NAICS) industry level. *4-digit NAICS* is the pointwise posterior median estimate using 4-digit NAICS industry-level data. *3-digit NAICS* is the pointwise posterior median estimate using data at the 3-digit NAICS industry level. The gray-shaded regions show NBER recessions.

Figure 7 – Estimated Response to 1 Percent Aggregate Shock by Level of Aggregation.



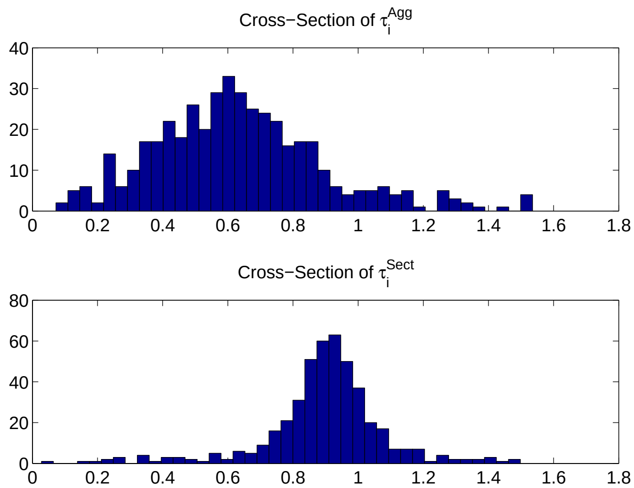
Notes: This figure plots impulse responses of sectoral investment at different levels of aggregation to a one percent aggregate shock in year zero. *6-digit NAICS* is the response in the median industry shown in Figure 3 using data at the 6-digit NAICS industry level. *4-digit NAICS* is the response in the median industry using 4-digit NAICS industry-level data. *3-digit NAICS* is the response in the median industry at the 3-digit NAICS industry level. See the notes to Figure 3 for further information.

Figure 8 – Estimated Response to 1 Percent Sector-Specific Shock by Level of Aggregation.



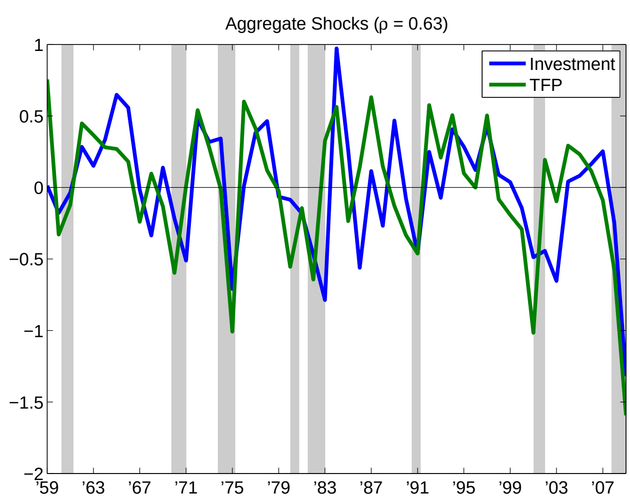
Notes: This figure plots impulse responses of sectoral investment at different levels of aggregation to a one percent sector-specific shock in year zero. *6-digit NAICS* is the response in the median industry shown in Figure 3 using data at the 6-digit NAICS industry level. *4-digit NAICS* is the response in the median industry using 4-digit NAICS industry-level data. *3-digit NAICS* is the response in the median industry at the 3-digit NAICS industry level. See the notes to Figure 4 for further information.

Figure 9 – Speed of Adjustment to Shocks Using Proxy Variables for Each Shock.



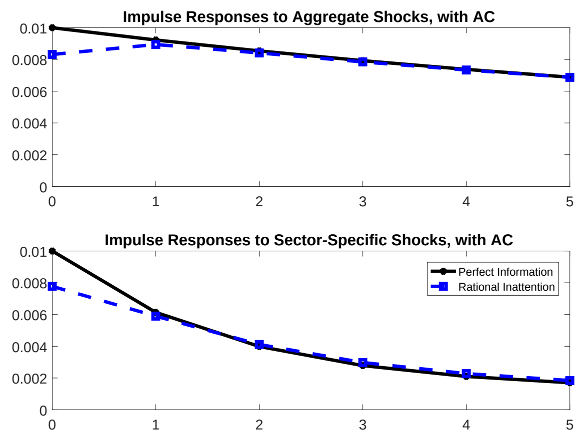
Notes: This figure depicts histograms of the speed of adjustment to aggregate shocks and sector-specific shocks using direct proxy variables for each shock. The proxy variables for each shock are measures of aggregate and sector-specific total factor productivity (TFP), respectively, constructed as described in the text. The top panel plots the cross-section of the speed of adjustment statistic for aggregate shocks, τ_i^{Agg}, the bottom panel the cross-section of the speed of adjustment statistic for sector-specific shocks, τ_i^{Sect}, both defined in Equation (6). Each panel trims the histogram at the maximum of the 95th percentiles of either the τ_i^{Agg} or the τ_i^{Sect}.

Figure 10 – Estimated Common Factor and Aggregate Total Factor Productivity.



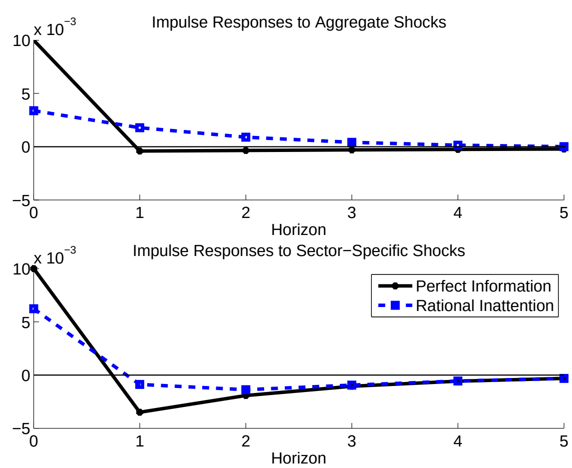
Notes: This figure plots the pointwise posterior median estimate of the common factor, x_t, in the dynamic factor model given by Equations (1)-(3). The model is estimated using Gibbs-sampling with a Metropolis step as described in the text. TFP is the first principal component of sectoral TFP growth rates constructed as described in the text. The correlation coefficient between x_t and TFP, ρ, is 0.63. The gray-shaded regions show NBER recessions.

