The Secular Stagnation of Investment?*

Germán Gutiérrez† Callum Jones‡ Thomas Philippon§

June 2018

Abstract

We argue that a secular decline in competition in the goods markets explains several macroeconomic puzzles, in particular low interest rates and weak corporate investment. Corporate investment in the U.S. is lower than what one would expect based on profitability, discount rates, or the market value of corporate assets (Q-theory). This investment gap only appears in concentrating industries and is driven by large and highly profitable firms. We explore the macro-economic consequences of this phenomenon in a DSGE model with time-varying parameters and an occasionally binding zero lower bound constraint on nominal interest rates (ZLB). We propose a novel estimation strategy by embedding information from the cross section of industries directly into the Kalman filter and by mixing structural estimation and instrumental variables. We show that the trend decrease in competition can explain the joint evolution of investment, Q, and the nominal interest rate. Absent the decrease in competition, we find that the U.S. economy would have escaped the ZLB by the end of 2010 and that the nominal rate would have been above 2% since 2015.

*The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

†New York University, ggutierr@stern.nyu.edu

‡International Monetary Fund, jonescallum@gmail.com

§New York University, CEPR and NBER, tphilipp@stern.nyu.edu

1 Introduction

In December 2008, the Federal Reserve lowered the federal funds rate to a target range of zero to 25 basis points. The U.S. economy has remained stuck at or near this zero lower bound (ZLB) on nominal rate of interest rates ever since. Our goal in this paper is to shed some light on why this has happened.

We relate low interest rates to the weakness of investment. Two important stylized facts have emerged in recent years regarding the U.S. business sector. The first fact is that concentration and profitability have increased across most U.S. industries, as shown by Grullon et al. (2016). Figure 1 shows the aggregate Lerner index (operating income over sales) across all Compustat firms along with the change in weighted average 8-firm concentration ratio in manufacturing and non-manufacturing industries.[1]

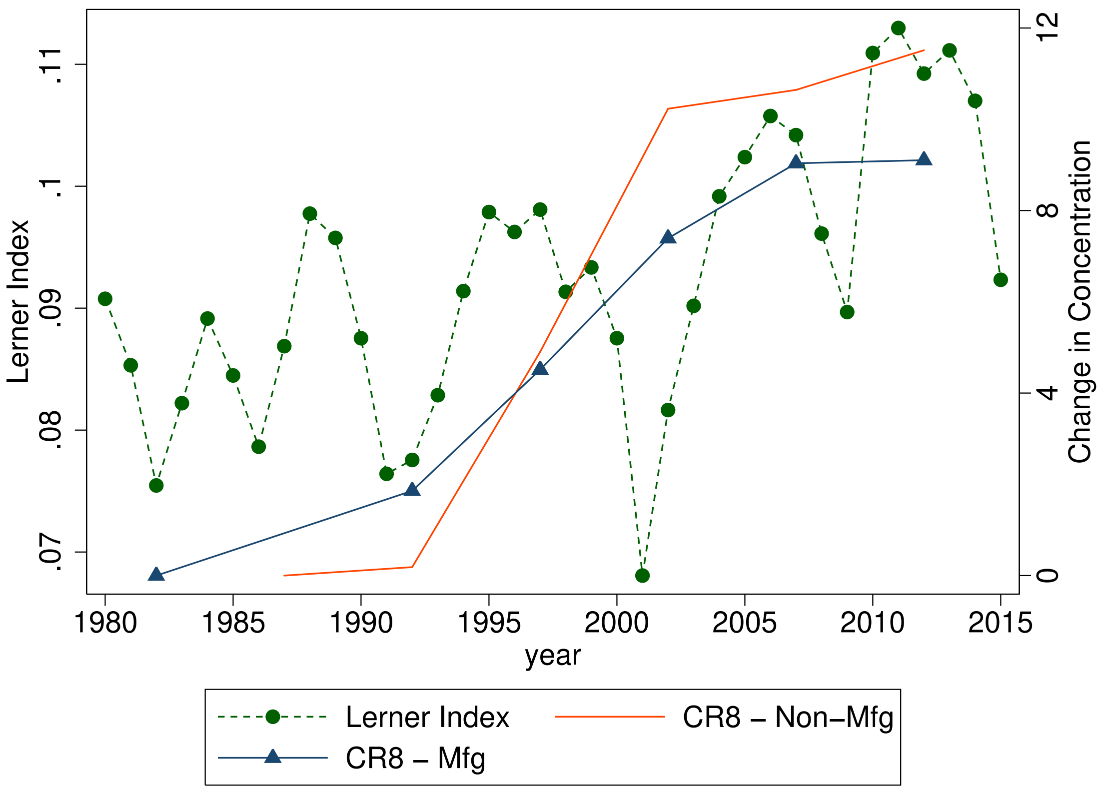
The second stylized fact is that business investment has been weak relative to measures of profitability, funding costs, and market values since the early 2000s. The top chart in Figure 2 shows the ratio of aggregate net investment to net operating surplus for the non financial business sector, from 1960 to 2015. The bottom chart shows the residuals (by year and cumulative) of a regression of net investment on (lagged) Q from 1990 to 2001. Both charts show that investment has been low relative to profits and Q since the early 2000's. By 2015, the cumulative under-investment is large, around 10% of capital.

While these two stylized facts are well established, their interpretation remains controversial. There is little agreement about the causes of these evolutions, and even less about their consequences. For instance, Furman (2015) and CEA (2016) argue that the rise in concentration suggests "economic rents and barriers to competition", while Autor et al. (2017) argue almost exactly the opposite: they think that concentration reflects "a winner take most feature" explained by the fact that "consumers have become more sensitive to price and quality due to greater product market competition." Network effects and increasing differences in the productivity of Information Technology could also increase the efficient scale of operation of the top firms, leading to higher concentration. The key point of these later explanations is that concentration reflects an efficient increase in the scale of operation. For short, we will refer to this hypothesis as the efficient scale hypothesis (henceforth EFS).

The evolution of profits and investment could also be explained by intangible capital deepening, as discussed in Alexander and Eberly (2016). More precisely, an increase in the (intangible)

[1]The appendix shows that alternate mark-up estimates, notably those based on Barkai (2017), yield similar results.

Figure 1: Concentration and Mark-ups



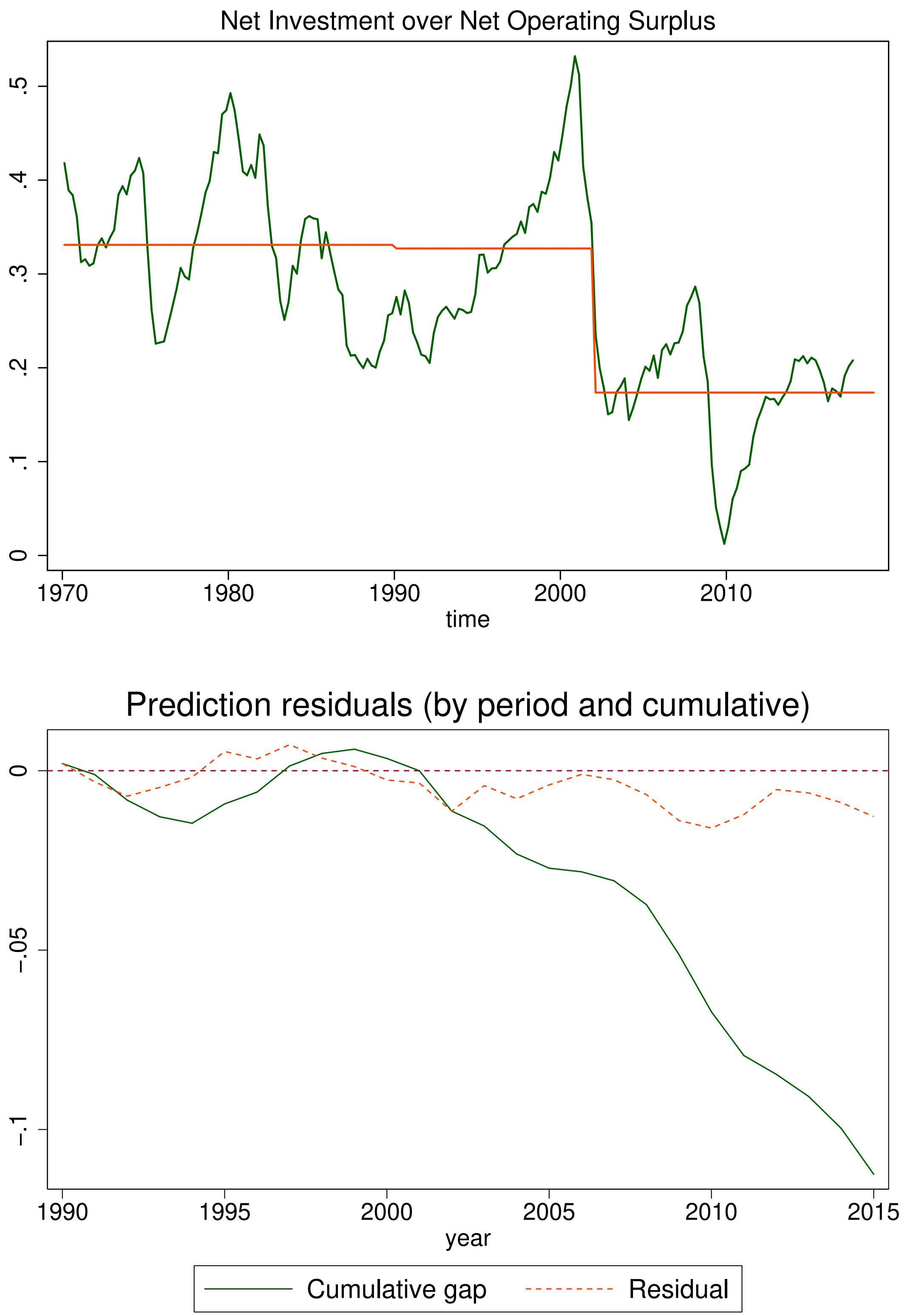
Notes: Lerner Index from Compustat, defined as operating income before depreciation minus depreciation divided by sales. 8-firm CR from Economic Census, defined as the market share (by sales) of the 8 largest firms in each industry. Data before 1992 based on SIC codes. Data after 1997 based on NAICS codes. Data for Manufacturing reported at NAICS Level 6 (SIC 4) because it is only available at that granularity in 1992. Data for Non-Manufacturing based on NAICS level 3 segments (SIC 2).

capital share together with a downward bias in our traditional measures of intangible investment could lead, even in competitive markets, to an increase in profits (competitive payments for intangible services) and a decrease in (measured) investment. We will refer to this hypothesis as the intangible deepening hypothesis (henceforth INTAN). Finally, trade and globalization can explain some of the same facts (Feenstra and Weinstein, 2017). Foreign competition can lead to an increase in measured (domestic) concentration (e.g. textile industry), and a decoupling of firm value from the localization of investment. We refer to this hypothesis as the globalization hypothesis (henceforth GLOBAL).[2]

Section 2 presents the relevant facts about the U.S. economy in recent years. Section 3 presents our benchmark model. We start from a standard DSGE model in which we allow for

[2]One could entertain other hypotheses – such as weak demand or credit constraints – but previous research has shown that they do not fit the facts. See Gutiérrez and Philippon (2017b) for detailed discussions and references.

Figure 2: Net Investment, Profits and Q-Residuals



Notes: Annual data from US Flow of Funds accounts. Net investment, net operating surplus for Non Financial Business sector; Q for Non Financial Corporate sector.

the possibility that the zero lower bound constraint on short term nominal rates binds. The most important feature of our model is a time-varying degree of competition in the goods market. The rational expectation equilibrium of the model is then represented by the time-varying function $x_t = \Psi_t(x_{t-1}, \mathbb{E}_t x_{t+1}, \epsilon_t)$, where x represents the state and ϵ the shocks.

An empirical contribution of our paper is that we construct an observable time series for the degree of competition that we feed in the model. Competition is not a residual that we obtain after fitting the macroeconomic data. It is an observable input that parameterizes the function Ψ_t above.

We solve for the path of the economy using the solution method and approach of Jones (2018). We use a Kalman filter and information about expected duration of the ZLB to back out the other shocks that drive the model (productivity, discount rate, risk premia). Our main finding is that time-varying competition has had a significant impact on macro-economic dynamics over the past 30 years. For instance, absent the decrease in competition since 2000, the nominal interest rate would have been just below 2 per cent per annum in 2015.

Literature A large and growing literature studies the consequences of a binding zero lower bound (ZLB) on the nominal rate of interest. Krugman (1998) and Eggertsson and Woodford (2003) argue that the ZLB can lead to a large drop in output. Lawrence Christiano (2011) show that the government spending multiplier can be large when the ZLB binds, suggesting a more important role for fiscal policy. Coibion et al. (2012) ask whether the risk of a binding ZLB should lead policy makers to increase the average rate of inflation. Swanson and Williams (2014) study the impact of the ZLB on long rates, that are more relevant for economic decisions.

Most studies of the liquidity trap are based on simple New-Keynesian models that abstract from capital accumulation. Fernández-Villaverde et al. (2015) study the exact properties of the New Keynesian model around the ZLB. In these models, consumption is depressed because the equilibrium interest rate is higher than the natural rate – the rate that would have cleared the asset market in the absence of price or wage rigidities. In most of the existing models, the ZLB episode is triggered by an increase in households' patience, that is, an increase in their subjective discount factor. Explicitly allowing for capital accumulation complicates matters, however, because changes in discount rates imply that consumption and investment move in opposite directions. The shock that triggers the ZLB episode is also a shock that reduces the real rate, and therefore encourages investment.

The ZLB has been proposed as an explanation for the slow recovery of most major economies following the financial crisis of 2008-2009. Summers (2013) argues that the natural rate of interest has become negative, thus creating the risk of a secular stagnation, an environment with low interest rates and output permanently below potential. Eggertsson and Mehrotra (2014) propose a model where secular stagnation can be triggered by a decrease in population growth, among other factors.

