DALLASFED Working Paper

Asymmetric Firm Dynamics under Rational Inattention

Anton Cheremukhin and Antonella Tutino

Federal Reserve Bank of Dallas
Research Department

Working Paper 1411

Asymmetric Firm Dynamics under Rational Inattention

Anton Cheremukhin, Antonella Tutino *

Abstract

We study the link between business failures, markups and business cycle asymmetry in the U.S. economy with a model of optimal firm exit under rational inattention. We show that the model's predictions of lagged, counter-cyclical and positively skewed markups together with counter-cyclical exit rates are consistent with the empirical evidence. Moreover, our model uncovers a new mechanism that links information processing with the business cycle. It predicts countercyclical attention to economic conditions consistent with survey evidence.

JEL: E32, D22, D21, D80, C63.

Keywords: Information, Markups, Exit Rates, Rational Inattention.

---

*Anton Cheremukhin: Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas TX 75201, chertosha@gmail.com, 214-922-6785. Antonella Tutino: Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas TX 75201, tutino.antonella@gmail.com, 214-922-6804. Corresponding authors: Anton Cheremukhin and Antonella Tutino. We thank the editor and six anonimous referees for their invaluable comments. We thank Christopher Nekarda and Valerie Ramey for kindly providing the data on markups. We are grateful to Susanto Basu, Simona Cociuba, John Duca, Evan Koenig, Giuseppe Moscarini, Ananth Ramanarayanan, Carlos Zarazaga and participants of various conference meetings for helpful comments. All errors are our own. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. First draft: January 14, 2011.

1 Introduction

A recurring feature of business activity in the U.S. economy is the asymmetry between slow expansions and fast contractions.[1] Three main features of the U.S. business cycle motivate our theory. First, the growth rate of GDP is negatively skewed while the growth rate of markups is positively skewed. Second, markups lag GDP. Finally, firms' exit rates are strongly countercyclical and skewed. The first two features are novel findings while the third is well established.[2] Taken together, these facts suggest a link between the asymmetry of business cycles, markups and firms' exit decisions. In this paper we provide a theoretical framework based on the rational inattention theory of Sims (2006) that reconciles these facts.

The paper's main contribution is to show theoretically that rationally inattentive decisions of entrepreneurs to keep producing or exit the market may be an important source of business cycle asymmetry and that introducing information frictions matters for business cycle analysis.

The model is built on the idea that entrepreneurs have limited cognitive ability to process all available information about the markets in which they operate. This information involves both economywide markup and demand for their business's output. Firms choose signals to monitor market conditions. Based on these signals, they decide whether to exit. Entrepreneurs' choice of signals influences their perception of market conditions and, in turn, their exit decisions. Individual exit decisions affect the number of market participants and through competitive pressure the economywide markup. Thus, individuals' information-processing choices affect aggregate outcomes.

[1]According to the NBER's business cycle dating committee, since 1900 the average length of expansions (11 quarters) has been three times longer than the average length of contractions (3.6 quarters).

[2]For instance, Jaimovich and Floetotto (2008) document that the number of firms is strongly procyclical, and entry and exit decisions account for a substantial fraction of jobs created and destroyed in the U.S. economy.

However, information choices are influenced by aggregate markups as well. When the economy is in an expansion and markups are slowly falling, firms choose to process little information and delay exit. The presence of slim markups and many firms that are no longer profitable in the economy leads to a recession during which a lot of firms exit the market simultaneously. In the aftermath of the recession, markups rise sharply and so does the attention of the incumbent firms. The model predicts lagged, counter-cyclical and positively skewed markups.

Information frictions based on rational inattention theory are the source of an endogenous skew in our model. Different from other information-based theories, rational inattention postulates a cost of processing information about economic conditions whose nature is cognitive. There are no frictions preventing firms from knowing their economic environment other than the information they are capable of processing. This friction is modelled as a fixed marginal cost of processing information associated with a Shannon's channel. The latter regulates the informativeness of signals about economic conditions that a firm chooses. The more information the signals contain, the greater the overall cognitive cost the firm incurs to process them. The key mechanism of the model exploits the fact that a rationally inattentive firm's attention varies with the relative value of information and its cost. In economic expansions, when economy-wide markups are low, the cost of choosing precise signals is higher than the benefits and firms optimally choose relatively uninformative signals. In economic contractions, when economy-wide markups are high, firms find it optimal to process progressively more information as the benefit of being relatively well informed is superior to the cost.

The specification of the cognitive nature of the attention cost together with the endogenously varying amount of attention is at the core of our model. This specification is strongly supported by experimental evidence in economics, psychology and neurology.[3] Our model also uncovers a novel mechanism through which information

[3] For experimental evidence on rational inattention, see Woodford (2013), Cheremukhin, Popova and Tutino (2014).

frictions affect the business cycle. This mechanism is at the core of the prediction that the amount of attention to economic conditions and the precision of forecasts increase sharply in recessions and fall slowly in expansions. Empirical studies of information rigidities support this prediction as shown, e.g., by Coibion and Gorodnichenko (2012) using the Survey of Professional Forecasters.

Using a rational inattention model implies solving a problem where the state variable, its evolution and the choice variable are all distributions. This computational complexity limits the number of firms we can analyze. However, evidence on the U.S. firm size distribution shows that it has fat tails, corroborating the modeling assumption of a finite set of firms. Moreover, we show that the granularity of the U.S. economy, as documented by Gabaix (2011) allows us to meaningfully capture the skewness of individual perceptions of shocks and to relate it to the observed asymmetric aggregate outcomes.

The way information aggregation works in our model is related to models of pure information externality. In the literature,[4] aggregation of information dispersed across many firms results in swift rushes of exits. Learning about the optimal stopping time of others is the key mechanism through which information is dispersed across the population. Much like Murto and Valimaki (2011), our model predicts exit waves due to information aggregation. However, in our framework the mechanism through which learning occurs comes from endogenous firms' choices of information structure.

The paper contributes to three strands of literature. First, we contribute to the literature aimed at explaining business cycle asymmetries.[5] Unlike most previous studies, the key element of our model is firms' endogenous choice of information structure

[4] See, inter alia, Chamley and Gale (1994) and its extension by Murto and Valimaki (2011).