Figure 11 – Model Responses to Aggregate and Sector-Specific Shocks



Notes: This figure depicts impulse responses of sectoral investment to aggregate and sector-specific shocks in the model with capital adjustment costs and two different information structures. The top panel shows the impulse response to aggregate shocks, the bottom panel the impulse response to sector-specific shocks. *Perfect Information* plots the response to a one percent innovation when decision-makers know the history of the economy up to and including period t. *Rational Inattention* depicts the response for the same shock when the information-flow constraint in Equation (17) is binding. The calibration and numerical solution of the model follows the description in the text.

Figure 12 – Model Responses to Shocks without Capital Adjustment Costs



Notes: This figure depicts impulse responses of sectoral investment to aggregate and sector-specific shocks in the model without capital adjustment costs and two different information structures. The top panel shows the impulse response to aggregate shocks, the bottom panel the impulse response to sector-specific shocks. *Perfect Information* plots the response to a one percent innovation when decision-makers know the history of the economy up to and including period t. *Rational Inattention* depicts the response for the same shock when the information-flow constraint in Equation (17) is binding. The calibration and numerical solution of the model follows the description in the text.

A Model with Investment Adjustment Costs

The purpose of this appendix is to investigate whether other, existing propagation mechanisms are consistent with my empirical findings. Following Christiano et al. (2005), many business cycle models assume convex costs in the growth rate of investment, so-called investment adjustment costs, so as to match the hump-shaped response of aggregate investment to macroeconomic shocks. This appendix outlines and calibrates a model with investment adjustment costs and perfect information. I use the model to study the responses of sectoral investment to aggregate and sector-specific shocks under this alternative propagation mechanism.

The model takes into account the effects of time aggregation on the estimated investment responses. Remember that the capital expenditure data in the estimation of the statistical model is at the yearly frequency. It is possible that the speed of adjustment following sector-specific shocks is faster (absent general equilibrium price responses, for instance) and that the response of sectoral investment is also hump-shaped at higher frequencies. In this case, time aggregation from quarterly to yearly frequency can obtain a monotonically decreasing response to sector-specific shocks. Therefore, I calibrate the model to the quarterly frequency and time-aggregate the theoretical investment responses to the yearly frequency.

My findings are as follows. In partial equilibrium, the effects of both aggregate and sector-specific shocks on sectoral investment are hump-shaped. In addition, if a household sector closes the model in general equilibrium, the response of sectoral investment to sector-specific shocks becomes relatively *more* hump-shaped in the sense that the peak response occurs after a longer period of time. Time aggregation does not change these results. Hence, under standard assumptions and using a standard calibration of the model parameters, a model with investment adjustment costs and perfect information has difficulties to explain my empirical findings.

A.1 Setup

The physical environment of the economy is the same as in Section 5, except that production units now face investment instead of capital adjustment costs.

The economy consists of a unit measure of sectors, which are each populated by a representative production unit indexed by i. Time is discrete. Production unit i operates the production function $Y_{it} = Z_t \mathcal{E}_{it} K_{it}^{\alpha}$, where K_{it} denotes the current stock of capital, Z_t and $\mathcal{E}_{it}$ are aggregate and sectoral total factor productivity (TFP), and α is a parameter.

Production units own the capital stock, which is specific to their sector. The law of motion for capital now reads $K_{it+1} = (1-\delta)K_{it} + \left(1 - S\left(\frac{I_{it}}{I_{it-1}}\right)\right) I_{it}$, where I_{it} is investment, δ denotes the rate of depreciation, and $S\left(\frac{I_{it}}{I_{it-1}}\right)$ are investment adjustment costs. The function S is monotonically increasing, convex, and satisfies $S(1) = S'(1) = 0$.

The sectoral and aggregate components of TFP each follow stationary Gaussian first-order autoregressive processes in logs: $\ln Z_t = \rho_z \ln Z_{t-1} + e_t$ and $\ln \mathcal{E}_{it} = \rho_\varepsilon \ln \mathcal{E}_{it-1} + v_{it}$, where the error terms are Gaussian white noise with distributions $e_t \sim \mathcal{N}(0, \sigma_e^2)$ and $v_{it} \sim \mathcal{N}(0, \sigma_v^2)$, respectively. The sector-specific shocks, v_{it}, are pairwise independent in the cross-section. Moreover, the v_{it} are independent of aggregate shocks, e_t.

Decision-makers in production units discount future profits between period t and period 0 using the stochastic discount factor $\beta^t \lambda_t$. Their profit maximization problem reads

$$\max_{\{K_{it+1}, I_{it}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \lambda_t \left[Z_t \mathcal{E}_{it} K_{it}^{\alpha} - I_{it} \right]$$

subject to the capital accumulation equation, the stochastic processes for aggregate and sector-specific TFP, and given an initial capital stock K_{i0}.

The household sector of this economy is deliberately simple. A representative household consumes, buys shares of production units, receives dividends, and trades in a risk-free bond. Market are complete. Households maximize lifetime utility, their instantaneous utility function is $U(C_t)$, and their discount factor is β.

Market clearing and aggregation require the following:

$$\int_0^1 Y_{it} di = C_t + \int_0^1 I_{it} di,$$
$$K_t = \int_0^1 K_{it} di.$$

Aggregate output equals consumption and aggregate investment expenditures. Aggregate capital equals the integral over each production's unit capital stock.

A.2 Solution and Calibration

I solve this model by taking a log-linear approximation to the household's and production unit's optimality conditions, the law of motion for capital, and the market clearing conditions.

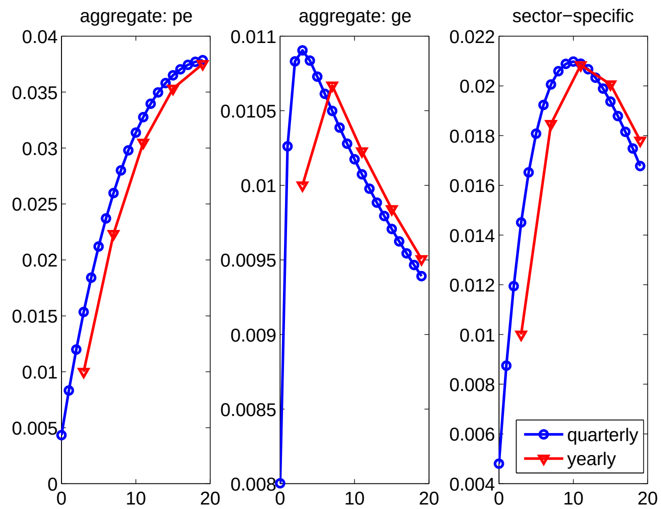
A period in the model now corresponds to a quarter. The calibration of the model's parameters is exactly the same as in Section 5, adjusted correspondingly to account for the change in frequency. The second derivative of the function S is set to 1.5. This value corresponds to the estimate by Altig et al. (2011).