Gutiérrez and Philippon (2017b), Alexander and Eberly (2016), and Lee et al. (2016) present recent firm and industry level evidence on investment and Q. There is a growing literature studying trends on competition, concentration, and entry. Davis et al. (2006) find a secular decline in job flows. They also show that much of the rise in publicly traded firm volatility during the 1990's is a consequence of the boom in IPOs, both because young firms are more volatile, and because they challenge incumbents. Decker et al. (2015) argue that, whereas in the 1980's and 1990's declining dynamism was observed in selected sectors (notably retail), the decline was observed across all sectors in the 2000's, including the traditionally high-growth information technology sector. Furman (2015) shows that "the distribution of returns to capital has grown increasingly skewed and the high returns increasingly persistent" and argues that it "potentially reflects the rising influence of economic rents and barriers to competition."[3] CEA (2016) and Grullon et al. (2016) are the first papers to extensively document the broad increases in profits and concentration. Grullon et al. (2016) also show that firms in concentrating industries experience positive abnormal stock returns and more profitable M&A deals. Blonigen and Pierce (2016) find that M&As are associated with increases in average markups. Dottling et al. (2017) find that concentration has increased in the U.S. while it has remained stable (or decreased) in Europe. Faccio and Zingales (2017) show that competition in the mobile telecommunication industry is heavily influenced by political factors, and that, in recent years, many countries have adopted more competition-friendly policies than the US. Autor et al. (2017) study the link between concentration and the labor share. An important issue in the literature is the measurement of markups and excess profits. The macroeconomic literature focuses on the cyclical behavior of markups (Rotemberg and Woodford, 1999; Nekarda and Ramey, 2013). Over long horizons, however, it is difficult to separate excess profits from changes in the capital share. **?** estimate markups using the ratio of sales to costs of goods sold, but in the long run this ratio depends on the share of intangible expenses, and the resulting markup does not directly provide a measure of market power. Barkai (2017), on the other hand, estimates directly the required return on capital and finds a significant increase in excess profits.

[3] Furman (2015) also emphasizes emphasizes the weakness of corporate fixed investment and points out that low investment has coincided with high private returns to capital, implying an increase in the payout rate (dividends and shares buyback).

Table 1: Current Account of Non financial Sector

		Value in 2014 ($ billions)		
Name	Notation	Corporate[1]	Non corporate[2]	Business[1+2]
Gross Value Added	$P_t Y_t$	$8,641	$3,147	$11,788
Net Fixed Capital at Rep. Cost	$P_t^k K_t$	$14,857	$6,126	$20,983
Consumption of Fixed Capital	$\delta_t P_t^k K_t$	$1,286	$297	$1,583
Net Operating Surplus	$P_t Y_t - W_t N_t - T_t^y - \delta_t P_t^k K_t$	$1,614	$1,697	$3,311
Gross Fixed Capital Formation	$P_t^k I_t$	$1,610	$354	$1,964
Net Fixed Capital Formation	$P_t^k (I_t - \delta_t K_t)$	$325	$56	$381

2 Empirical Evidence

2.1 Aggregate Evidence

Table 1 summarizes some facts about the balance sheet and current account of the non financial business sector.

Figure 3 shows the net investment rate and the net operating return on capital of the non financial corporate, non financial non corporate and non financial business sector, defined as net operating surplus over the replacement cost of capital:

$$\text{Net Operating Return} = \frac{P_t Y_t - \delta_t P_t^k K_t - W_t N_t - T_t^y}{P_t^k K_t}$$

The operating return fluctuates significantly but appears to be stationary. For corporates, the yearly average from 1971 to 2014 is 10%, with a standard deviation of only one percentage point. The minimum is 8.1% and the maximum 12.6%. In 2014, the operating return was 11.3%, close to the historical maximum. A striking feature is that the net operating margin was not severely affected by the Great Recession, and has been consistently near its highest value since 2010 for both Corporates and Non corporates.

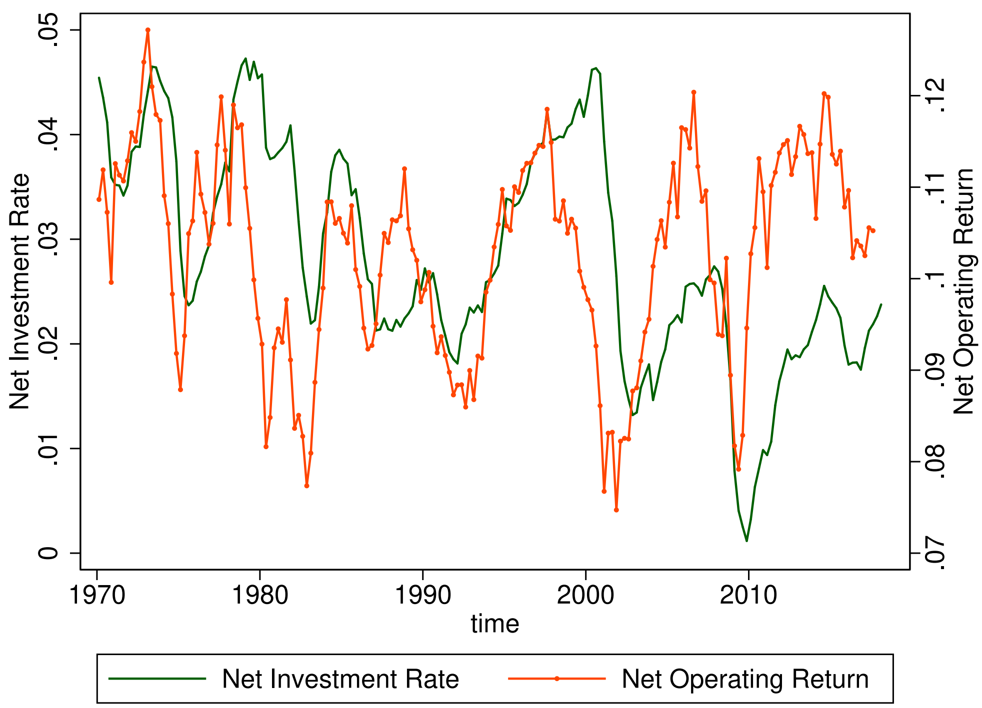
Firms are (very) profitable but they do not invest the same fraction of their operating returns as they used to. Figure 4 shows the ratio of net investment to net operating surplus for the non financial business sector:

$$^{NI}/_{OS} = \frac{P_t^k (I_t - \delta_t K_t)}{P_t Y_t - \delta_t P_t^k K_t - W_t N_t - T_t^y}$$

The average of the ratio between 1970 and 1999 is 32%. The average of the ratio from 2000 to 2015 is only 20.5%.[4] Current investment is low relative to operating margins. Similar patterns are observed when separating corporates and non corporates.

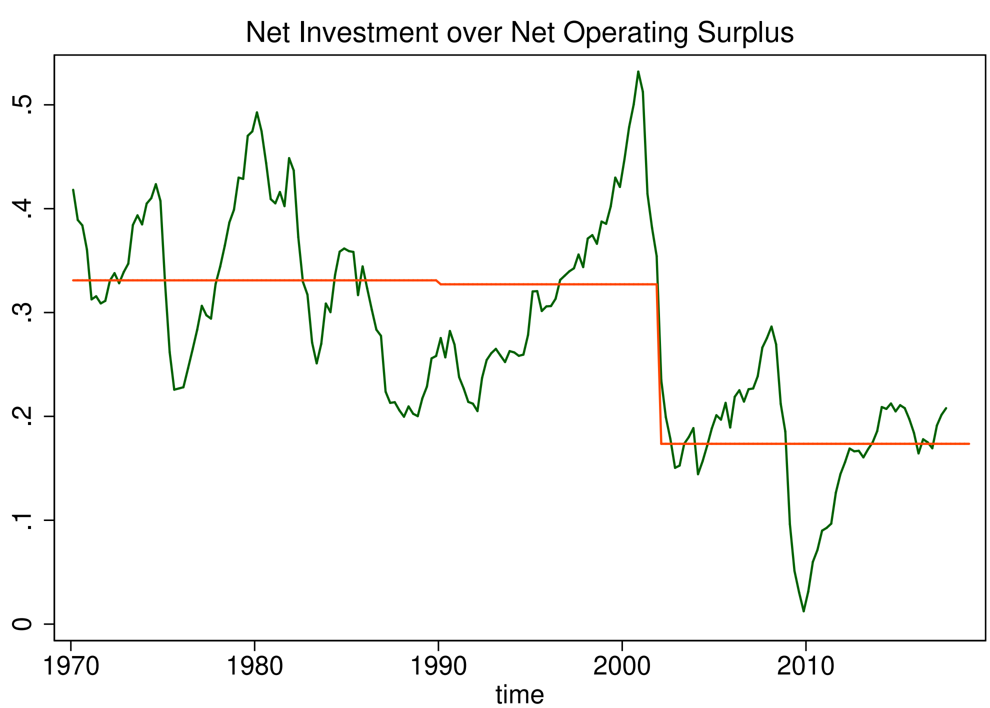
[4] Note that 2002 is used for illustration purposes only. It was chosen based on graphically, not based on a formal statistical analysis.

Figure 3: Net Investment Rate and Net Operating Return



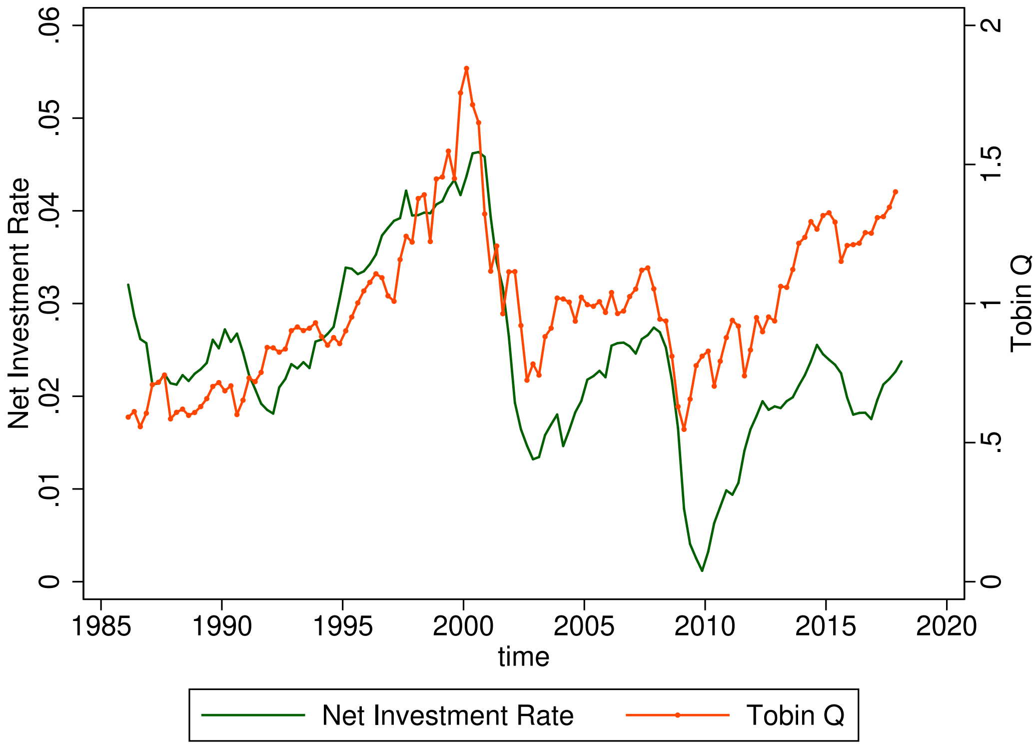
Note: Quarterly data for Non financial Businesses.

Figure 4: Net Investment Relative to Net Operating Surplus



Note: Quarterly data for Non financial Businesses.

Figure 5: Investment and Q



Note: Quarterly data. Q for Non Financial Corporate sector (data for Non Corporate sector not available)

Finally, investment is low relative to Q. The issue with Q is that it is not stationary and shows a significant change in its mean between the 1970s and 1980s. Figure 5 shows investment and Q starting in 1985. Q in 2015 is about the same as it was in 1998, yet the net investment rate is only barely more than 2% against almost 4% in 1998.