[5] See Shleifer (1986), Zeira (1994), Matsuyama (1999), Francois and Lloyd-Ellis (2003) for examples of variations in monopoly power under uncertainty and McKay and Reis (2008) and Abbritti and Fahr (2013) for asymmetries of wages, price and output over the business cycle. For asymmetries related to information, see Jovanovic (2006) and Van Nieuwerburgh and Veldkamp (2006).

implied by the presence of cognitive costs and its influence on exit decisions.

Second, we contribute to the literature studying the effects that information processing and belief formation have on aggregate fluctuations[6] and coordination in the presence of externalities.[7] Unlike models of coordination games, our model explicitly maps firms' perceptions of the state of the economy into exit decisions through choice of information structure. In this respect, the paper closely relates to the rational inattention framework proposed by Sims (2006).[8]

Finally, our mechanism generates counter-cyclical variations in profit margins.[9] Unlike the mechanisms described in the literature, most of the business cycle adjustment in our model occurs on the exit (rather than entry) margin. This is consistent with evidence on the behavior of establishment entry and exit rates in the U.S. economy, which this paper summarizes.

The paper is organized as follows. Section 2 presents empirical regularities in the U.S. economy regarding markups and exit rates. In Section 3 we describe the primitives of the model together with the information structure and clarify the nature of the attention cost. We complete the section by stating the problem of the firm. Section 4 illustrates the mechanism of the theoretical model and its predictions. We test these predictions against U.S. data. We conclude the section by discussing sensitivity of our findings to the core modelling assumptions. Section 5 concludes with potential policy implications of our business cycle mechanism. Appendices 1-9 contain additional

[6] See, e.g., Mankiw and Reis (2002) and Lorenzoni (2009).

[7] See Angeletos and Pavan (2007), Hellwig and Veldkamp (2009), Myatt and Wallace (2012).

[8] Unlike Mackowiak and Wiederholt (2009), who apply this framework to analyze price stickiness, we do not rely on Gaussian distributions. Instead, the optimal joint distribution of attention is fully endogenous. This approach has already proven useful in Tutino (2013)'s analysis of asymmetries in consumption-savings decisions.

[9] The idea of competitive wars was introduced by Rotemberg and Saloner (1986). Bilbiie, Ghironi and Melitz (2012) relate variations in markups to endogenous variations in entry and product variety, while Edmond and Veldkamp (2009) analyze the interplay of variations in the degree of heterogeneity and counter-cyclical markups.

details.

2 Empirical Evidence

In the data, economic expansions are often linked with declining profit margins. In recessions, declines in real GDP are normally followed by bursts of firm exit and increased markups. These patterns suggest that analyzing the interaction of output, markups and firm exit is important for understanding the business cycle, especially business cycle asymmetries. In order to better motivate our modeling assumptions, we describe in this section the business cycle properties of GDP, markups, exit and entry rates of firms in the U.S. economy. Some of their properties are well known, while others are new to the literature.

To describe the first two facts, we utilize quarterly data on U.S. real GDP from the BEA and quarterly data on markups for the U.S. economy from 1948:1 to 2010:4 constructed by Nekarda and Ramey (2013).[10] The broad question of interest is the dynamic interaction between GDP and economy-wide markups.

Fact 1. *Markups lag the business cycle. Lagged markups are counter-cyclical.*

In Table 1 we present tests of Granger causality for real GDP and markups. We present results for detrended levels and for growth rates.[11] Table 1 demonstrates that real GDP Granger causes markups, but markups do not Granger cause GDP.[12] That is, observations of GDP have predictive power for markups, but observations of markups

[10] We show in Appendix 2 that using alternative measures of markups, e.g., the ratio of revenues to employee compensation, leads to similar results.

[11] To detrend, we subtract a quadratic trend. Growth rates are changes of logs. Subtracting the mean, a linear trend, or the CBO potential GDP, and combinations of these options - all yield similar results.

[12] The hypothesis that variations in detrended real GDP (GDP growth rates) do not Granger cause variations in (growth rates of) markups is rejected. The opposite hypotheses cannot be rejected.

Number of lags	1	2	3	4	5	6
Levels:	GDP Granger causes markups					
F-statistic	28.3***	13.8***	8.3***	6.4***	7.3***	6.0***
	Markups do not Granger cause GDP					
F-statistic	14.1***	1.1	0.8	0.7	0.9	0.8
Growth rates:	GDP Granger causes Markups					
F-statistic	10.8***	8.6***	8.7***	11.4***	8.2***	6.2***
	Markups do not Granger cause GDP					
F-statistic	0.2	2.5*	1.4	1.4	0.9	2.1*

Frequency: Quarterly 1948:1-2010:4. Observations: 251. *, ** and *** show significance at 10%, 5% and 1%. Sources: NIPA; Nekarda and Ramey (2013).

Table 1: **Cyclical Properties of Markups**

have no predictive power for GDP. This implies, that markups have a cyclical component that lags the business cycle.

Figure 1 exposes the dynamic relationship between GDP and markups by reporting correlations of their leads and lags for de-trended levels and growth rates plotted against quarters (on the x-axis). De-trended levels and growth rates corresponding to positive quarters represent leads whereas those corresponding to negative quarters represent lags. As Nekarda and Ramey (2013) remark, Figure 1 can be interpreted in two ways. First, markups peak well before the peak in GDP. Second, peaks in markups follow declines in GDP. Evidence on Granger causality in Table 1 indicates that only the second interpretation is statistically meaningful. Hence, markups are counter-cyclical lagging the business cycle by a few quarters. To our knowledge, the differentiation between these lead-lag relationships is new to the literature.

Fact 1 indicates that monopoly power plays an important role in business cycles. This is consistent with a large body of theoretical work,[13] which stresses the importance

[13] See Bilbiie Ghironi and Melitz (2012), Jaimovich (2007) and Jaimovich and Floetotto (2008).