To aggregate the investment responses over time, I use the fact that $i_y = \frac{1}{4}\left(i_{q1} + i_{q2} + i_{q3} + i_{q4}\right)$. That is, the log-deviation of investment from its non-stochastic steady state at the yearly frequency equals the yearly average log-deviation of investment from its non-stochastic steady state at the quarterly frequency.

A.3 Results

Figure 13 shows the effects of aggregate and sector-specific shocks on sectoral investment in the model with investment adjustment costs and perfect information. The left panel shows the responses to aggregate shocks in the partial equilibrium version of the model (that is, with the real interest rate fixed at its steady-state value). The middle panel

Figure 13 – Investment Responses in Model with Investment Adjustment Costs.



depicts the effects of aggregate shocks in general equilibrium. The right panel graphs the responses of sectoral investment to sector-specific shocks. In each panel, blue lines with circles show the model response of sectoral investment at the quarterly frequency, while red lines with triangles correspond to the model responses time-aggregated to the yearly frequency.

At the quarterly frequency, the response of sectoral investment to both aggregate and sector-specific shocks is slowly building over time and the peak response does not occur on impact. In either case, production units must pay investment adjustment costs and abrupt investment growth is extremely costly. Aggregate shocks are more persistent than sector-specific shocks, which is why in partial equilibrium decision-makers find it optimal to smooth their investment expenditure over a longer time period of time. As a result, the peak response following aggregate shocks occurs later. In general equilibrium, the real interest rate decreases because the supply of funds increases stronger than investment demand, because the latter is constrained by investment adjustment costs. The rate reduction makes capital today more valuable and decision-makers find it optimal to front-load their investment spending.

Time aggregation from the quarterly to the yearly frequency does not change these findings. Note that, in general equilibrium, the response following sector-specific shocks is actually more hump-shaped in the sense that the peak response occurs later. I conclude that a model with investment adjustment costs and perfect information has difficulties to explain the discrepancy in the empirical responses of sectoral investment to differential shocks, at least under standard assumptions and under the standard calibration used in this exercise.

B Aggregate Shocks and the Relative Price of Investment

This section tests whether the macroeconomic shock estimated in the statistical model of this paper is correlated with the relative price of investment in the manufacturing sector. Fiori (2012) formulates an alternative model that is also consistent with the observed hump-shape response of aggregate investment. In his model, rapid output expansion in the investment good producing sector is costly. In response to aggregate shocks, the relative price of investment increases, initially depressing demand for investment goods in all other sectors of the economy. As the supply of investment goods increases over time, the relative price of investment falls and investment demand in the rest of the economy picks up. Aggregation across all sectors in the economy obtains a hump-shaped response of aggregate investment to macroeconomic shocks.

In order to evaluate this alternative model, I test one of its key predictions in the data: movements in the relative price of investment in response to macroeconomic shocks. To this end, I estimate (by ordinary least squares) a bivariate vectorautoregression (VAR) and test for Granger-causality of the macroeconomic factor for the relative price of investment in the manufacturing sector. The VAR contains three lags.

For simplicity, I use the pointwise posterior median estimate of the macroeconomic factor (depicted in Figure 6). The relative price of investment in the manufacturing industry corresponds to the ratio of the deflators for investment and gross domestic product. I work with two measures of the deflator for investment. The first measure uses aggregate manufacturing investment data while the second measure is based on the aggregated micro data.[20]

At the 5% significance level, the macroeconomic factor is not Granger-causal for the relative price of investment in the manufacturing sector for neither measure of the latter. Hence, there is no evidence that macroeconomic shocks are followed by movements in the relative price of investment, one of the key predictions of the model by Fiori (2012).

C Econometric Appendix

This appendix provides further details about the statistical model for the sectoral capital expenditure data. For the reader's convenience, I first restate the dynamic factor model from Section 2. Next, I describe identification of the unobserved factors and the unobserved loadings. The appendix then moves on to explain the estimation methodology, which closely follows Del Negro and Schorfheide (2011). Specifically, I use the Gibbs sampling algorithm to sample from the joint posterior of the factors and the model's parameters. This algorithm draws alternately from their respective conditional distributions to generate a sample from the joint distribution. I lay out the priors and write down the conditional posterior densities. Importantly, I do not condition on initial observations but use the full conditional distributions in the Gibbs sampling algorithm. A minor difference between this paper and the estimation methodology by Del Negro and Schorfheide (2011) is that I switch the ordering of conditional distributions in the algorithm. In particular, I first sample from the conditional posterior density of the factors and then from the conditional posterior density of the model's parameters. The appendix concludes by describing how I initialize the Gibbs sampling algorithm.

[20] See Footnote 11 for data sources of aggregate manufacturing data used in this exercise.

Model Consider the dynamic factor model

$$x_t = F(\ell)x_{t-1} + v_t, \qquad v_t \sim i.i.d.\,\mathcal{N}(0, Q) \tag{24}$$

$$y_{it} = H_i x_t + w_{it} \tag{25}$$

$$w_{it} = D_i(\ell)w_{it-1} + u_{it}, \qquad u_{it} \sim i.i.d.\,\mathcal{N}(0, R_i) \tag{26}$$

where y_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$, denotes the standardized period t sector i log change of real investment, x_t is an unobserved factor, the H_i are factor loadings, and the w_{it} are sector-specific components. Both x_t and w_{it} follow AR processes, $F(\ell)$ and $D_i(\ell)$ denote lag polynomials of order three, and v_t and the u_{it} are Gaussian white noise with variance Q and R_i, respectively. Assume that the u_{it} are pairwise independent and uncorrelated with v_t.

Identification Stacking Equation (25) over all i gives

$$y_t = Hx_t + w_t \tag{27}$$

where y_t, w_t, and H are column vectors of length n. Because the factor and the loadings are unobserved, their sign and scale are not identified from the data. Therefore, I assume that the first element in H is positive and that Q in Equation (24) is a known constant. These assumptions are standard in the literature on dynamic factor models and uniquely identify the space spanned by the factors.