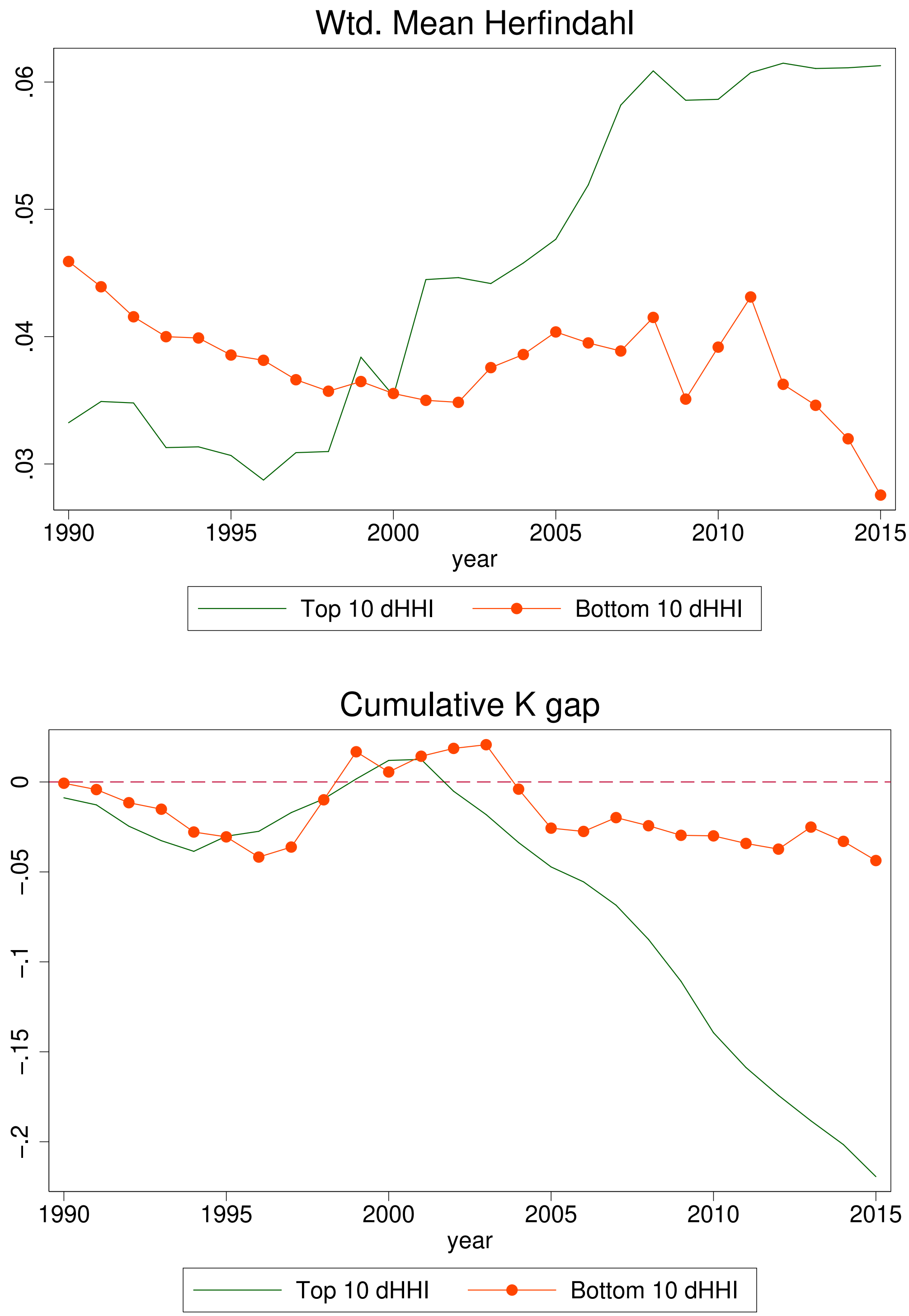
2.2 Firm and Industry Evidence

This section shows why it is critical to understand the dynamics of concentrating industries, and within industries, of the leading firms.

Fact 1: The Investment Gap Comes from Concentrating Industries. Figure 6 shows that the capital gap is coming from concentrating industries.[5] The solid (dotted) line plots the

[5] We define concentrating industries based on the relative change in import adjusted Herfindahls from 2000 to 2015. The top 10 concentrating industries include Arts, Health other, Inf. motion, Inf. publish and software, Inf Telecom, Transp pipeline, Transp truck, Min exOil, Retail trade, Transp_air. We exclude Agriculture because Compustat provides limited coverage for this industry.

Figure 6: Cumulative Capital Gap for Concentrating and Non-Concentrating Industries



Notes: Annual data. Left plot shows the weighted average import adjusted Herfindahl for the 10 industries with the largest and smallest relative change in import-adjusted Herfindahl. Right plot shows the cumulative implied capital gap (as percent of capital stock) for the corresponding industries. See text for details.

implied capital gap relative to Q for the top (bottom) 10 concentrating industries. For each group, the capital gap is calculated based on the cumulative residuals of separate industry-level regressions of net industry investment from the BEA on our measure of (lagged) industry Q from Compustat.[6] The Herfindahl index for the bottom 10 turns out to be rather stable over time, and investment remains largely in line with Q for this group.

Fact 2: Industry Leaders Account for the Increased Profit Margins and for the Investment Gap. In Table 2 (see also Appendix Figure **??**), we define leaders by constant shares of market value to ensure comparability over time.[7] Capital K includes intangible capital as estimated by Peters and Taylor (2016). Table 2 shows that the leaders' share of investment and capital has decreased, while their profit margins have increased.

Table 2 suggests that leaders are responsible for most of the decline in investment relative to profits. To quantify the implied capital gap, Figure 7 plots the percentage increase in the capital stock of the U.S. non-financial private sector assuming that Compustat leaders continued to invest 35% of CAPX plus R&D from 2000 onward, while the remaining groups invested as observed. The capital stock would be ~3.5% higher under the counter-factual. This is a large increase considering that our Compustat sample accounts for about half of investment (see the Appendix for details) and that the average annual net investment rate for the U.S. Non Financial Business sector has been less than 2% since 2002.

3 Benchmark Model

We use a standard DSGE model with capital accumulation, nominal rigidities, and time varying competition in the goods markets. For simplicity, we separate firms into capital producers – who lend their capital stock – and good producers – who hire capital and labor to produce goods and services.

[6]To be specific, each line is computed as follows: we first compute the residuals from separate industry-level regressions of net investment on (lagged) mean industry Q, from 1990 to 2001. Then, we average yearly residuals across the industries with the ten largest and ten smallest relative changes in import-adjusted Herfindahls from 2000 to 2015. Last, we compute the cumulative capital gap by adding residuals from 1990 to 2015, accounting for depreciation.

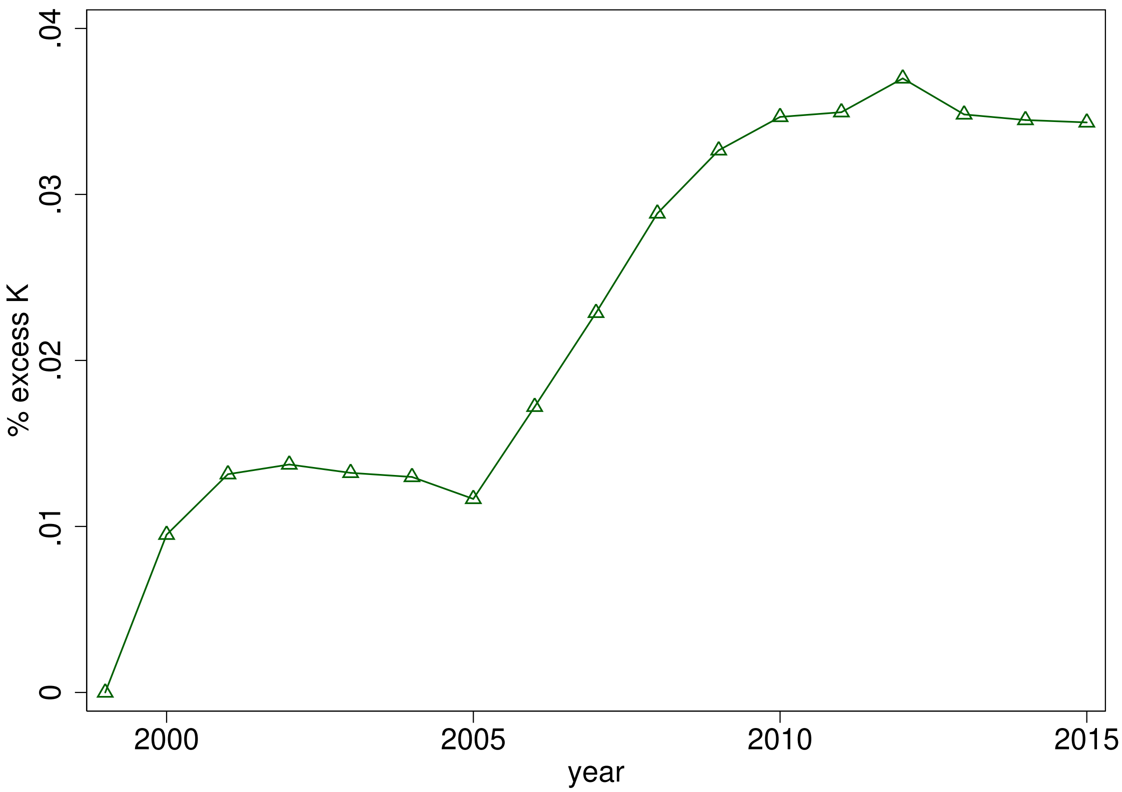
[7]OIBDP shares are stable which is consistent with stable shares of market value and stable relative discount factors. Because firms are discrete, the actual share of market value in each grouping varies from year to year. To improve comparability, we scale measured shares as if they each contained 33% of market value.

Table 2: Investment, Capital and Profits by Leaders and Laggards

Table shows the average value of a broad set of investment, capital and profitability measures by time period and market value. Leaders (laggards) include the firms with the highest (lowest) MV that combined account for 33% of MV within each industry and year. Annual data from Compustat. Lerner Index defined as $(OIBDP - DP)/SALE$.

	1980-1995			**1996-2015**			**Difference**		
	Leaders 0-33 pct	**Mid** 33-66 pct	**Laggards** 66-100 pct	**Leaders** 0-33 pct	**Mid** 33-66 pct	**Laggards** 66-100 pct	**Leaders** 0-33 pct	**Mid** 33-66 pct	**Laggards** 66-100 pct
Share of OIBDP	0.33	0.34	0.33	0.33	0.33	0.33	**0.00**	0.00	0.00
Share of CAPX + R&D	0.35	0.32	0.34	0.29	0.32	0.39	**-0.06**	0.01	0.05
Share of PP&E	0.32	0.33	0.34	0.30	0.31	0.39	**-0.02**	-0.02	0.05
Share of K	0.31	0.33	0.36	0.27	0.33	0.39	**-0.03**	0.00	0.04
(CAPX+R&D)/OIBDP	0.71	0.67	0.71	0.52	0.57	0.70	**-0.20**	-0.10	-0.01
Lerner Index	0.10	0.09	0.07	0.13	0.11	0.07	**0.02**	0.02	0.001

Figure 7: Implied Gap in K due to Leader Under-Investment



Notes: Annual data. Figure shows the cumulative implied excess capital (as percent of total U.S. capital stock for the industries in our sample) assuming Compustat leaders continue to account for 35% of CAPX and R&D investment from 2000 onward. Non-leaders assumed to maintain their observed invest levels. Excess investment assumed to depreciate at the US-wide depreciation rate. US-wide capital and depreciation data from BEA.

3.1 Capital Producer's Problem

Consider a firm that accumulates capital K to maximize its market value, taking as given the economy's pricing kernel Λ. Management chooses employment and investment to maximize firm value. Let V_t denote the cum-dividend value (i.e., at the beginning of time t, before dividends are paid):

$$V_t = \sum_{j=0}^{\infty} \Lambda_{t,t+j} Div_{t+j} \tag{1}$$

where Div_t are the distributions to the firm's owner. Capital accumulates as

$$K_{t+1} = (1 - \delta_t) K_t + I_t \tag{2}$$

Let R_k be the real rental rate, I_t gross investment, and P_t^k be the (real) price of investment goods. Investment is subject to convex adjustment costs a la Lucas and Prescott (1971) and we ignore taxes so

$$Div_t = R_{k,t} K_t - P_{k,t} I_t - \frac{\varphi_k}{2} P_{k,t} K_t \left(\frac{I_t}{K_t} - \delta_t \right)^2 . \tag{3}$$

where the depreciation rate δ_t can be time varying (to match the data). The firm's problem is to maximize (1) subject to (2) and (3). We can write this as a dynamic programming problem

$$V_t (K_t) = \max_{I_t} Div_t + \mathbb{E}_t [\Lambda_{t+1} V_{t+1} (K_{t+1})]$$

Given our homogeneity assumptions, it is easy to see that the value function is homogeneous in K. We can then define

$$\mathcal{V}_t \equiv \frac{V_t}{K_t}$$

and net investment

$$x_t \equiv \frac{I_t}{K_t} - \delta = \frac{K_{t+1} - K_t}{K_t}$$

Then we have

$$\mathcal{V}_t = \max_x R_{k,t} - P_{k,t} (x_t + \delta_t) - \frac{\varphi_k}{2} P_{k,t} x^2 + (1 + x_t) \mathbb{E}_t [\Lambda_{t+1} \mathcal{V}_{t+1}]$$

The first order condition for the net investment rate is

$$P_{k,t} (1 + \varphi_k x_t) = \mathbb{E}_t [\Lambda_{t+1} \mathcal{V}_{t+1}]$$

which we can write as a q-investment equation

$$x_t = \frac{1}{\varphi_k} \left(Q_t^k - 1 \right)$$

where

$$Q_t^k \equiv \frac{\mathbb{E}_t\left[\Lambda_{t+1}\mathcal{V}_{t+1}\right]}{P_t^k} = \frac{\mathbb{E}_t\left[\Lambda_{t+1}V_{t+1}\right]}{P_t^k K_{t+1}}$$

is Tobin's Q, i.e. the market value of the firm divided by the replacement cost of capital, all measured at the end of the period. We index it by k to distinguish it from the aggregate, measured Q which includes also the rents of the final producers. Tobin's Q satisfies the recursive equation

$$Q_t^k = \mathbb{E}_t\left[\frac{\Lambda_{t+1}}{P_t^k}\left(R_{k,t+1} + P_{k,t+1}\left((1+x_{t+1})\,Q_{t+1}^k - x_{t+1} - \delta_{t+1} - \frac{\varphi_k}{2}x_{t+1}^2\right)\right)\right]$$

which, given the FOC, can be written as

$$Q_t^k = \mathbb{E}_t\left[\frac{\Lambda_{t+1}}{P_t^k}\left(R_{k,t+1} + P_{t+1}^k\left(Q_{t+1}^k - \delta_{t+1} + \frac{1}{2\varphi_k}\left(Q_{t+1}^k - 1\right)^2\right)\right)\right]$$

In the logic of the theory, Q_t is the discounted value of operating returns $R_{k,t+1}$, plus future Q net of depreciation, plus the option value of investing more when Q is high, and less when Q is low.