	GDP growth	Markup growth	Exit rate	Entry rate
Volatility	1.0%	0.82%	5.3%	4.0%
Skewness	-1.07***	1.05***	1.12***	-0.09
		Cross-corellations		
Markup growth	−.24**			
Exit rate	−.49***	+.19*		
Entry rate	+.23**	−.21**	−.42***	

Frequency: Quarterly 1992:3-2010:4. Observations: 71. *, **, *** show significance at 10%, 5%, 1%.

GDP and Markup stand for log changes in GDP and 3-quarter lead of markups respectively.

Sources: NIPA; Nekarda and Ramey (2013); Business Employment Dynamics.

Table 2: **Markups, Exit and Entry rates in the U.S.**

of markup variations for understanding business cycles. In the theoretical model, this fact motivates the assumption that the total number of products offered generates an aggregate demand externality by affecting the markup of each firm.

Fact 2. *Markups show a strong positive growth rate asymmetry, the opposite of real GDP.*

One way to illustrate the growth rate asymmetry of a variable is to compute the skewness of the distribution of its growth rates. A positive value would indicate that a variable makes more large positive jumps than negative ones (and vice versa). The first two rows of Table 2 support the finding that lagged markups are counter-cyclical and highly asymmetric by illustrating the sizeable skewness of log changes in GDP and markups. We find that log changes in markups have a skewness of 1.05 and log changes in real GDP have a skewness of -1.07. While the fact that real GDP growth is negatively skewed has been widely documented in the literature, the strong positive skewness of markup growth is new to the literature.

Fact 2 states that markup variations are highly asymmetric. The dates of the biggest changes indicate that markups rise steeply following recessions and fall gradually in expansions. This new finding suggests that understanding changes in market power is crucial to understanding business cycle asymmetries. While different theories (such as New-Keneyesian models with sticky prices) could possibly account for the cyclicality of markups at different leads and lags, in order to account for Fact 2, a compelling theory of firm dynamics ought to produce an endogenously skewed distribution of changes in markups.

Modelwise, Facts 1 and 2 naturally lead us towards a theory of lagged countercyclical markups. This theory implies that markups rise sharply in the aftermath of a recession due to a decline in the number of competitors, and then fall gradually in a boom as new businesses populate the economy. This interpretation is consistent with the pattern of cross-correlations of growth rates of GDP and markups, and with correlations of de-trended levels depicted in Figure 1.

Variations in the number of firms are a major contributor to variations in markups through their effect on market tightness. The pro-cyclicality of the number of firms is well documented.[14] However, in modeling the relationship between the number of firms and markups the literature has either abstracted from explicitly modeling entry and exit, or focused only on entry of new firms.[15] To better understand the asymmetry of markups and to inform our modeling choices we further explore the cyclical and asymmetric properties of both exit and entry of firms.

To describe the properties of firm entry and exit rates we use quarterly data on the number of opening and closing establishments in the U.S. economy reported by the Business Employment Dynamics survey. The last four columns of Table 2 illustrate the next two observations.

[14] See, inter alias Chatterjee and Cooper (1998), Devereux et. al. (1996) and Jaimovich and Floetotto (2008).

[15] See, inter alias, Bilbiie Ghironi and Melitz (2012), Jaimovich (2007), and the subsequent literature.

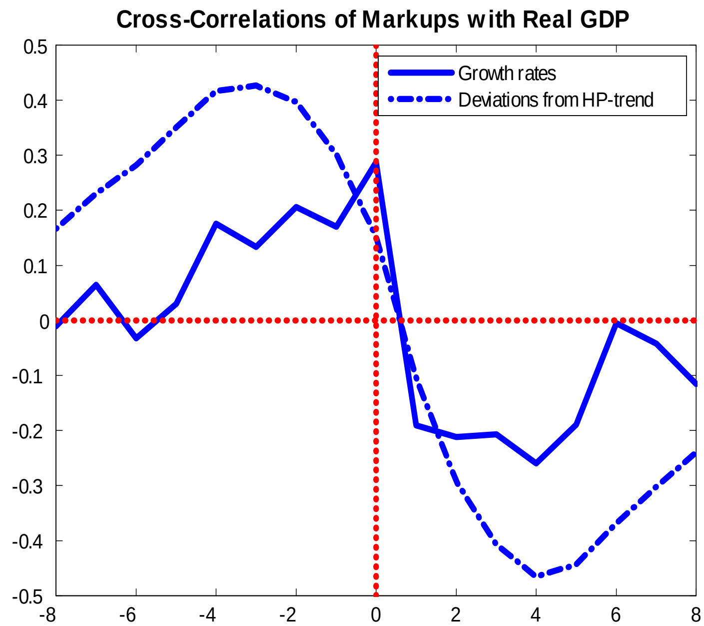

FIGURE 1. **Cyclical Properties of Markups.**

Sources: NIPA, Nekarda and Ramey (2013).

First, we find that establishment exit rates are more volatile than entry rates. The average volatility of establishment exit rates in the U.S. economy in the last twenty years has been 5.3%, about 30% higher than the volatility of establishment entry rates at 4.0%.

Second, we note a strong asymmetry between the cyclical behavior of exit and entry rates. While establishment exit rates are strongly counter-cyclical and highly positively skewed, establishment entry is pro-cyclical and symmetric. More specifically, the correlation of the establishment exit rate with real GDP growth in the U.S. economy in the last twenty years has been -0.49, with an exit rate skewness at 1.12. Table 2 also supports the finding that lagged markups are positively related to exit, and negatively related to entry.

We illustrate our findings and provide extensive robustness checks of the results in Appendix 2. We find that all alternative measures of markups we consider have positively skewed growth rates, are positively correlated with exit rates, and negatively correlated with entry rates. These results are robust to different detrending procedures that remove secular trends from entry and exit rates and their proxies.

The empirical evidence informs the main assumptions of our theoretical framework. First, the fact that makups are countercyclical and lag GDP leads us to the assumption that the variety of goods each firm produces affects individual markups as well as the economy-wide markup. Second, the link between asymmetric behavior of GDP, markups and firm exit suggests a departure from standard theories of expectation formation that typically generate symmetric distributions. Rational inattention theory is better suited for that. Third, empirical properties of entry and exit suggest that the exit margin is far more important in determining the aggregate markup in the U.S. economy. We focus on the exit margin and assume that new firms enter at a constant rate.