Priors The prior distribution for the coefficients of $F(\ell)$ is $\mathcal{N}(\phi_0, \Phi_0^{-1})I_{S_F}$, where $\mathcal{N}$ denotes the multivariate Normal distribution with mean ϕ_0 and second moment Φ_0^{-1}, and I_{S_F} is an indicator function for stationary of x_t implied by $F(\ell)$. Similarly, the prior for the coefficients of $D_i(\ell)$ is $\mathcal{N}(\theta_0, \Theta_0^{-1})I_{S_D}$. I choose prior means ϕ_0 and θ_0 equal to column vectors of zeros of length three. The prior precisions are small but increase with lag length as in the case of the Minnesota prior. In particular, following Robertson and Tallman (1999), I set the lag l prior precisions implied by Φ_0 and Θ_0 equal to $(\exp(cl - c))^{-1}$, where c matches a quarterly harmonic decay rate at lag three. The prior for each R_i is $\mathcal{IG}(\nu_0/2, \delta_0/2)$, where $\mathcal{IG}$ denotes the inverse gamma distribution. Following Otrok and Whiteman (1998), I set $\nu_0 = 6$ and $\delta_0 = 0.001$, which implies a diffuse prior distribution. Finally, the prior on each loading H_i is $\mathcal{N}(\beta_0, B_0^{-1})$. I choose $\beta_0 = 0$ and $B_0 = 1$.

Sample factors, conditional on parameters and data In general, let p_x and p_w denote the order of the lag polynomials $F(\ell)$ and $D_i(\ell)$, respectively. To sample from the conditional posterior density of the factors given the parameters and the data, I follow Carter and Kohn (1994). Given $D_i(\ell)$ and H_i, define $y_{it}^* = (1 - D_i(\ell)L)y_{it}$ and the lag polynomial $h_i^*(\ell) = (1 - D_i(\ell)L)H_i$ of order p_w and, using Equation (26), rewrite Equation (25) as $y_{it}^* = h_i(\ell)^* x_t + u_{it}$. Let H_i^* the $(p_w + 1) \times 1$ column vector which stacks all the coefficients of $h_i^*(\ell)$ and define the $(p_w + 1) \times 1$ column vector $x_t^* = \begin{bmatrix} x_t & x_{t-1} & \dots & x_{t-p_w} \end{bmatrix}^T$. Thus, we can express the equation for y_{it}^* as $y_{it}^* = H_i^{*T} x_t^* + u_{it}$. Stacking each of these n equations, we can write down the state-space representation:

$$x_t^* = F^* x_{t-1}^* + v_t^* \tag{28}$$

$$y_t^* = H^* x_t^* + u_t \tag{29}$$

where v_t^* is the $(p_w + 1) \times 1$ vector $v_t^* = \begin{bmatrix} v_t & 0 & \dots & 0 \end{bmatrix}^T$, H^* is an $n \times (p_w + 1)$ matrix, and F^* is the $(p_w + 1) \times (p_w + 1)$ matrix

$$F^* = \begin{bmatrix} F & 0_{1 \times ((p_w+1) - p_x)} \\ \mathcal{I}_{p_w} & 0_{p_w \times 1} \end{bmatrix} \tag{30}$$

where F is the $1 \times p_x$ row vector which corresponds to the first row of the companion form matrix of $F(\ell)$. Note that this notation assumes $p_w + 1 \geq p_x$ and that Equation (29) starts from $t = p_w + 1$ instead of $t = 1$ because $y^*_0, \ldots, y^*_{-p_w+1}$ are unobserved. The variance-covariance matrix of v^*_t, Q^*, is $(p_w + 1) \times (p_w + 1)$, the first element on the main diagonal corresponds to Q, and all other elements equal zero. The variance-covariance matrix of u_t is given by $R = \text{diag}(R_1, \ldots, R_n)$. Conditional on F^*, Q^*, H^*, R, and the data, the Carter and Kohn (1994) simulation smoother draws a whole sample of the x_t, $t = p_w + 1, \ldots, T$, from the corresponding conditional posterior density function. For the sake of brevity, I omit the conditioning arguments below. Let $\widetilde{F}^*$ denote the first row of F^*. Following Kim and Nelson (1999), I recursively sample from the conditional distributions $x^*_T \sim \mathcal{N}(x^*_{T|T}, P_{T|T})$ and $x^*_t \mid x_{t+1} \sim \mathcal{N}(x^*_{t|t,x_{t+1}}, P_{t|t,x_{t+1}})$, $t = T-1, \ldots, p_w + 1$, where

$$x^*_{t|t,x_{t+1}} = x^*_{t|t} + P_{t|t}\widetilde{F}^{*T}(\widetilde{F}^* P_{t|t}\widetilde{F}^{*T} + Q)^{-1}(x_{t+1} - \widetilde{F}^* x^*_{t|t}) \tag{31}$$

$$P_{t|t,x_{t+1}} = P_{t|t} - P_{t|t}\widetilde{F}^{*T}(\widetilde{F}^* P_{t|t}\widetilde{F}^{*T} + Q)^{-1}\widetilde{F}^* P_{t|t} \tag{32}$$

and $x^*_{t|t}$ and $P_{t|t}$ are the conditional mean and the conditional variance of x^*_t obtained from Kalman filtering. The first element of each draw x^*_t corresponds to a draw of x_t.

Following Del Negro and Otrok (2008), I use the density of $x^*_{p_w}$ conditional on the model's parameters and the data to initialize the Kalman filter. Specifically, rewrite Equation (27) as

$$y_t = \overbrace{\begin{bmatrix} H & 0_{n \times p_w} \end{bmatrix}}^{\equiv \tilde{H}} x^*_t + w_t \tag{33}$$

and substitute $x^*_t = (F^*)^t x^*_0 + \sum_{j=0}^{t-1} (F^*)^j v^*_{t-j}$ for x^*_t. Stacking the first p_w observations gives

$$\underbrace{\begin{bmatrix} y_{p_w} \\ \vdots \\ y_1 \end{bmatrix}}_{\equiv y^{p_w \ldots 1}} = \underbrace{\begin{bmatrix} \tilde{H}(F^*)^{p_w} \\ \vdots \\ \tilde{H}(F^*) \end{bmatrix}}_{\equiv A} x_0^* + \overbrace{\begin{bmatrix} \tilde{H} & \tilde{H}F^* & \cdots & \tilde{H}(F^*)^{p_w-1} \\ 0_{n \ \times (p_w+1)} & \tilde{H} & \cdots & \tilde{H}(F^*)^{p_w-2} \\ \vdots & & \ddots & \vdots \\ 0_{n \ \times (p_w+1)} & \cdots & \cdots & \tilde{H} \end{bmatrix}}^{\equiv B} \overbrace{\begin{bmatrix} v_{p_w}^* \\ \vdots \\ v_1^* \end{bmatrix}}^{\equiv (v^*)^{p_w \ldots 1}} + \overbrace{\begin{bmatrix} w_{p_w} \\ \vdots \\ w_1 \end{bmatrix}}^{\equiv w^{p_w \ldots 1}} \tag{34}$$

$$x_{p_w}^* = (F^*)^{p_w} x_0^* + \underbrace{\begin{bmatrix} \mathcal{I}_{(p_w+1)} & F^* & \cdots & (F^*)^{p_w-1} \end{bmatrix}}_{\equiv C} (v^*)^{p_w \ldots 1} \tag{35}$$