3.2 Households

We know introduce the familiar elements to close the model. We assume a balanced growth path with deterministic labor augmenting technological progress at rate $\mathsf{\bar{g}}$. Households maximize lifetime utility

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{C_t^{1-\gamma}}{1-\gamma} - (1+\mathsf{\bar{g}})^{(1-\gamma)t}\,\frac{N_t^{1+\varphi}}{1+\varphi}\right)\right],$$

subject to the budget constraint

$$S_t + P_t C_t \leq \tilde{R}_t S_{t-1} + W_t N_t,$$

where W_t is the nominal wage and $\tilde{R}_t$ is the (random) nominal gross return on savings from time $t-1$ to time t. The trend growth term $(1+\mathsf{\bar{g}})^{(1-\gamma)t}$ is simply there to ensure balanced growth. The household's real pricing kernel is

$$\Lambda_{t,t+j} = \beta^j\left(\frac{C_t}{C_{t+j}}\right)^{\gamma}$$

By definition of the pricing kernel, nominal asset returns must satisfy

$$\mathbb{E}_t\left[\Lambda_{t+1}\frac{P_t}{P_{t+1}}\tilde{R}_{t+1}\right] = 1$$

Wage setting Wage setting takes place as in the standard NK model. The wage reset at time t, W_t^*, solves

$$\mathbb{E}_t \sum_{k=0}^{\infty} (\beta\vartheta_w)^k N_{l,t+k} C_{t+k}^{-\gamma} \left(\frac{1-\epsilon_w}{P_{t+k}} + \epsilon_w \frac{\mathtt{MRS}_{t+k}}{W_t^*} \right)$$

where we define the marginal rate of substitution as

$$\mathtt{MRS}_{l,t+k} \equiv N_{l,t+k}^{\varphi} C_{t+k}^{\gamma}.$$

3.3 Price Setting

Firms have a Cobb-Douglass production function with stationary TFP shocks A_t and labor augmenting technology

$$Y_t = A_t K_t^{\alpha} \left((1+\bar{\mathtt{g}})^t N_t \right)^{1-\alpha} - (1+\bar{\mathtt{g}})^t \Phi$$

where Φ is a fixed cost of production, which ensures free entry despite monopoly rents. Firms take the wage and the rental rate as given when they hire capital and labor. The average cost of production is given by

$$\begin{aligned} &\min {}^{W}/_{P} N + R_k K \\ &s.t. \\ &Y = A K^{\alpha} N^{1-\alpha} \end{aligned}$$

The Cobb-Douglass function, like any CRS function, leads to a constant marginal cost. Taking into account the fixed cost, we get that the average cost is $\mathtt{MC}_t Y_t + (1+\bar{\mathtt{g}})^t \Phi$, where the real marginal cost is

$$\mathtt{MC}_t = \frac{1}{A_t} \left(\frac{R_{k,t}}{\alpha} \right)^{\alpha} \left(\frac{\frac{W_t}{(1+\bar{\mathtt{g}})^t P_t}}{1-\alpha} \right)^{1-\alpha}$$

Cost minimization implies that all firms choose the same (optimal) capital labor ratio

$$\frac{\alpha}{1-\alpha} \frac{N_t}{K_t} = \frac{R_{k,t}}{{}^{W_t}/_{P_t}}$$

Firms set prices a la Calvo (with indexation on average inflation). The main departure from the standard model is that competition in the goods market varies over time. In the standard model, competition is characterized by the elasticity of substitution between goods, ϵ. In our baseline model, we simply assume that this elasticity varies over time. Then the price reset at time t, $P_{i,t}^*$, solves

$$\mathbb{E}_t \left[\sum_{k=0}^{\infty} \vartheta^k \Lambda_{t,t+k} Y_{i,t+k} \left(1 - \varepsilon_{t+k} + \varepsilon_{t+k} \frac{P_{t+k}}{P_{i,t}^*} \mathtt{MC}_{t+k} \right) \right] = 0$$

We consider different models of imperfect competition in extensions of the basic model.

4 Equilibrium

4.1 Detrended Model

The model defined above has a trend. We are going to write the equilibrium conditions of the detrended model. To avoid heavy notations, I do not change the symbols, since it is obvious which variables have trends

$$K_t := \frac{K_t}{(1+\mathtt{\bar{g}})^t}$$

and similarly for the other trending variables: $Y, I, \frac{W}{P}, \mathtt{MRS}$. The detrended model is therefore

$$\begin{aligned}
Y_t &= A_t K_t^{\alpha} N_t^{1-\alpha} - \Phi_t \\
Y_t &= C_t + P_{k,t} I_t + \frac{\varphi_k}{2} P_{k,t} K_t \left(\frac{I_t}{K_t} - \delta_t\right)^2 \\
(1+\mathtt{\bar{g}}) K_{t+1} &= (1-\delta_t) K_t + I_t \\
\frac{N_t}{1-\alpha} \frac{W_t}{P_t} &= R_{k,t} \frac{K_t}{\alpha} \\
\mathtt{MC}_t &= \frac{1}{A_t} \left(\frac{R_{k,t}}{\alpha}\right)^{\alpha} \left(\frac{W_t/P_t}{1-\alpha}\right)^{1-\alpha} \\
\mathtt{MRS}_t &= N_t^{\varphi} C_t^{\gamma} \\
\Lambda_{t+1} &= \beta (1+\mathtt{\bar{g}})^{-\gamma} \left(\frac{C_t}{C_{t+1}}\right)^{\gamma} \\
\frac{I_t}{K_t} - \delta_t &= \frac{1}{\varphi_k} (Q_t - 1) \\
Q_t^k &= \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{P_{k,t}} \left(R_{k,t+1} + P_{k,t+1} \left(Q_{t+1}^k - \delta_{t+1} + \frac{1}{2\varphi_k} \left(Q_{t+1}^k - 1\right)^2\right)\right)\right]
\end{aligned}$$

Finally, the measured value of Q includes the rents of the final producers

$$Q = Q_t^k + \frac{\mathbb{E}_t \left[\Lambda_{t+1} V_{t+1}^{\epsilon}\right]}{P_t^k K_{t+1}}$$

where the value of the final producers is

$$V_t^{\epsilon} = P_t Y_t (1 - \mathtt{MC}_t) - \Phi_t + \mathbb{E}_t \left[\Lambda_{t+1} V_{t+1}^{\epsilon}\right]$$

This theoretical Q is the one that we can compare to Tobin's Q in the data.

Finally, we need to specify a policy rule for the central bank, taking into account the zero lower bound on nominal interest rates. We assume that monetary policy follows a Taylor rule for the nominal interest rate

$$i_t^* = -\log(\beta) + \phi_p \pi_t^p + \phi_w \pi_t^w + \phi_y (n - \bar{n})$$

but the actual short rate is constrained by the zero lower bound

$$i_t = \max(0; i_t^*)$$

We discuss the issue of forward guidance in the estimation section.

4.2 Shocks

We introduce the following shocks to the model (in logs):

- Productivity shock:

$$a_t = \rho_a a_{t-1} + \epsilon_{a,t}$$

- Discount rate shock to the pricing kernel

$$\begin{aligned} \lambda_{t+1} &= \log\beta - \gamma\left(c_{t+1} - c_t\right) + \zeta_t^d \\ \zeta_t^d &= \rho_d \zeta_{t-1}^d + \epsilon_t^d \end{aligned}$$

- A shock to the valuation of corporate assets

$$\begin{aligned} q_t^k &= \mathbb{E}_t\left[\lambda_{t+1} + \log\left(r_{t+1}^k + q_{t+1} + 1 - \delta + \frac{1}{2\gamma} q_{t+1}^2\right)\right] + \zeta_t^q \\ q_t^\epsilon &= \mathbb{E}_t\left[\lambda_{t+1} + v_{t+1}^\epsilon - k_{t+1}\right] + \zeta_t^q \\ \zeta_t^q &= \rho_d \zeta_{t-1}^q + \epsilon_t^q \end{aligned}$$

The discount rate shock will help us account for the sharp drop in risk free rates during the Great Recession, as in the standard NK model. The valuation shock is a risk premium shock that applies to corporate (risky) assets. It is important to account for time varying-risk aversion and expected returns.

The novel part of our model is that competition varies over time. In our benchmark model we capture this idea with a time varying elasticity ε_t. We assume that it follows a random walk, so at any point in time, the agents in our model anticipate that competition will (on average) remain at its current level.

- Time-varying elasticity of substitution between goods

$$\varepsilon_t = \varepsilon_{t-1} + \epsilon_t^\varepsilon$$

A crucial point of our analysis is that ε is not a free series of shocks. We *measure* it in the data, as explained in the next section. We solve for the path of the economy using the solution method and approach of Jones (2018), as explained in the Appendix.

5 Industry Model and Calibration

We want to use the evidence in Gutiérrez and Philippon (2017b) to calibrate and estimate our model. Their evidence is cross-sectional, based on heterogeneity across industries (and firms), so we need to extend the model to obtain a realistic mapping.

5.1 Theory

In the standard model C is an index of goods

$$C_t \equiv \left(\int_0^1 C_{j,t}^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}, \tag{4}$$

where ϵ is the elasticity of substitution between goods. Utility maximization implies that the relative demand of any two goods satisfies $\frac{C_{i,t}}{C_{j,t}} = \left(\frac{P_{i,t}}{P_{j,t}}\right)^{-\epsilon}$. This then implies the existence of a price index, defined by

$$P_t \equiv \left(\int_0^1 P_{j,t}^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}, \tag{5}$$

such that consumption expenditures are $P_t C_t = \int_0^1 P_{j,t} C_{j,t} dj$, and the demand curves are simply $C_{j,t} = \left(\frac{P_{j,t}}{P_t}\right)^{-\epsilon} C_t$. Now we want to think of $j \in [0,1]$ as industries, and each industry is populated by firms $i \in [0,1]$ (so technically a firm is point $(i,j) \in [0,1]^2$):

$$C_{j,t} = \left(\int_0^1 C_{i,j,t}^{\frac{\epsilon_j-1}{\epsilon_j}} di \right)^{\frac{\epsilon_j}{\epsilon_j-1}}$$

Firm i in industry j takes $Y_{j,t} = C_{j,t}$ as given and sets its price to maximize

$$\mathbb{E}_t \left[\sum_{k=0}^{\infty} \vartheta^k \Lambda^{\$}_{t,t+k} \left(P^*_{i,j,t} Y_{i,j,t+k} - W_{t+k} N_{i,j,t+k} - R^k_{t+k} K_{i,j,t+k} \right) \right]$$

subject to the demand curve $Y_{i,j,t+k} = \left(\frac{P^*_{i,t}}{P_{t+k}}\right)^{-\epsilon} Y_{j,t+k}$. Since all the firms face the same factor prices, they will have the same marginal cost

$$\mathtt{MC}_t = \frac{1}{A_t} \left(\frac{R_{k,t}}{\alpha} \right)^{\alpha} \left(\frac{\frac{W_t}{(1+\bar{\mathrm{g}})^t P_t}}{1-\alpha} \right)^{1-\alpha}$$

and they will choose the same (optimal) capital labor ratio

$$\frac{\alpha}{1-\alpha} \frac{N_t}{K_t} = \frac{R_{k,t}}{W_t/P_t}$$

Firms set prices a la Calvo with indexation on average inflation, so the price reset at time t, $P^*_{i,t}$, solves

$$\mathbb{E}_t\left[\sum_{k=0}^{\infty}\vartheta^k\Lambda_{t,t+k}Y_{i,j,t+k}\left(1-\epsilon_j+\epsilon_j\frac{P_{j,t+k}}{P^*_{i,j,t}}\frac{P_{t+k}}{P_{j,t+k}}\mathtt{MC}_{t+k}\right)\right]=0$$

In steady state, we have $\frac{P_{j,t}}{P^*_{i,j,t}}=1$ and

$$\frac{P_j}{\bar{P}}=\mu_j\mathtt{MC}$$

where$\mu_j\equiv\frac{\epsilon_j}{\epsilon_j-1}$. Therefore since$\bar{P}\equiv\left(\int_0^1 P_j^{1-\epsilon}dj\right)^{\frac{1}{1-\epsilon}}$ we have

$$\mathtt{MC}\equiv\left(\int_0^1\left(\mu_j\right)^{1-\epsilon}dj\right)^{\frac{-1}{1-\epsilon}}$$

and

$$C_j=\left(\mu_j\mathtt{MC}\right)^{-\epsilon}C$$

Since $C_j=Y_j=AK_j^{\alpha}N_j^{1-\alpha}$ and all the firms and industries use the same factor intensities we can write

$$Y_j=AK_j\left(\frac{N}{K}\right)^{1-\alpha}$$

and we get in the cross-section

$$\log K_j=cte-\epsilon\log\mu_j$$

5.2 Data

Using the data of Gutiérrez and Philippon (2017b), we measure the concentration ratio, $\chi_{j,t}$ as the share of sales by the top 8 firms in the industry (concentration ratio). In a panel regression across industries, including time fixed-effects and industry fixed-effects, we find

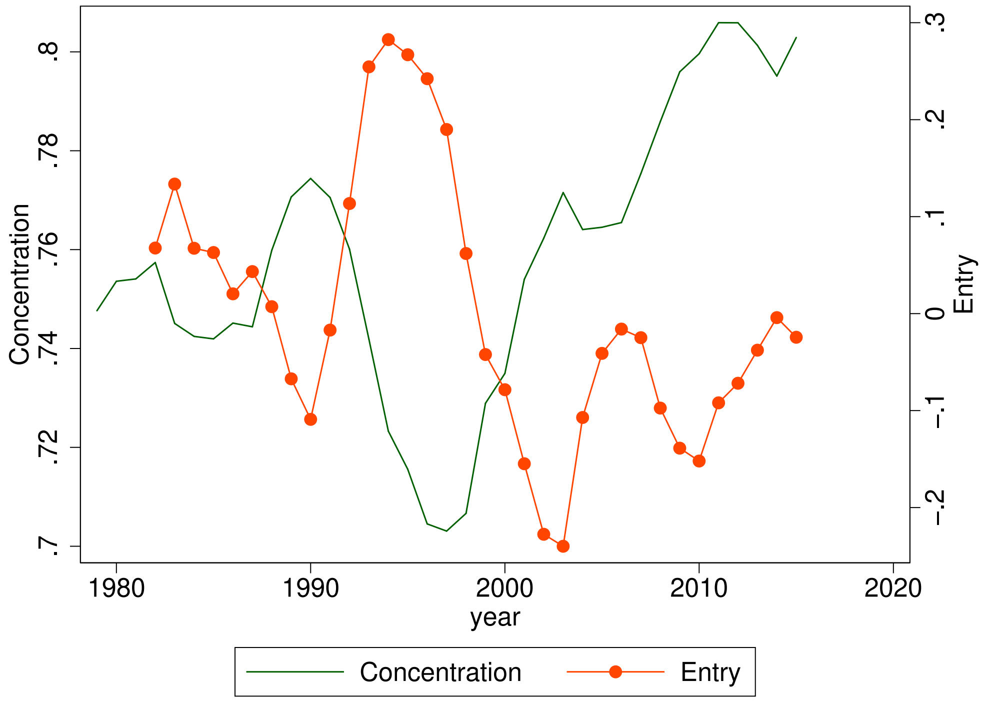
$$\log K_{j,t}=-1.3\chi_{j,t}+...$$

If the relative concentration ratio of an industry increases by 1%, its relative capital stock decreases by 1.3%. This result holds for various types of investment goods, for various controls, and also when instrumenting for the degree of competition.