3 Model

In this section, we construct the simplest possible economy where variations in the number of firms in the market induce an aggregate demand externality. As we shall see in our model, variations in the number of firms are determined by firms' decisions to exit conditional on their endogenous choice of information. Entry is exogenous.[16]

Our model economy could be thought of as a single sector version of the economy described in Jaimovich (2007). The main difference comes from our focus on modeling separately the entry and exit decisions of firms. We replace the assumption that the number of firms is a jump variable determined by a zero profit condition with a setup where incumbent firms can choose whether to exit the market. For simplicity, we disregard capital as a factor of production. In our economy the stock of available production capacity is a counterpart of capital. We show in Appendix 3 that this setup

[16] We defer to section 4.4 for a discussion of sensitivity of the results to this assumption.

is equivalent to a setup with imperfect substitution between goods in the presence of fixed costs.

Thus, apart from the determination of the number of firms, our economy is an exact counterpart of the economy analyzed by Jaimovich (2007) for the case when the parameter of capital share approaches 0, the parameter of the elasticity of substitution equals 1, the parameter of the elasticity of labor supply equals 0, and there is a single representative sector in the economy. This case satisfies the sufficient condition for existence and uniqueness of a steady-state and the necessary condition for multiple equilibria.

3.1 Primitives

Time is discrete and continues forever, $t = 1, .., \infty$. In each time period, the model economy consists of a representative household and K_t firms. We denote each firm with the subscript i, where $i = 1, ..., K_t$. Firm profit, denoted by π_{it}, is derived from producing and selling differentiated product q_{it} at price p_{it} net of the wage bill, w_t. For simplicity, we assume that there are no strategic interactions across firms.

Firm i's profit function amounts to:

$$\pi_{it} = p_{it} q_{it} - w_t l_{it}. \tag{1}$$

Firms use identical production functions, which are linear in labor inputs, l_{it}, and have a capacity constraint:

$$q_{it} = A l_{it} \leq A, \tag{2}$$

where A is total production capacity. Within each period, firms either operate at full capacity, or do not operate.

The representative household trades off leisure for consumption, maximizing a standard utility function:

$$\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\gamma} - 1}{1-\gamma} - L_t \right), \tag{3}$$

with respect to the supply of labor, L_t, and a consumption aggregator, C_t. In (3), γ denotes the coefficient of relative risk aversion. The consumption aggregator weights differentiated products, q_{it}, by their tastes, v_{it}:

$$C_t = \sum_{i=1}^{K_t} v_{it} q_{it}. \tag{4}$$

The household owns all the firms in the economy. It spends wage and profit income on contemporaneous consumption, maximizing utility subject to a budget constraint:

$$\sum_{i=1}^{K_t} p_{it} q_{it} = w_t L_t + \sum_{i=1}^{K_t} \pi_{it}. \tag{5}$$

Maximization yields the following first-order condition, which determines the demand curve for each good indexed by i:

$$p_{it} = w_t C_t^{-\gamma} v_{it}, \tag{6}$$

which is driven by variations in idiosyncratic tastes v_{it}. Variations in tastes are the only source of uncertainty in the economy.

Prices in the economy are determined according to a Walrasian equilibrium among the incumbent firms in period t.[17] We define a consumption price index as follows:[18]

$$P_t = \frac{1}{C_t} \sum_{i=1}^{K_t} p_{it} q_{it} = w_t C_t^{-\gamma}. \tag{7}$$

[17] We postulate that prices are set by an intermediary whose sole purpose is to acknowledge v_{it} and set the price accordingly. Profits resulting from sales of the products are passed onto the firm. We discuss implications of alternative assumptions in Section 4.

[18] To obtain (7), we start from the definition of the price index, $P_t C_t = \sum_{i=1}^{K} p_{it} q_{it}$. Substituting (6) into this definition and using (4) yield the desired result. More details are in Appendix 3.

Let the wage, w_t, be the numeraire. The expression for firms' profits then simplifies to:

$$\pi_{it} = \mu_{t+1} v_{it} - 1, \tag{8}$$

where the markup, μ_{t+1} characterizes the aggregate state of the economy:

$$\mu_{t+1} = AP_t = A^{1-\gamma} \left(\sum_{i=1}^{K_t} v_{it} \right)^{-\gamma}. \tag{9}$$

We will explain the $t+1$ subscript on the markup μ when we discuss the timing of the model. For now, we shall note that even ruling out strategic interactions among firms, eq. (9) makes clear that changes in the number of firms (K_t) and consumer tastes (v_{it}) affect the market through their impact on the degree of competition. As a result, they generate a negative demand externality for the incumbent firms.

For tractability we abstract from endogenous variations in the entry margin and focus on the exit decision.[19] We assume that on average λ new firms arrive every period. Because this number may not be round, we assume that the number of entrants is drawn from a Poisson distribution with parameter λ:

$$f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{10}$$

New entrants receive the highest possible value of taste of 1. During the life of a firm the evolution of taste, $v_{i,t}$, for its product is described by the following curvature function and transition rule:

$$v_{it} = e^{-g x_{it}}, \tag{11}$$

$$x_{it} | x_{it-1} = \begin{cases} x_{it-1} + \Delta, & 1 - \varphi \\ U[0, \bar{x}], & \varphi \end{cases}, \tag{12}$$

[19] We discuss alternative assumptions about entry and their implications in Section 4.

where g is a scale parameter and x_{it} denotes the distance of firm i from the frontier. Figure 2 illustrates a typical path of v_{it}.