The joint distribution of the p_w initial observations of the data and the $(p_w + 1)$ initial observations of the factors, conditional on the data, therefore reads

$$\begin{pmatrix} y^{p_w \ldots 1} \\ x_{p_w}^* \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} AE\{x_0^*\} \\ (F^*)^{p_w} E\{x_0^*\} \end{pmatrix}, \begin{pmatrix} A\Sigma_{x_0^*} A^T + B\Sigma_{(v^*)^{p_w \ldots 1}} B^T + \Sigma_{w^{p_w \ldots 1}} & \bullet \\ (F^*)^{p_w} \Sigma_{x_0^*} A^T + C\Sigma_{(v^*)^{p_w \ldots 1}} B^T & (F^*)^{p_w} \Sigma_{x_0^*} ((F^*)^{p_w})^T + C\Sigma_{(v^*)^{p_w \ldots 1}} C^T \end{pmatrix} \right)$$

where $E\{x_0^*\}$ and $\Sigma_{x_0^*}$ are the unconditional mean and variance covariance matrix of x_0^*, respectively, $\Sigma_{(v^*)^{p_w \ldots 1}}$ denotes the variance covariance matrix of $(v^*)^{p_w \ldots 1}$, and $\Sigma_{w^{p_w \ldots 1}}$ is the variance covariance matrix of $w^{p_w \ldots 1}$.

From the properties of the multivariate normal distribution, it follows that $x_{p_w}^* \mid y^{p_w \ldots 1} \sim \mathcal{N}$ with first and second moment given by

$$\begin{aligned} E\{x_{p_w}^* \mid y^{p_w \ldots 1}\} = {} & (F^*)^{p_w} E\{x_0^*\} + ((F^*)^{p_w} \Sigma_{x_0^*} A^T + C\Sigma_{(v^*)^{p_w \ldots 1}} B^T) \\ & (A\Sigma_{x_0^*} A^T + B\Sigma_{(v^*)^{p_w \ldots 1}} B^T + \Sigma_{w^{p_w \ldots 1}})^{-1} (y^{p_w \ldots 1} - AE\{x_0^*\}) \end{aligned} \tag{36}$$

$$\begin{aligned} V\{x_{p_w}^* \mid y^{p_w \ldots 1}\} = {} & ((F^*)^{p_w} \Sigma_{x_0^*} ((F^*)^{p_w})^T + C\Sigma_{(v^*)^{p_w \ldots 1}} C^T) - ((F^*)^{p_w} \Sigma_{x_0^*} A^T + C\Sigma_{(v^*)^{p_w \ldots 1}} B^T) \\ & (A\Sigma_{x_0^*} A^T + B\Sigma_{(v^*)^{p_w \ldots 1}} B^T + \Sigma_{w^{p_w \ldots 1}})^{-1} ((F^*)^{p_w} \Sigma_{x_0^*} A^T + C\Sigma_{(v^*)^{p_w \ldots 1}} B^T)^T \end{aligned} \tag{37}$$

where $\Sigma_{(v^*)^{p_w\dots 1}} = \mathcal{I}_{p_w} \otimes Q^*$. To work out $\Sigma_{w^{p_w\dots 1}}$, rewrite the process for w_t in companion form

$$\begin{bmatrix} w_t \\ \vdots \\ w_{t-p_w+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \text{diag}(D_1) & \text{diag}(D_2) & \cdots & \text{diag}(D_{p_w}) \\ \mathcal{I}_n & & \cdots & 0_n \\ \vdots & \ddots & & \vdots \\ 0_n & \cdots & \mathcal{I}_n & 0_n \end{bmatrix}}_{\equiv D} \begin{bmatrix} w_{t-1} \\ \vdots \\ w_{t-p_w} \end{bmatrix} + \begin{bmatrix} u_t \\ \vdots \\ 0_n \end{bmatrix} \tag{38}$$

where $\text{diag}(D_i)$ is a $n \times n$ diagonal matrix with the coefficients on the ith lag for each sector on the main diagonal and $u_t \sim \mathcal{N}(0_n, R)$. Hence, under stationarity, we have

$$\text{vec}(\Sigma_{w^{p_w\dots 1}}) = (\mathcal{I}_{(np_w)^2} - D \otimes D)^{-1} \text{vec}\left(\begin{bmatrix} R & \cdots & 0_n \\ \vdots & \ddots & \vdots \\ 0_n & \cdots & 0_n \end{bmatrix}\right) \tag{39}$$

Finally, under stationarity of the factors, $E\{x_0^*\} = 0_{(p_w+1)\times 1}$ and $\text{vec}(\Sigma_{x_0^*}) = (\mathcal{I}_{(p_w+1)^2} - F^* \otimes F^*)^{-1} \text{vec}(Q^*)$. For numerical robustness, I use the method by Bai and Wang (2015) to compute the conditional variance covariance matrix.

To initialize the Kalman filter in the Carter and Kohn (1994) simulation smoother, I use the conditional mean $F^* E\{x_{p_w}^* \mid y^{p_w\dots 1}\}$ and conditional variance $F^* V\{x_{p_w}^* \mid y^{p_w\dots 1}\}(F^*)^T + Q^*$. The p_w initial observations of x_t are drawn from $x_{p_w}^* \mid y^{p_w\dots 1} \sim \mathcal{N}$ with first and second moment given by Equation (36) and (37), respectively. The last element of x_{p_w}, x_0, is discarded.