To match the model and the data, we need to specify the elasticity of substitution between industries, ϵ. Following the trade literature, we take as a benchmark a value of 1 for the elasticity of substitution between broad classes of goods. Hence we have $\log\mu_{j,t}\approx 1.3\chi_{j,t}$. In the aggregate, we can calibrate the evolution of competition as

$$\log\bar{\mu}_t=\log\frac{\epsilon_t}{\epsilon_t-1}\approx 1.3\bar{\chi}_t$$

Figure 8: Entry and Concentration



Note: Annual data for Non financial Businesses (Corporate and Non corporate).

Figure 8 shows the series for the average concentration across industries in our sample. Concentration decreased in the 1990s, and increased in the 2000s. We also include a measure of entry, the 3-year log-change in the number of firms.

6 Simulation Result

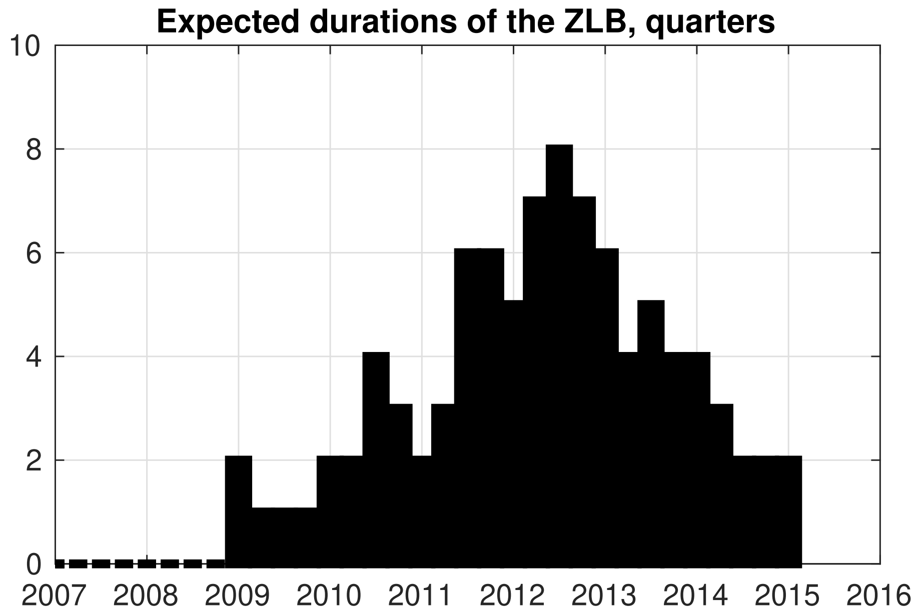
6.1 Estimation of the Model

The parameters of the model are calibrated in the standard way. We perform a simulation over the period 1986:1 to 2015:1. We use as data consumption, the net investment rate, the nominal interest rate, and the expected duration of the ZLB obtained from Federal Funds futures and Morgan Stanley. The persistence and size of the shock processes are estimated using maximum likelihood with data from 1986Q1 to 2015Q1. The remaining parameters are calibrated to standard values. Our data includes:

$$Data = \left(\log\left(C_t\right); \frac{I_t}{K_t} - \delta_t; \log\left(1 + r_t^{3m}\right); T_t\right)_{t=[1986:1;2015:1]}$$

where C_t is real consumption per capita, r_t^{3m} is the 3-month Treasury Bill rate, T_t is the expected duration of the ZLB, and the other variables are as defined earlier. We use the Kalman filter

Figure 9: Expected Duration of ZLB



to recover the three unobserved shocks introduced above:

$$Shocks = \left(a_t, \zeta_t^d, \zeta_t^q\right)$$

A critical issue is the presence of the ZLB. It implies that the short rate becomes uninformative when it reaches 0. For any time t where $i_t = 0$, what matters for agents in the model is the expected duration of the ZLB episode, which we call T_t. So what enters the Kalman filter in period t is either i or T, whichever is strictly positive.

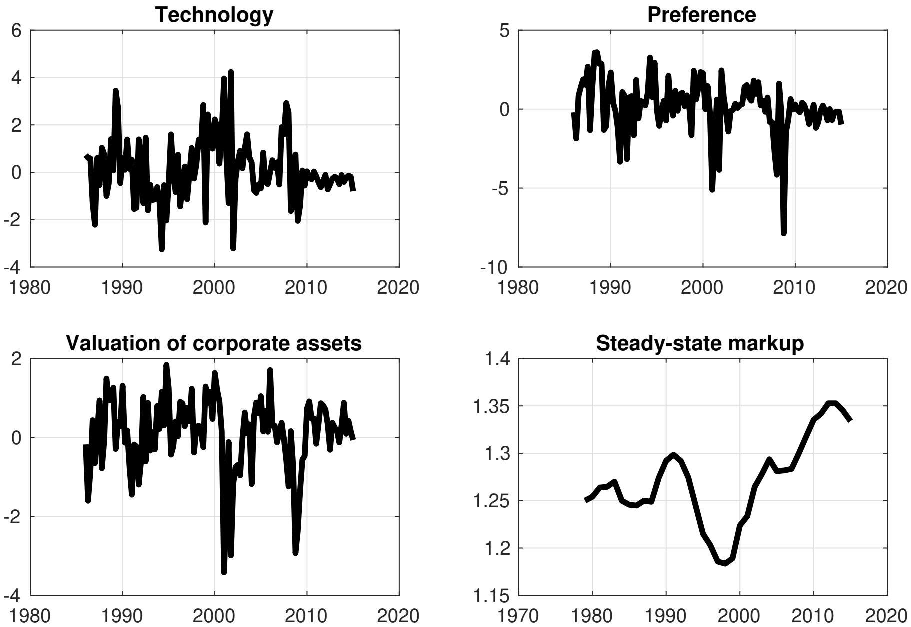
There are several ways to construct T_t. We use a measure constructed by Morgan Stanley from the Fed Funds Futures contracts. Figure 9 presents our series for T_t, based on i^*. In 2013, agents in the model anticipate the ZLB to last about two years. By 2015, the agents anticipate a lift off in the near future.

Once we have chosen a particular series for the ZLB durations T_t, we can recover the three shocks following the methodology described in Jones (2018). Figure 10 presents the shocks. There is a large innovation to the discount rate around the time of the Great Recession. Our main interest is in the time varying markup. It increases substantially from a markup of 20% to one of 35% from 2000 to 2011.

6.2 Counter-Factual

We now present our results from the model. The observed, filtered, and counterfactual levels of the real variables are presented in Figure 11. Absent the decline in the elasticity of substitution,

Figure 10: Model-Implied Shocks



Notes: Quarterly data, shocks in units of standard deviation.

the levels of log output, log consumption, and log capital would have been significantly higher from 2000 onwards.

The counterfactual paths for inflation and the nominal interest rate are presented in Figure 12. The observed inflation rate (CPI inflation) is plotted alongside the prediction of inflation from the full model, and the counterfactual path of inflation when the shocks to the elasticity of substitution are turned off. Absent the changes in the elasticity, the nominal interest rate would increased to just less than 2 per cent per annum in 2015. This is primarily a result of responding to the higher counterfactual level of output when goods markets are more competitive without the trend increase in the steady-state markup. That is, the increase in aggregate demand would have had a large impact on the equilibrium rate.

7 Firm Entry and Expectation Shocks

In this section, we use a richer model of firm entry across industries and describe our approach to use industry-level concentration and profitability data directly, together with aggregate data, to understand the aggregate consequences of a decline in competition. As discussed in Gutiérrez and Philippon (2017a), an identification challenge arises in the use of industry-level concentration and investment data, where the presence of anticipated demand shocks can give rise to a correlation between concentration and investment which is unrelated to changes in the compet-

Figure 11: Counter-Factual Series

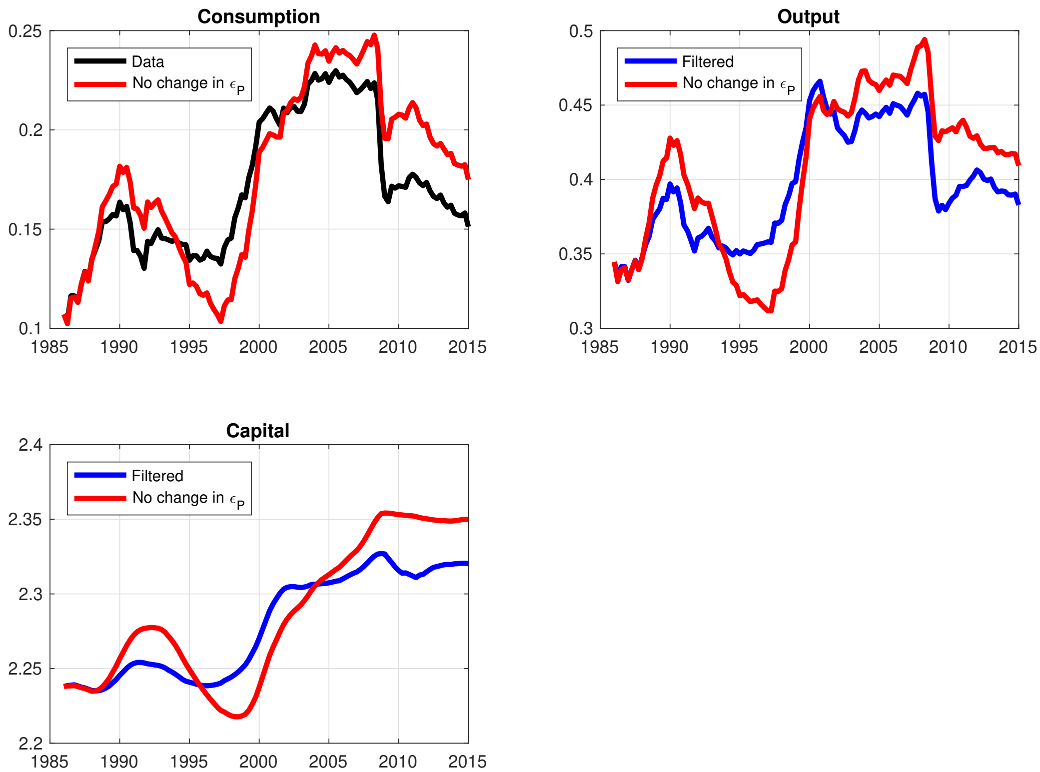
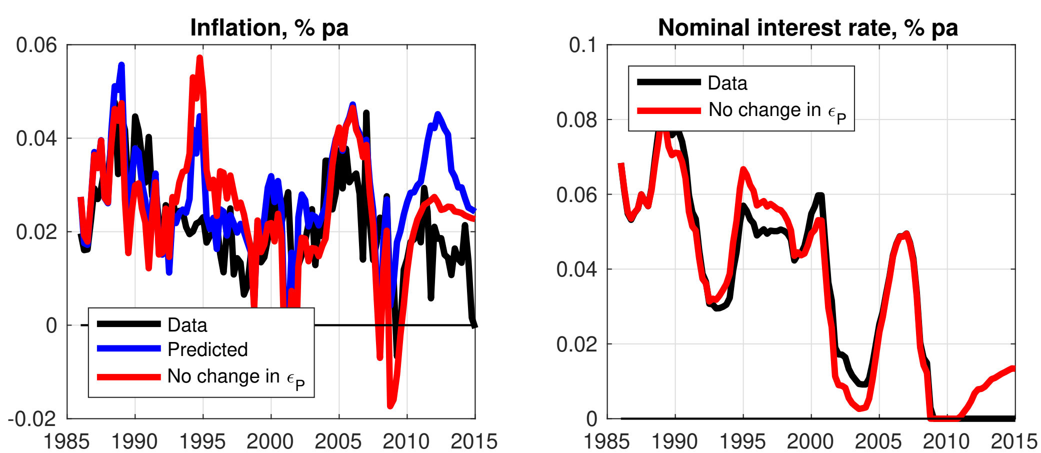


Figure 12: Counter-Factual Inflation and Short Rate



itive environment. We address this identification challenge by estimating anticipated changes in the growth rate of industry-specific TFP.

7.1 Model of Firm Entry

To our industry-level model we add firm entry in the goods-producing sector.

Goods-producing firms are organized by industries s that take the aggregate price index and factor prices as given. The final good is a composite of industry-level outputs aggregated by a perfectly competitive final goods firm:

$$Y_t = \left(\int_0^1 Y_t(s)^{\frac{\sigma-1}{\sigma}} \mathrm{d}s \right)^{\frac{\sigma}{\sigma-1}}$$

where σ is the elasticity of demand across industries. The final good is used for consumption and for investment.