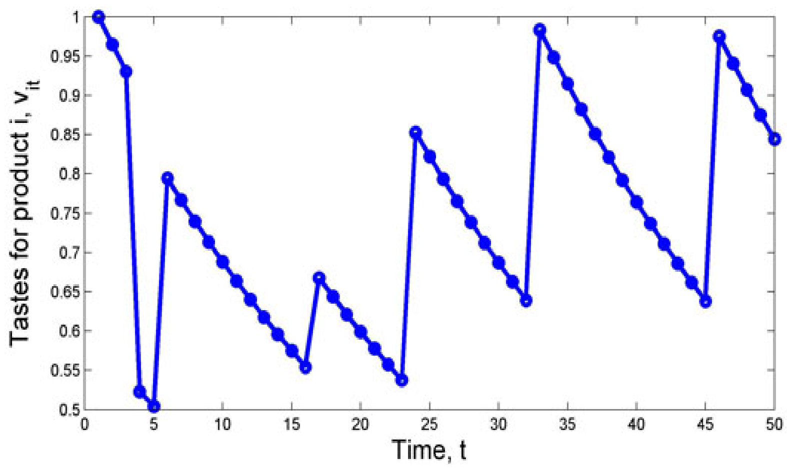

FIGURE 2. **Sample path of v_{it}.**

As new firms enter the market, the distance, x_{it}, of an existing firm from the frontier increases with a drift parameter, Δ, which is related to the number of entrants, λ, through the entry rate, s:

$$s = \frac{\Delta}{\bar{x}} = \frac{\lambda}{\bar{K}}, \tag{13}$$

where $\bar{K}$ denotes the average number of firms. The transition rule in equation (12) captures the idea that entry of λ new products makes older products less desirable, shifting down consumers' relative taste for them by Δ. To make the distribution stationary, we assume that each firm can innovate with probability φ, in which case the taste for the product is drawn from a uniform distribution. Parameter $1 - \varphi$ is related to the persistence of idiosyncratic tastes.

We assume that incumbent firms decide whether to exit or stay in each period. If a firm stays, it produces according to (2) to satisfy the demand for its product and obtains profits (8). If a firm exits, the firm never reenters the market and receives a continuation value of zero. The timing of events is as follows: 1) new entrants arrive; 2) nature determines tastes; $\{v_{it},\ i = 1, ..., K_t\}$; 3) firms process information and form beliefs; 4) based on this information each firm decides whether to stay in the market or exit; 5) the combination of exit decisions determines aggregate variables and profits of individual firms in general equilibrium according to equation (9).

This timing implies that exit decisions are made simultaneously by various firms, so the choices of others are not yet known when a firm makes its own exit decision. Because of this timing structure, last period's markup, μ_t, is the contemporaneous aggregate state of the economy.

Let $e_{it} \in (\{0\}, \{1\})$ denote the exit decision of firm i in period t. Each firm maximizes the expected discounted profits by choosing whether to exit ($e_{it} = 1$) or stay ($e_{it} = 0$). Firms base their decisions on the information available to them. The information structure of incumbent firms is the central element of this model. We turn to it next.

3.2 Information structure

In this subsection, we contrast the outcome of the incumbent firm's problem under full information with the outcome of the problem where incumbent firms face information-processing constraints. The way in which firms acquire and process information has a non-trivial effect on the aggregate behavior of the model.

We use as a benchmark the model with full information. Under full information, at time t an incumbent firm i observes the taste for its own product, v_{it}, as well as tastes for all the products offered in the market, v_{jt}, $j \neq i$, and the number of firms competing in the market at period $t-1$. Note that, for given A, knowledge of the number of competitors and demand conditions implies knowledge of μ_t from eq. (9). Let $S_{it} = \{v_{it}, \mu_t\}$ be the vector of states that determines an incumbent firm's decision to exit at time t, $e_{it}(S_{it})$. We assume that incumbent firms know the law of motion of S_{it}, denoted by $T(S_{it+1}|S_{it})$.

We characterize the solution of the full information problem via value iteration of the following Bellman equation:

$$V_{it}\left(S_{it}\right)=\max_{e_{it}\left(S_{i}\right)}\left\{E_{t}\left[\pi_{it}+\beta V_{t+1}\left(S_{it+1}\right)|S_{it}\right],0\right\}. \tag{14}$$

The outcome of the full information model is characterized by mild symmetric fluctuations in the number of firms operating each period.[20] This finding is due to firms' knowledge of demand conditions (their own as well as their competitors'), which is used to optimally time their exit. Thus, the full information outcome produces symmetric business-cycle fluctuations that are at odds with the properties of business cycles in the U.S. economy.

To better account for these, we propose a model with information frictions based on rational inattention theory (Sims 2003, 2006). We find that the outcome of the model with rationally inattentive firms much better fits the evidence on the asymmetry of exit rates and markups over the business cycle.

Before turning to the formal statement of the rational inattention model in Section 3.4, we describe in Section 3.3 the nature and structure of the information-processing costs under rational inattention. In the same section we argue on the plausibility of our cost specification with respect to other models of information frictions.

3.3 Nature and structure of the attention cost: cognitive cost and elastic capacity

Suppose an economic environment can be described by the joint behavior of two stochastic processes: an idiosyncratic variable, v_{it}, reflecting a firm's own demand conditions, and an aggregate variable, μ_t, reflecting market tightness. This environment can be summarized by a state variable $S_{it}=\{v_{it},\mu_t\}$ in period t. Consider a firm

[20]The full-information economy exhibits aggregate fluctuations simply because the number of firms is finite and idiosyncratic shocks do not wash away. See Appendix 5 for a comparison between the predictions of the full information model and those of the rational inattention model.

that needs to track S_{it} to time its exit. It has a prior belief on S_{it} defined as the distribution $g_t(S_{it})$. The firm maps information about the state variable into exit decisions according to a deterministic rule.[21] It chooses to keep operating, $e_i(g_t(S_{it})) = 0$, if its expected profits, $E_t\pi_t$, exceed a given threshold $\bar{\pi}$, otherwise it shuts down. Under rational inattention theory, the firm has a limited time and attention span to track S_{it} precisely. Aware of its limits, the firm chooses the joint distribution of the state of the economy S_{it} and its profits π_{it}, defined as $p(S_{it}, \pi_{it})$, subject to the constraint that current profits can provide limited information about S_{it}. Choosing $p(S_{it}, \pi_{it})$ implies that firms can compute conditional expectations of the state for each realization of profits. Hence, choosing $p(S_{it}, \pi_{it})$ is akin to having firms optimally choose signals about the state, subject to the constraint that the signal cannot be too informative. Moreover, the optimal choice of $p(S_{it}, \pi_{it})$ impacts the firm's beliefs about the state and profits and, hence, its exit decision.