Sample parameters of state equation, conditional on parameters in observation equation, factors and data Abusing notation, write Equation (24) in companion form $x_t^* = F^* x_{t-1}^* + v_t^*$ where F^* denotes the $p_x \times p_x$ companion form matrix of $F(\ell)$ and $v_t \sim \mathcal{N}(0_{p_x}, Q^*)$. Suppose that this process is stationary and that the initial observation $x_0^* =$

$\begin{bmatrix} x_0 & x_{-1} & \dots & x_{-p_x+1} \end{bmatrix}^T$ is drawn from the stationary distribution $x_0^* \sim \mathcal{N}(0_{p_x}, Q\Sigma_x)$ where $\text{vec}(\Sigma_x) = (\mathcal{I}_{p_x^2} - F^* \otimes F^*)^{-1} + \text{vec}(e_1(p_x)e_1(p_x)^T)$ with $e_1(p_x) = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T$ denoting the $p_x \times 1$ unit vector. Let e the $T - p_x \times 1$ column vector containing x_t, $t = p_x + 1, \dots, T$ and E the $T - p_x \times p_x$ matrix with tth row given by $\begin{bmatrix} x_{t-1} & \dots & x_{t-p_x} \end{bmatrix}$. Given Q, H, R, and the data, Chib and Greenberg (1994) show that the full conditional posterior of the parameters of the lag polynomial $F(\ell)$ is given by $F \propto \Psi_F(F) \times \mathcal{N}(\hat{\phi}, \Phi_n^{-1}) I_{S_F}$, where $\hat{\phi} = \Phi_n^{-1}(\Phi_0\phi_0 + Q^{-1}E^T e)$, $\Phi_n = (\Phi_0 + Q^{-1}E^T E)$, and

$$\Psi_F(F) = |\Sigma_x(F)|^{-1/2} \exp[-\frac{1}{2Q} x_0^T \Sigma_x^{-1}(F) x_0] \tag{40}$$

To sample from the conditional distribution, Chib and Greenberg (1994) use a Metropolis-Hastings step. That is, in the jth iteration of the Gibbs sampler, I generate a candidate draw F' from the distribution $\mathcal{N}(\hat{\phi}, \Phi_n^{-1}) I_{S_F}$ and use it for the next iteration with probability $min(\Psi_F(F')/\Psi_F(F^{(j-1)}), 1)$. With probability $(1 - min(\Psi_F(F')/\Psi_F(F^{(j-1)}), 1))$, I retain the current value $F^{(j-1)}$.

Sample parameters of observation equation, conditional on factors and data To sample from the conditional posterior density of the observation equation's parameters, note that the Equations (25) are independent regressions with AR(p_w) errors, given the factor (Otrok and Whiteman, 1998). I follow the method by Chib and Greenberg (1994) to sample from the posterior equation-by-equation.

Write Equation (26) in companion form $w_{it}^* = D_i^* w_{it-1}^* + u_{it}^*$, where D_i^* denotes the $p_w \times p_w$ companion form matrix of $D_i(\ell)$, and $u_{it}^* \sim \mathcal{N}(0_{p_w}, R_i^*)$, $R_i^* = \text{diag}(R_i, 0, \dots, 0)$. Suppose that this process is stationary and that the initial observation $w_0^* = \begin{bmatrix} w_0 & w_{-1} & \dots & w_{-p_w+1} \end{bmatrix}^T$ is drawn from the stationary distribution $w_0^* \sim \mathcal{N}(0_{p_w}, R_i\Sigma_w)$, where $\text{vec}(\Sigma_w) = (\mathcal{I}_{p_w^2} - D_i^* \otimes D_i^*)^{-1} + \text{vec}(e_1(p_w)e_1(p_w)^T)$ with $e_1(p_w) = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T$ denoting the $p_W \times 1$ unit vector. Let $y_{i1}^* = P^{-1}y_{i1}$, $x_1^* = P^{-1}x_1$, where P solves $PP^T = \Sigma_w$. Define y_{i2}^* and x_2^* with typical element $(1 - D_i(\ell)L)y_{it}$ and $(1 - D_i(\ell)L)x_t$, $t = p_w + 1, \dots, T$, respectively.

Stacking all transformed observations gives $y^* = \left[y_{i1}^{*T} \; y_{i2}^{*T} \right]^T$ and $x^* = \left[x_1^{*T} \; x_2^{*T} \right]^T$. Let $e_t = y_{it} - H_i x_t$ and define $e = \left[e_{p_w+1} \cdots e_T \right]^T$ and the $T - p_w \times p_w$ matrix E with typical row given by $\left[e_{t-1} \cdots e_{t-p_w} \right]^T$, $t = p_w, \ldots, T$. Chib and Greenberg (1994) give the full conditional posterior densities

$$H_i \mid R_i, D_i(\ell) \sim \mathcal{N}(B_n^{-1}(B_0\beta_0 + R_i^{-1}X^{*T}y_i^*), B_n^{-1}), \tag{41}$$

$$R_i \mid H_i, D_i(\ell) \sim \mathcal{IG}((v_o + n)/2, (\delta_0 + d_1)/2), \tag{42}$$

$$D_i(\ell) \mid H_i, R_i \quad \propto \Psi_D(D_i) \times \mathcal{N}(\hat{\theta}, \Theta_n^{-1}) I_{S_{D_i}}, \tag{43}$$

where $B_n = B_0 + R_i^{-1}X^{*T}X^*$, $\hat{\theta} = \Theta_n^{-1}(\Theta_0\theta_0 + R_i^{-1}E^T e)$, $\Theta_n = (\Theta_0 + R_i^{-1}E^T E)$, $d_1 = \|y^* - X^*\beta\|^2$, and

$$\Psi_D(D_i) = |\Sigma_y(D_i)|^{-1/2} \exp\left[-\frac{1}{2R_i}(y_1 - X_1\beta)^T \Sigma_y^{-1}(D_i)(y_1 - X_1\beta)\right] \tag{44}$$

To sample from the conditional posterior of $D_i(\ell)$, Chib and Greenberg (1994) use a Metropolis-Hastings step. That is, in the jth iteration of the Gibbs sampler, I generate a candidate draw D_i' from the distribution $\mathcal{N}(\hat{\theta}, \Theta_n^{-1}) I_{S_D}$ and use it for the next iteration with probability $min(\Psi_D(D_i')/\Psi_D(D_i^{(j-1)}), 1)$. With probability $(1 - min(\Psi_D(D_i')/\Psi_D(D_i^{(j-1)}), 1))$, I retain the current value $D_i^{(j-1)}$.

Initialization In order to initialize the Gibbs sampling algorithm, I use the first principal component of the data to obtain an estimate for the factor. Given this estimate, I run an OLS regression on its own p_x lags to initialize $F(\ell)$. I compute the variance of the error term of this regression and use it throughout as the constant (by assumption) value of Q. For each H_i, I obtain the OLS estimate from a regression of y_{it} on the principal components factor estimate. On the residuals of this regression, I run an OLS regression on its own p_w lags to initialize the $D_i(\ell)$. Using the residuals of this regression in turn, I compute their variance to set the initial value of R_i.

The Gibbs sampling algorithm Using the initial values for the model's parameters described in the previous paragraph, I sample the factors using their conditional posterior density from above. Next, I first draw the parameters of state equation and then the parameters of the observation equation from their respective conditional posterior density as explained in this appendix. Using the parameter draws from this iteration, I repeat the algorithm and sample the factors again. In total, I run 20,000 iterations and discard the first 5,000 draws to ensure that the algorithm has converged to its ergodic distribution.