Each industry is populated by firms i who have pricing and production decisions. They face the industry demand curve:

$$Y_t(s) = \left(\frac{P_t(s)}{P_t} \right)^{-\sigma} Y_t.$$

where $P_t(s)$ is the industry level price index, σ is the elasticity of substitution between industry goods, and P_t is the price index, defined as

$$P_t = \left(\int_0^1 P_t^{1-\sigma}(s) \mathrm{d}s \right)^{\frac{1}{1-\sigma}}$$

The firms' i output is aggregated into an industry s output by perfectly competitive firms:

$$Y_t(s) = \left(\int_0^{N_{t-1}(s)} y_{i,t}^{\frac{\epsilon(s)-1}{\epsilon(s)}}(s) \mathrm{d}i \right)^{\frac{\epsilon(s)}{\epsilon(s)-1}}.$$

where $N_{t-1}(s)$ is the number of firms in industry s active at time period t, where the number of firms active in period t is determined in period $t-1$, as described below. The industry price index is an aggregate of firm level price choices:

$$P_{j,t} = \left(\int_0^{N_{t-1}(s)} p_{i,t}^{1-\epsilon_j}(s) \mathrm{d}i \right)^{\frac{1}{1-\epsilon(s)}}.$$

The timing of a firm's life in an industry j is:

- In $t-1$, firms pay the industry-specific fixed cost $\frac{1}{2}\kappa_{t-1}^e(s)(1+g(s))^t$ to become active. The fixed cost $\kappa_{t-1}^e(s)(1+g(s))^t$ has an industry-specific trend reflecting industry-specific trend productivity growth.

- In period t, active firms, indexed by $i \in [0, N_{t-1}(s)]$, make the following decisions: they rent capital $k_{i,t}(s)$ from capital-producing firms, and labor $\ell_{i,t}(s)$ from households, and produce the industry-specific output $y_{i,t}(s) = A_t(s) k_{i,t}^{\alpha}(s) \left((1+g(s))^t \ell_{i,t}(s)\right)^{1-\alpha}$, where $A_t(s)$ is a stationary industry-specific technology. They sell the output $y_{i,t}(s)$ at price $p_{i,t}(s)$ to the industry wide aggregating firm.

The active firm's problem in period t is to choose $k_{i,t}(s)$, $\ell_{i,t}(s)$ and $p_{i,t}(s)$ to maximize profits, subject to the firm's demand curve and the production function, taking rental rates w_t and $R_{k,t}$ as given. So

$$\max_{k_{i,t}(s), \ell_{i,t}(s), p_{i,t}(s)} p_{i,t}(s) y_{i,t}(s) - \frac{W_t}{P_t} \ell_{i,t}(s) - R_{k,t} k_{i,t}(s)$$

subject to the production function

$$y_{i,t}(s) = A_t(s) k_{i,t}^{\alpha}(s) \left((1+g(s))^t \ell_{i,t}(s)\right)^{1-\alpha}$$

and the demand curve

$$y_{i,t}(s) = \left(\frac{p_{i,t}(s)}{P_t(s)}\right)^{-\epsilon(s)} Y_t(s).$$

Substituting out for the demand $y_{i,t}(s)$, the solution to the pricing problem is to set a fixed markup over marginal cost $\chi_t(s)$, the multiplier on the production function:

$$p_{i,t}(s) = \mu(s) \chi_t(s)$$

where $\mu(s) = \frac{\epsilon(s)}{\epsilon(s)-1}$. Factor choices in the firm's problem imply

$$k_{i,t}(s) = \alpha \frac{\chi_t(s)}{R_{k,t}} y_{i,t}(s)$$

and for labor

$$\ell_{i,t}(s) = (1-\alpha) \frac{\chi_t(s)}{W_t/P_t} y_{i,t}(s)$$

These factor choices imply that all goods-producing firms choose the same capital labor ratio $\frac{R_{k,t}}{W_t/P_t} = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{\ell_{i,t}(s)}{k_{i,t}(s)}\right)$. Firms' marginal costs depend on factor prices and sectoral productivity:

$$\chi_t(s) = \frac{1}{A_t(s)} \left(\frac{R_{k,t}}{\alpha}\right)^{\alpha} \left(\frac{\frac{1}{(1+g(s))^t} W_t/P_t}{1-\alpha}\right)^{1-\alpha}$$

Given $p_{i,t}(s) = \mu(s)\chi_t(s)$,

$$P_t(s) = \left(\int_0^{N_{t-1}(s)} p_{i,t}^{1-\epsilon(s)}(s) \mathrm{d}i\right)^{\frac{1}{1-\epsilon(s)}} = \mu(s) \chi_t(s) N_{t-1}^{\frac{1}{1-\epsilon(s)}}(s)$$

And since all firms in an industry s set the same price, they choose the same output, so

$$Y_t(s) = \left(\int_0^{N_{t-1}(s)} y_{i,t}^{\frac{\epsilon(s)-1}{\epsilon(s)}}(s)\mathrm{d}i \right)^{\frac{\epsilon(s)}{\epsilon(s)-1}} = y_{i,t}(s)\left(N_{t-1}(s)\right)^{\frac{\epsilon(s)}{\epsilon(s)-1}}$$

The economy's price index is

$$P_t = \left[\int_0^1 (P_t(s))^{1-\sigma}\mathrm{d}s \right]^{\frac{1}{1-\sigma}}$$

which firms in industry s take as given.

The firm entry condition determines the number of firms per industry. At the end of time period $t-1$, a representative entrepreneur in industry s pays the entry cost $\frac{1}{2}\kappa_{t-1}^e(s)(1+g(s))^t$ to earn the expected return from having active goods-producing firms in industry s, which is the expected profit.

The standard free-entry condition implies $N_{t-1}^{\frac{\epsilon(s)-\sigma}{\epsilon(s)-1}}(s)\mathbb{E}_{t-1}\tilde{R}_t = \frac{\mu^{-\sigma}(s)(\mu(s)-1)}{\kappa_{t-1}^e(s)}\mathbb{E}_{t-1}\left[\chi_t^{1-\sigma}(s)\frac{Y_t}{P_t^{-\sigma}}\right]$, so that, if the number of firms is increasing and expected to keep doing so, then the number of new firms that entrepreneurs establish is lower compared to the number of new firms estalibhsed under the standard free-entry condition. Within an industry, all firms have the same market share $\frac{1}{N_{t-1}(s)}$ at time t, so that the Herfindahl index is $\frac{1}{N_{t-1}(s)}$. In steady-state, for industries to have steady-state number of firms, then $g(s) = \bar{g}$ for all industries.

7.2 Estimation Approach

Our approach is motivated by the work of Mian and Sufi, and Jones, Midrigan and Philippon (2018), and Beraja, Hurst and Ospina (2015) who illustrate the value of using regional variation in identifying key parameters. We use the methodology developed in Jones, Midrigan and Philippon that exploits the structure of the model and allows us to combine industry-level and aggregate data to form the likelihood function.

7.2.1 The Likelihood Function

In the Appendix, we show that the state-space representation of the piece-wise linear approximation of our model discussed is:

$$\mathbf{x}_t = \mathbf{J}_t + \mathbf{Q}_t\mathbf{x}_{t-1} + \mathbf{G}_t\epsilon_t,$$

where $\mathbf{x}_t$ collects the endogenous variables, both industry and aggregate and ϵ_t collects the shocks, both industry and aggregate.

The conventional approach to estimating the model would be to write down a likelihood function that directly uses industry and aggregate data and the above solution, however the nonlinearity induced by the ZLB and the large number of industries and associated state variables makes this approach infeasible.

As a result, we follow Jones, Midrigan and Philippon (2018) and formulate an alternative approach to constructing the likelihood function that exploits *relative* variation across industries outcomes. This allows us to separate the likelihood into an industry-level component and an aggregate component.

Formally, let $\mathbf{x}_t^j$ denote the vector of variables for each industry j, expressed in log-deviations from the steady state. Given our piece-wise linear approximation, we can write:

$$\mathbf{x}_t^j = \underbrace{\mathbf{Q}^s \mathbf{x}_{t-1}^j + \mathbf{G}^s \epsilon_t^j}_{\text{industry-level component}} + \underbrace{\mathbf{J}_t^a + \mathbf{Q}_t^a \mathbf{x}_{t-1}^* + \mathbf{G}_t^a \epsilon_t^*}_{\text{aggregate component}}. \tag{6}$$

Here, $\mathbf{Q}^s$ and $\mathbf{G}^s$ encodes how industry j's variables depend on its own state variables and industry-specific shocks ϵ_t^j, while the vector $\mathbf{x}_t^*$ collects the aggregate variables and evolves according to:

$$\mathbf{x}_t^* = \mathbf{J}_t^* + \mathbf{Q}_t^* \mathbf{x}_{t-1}^* + \mathbf{G}_t^* \epsilon_t^*. \tag{7}$$

Here, ϵ_t^* are the aggregate shocks. In contrast, the matrix of coefficients $\mathbf{Q}^s$ and $\mathbf{G}^s$ multiplying the island-level variables is time-invariant.

From the perspective of those in an individual industry, aggregate shocks and the ZLB do not change how that industry responds to its own history of idiosyncratic shocks.

Given this structure our model, letting $\mathbf{x}_t = \int \mathbf{x}_t^j \mathrm{d}j$ denote the economy-wide average of the island-level variables, the deviation of island-level variables from their economy-wide averages,

$$\hat{\mathbf{x}}_t^j = \mathbf{x}_t^j - \mathbf{x}_t, \tag{8}$$

can be written as a time-invariant function of island-level variables alone:

$$\hat{\mathbf{x}}_t^j = \mathbf{Q}^s \hat{\mathbf{x}}_{t-1}^j + \mathbf{G}^s \epsilon_t^j, \tag{9}$$

where we use the assumption $\int \epsilon_t^j \mathrm{d}j = 0$, that island-level shocks have zero mean in the aggregate.

We use the representation in (7) and (9), rather than the computationally infeasible representation in (6) to estimate the model using state-level and aggregate U.S. data.

To make the model solution and data comparable, we express the industry-level data series relative to their respective aggregate series, by subtracting a full set of time effects, one for each year and each variable. We also subtract a state-specific fixed effect. The resulting series for each industry are plotted in Figures 13 to 15.