While the firm's problem under rational inattention can be described as a signal extraction problem common in many information-based setups, the nature and the structure of attention costs in this theory is fundamentally different.

Models with information frictions postulate a monetary or physical cost of acquiring and processing information. In contrast, the costs under rational inattention theory are *cognitive* in nature: they represent the intellectual effort of the firm's owners to not only analyze the information (fully and freely available) about the economic environment but also to map that information into an economic decision of whether to remain in the market.

In modeling a firm's ability to process information, we posit a linear structure of the attention cost: $\theta * \kappa_t$, where the shadow cost of processing information, θ, is expressed in utils (marginal profits) per bit, and information capacity, κ_t, is expressed in bits. Information capacity represents the informativeness of the signal the firm

[21] Section 4.4 discusses the implication of this assumption on the mechanism and results.

chooses, $p\left(S_{it},\pi_{it}\right)$, for a given prior $g_t\left(S_{it}\right)$. The definition of κ_t is given by Shannon's mutual information between the decision variable and the state variable denoted as $\mathcal{I}\left(p\left(S_{it},\pi_{it}\right)\right)=\kappa_t$.[22]

Some models of rational inattention[23] postulate that information capacity is fixed each period, i.e., $\kappa_t=\bar{\kappa}$ for all t. Except for the cognitive nature of the attention cost, this way of modeling information frictions is observationally equivalent to information-based models with bounded rationality. In contrast, our specification of attention costs with *elastic* capacity allows for the possibility that firms may choose to vary the amount of information processed during their time in the market. Experimental evidence in economics and psychology[24] documents the relevance of cognitive costs and it strongly supports our specification of linear attention costs and elastic capacity. Moreover, empirical investigations on information rigidities reveal that agents vary their attention to economic conditions over the business cycle.[25] These investigations show that agents pay more attention in recessions than in expansions. The insight from these studies that informs our modeling choice of a cognitive cost with elastic capacity is that the value of information about economic conditions for an incumbent firm varies over time. This observation will be key for interpreting the results of Section 4. We turn next to the formal description of the model under rational inattention.

[22] Note that since entry is a deterministic function of profits, $S_i\rightarrow\pi_i\rightarrow e_i\left(g\left(S_i\right)\right)$, a trivial application of the data processing inequality yields: $\mathcal{I}\left(S_i,e_i\left(g\left(S_i\right)\right)\right)=\mathcal{I}\left(S_i,\pi_i\right)$. See Cover and Thomas, §2.10. See section 4.4 for further discussion on how this assumption impacts the results.

[23] See, e.g., Mackowiak and Wiederholt (2009).

[24] For examples of experimental evidence corroborating the nature and structure of the attention cost in psychology see, e.g., Kahneman (1973) and Sperling and Dosher (1986). For experimental evidence on rational inattention, see Cheremukhin, Popova and Tutino (2014).

[25] See, inter alias, Gorodnichenko, Coibion (2012) that provide evidence of U.S. business forecasts using Survey of Professional Forecaster (SPF) as a proxy for information rigidities, and Loungani, Stekler and Tamirisa (2013) that use surveys of professional forecasters in 46 countries.

3.4 Rational Inattention model

We assume that each firm knows the process characterizing the exogenous entry of new firms and the law of motion of the state vector, $S_{it} = \{\mu_t, v_{it}\}$, which we define as the transition function, $\tilde{T}(:) \equiv \tilde{T}(S_{it+1}; S_{it}|\pi_{it})$.

The state variable of the model is the joint distribution of $\{\mu_t, v_{it}\}$, $g_t(S_{it})$. As stated before, we postulate an environment where the exit decision is a deterministic function of the optimal signlas on profits.[26] Also, we assume away strategic interactions among firms.

Firm i solves the Bellman program:[27]

$$V(g_t(S_{it})) = \max_{e_i(g(S_{it}))} \{EJ_{it}, 0\}, \tag{15}$$

where

$$EJ_{it} \equiv \max_{p(S_{it},\pi_{it})} \int [\pi_{it}(S_{it}) - \theta\kappa_t + \beta V_{t+1}(g_{t+1}(S_{t+1}))]\, p(S_{it}, \pi_{it})\, d\pi_{it} dS_{it}, \tag{16}$$

subject to the information constraint

$$\kappa_t = \int p(S_{it}, \pi_{it}) \log \left(\frac{p(S_{it}, \pi_{it})}{\left(\int p(S_{it}, \pi_{it})\, dS_{it}\right) g(S_{it})} \right) d\pi_{it} dS_{it}, \tag{17}$$

and the updating rule for perception

$$g_{t+1}(S_{it+1}|\pi_{it} = \bar{\pi}_t) = \int \tilde{T}(S_{it+1}; S_{it}|\pi_{it} = \bar{\pi}_t)\, p(S_{it}|\pi_{it} = \bar{\pi}_t)\, dS_{it}, \tag{18}$$

$$g_0(S_0) \text{ given.} \tag{19}$$

Equation (15) is the value function of the firm, which is the maximum between the outside option of zero, if the firm decides to exit, and the expected discounted value of profits (16), if the firm decides to operate. The value function in (16) combines the

[26] Section 4.4 argues that such an assumption does not significantly affect the mechanism and the results.

[27] The statement of the model under rational inattention follows Sims (2006) and Tutino (2013).

expected value of profits this period, $\pi_{it}\left(S_{it}\right)$, and the expected value of future periods, $V_{t+1}\left(g_{t+1}\left(S_{t+1}\right)\right)$, discounted at rate β. The maximization is over the joint distribution $p\left(S_{it},\pi_{it}\right)$ which is also the metric under which firm i defines its own expectations.

The maximization is constrained by the Shannon's processing capacity, (17), which is a function of the optimal choice of the firm, $p\left(S_{it},\pi_{it}\right)$, and the prior $g_t\left(S_{it}\right)$. The interpretation of this constraint has been discussed in the previous subsection. Here we recall that θ is the shadow cost of processing information associated with information capacity κ_t defined by equation (17).