D Modeling Limited Attention

This appendix provides further details on how I model limited attention of decision-makers in firms. Following Sims (2003), I assume that limited attention is a constraint on uncertainty reduction, where uncertainty is measured by entropy. Entropy is a measure of uncertainty from information theory, defined as

$$H(\mathbf{X}) = -E\left\{\log_2\left(p\left(\mathbf{X}\right)\right)\right\},$$

where $\mathbf{X}$ is a random vector. For example, if $\mathbf{X}$ is a $T \times 1$ multivariate normal random vector with variance-covariance matrix Σ, then it has entropy

$$H(\mathbf{X}) = \frac{1}{2}\log_2\left[(2\pi e)^T \det \Sigma\right].$$

Similarly, given two $T \times 1$ multivariate normal random vectors $\mathbf{X}$ and $\mathbf{Y}$, the conditional entropy of $\mathbf{X}$ given $\mathbf{Y}$ is

$$H(\mathbf{X}|\mathbf{Y}) = \frac{1}{2}\log_2\left[(2\pi e)^T \det \Sigma_{X|Y}\right],$$

where $\Sigma_{X|Y}$ denotes the conditional variance-covariance of $\mathbf{X}$ given $\mathbf{Y}$.

Define uncertainty reduction as

$$I(\mathbf{X};\mathbf{Y}) = H(\mathbf{X}) - H(\mathbf{X}|\mathbf{Y}).$$

This measure is also called mutual information. It quantifies by how much uncertainty about $\mathbf{X}$ reduces having observed $\mathbf{Y}$. If $\{X_t\}_{t=0}^{\infty}$ and $\{Y_t\}_{t=0}^{\infty}$ are two stochastic processes, we can define the average per-period uncertainty reduction

$$\mathcal{I}(\{X_t\};\{Y_t\}) = \lim_{T\to\infty}\frac{1}{T}\left(H(X_1,\ldots,X_T) - H(X_1,\ldots,X_T|Y_1,\ldots,Y_T)\right).$$

E Derivation of Objective Function

This appendix derives the expected discounted sum of losses in profit when the actual investment decisions given less than perfect information deviate from the profit-maximizing investment decisions under perfect information given in Equation (13). The derivation closely follows Maćkowiak and Wiederholt (2015, Appendix D), which contains more details.

First, express the period profit function in log-deviations from the non-stochastic steady state, multiply by β^t, and sum over all periods from 0 to ∞. Let g denote this functional, and let $\tilde{g}$ denote the second-order Taylor expansion to g around the non-stochastic steady state.

Second, let $y_{it} = \begin{pmatrix} z_t & \varepsilon_{it} \end{pmatrix}^T$ denote the vector of shocks in period t. Conditional on production unit i's information in period -1, compute the second-order Taylor approximation to the expected discounted sum of profits around the non-stochastic steady state. This approximation gives

$$E_{i,-1}\left\{\tilde{g}\left(k_{i0}, k_{i1}, y_{i0}, k_{i2}, y_{i1}, k_{i3}, y_{i2}, \ldots\right)\right\} = E_{i,-1}\left\{\begin{array}{c} g(0,0,0,0,0,0,0,\ldots) \\ +\sum_{t=0}^{\infty}\beta^t\left(\begin{array}{c} h_k k_{it+1} + h_y^T y_{it} \\ +\frac{1}{2}k_{it+1}H_{k,-1}k_{it} + \frac{1}{2}H_{k,0}k_{it+1}^2 + \frac{1}{2}k_{it+1}H_{k,1}k_{it+2} \\ +\frac{1}{2}k_{it+1}H_{ky,1}y_{it+1} + \frac{1}{2}y_{it}^T H_{y,0}y_{it} + \frac{1}{2}y_{it}^T H_{yk,-1}k_{it} \end{array}\right) \\ +\beta^{-1}\left(h_{-1}k_{i0} + \frac{1}{2}H_{-1}k_{i0}^2 + \frac{1}{2}k_{i0}H_{k,1}k_{i1} + \frac{1}{2}k_{i0}H_{ky,1}y_{i0}\right) \end{array}\right\}, \tag{45}$$

where $E_{i,-1}$ denotes the expectation operator conditional on production unit i's information in period -1 and lower-case letters denote log-deviations from the non-stochastic steady state, for example $k_{it+1} = \ln K_{it+1} - \ln \overline{K}$. Moreover, $\beta^t h_k$ is the first derivative of g with respect to k_{it+1}, $\beta^t h_y$ is the vector of first derivatives of g with respect to y_{it}, $\beta^t H_{k,\tau}$ denotes the second derivative of g with respect to k_{it+1} and $k_{it+1+\tau}$, $\beta^t H_{y,0}$ denotes the matrix of second derivatives of g with respect to y_{it}, $\beta^t H_{ky,1}$ denotes the vector of second derivatives of g with respect to k_{it+1} and y_{it+1}, and $\beta^t H_{yk,-1}$ denotes the vector of second derivatives of g with respect to y_{it} and k_{it}. Similarly, $\beta^{-1} h_{-1}$ and $\beta^{-1} H_{-1}$ are the first and second derivative of g with respect to k_{i0}, respectively. All first and second derivatives appearing in Equation (45) are evaluated at the non-stochastic steady state. Because for all $t \geq 0$ the first derivatives of g with respect to k_{it+1} and y_{it+1} depend only on k_{it}, k_{it+1}, k_{it+2}, y_{it+1} and k_{it}, y_{it}, respectively, and because the first derivative of g with respect to k_{i0} depends only on k_{i0}, k_{i1}, y_{i0}, Equation (45) contains only certain second-order terms.

Third, define $v_t^T = \left(k_{it+1}, y_t^T, 1\right)$ and suppose there exist two constants $\delta < (1/\beta)$ and $A \in \mathbb{R}$ such that, for all m and n, for each period $t \geq 0$, and for $\tau = 0, 1$, the following regularity conditions hold:

$$E_{i,-1}\left\{k_{i0}^2\right\} < \infty \tag{46}$$

$$E_{i,-1}\left\{\left|k_{i0} v_{n,0}\right|\right\} < \infty \tag{47}$$

$$E_{i,-1}\left\{\left|v_{m,t} v_{n,t+\tau}\right|\right\} < \delta^t A \tag{48}$$

where $v_{i,t}$ denotes the ith element of v_t.[21] The conditions in Equations (47) and (48) allow to rewrite the expectation of every infinite sum appearing on the right-hand side of Equation (45) as the infinite sum of expectations and imply that each of these infinite sums of expectations converges to an element in $\mathbb{R}$. In conjunction with the fact that

[21] Maćkowiak and Wiederholt (2015, Appendix D) assume similar regularity conditions.