Figure 13: Relative Q Across Industries

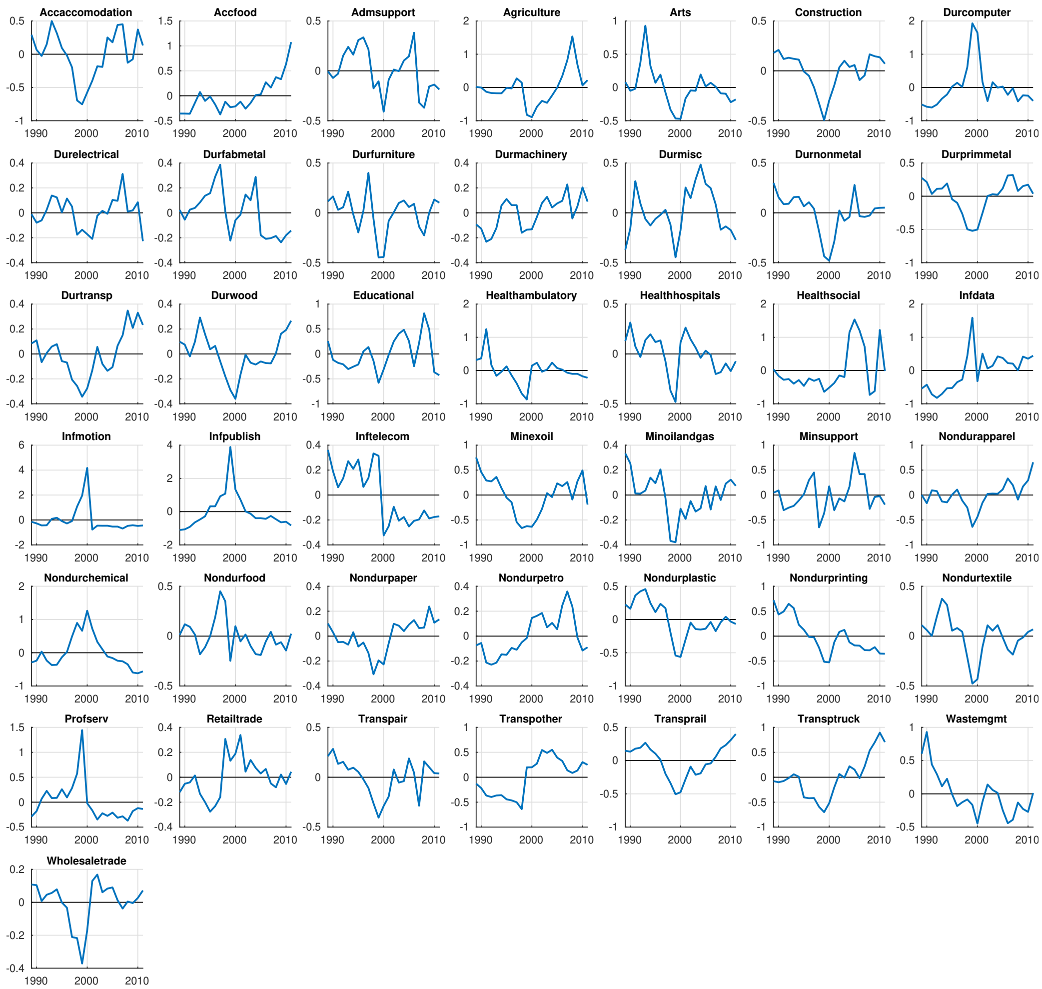


Figure 14: Relative Concentration Ratio Across Industries

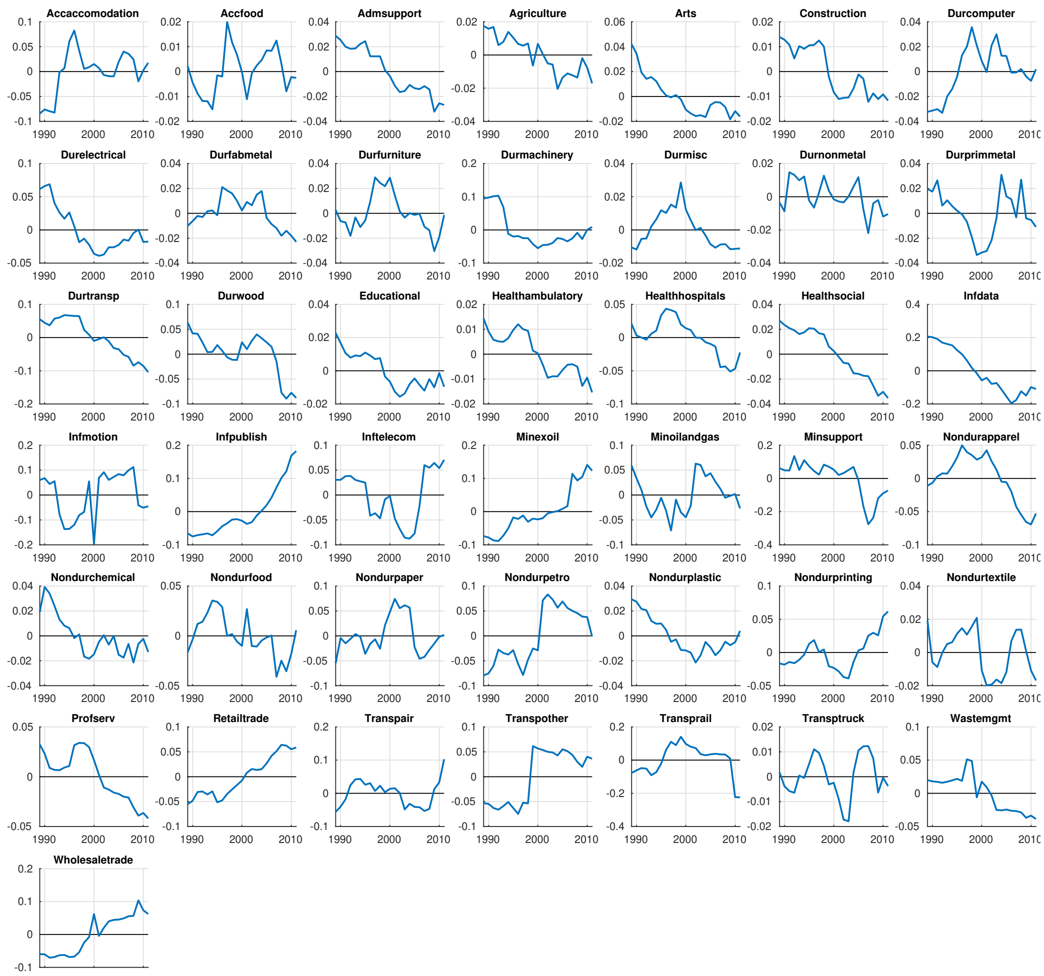
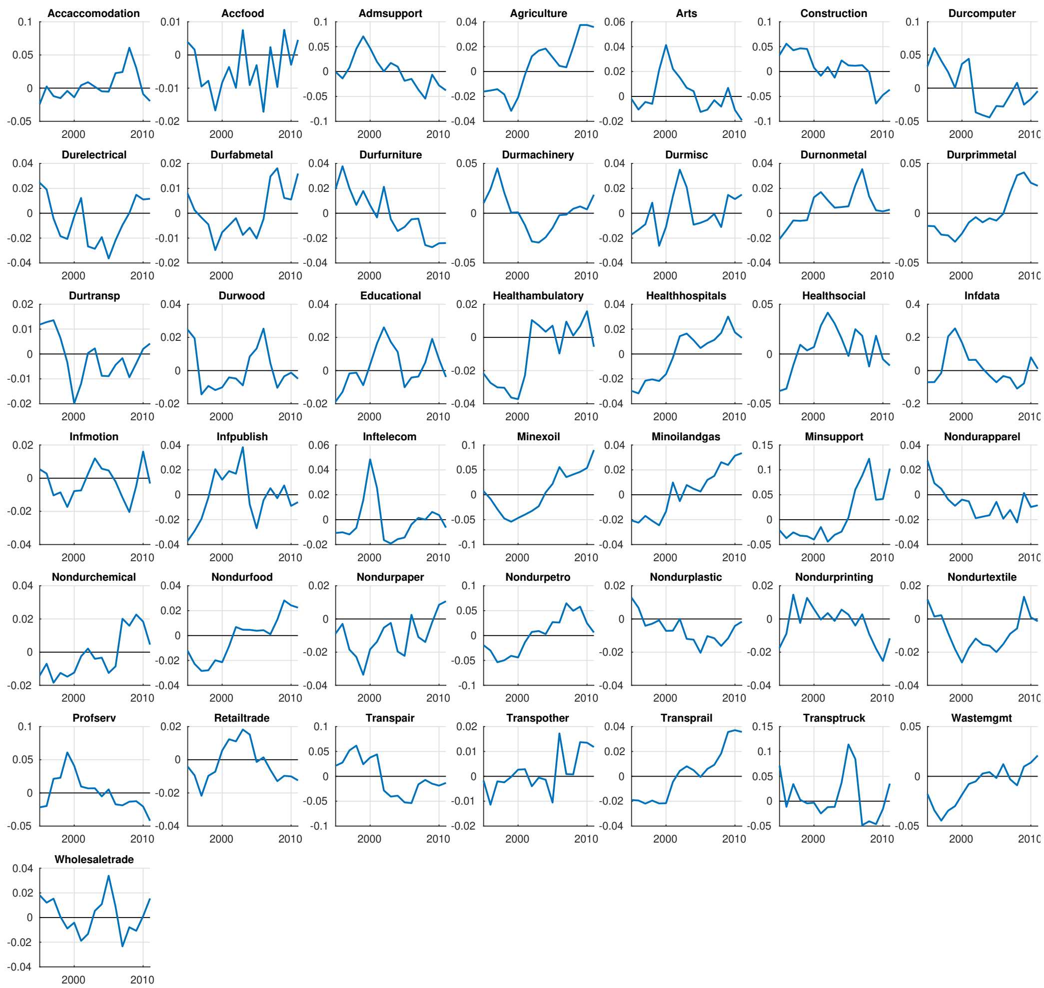


Figure 15: Relative Net Investment Across Industries



8 Conclusions

We have studied a New Keynesian model in which shocks occasionally trigger the zero lower bound on interest rates. We find that the slow recovery of the U.S. economy is not driven by weak consumption and depressed asset prices as the standard liquidity trap theory would predict. Instead, the slow recovery is explained by an apparent lack of willingness of businesses to invest despite favorable economics conditions, i.e. despite historically high profit margin, high asset prices and low funding costs. This investment gap, in turn, is explained by a decreased in competition in the goods market.

References

Alexander, L. and J. Eberly (2016). Investment hollowing out. *Working Paper*.

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. V. Reenen (2017, jan). Concentrating on the fall of the labor share. Technical report, mimeo.

Barkai, S. (2017). Declining labor and capital shares. University of Chicago.

Blonigen, B. A. and J. R. Pierce (2016, October). Evidence for the effects of mergers on market power and efficiency. Working Paper 22750, National Bureau of Economic Research.

CEA (2016, April). Benefits of competition and indicators of market power. Issue Brief.

Coibion, O., Y. Gorodnichenko, and J. Wieland (2012). The optimal inflation rate in new keynesian models: Should central banks raise their inflation targets in light of the zlb? *Review of Economic Studies 79*, 1371–1406.

Davis, S., J. Haltiwanger, R. Jarmin, and J. Miranda (2006). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.), *Macroeconomics Annual*, Cambridge. NBER.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2015, November). Where has all the skewness gone? the decline in high-growth (young) firms in the u.s. Working Papers 15-43, Center for Economic Studies, U.S. Census Bureau.

Dottling, R., G. Gutiérrez, and T. Philippon (2017). Is there an investment gap in advanced economies? if so, why? mimeo.

Eggertsson, G. and N. Mehrotra (2014). A model of secular stagnation. NBER WP 20574.

Eggertsson, G. and M. Woodford (2003). The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity 1*, 139–234.

Faccio, M. and L. Zingales (2017, January). Political determinants of competition in the mobile telecommunication industry. Working Paper 23041, National Bureau of Economic Research.

Feenstra, R. C. and D. E. Weinstein (2017). Globalization, markups and u.s. welfare. *Journal of Political Economy*.

Fernández-Villaverde, J., G. Gordon, P. Guerrón-Quintana, and J. F. Rubio-Ramírez (2015). Nonlinear adventures at the zero lower bound. *Journal of Economic Dynamics and Control 57*, 182–204.

Furman, J. (2015). Business investment in the united states: Facts, explanations, puzzles, and policies. Technical report, CEA.

Grullon, G., Y. Larkin, and R. Michaely (2016). Are us industries becoming more concentrated? Technical report, Working paper.

Gutiérrez, G. and T. Philippon (2017a). Declining competition and investment in the u.s.

Gutiérrez, G. and T. Philippon (2017b). Investment-less growth: An empirical investigation. *Brookings Papers on Economic Activity Fall*.

Jones, C. (2017). Unanticipated shocks and forward guidance at the zero lower bound. mimeo NYU.

Jones, C. (2018). Aging, secular stagnation and the business cycle. Working Paper NYU.

Krugman, P. (1998). It's baaack: Japan's slump and the return of the liquidity trap. *Brookings Papers on Economic Activity 2*, 137–205.

Lawrence Christiano, Martin Eichenbaum, S. R. (2011). When is the government spending multiplier large? *Journal of Political Economy 119*(1), 78–121.

Lee, D. W., H.-H. Shin, and R. M. Stulz (2016, September). Why does capital no longer flow more to the industries with the best growth opportunities? *Fisher College of Business Working Paper 2016*(15).

Lucas, R. E. and E. C. Prescott (1971). Investment under uncertainty. *Econometrica 39*, 659–682.

Nekarda, C. and V. Ramey (2013, jun). The cyclical behavior of the price-cost markup. Technical report, NBER Working paper.

Peters, R. H. and L. A. Taylor (2016). Intangible capital and the investment-q relation. *Journal of Financial Economics*.

Rotemberg, J. and M. Woodford (1999). The cyclical behavior of prices and costs. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1B. Elsevier Science, North Holland.

Summers, L. H. (2013). Why stagnation might prove to be the new normal. Financial Times.

Swanson, E. and J. Williams (2014). Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review 104*, 3154–3185.

Appendix

A Model

A.1 Equilibrium Conditions

We have 12 real unknowns $\{Y_t, N_t, C_t, I_t, K_{t+1}, {}^{W_t}/_{P_t}, R_{k,t}, \mathtt{MC}_t, \mathtt{MRS}_t, \Lambda_{t+1}, R_t, Q_t\}$ and 10 equations:

$$
\begin{aligned}
Y_t &= A_t K_t^{\alpha}\left(\left(1+\bar{\mathtt{g}}\right)^t N_t\right)^{1-\alpha} - \left(1+\bar{\mathtt{g}}\right)^t \Phi \\
Y_t &= C_t + P_{k,t} I_t + \frac{\varphi_k}{2} P_{k,t} K_t \left(\frac{I_t}{K_t} - \delta_t\right)^2 \\
K_{t+1} &= \left(1-\delta_t\right) K_t + I_t \\
\frac{N_t}{K_t} &= \frac{1-\alpha}{\alpha} \frac{R_{k,t}}{{}^{W_t}/_{P_t}} \\
\mathtt{MC}_t &= \frac{1}{A_t}\left(\frac{R_{k,t}}{\alpha}\right)^{\alpha}\left(\frac{\frac{W_t}{(1+\bar{\mathtt{g}})^t P_t}}{1-\alpha}\right)^{1-\alpha} \\
\mathtt{MRS}_t &= \left(1+\bar{\mathtt{g}}\right)^{(1-\gamma)t} N_t^{\varphi} C_t^{\gamma} \\
\Lambda_{t+1} &= \beta\left(\frac{C_t}{C_{t+1}}\right)^{\gamma} \\
\frac{I_t}{K_t} - \delta_t &= \frac{1}{\varphi_k}\left(Q_t - 1\right) \\
1 &= \mathbb{E}_t\left[\Lambda_{t+1}\frac{P_t}{P_{t+1}} R_t\right] \\
Q_t^k &= \mathbb{E}_t\left[\frac{\Lambda_{t+1}}{P_t^k}\left(R_{k,t+1} + P_{t+1}^k\left(Q_{t+1}^k - \delta_{t+1} + \frac{1}{2\varphi_k}\left(Q_{t+1}^k - 1\right)^2\right)\right)\right]
\end{aligned}
$$

The extra two equations to close the model depend on the frictions that we consider.

- With competitive goods markets and flexible prices, the price of output must equal the marginal cost: $\mathtt{MC}_t = 1$;
- Without frictions in the labor market, the real wage must equal the marginal rate of substitution: $\mathtt{MRS}_t = {}^{W_t}/_{P_t}$;
- NK models introduce markups and frictions in both markets and use the equilibrium conditions presented above.