Equation (18) represents the law of motion of the state $g_t\left(S_{it}\right)$, i.e. the posterior $g_{t+1}\left(S_{it+1}\right)$ updated using Bayes' rule. Given a realization of profits, $\pi_{it}=\hat{\pi}_t$, the expression in (18) convolves the stochastic knowledge of the law of motion of S_{it} summarized by the transition function $\tilde{T}\left(:\right)$ with the optimal strategy that led to $\hat{\pi}_t$, i.e., $p\left(S_{it}|\pi_{it}=\hat{\pi}_t\right)$. Finally, (19) provides the initial condition of the problem. Additionally, we require the optimal $p\left(S_{it},\pi_{it}\right)$ to belong to $\mathcal{D}\left(S_{it},\pi_{it}\right)$, that is the space of all the distributions for which:

$$p\left(S_{it},\pi_{it}\right)\geq 0,\ \forall\pi_{it},S_{it}, \tag{20}$$

$$\int\int p\left(S_{it},\pi_{it}\right)d\pi_{it}dS_{it}=1, \tag{21}$$

$$\int p\left(S_{it},\pi_{it}\right)d\pi_{it}=g_t\left(S_{it}\right), \tag{22}$$

where (20) and (21) constrain $p\left(S_{it},\pi_{it}\right)$ to be positive and to sum to one, respectively. Equation (22) represents the constraint that the joint distribution of the state and profits needs to be consistent with the prior belief about the state.

An equilibrium of this economy is a combination of optimal signals $p\left(S_{it},\pi_{it}\right)$, an exit rule $e_i\left(g_t\left(S_{it}\right)\right)$, a law of motion $\tilde{T}\left(:\right)$, prices $\{p_{it},P_t\}$ and allocations $\{q_{it},C_t,l_{it},E_t\}$ such that (i) signals and exit rules solve the firm's problem (15)-(22) given the law of motion, (ii) allocations are optimal given prices and prices clear markets as described by equations (4)-(9), and (iii) the law of motion is consistent with the combination of firms' choices.

Though all the variables are defined on a compact support, decision rules are not necessarily continuous because the exit decisions are discontinuous. This discontinuity prevents a general proof of existence and uniqueness. However, we prove in Appendix 8 that the problem of the firm is a contraction mapping. Hence it has a unique solution given the law of motion. Any solution of the firm maps uniquely into allocations, prices and a law of motion.

Solving for the equilibrium of the model requires equating the economy-wide supply obtained by aggregating the solution of the problem (15)-(22) for each firm to the economy-wide demand in (4).

To find an equilibrium, we solve for the fixed point of the tuple: $\{\tilde{T}\left(:\right), p\left(S_{it}, \pi_{it}\right), e_i\left(g_t\left(S_{it}\right)\right)\}$, such that the law of motion $\tilde{T}\left(:\right)$ is the outcome of exit decisions $e\left(:\right)$ based on the attention allocation solution $p\left(:\right)$, and the attention allocation is optimal given the law of motion.[28]

Iterations between the solution of the firm's problem and simulations of the economy show that convergence to a fixed point is relatively quick. Moreover, our confidence in the existence and uniqueness of a fixed point in practice is reassured by the fact that significant variations in starting points for the law of motion do not lead us to different equilibria.

Note also, that the problem of the firm without information processing constraints described by equation (14) is a special case of the constrained version when $\theta = 0$. Therefore, the information processing constraint is the only source of any differences between the two models we consider.

[28]We approximated the law of motion using a first-order Markov chain. For a description of the pseudo-code that we used to find the equilibrium see Appendix 9.

3.5 Number of firms

A dynamic rational inattention model demands a tall computational burden: each firm's state is a distribution, $g_t(S_{it})$, each firm's decision variable is a distribution, $p(S_{it}, \pi_{it})$, each firm's updating equation is a distribution, $g_{t+1}(S_{it+1}|\pi_{it} = \hat{\pi})$. As a result of the computational complexity, we are limited in the number of firms that we can solve for.

However, Gabaix (2011) shows that because of the fat tailed distribution of firms by size in the U.S. economy, aggregate fluctuations can be well described by a handful of firms. When the distribution of firm sizes is Pareto, the speed of decay of idiosyncratic fluctuations is $\ln N$ instead of $N^{1/2}$. This makes a huge difference, as 10^6 firms in a world with a symmetric size distribution would be equivalent to an economy with on the order of 10^2 firms in a world with a Pareto firm size distribution. This implies that we can get a good idea of the behavior of our informationally unconstrained economy in a granular world by increasing the number of firms to 100.

The second important fact to note is that skewness (our prefered measure of asymmetry) is a normalized variable, which does not decay with the law of large numbers when you aggregate idiosyncratic decisions of firms. Using the same method Gabaix used to derive properties of standard deviations, we derive properties of skewness in Appendix 4. This derivation shows, that the number of firms has a very limited effect on the asymptotic skewness of the distribution of GDP growth rates.

Note that the granularity result makes combinations of idiosyncratic taste shocks act as an aggregate demand shock. Gabaix (2011) shows that an appropriate degree of granularity in the economy makes it possible for idiosyncratic shocks alone to generate aggregate fluctuations comparable in magnitude to observed business cycles movements in the U.S.

4 Results

This section consists of three parts. In Section 4.1, we discuss our calibration and the numerical solution of the rational inattention model from the previous section. Section 4.2 illustrates the mechanism of the model and its main findings. It also shows how the mechanism helps the model match some properties of the data. In Section 4.3, we compare the outcome of the model with information frictions to evidence of business cycle asymmetries in U.S. data. Finally, in Section 4.4 we discuss the sensitivity of the model's predictions to alternative assumptions. Appendix 5 compares the outcomes of the rational inattention model with the outcomes of a full information model and Appendixes 6 and 7 provide alternative calibrations of key model parameters and of the number of firms, respectively.

4.1 Calibration

We focus on business cycle asymmetries, so we are mostly interested in skewness and correlations.[29]

Each time period is a quarter. This choice determines the discount factor, β, at 0.99 and the entry rate, s, at 5%, the average fraction of opening establishments among total private sector establishments in the U.S. in a given quarter, as measured by Business Employment Dynamics (BED). We set the curvature of utility, γ, close to unity, which implies logarithmic utility, consistent with a balanced growth path. We fix the grid size for $x_{i,t}$ to the unit interval and set $\bar{x}$ at 0.9. We set the scale of the idiosyncratic component of tastes, g, to 0.8, which implies an average markup of 90% - the mean of the marginal price-cost markup in the U.S. economy over the last 50 years (see Nekarda and Ramey (2013)).