$H_{k,1} = \beta H_{k,-1}$ and $H_{ky,1} = \beta H^T_{yk,-1}$, one can rewrite Equation (45) as

$$\begin{aligned}
&E_{i,-1}\left\{\tilde{g}\left(k_{i0}, k_{i1}, y_{i0}, k_{i2}, y_{i1}, \ldots\right)\right\} \\
= {} & g(0,0,0,0,0,\ldots) + \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{h_k k_{it+1}\right\} + \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{h_y^T y_{it}\right\} \\
&+ \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\frac{1}{2} H_{k,0} k_{it+1}^2\right\} + \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\frac{1}{2} k_{it+1} H_{k,1} k_{it+2}\right\} \\
&+ \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{k_{it+1} H_{ky,1} y_{it+1}\right\} + \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\frac{1}{2} y_{it}^T H_{y,0} y_{it}\right\} \\
&+ \beta^{-1} E_{i,-1}\left\{h_{-1} k_{i0} + \frac{1}{2} H_{-1} k_{i0}^2 + k_{i0} H_{k,1} k_{i1} + k_{i0} H_{ky,1} y_{i0}\right\}. \qquad (49)
\end{aligned}$$

The regularity conditions in Equations (46) and (47) imply that the term in the last line on the right-hand side of Equation (49) is finite.

Fourth, define the stochastic process $\{k_{it+1}^*\}_{t=-1}^{\infty}$ for the profit-maximizing capital choice under perfect information satisfying the following properties: (i) $k_{i0}^* = k_{i0}$, (ii) in each period $t \geq 0$, k_{it+1}^* satisfies

$$E_t\left\{h_k + H_{k,-1} k_{it}^* + H_{k,0} k_{it+1}^* + H_{k,1} k_{it+2}^* + H_{ky,1} y_{it+1}\right\} = 0, \qquad (50)$$

where E_t denotes the the expectation operator conditioned on the history of the economy up to and including period t, and (iii) the vector v_t with $k_{it+1} = k_{it+1}^*$ satisfies the regularity conditions given in Equations (46)-(48).

Fifth, multiply Equation (50) by $(k_{it+1} - k_{it+1}^*)$, use the law of iterated expectations, and rearrange. This gives

$$\begin{aligned}
&E_{i,-1}\left\{\left(k_{it+1} - k_{it+1}^*\right)\left(h_k + H_{ky,1} y_{it+1}\right)\right\} \\
&= -E_{i,-1}\left\{\left(k_{it+1} - k_{it+1}^*\right)\left(H_{k,-1} k_{it}^* + H_{k,0} k_{it+1}^* + H_{k,1} k_{it+2}^*\right)\right\}, \qquad (51)
\end{aligned}$$

a useful result in the next step.

Sixth, compute the expected discounted sum of losses in profit when the actual investment decisions given less than perfect information deviate from the profit-maximizing investment decisions under perfect information, recall that $k_{i0}^* = k_{i0}$, note that $h_k = 0$, and use the result in Equation (51) to obtain

$$\begin{aligned}
& E_{i,-1}\left\{\tilde{g}\left(k_{i0}, k_{i1}, y_{i0}, k_{i2}, y_{i1}, \ldots\right)\right\} - E_{i,-1}\left\{\tilde{g}\left(k_{i0}^*, k_{i1}^*, y_{i0}, k_{i2}^*, y_{i1}, \ldots\right)\right\} \\
= & \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\frac{1}{2} H_{k,0} k_{it+1}^2 - \frac{1}{2} H_{k,0} k_{it+1}^{*2} + k_{it+1} H_{k,1} k_{it+2} - k_{it+1}^* H_{k,1} k_{it+2}^*\right\} \\
& - \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\left(k_{it+1} - k_{it+1}^*\right)\left(H_{k,-1} k_{it}^* + H_{k,0} k_{it+1}^* + H_{k,1} k_{it+2}^*\right)\right\} \\
& + \beta^{-1} E_{i,-1}\left\{k_{i0} H_{k,1}\left(k_{i1} - k_{i1}^*\right)\right\}
\end{aligned} \tag{52}$$

The conditions in Equations (46) and (48) and the definition of the process $\left\{k_{it+1}^*\right\}_{t=-1}^{\infty}$ imply that $\sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{k_{it+1} k_{it+\tau}^*\right\}$, $\tau = -1, 0, 1$, converges to an element in $\mathbb{R}$. Together with the fact that $H_{k,1} = \beta H_{k,-1}$, $k_{i0}^* = k_{i0}$ and after rearranging we have

$$\begin{aligned}
& E_{i,-1}\left\{\tilde{g}\left(k_{i0}, k_{i1}, y_{i0}, k_{i2}, y_{i1}, \ldots\right)\right\} - E_{i,-1}\left\{\tilde{g}\left(k_{i0}^*, k_{i1}^*, y_{i0}, k_{i2}^*, y_{i1}, \ldots\right)\right\} \\
= & \sum_{t=0}^{\infty} \beta^t E_{i,-1}\left\{\frac{1}{2} H_{k,0}\left(k_{it+1} - k_{it+1}^*\right)^2 + \left(k_{it+1} - k_{it+1}^*\right) H_{k,1}\left(k_{it+2} - k_{it+2}^*\right)\right\}
\end{aligned} \tag{53}$$

Seventh, compute the first and second derivatives appearing in Equations (50) and (53). These are:

$$h_k = 0 \tag{54}$$

$$H_{k,0} = \overline{K}\left[-\gamma + \beta\left(\alpha(\alpha-1)\overline{K}^{\alpha-1} - \gamma\right)\right] \tag{55}$$

$$H_{k,1} = \beta\gamma\overline{K} \tag{56}$$

$$H_{k,-1} = \gamma\overline{K} \tag{57}$$

$$H_{ky,1} = \begin{bmatrix} \beta\alpha\overline{K}^{\alpha} & \beta\alpha\overline{K}^{\alpha} \end{bmatrix} \tag{58}$$

Eighth, solve for the profit-maximizing investment decision under perfect information by substituting Equations (54)-(58) into Equation (50) and rearrange to arrive at:

$$k^*_{it+1} = \frac{\gamma k^*_{it} + \beta E_t \left\{ \gamma k^*_{it+2} + \alpha \overline{K}^{\alpha-1}(z_{t+1} + \varepsilon_{it+1}) \right\}}{\gamma + \beta\gamma - \beta\alpha(\alpha - 1)\overline{K}^{\alpha-1}} \tag{59}$$