A.2 Steady State

To compute the steady state, we normalize $A = 1$. As usual in this class of model with a representative saver, the discount rate is pinned down by the rate of time preference: $\Lambda = \beta(1+\bar{\mathtt{g}})^{-\gamma}$. Constant capital requires $\frac{I}{K} = \delta + \bar{\mathtt{g}}$ and thus

$$Q^k = 1 + \bar{\mathtt{g}}\varphi_k.$$

Since Q is constant in steady state, we have $Q^k = \Lambda\left(\frac{R_k}{P^k} + Q^k - \delta + \frac{\bar{\mathtt{g}}^2\varphi_k}{2}\right)$.[8] This pins down the required rental rate R_k as a function of discounting and the technology to produce capital goods (P_k, φ_k):

$$\frac{R_k}{P_k} = \delta + \left(\frac{1}{\Lambda} - 1\right) Q^k - \frac{\bar{\mathtt{g}}^2\varphi_k}{2}$$

Monopoly pricing implies a markup of price over marginal cost:

$$\mathtt{MC} = \frac{\epsilon_p - 1}{\epsilon_p}$$

This pins down the real wage:

$$\frac{W}{P} = (1-\alpha)\left(\frac{\epsilon_p - 1}{\epsilon_p}\right)^{\frac{1}{1-\alpha}}\left(\frac{R_k}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}$$

Labor demand then pins down the ratio of labor to capital

$$\frac{N}{K} = \left(\frac{\epsilon_p}{\epsilon_p - 1}\frac{R_k}{\alpha}\right)^{\frac{1}{1-\alpha}}$$

which is the standard MPK condition adjusted for the markup. From the capital labor ratio we get the capital output ratio:

$$\frac{Y + \Phi}{K} = \left(\frac{N}{K}\right)^{1-\alpha}$$

Now we need to think about the fixed cost and the valuation of firms. The valuation of rents is

$$V_t^{\epsilon} = P_t Y_t (1 - \mathtt{MC}_t) - \Phi_t - + \mathbb{E}_t\left[\Lambda_{t+1} V_{t+1}^{\epsilon}\right]$$

so in steady state with $P = 1$, we have

$$V^{\epsilon} = \frac{Y_t(1 - \mathtt{MC}_t) - \Phi}{1 - \Lambda}$$

We think of a model where corporate value must cover fixed costs with some excess premium, so

$$\Phi = \frac{\phi}{\epsilon_p} Y$$

[8]The last term $\frac{\mathtt{g}^2\varphi_k}{2}$ is small since $\mathtt{g}$ is a small number. For instance, with annual data, we would get $\mathtt{g} = 2\%$ and with adjustment costs of 10 (an upper bound), this term is only 2%.

so

$$V^{\epsilon} = \frac{(1-\phi)\, Y_t\, (1 - \mathtt{MC}_t)}{1-\Lambda}$$

and

$$\frac{Y}{K}\left(1 + \frac{\phi}{\epsilon_p}\right) = \left(\frac{N}{K}\right)^{1-\alpha}$$

Then $Y = C + I + \frac{\varphi_k}{2} K \left(\frac{I}{K} - \delta\right)^2$ implies

$$\left(\frac{N}{K}\right)^{1-\alpha} - \frac{\Phi}{K} = \frac{C}{K} + \delta + \bar{\mathtt{g}} + \frac{\varphi_k}{2}\bar{\mathtt{g}}^2 \Longrightarrow \frac{C}{K} = \frac{1}{1+\frac{\phi}{\epsilon_p}}\left(\frac{N}{K}\right)^{1-\alpha} - \delta - \bar{\mathtt{g}} - \frac{\varphi_k}{2}\bar{\mathtt{g}}^2$$

which also means $\alpha \frac{C}{K} = \frac{1-\beta}{\beta} + (1-\alpha)\,\delta - \alpha\bar{\mathtt{g}} - \alpha\frac{\varphi_k}{2}\bar{\mathtt{g}}^2$. And the labor supply condition, with the wage markup, pins down K

$$K = \left(\frac{\epsilon_w - 1}{\epsilon_w}\frac{W}{P}\left(\frac{C}{K}\right)^{-\gamma}\left(\frac{N}{K}\right)^{-\varphi}\right)^{\frac{1}{\varphi+\gamma}}$$

Which we can use to get steady state employment

$$N = \frac{N}{K} \times K$$

A.3 Methodology

The model is approximated subject to an unanticipated trend in the elasticity of substitution between intermediate goods. The nominal interest rate is also subject to the zero lower bound. This section describes the methodology used.

First, consider the time-invariant approximation of a rational-expectations model of the form $x_t = \Psi(x_{t-1}, \mathbb{E}_t x_{t+1}, \varepsilon_t)$ where x_t is the vector of model variables (state and jump), and ε_t is a vector of exogenous unanticipated shocks whose stochastic properties are known. The well-known rational expectations approximation of the model, linearized around its steady state, is written as:

$$\mathbf{A}x_t = \mathbf{C} + \mathbf{B}x_{t-1} + \mathbf{D}\mathbb{E}x_{t+1} + \mathbf{F}\epsilon_t \tag{10}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{F}$ are matrices that encode the structural equations of the model. A solution is:

$$x_t = \mathbf{J} + \mathbf{Q}x_{t-1} + \mathbf{G}\epsilon_t$$

where $\mathbf{J}$, $\mathbf{Q}$, and $\mathbf{G}$ are conformable matrices which are functions of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{F}$.

When agents in the model have time-varying beliefs about the evolution of the model's structural parameters, then: $x_t = \Psi_t(x_{t-1}, \mathbb{E}_t x_{t+1}, \varepsilon_t)$. Denote the corresponding structural matrices for the model linearized at each point in time around the steady state corresponding to the time t structural parameters by $\mathbf{A}_t$, $\mathbf{B}_t$, $\mathbf{C}_t$, $\mathbf{D}_t$, and $\mathbf{F}_t$. A solution to the problem with time-varying structural matrices exists if agents in the model expect the structural matrices to be fixed in the future at values which are consistent with a time-invariant equilibrium **?**.[9] In this case, the solution has a time-varying VAR representation:

$$x_t = \mathbf{J}_t + \mathbf{Q}_t x_{t-1} + \mathbf{G}_t \epsilon_t \tag{11}$$

where $\mathbf{J}_t$, $\mathbf{Q}_t$, and $\mathbf{G}_t$ are conformable matrices which are functions of the evolution of beliefs about the time-varying structural matrices $\mathbf{A}_t$, $\mathbf{B}_t$, $\mathbf{C}_t$, $\mathbf{D}_t$, and $\mathbf{F}_t$, satisfying the recursion:

$$\begin{aligned}
\mathbf{Q}_t &= \left[\mathbf{A}_t - \mathbf{D}_t \mathbf{Q}_{t+1}\right]^{-1} \mathbf{B}_t \\
\mathbf{J}_t &= \left[\mathbf{A}_t - \mathbf{D}_t \mathbf{Q}_{t+1}\right]^{-1} (\mathbf{C}_t + \mathbf{D}_t \mathbf{J}_{t+1}) \\
\mathbf{G}_t &= \left[\mathbf{A}_t - \mathbf{D}_t \mathbf{Q}_{t+1}\right]^{-1} \mathbf{E}_t
\end{aligned}$$

where the final structures $\mathbf{Q}_T$ and $\mathbf{J}_T$ are known and computed from the time invariant structure above under the terminal period's structural parameters. The solution shows that the law of motion for the model's state variables at a time period t depends on the full anticipated path of the structural matrices.

Once the model is in the time-varying VAR representation, then it is straightforward to express the model in its state-space representation and to use the Kalman filter.

A.4 The zero lower bound

The occasionally binding zero lower bound constraint is implemented using solution (12) with a regime-switching algorithm, where the two regimes are the zero lower bound regime and a Taylor-rule policy regime (for full details, see Jones, 2017). Agents have rational expectations over which of the two regimes will apply at each point in time. The algorithm iterates on the forecast time periods that the zero lower bound regime applies. To obtain the time-varying representation (11) that reflects an expected duration of the zero lower bound at each point in time, the method iterates backwards through the model's structural equations starting from the system (12) that arises at the expected exit from the zero lower bound regime.

The zero lower bound duration that agents expect is not constrained to be the same duration as that implied by structural shocks. In this case, the central bank has actively extended the zero

[9] Also see Jones (2015) and Guerieri and Iacoviello (2015), who apply this procedure to approximating models with occasionally binding constraints quickly and efficiently.

lower bound duration through a policy of calendar-based forward guidance. In the estimation, these expected zero lower bound durations are set to those implied by Federal Funds futures data. This ensures forward guidance policy over the post-2009 period is taken into account.

B Solution for Regime Switches

B.1 Time-invariant Solution

A rational expectations model is $x_t = \Psi\left(x_{t-1}, \mathbb{E}_t x_{t+1}, w_t\right)$. Linearize the model around a non-stochastic steady state to get:

$$\mathbf{A}x_t = \mathbf{C} + \mathbf{B}x_{t-1} + \mathbf{D}\mathbb{E}_t x_{t+1} + \mathbf{E}w_t. \tag{12}$$

In an economy where all agents know the regime and expectations are formed under that regime, the solution is a reduced-form VAR:

$$x_t = \mathbf{J} + \mathbf{F}x_{t-1} + \mathbf{G}w_t. \tag{13}$$

where $\mathbf{J}$, $\mathbf{F}$ and $\mathbf{G}$ are conformable matrices which are functions of the structural matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ and $\mathbf{E}$. The matrix $\mathbf{F}$ can be solved by iterating on the quadratic expression:

$$\mathbf{F} = \left[\mathbf{A} - \mathbf{D}\mathbf{F}\right]^{-1}\mathbf{B}. \tag{14}$$

With $\mathbf{F}$ in hand, we compute $\mathbf{J}$ and $\mathbf{G}$ with:

$$\mathbf{J} = \left[\mathbf{A} - \mathbf{D}\mathbf{F}\right]^{-1}\left(\mathbf{C} + \mathbf{D}\mathbf{J}\right) \tag{15}$$

$$\mathbf{G} = \left[\mathbf{A} - \mathbf{D}\mathbf{F}\right]^{-1}\mathbf{E}. \tag{16}$$

B.2 Uncertain Future Regime

The idea will be to take multiple regimes, piece them together with time-varying weights which agents could learn about through the data. Suppose there is one regime which is driving the observables (denoted by the starred system):

$$\mathbf{A}^* x_t = \mathbf{C}^* + \mathbf{B}^* x_{t-1} + \mathbf{D}^*\mathbb{E}_t x_{t+1} + \mathbf{E}^* w_t. \tag{17}$$

Suppose agents in the economy are uncertain about which regime (starred or not starred) will arise tomorrow, and instead believes that the starred regime will arise with probability γ, with the non-starred regime with the residual probability $1 - \gamma$. Seek a solution of the form:

$$x_t = \widetilde{\mathbf{J}} + \widetilde{\mathbf{F}}x_{t-1} + \widetilde{\mathbf{G}}w_t. \tag{18}$$

For this to be a solution, expectations must satisfy:

$$\mathbb{E}_t x_{t+1} = \gamma \mathbf{J}^* + (1-\gamma)\mathbf{J} + \gamma \mathbf{F}^* x_t + (1-\gamma)\mathbf{F} x_t. \tag{19}$$

Substituting this into the system (17) yields:

$$\mathbf{A}^* x_t = \mathbf{C}^* + \mathbf{B}^* x_{t-1} + \mathbf{D}^* \left(\gamma \mathbf{J}^* + (1-\gamma)\mathbf{J} + \gamma \mathbf{F}^* + (1-\gamma)\mathbf{F}\right) x_t + \mathbf{E}^* w_t. \tag{20}$$

Rearranging, we find that:

$$\widetilde{\mathbf{J}} = \left[\mathbf{A}^* - \mathbf{D}^* \left(\gamma \mathbf{J}^* + (1-\gamma)\mathbf{J} + \gamma \mathbf{F}^* + (1-\gamma)\mathbf{F}\right)\right]^{-1} \mathbf{C}^* \tag{21}$$

$$\widetilde{\mathbf{F}} = \left[\mathbf{A}^* - \mathbf{D}^* \left(\gamma \mathbf{J}^* + (1-\gamma)\mathbf{J} + \gamma \mathbf{F}^* + (1-\gamma)\mathbf{F}\right)\right]^{-1} \mathbf{B}^* \tag{22}$$

$$\widetilde{\mathbf{G}} = \left[\mathbf{A}^* - \mathbf{D}^* \left(\gamma \mathbf{J}^* + (1-\gamma)\mathbf{J} + \gamma \mathbf{F}^* + (1-\gamma)\mathbf{F}\right)\right]^{-1} \mathbf{E}^*. \tag{23}$$

More generally, this solution can be used to approximate other uncertainties, such as how long the ZLB might bind in future periods.

If γ is time-varying, then the solution in (12) is also time-varying.

B.3 Time-varying Solution

Now consider the case where the model in hand has time-varying structural parameters:

$$\mathbf{A}_t x_t = \mathbf{C}_t + \mathbf{B}_t x_{t-1} + \mathbf{D}_t \mathbb{E}_t x_{t+1} + \mathbf{E}_t w_t. \tag{24}$$

When all agents know the evolution of the structure of the economy up to a period T, after which the economy structure is assume to be time-invariant, the solution is a reduced-form VAR with time-varying matrices:

$$x_t = \mathbf{J}_t + \mathbf{F}_t x_{t-1} + \mathbf{G}_t w_t. \tag{25}$$

The coefficient matrices satisfy the recursion:

$$\mathbf{F}_t = \left[\mathbf{A}_t - \mathbf{D}_t \mathbf{F}_{t+1}\right]^{-1} \mathbf{B}_t \tag{26}$$

$$\mathbf{J}_t = \left[\mathbf{A}_t - \mathbf{D}_t \mathbf{F}_{t+1}\right]^{-1} \left(\mathbf{C}_t + \mathbf{D}_t \mathbf{J}_{t+1}\right) \tag{27}$$

$$\mathbf{G}_t = \left[\mathbf{A}_t - \mathbf{D}_t \mathbf{F}_{t+1}\right]^{-1} \mathbf{E}_t. \tag{28}$$

where the final structures $\mathbf{F}_T$ and $\mathbf{J}_T$ are known and computed from the time invariant structure above.