[29] Since these second and third moments are normalized and do not decay with the law of large numbers, we do not attempt to follow either Gabaix (2011)'s or Jaimovich (2007)'s calibrations targeted at second moments. This makes our exercise much simpler and results significantly more transparent.

We set the probability of innovation, φ, to 0.8, which in our view captures well the dynamic and unpredictable nature of tastes for particular products. We set this parameter in the ballpark of the numbers from Cooper, Haltiwanger and Willis (2007) who estimate the autocorrelation of establishment-specific profitability shocks to be 0.33, with the standard deviation of these shocks to be as large as 0.23. [30]

We are not aware of direct evidence on firm dynamics that would allow us to pin down the shadow cost of information, θ.[31] We set the cost of information to 0.01, which implies that the total shadow cost of information varies in the range from 10% to 20% of average profits in the dynamic equilibrium.

Finally, we set production capacity, A, targeting the average number of firms at 15.

We set the grid for the markup, μ_t, to equi-spaced intervals between 1.15 and 2.77. Because of the computational intensity of the model with inattention, we use a relatively coarse 20-point grid. We set the length of simulations to 250 periods, from which we discard the first 50. Thus, the total history from which firms can learn is comparable to the length of available U.S. data.

In each case, we solve for a fixed point of the mapping between the exit rule $e_{i,t}(:)$ and the transition rule $T(:)$. We switch recursively between finding the solution of the problem of the firm by value function iteration, and simulating the model using the solution to obtain the law of motion. Table 3 summarizes the calibration for the numerical algorithm. Note that for the information-constrained model the joint distribution $g(v_i, \mu)$ has been constructed so that the points on the simplex have marginal mean and standard deviation that reflect properties of the empirical distribution. Then, the transition function convolves the transition properties of $g(\mu, x_i)$ for each

[30] We think that indirect estimates of Cooper, Haltiwanger and Willis (2007) are best suited for our calibration because they cover firms in all sectors of the economy and are computed at quarterly frequencies.

[31] For laboratory evidence on the size of information costs, θ, see Cheremukhin, Popova and Tutino (2014).

SYMBOL	PARAMETER	VALUE
β	Time discount factor	0.99
γ	CRRA coefficient	0.95
g	Scale distance	0.8
s	Entry rate	0.05
φ	Probability of innovation	0.8
θ	Cost of processing information	0.01
π_i	Grid for profits	$[-0.48, 1.78]$
μ	Grid for markup	$[1.15, 2.77]$
x_i	Grid for distance	$[0.0, 1.0]$

Table 3: **Numerical approximation.**

possible value of profits to assign a distribution for next period values of the state, $T(.) \equiv T(\mu', v_i' | \mu, v, \pi)$.[32]

4.2 Mechanism and Findings of the Rational Inattention Model

In this subsection, we discuss the mechanism of the rational inattention model. First, we illustrate the mechanism by analyzing the optimal choices of individual firms. Then we describe the implications of choices of individual firms for aggregate outcomes. These findings allow us to compare the aggregate behavior of our model with business cycle asymmetries in the U.S. data in Section 4.3.

[32] It is important to note that the solution of the model is extremely computationally intensive. Even using advanced programming techniques on a powerful computational cluster, a solution for a single calibration of the overly simplified model with a relatively small number of firms, takes about a week. Thus, the computational intensity places significant restrictions on the scope of our analysis. To explore the sensitivity of our mechanism to the modeling assumptions we do robustness checks by varying several key parameters in Appendix 6-7.

4.2.1 Attention allocation and the profit cycle of an individual firm

We start by describing the properties of the solution of a representative firm's problem. The left panel of Figure 3 shows information capacity, κ, as a function of the firm's realized profits, $\hat{\pi}_i$. The black solid line indicates the total capacity acquired by the firm conditional on profits. The blue dashed line and the green star-dashed line are the components of conditional information capacity spent on forming the perception of the aggregate state, μ, and the idiosyncratic state, v_i, respectively.[33] As stated in Section 3, we use markups as the aggregate state captures tightness in the market in which the firm operates.

Figure 3 demonstrates that a rationally inattentive firm optimally chooses to vary the amount of information over its profit cycle. In particular, the firm chooses a more precise signal about economic conditions when its profits are high than when its profits are low. The rationale for this finding is as follows: When profits are low, it does not pay off to process much information about economic conditions as low levels of profits trigger exit. When profits are high, the firms optimally decide to acquire a more informative signal about economic conditions as misjudging the economic environment

[33] The conditional capacities are computed as follows. We evaluate overall capacity conditional on $\hat{\pi}_i = (\{-0.05\}, \{0\}, \{0.05\}, \{0.35\}, \{1\}, \{1.5\})$ denoted $\kappa|\pi_i$ using the optimal solution $p^*(S_i, \pi_i = \hat{\pi}_i)$ of (15)-(19). The average is taken over all priors $g(S_i)$ in the simplex. From $\kappa|\pi_i$ we integrate out the component allocated to the idiosyncratic state to obtain the conditional capacity of the aggregate state $\kappa_\mu|\pi_i$. Similarly, the capacity allocated to idiosyncratic state $\kappa_{v_i}|\pi_i$ is computed by integrating out the aggregate state. The cross-correlation between μ and v_i in the optimal signal, conditional on profits, is computed as:

$$\rho^*_{\mu,v_i}|\pi_i = Cov(v_i, \mu|\pi_i = \hat{\pi}_i) / \hat{\sigma}^\pi_{v_i}\hat{\sigma}^\pi_\mu$$

where $\hat{\sigma}^\pi_X$ for $X = (v_i, \mu)$ denotes:

$$\hat{\sigma}^\pi_X = \sqrt{Var(X|\pi_i = \hat{\pi}_i)}.$$

at that point of its life-cycle implies a bigger loss of profits for the firm. Hence, the value of information and, as a result, the optimal amount of total information, both increase with the firm's profits.

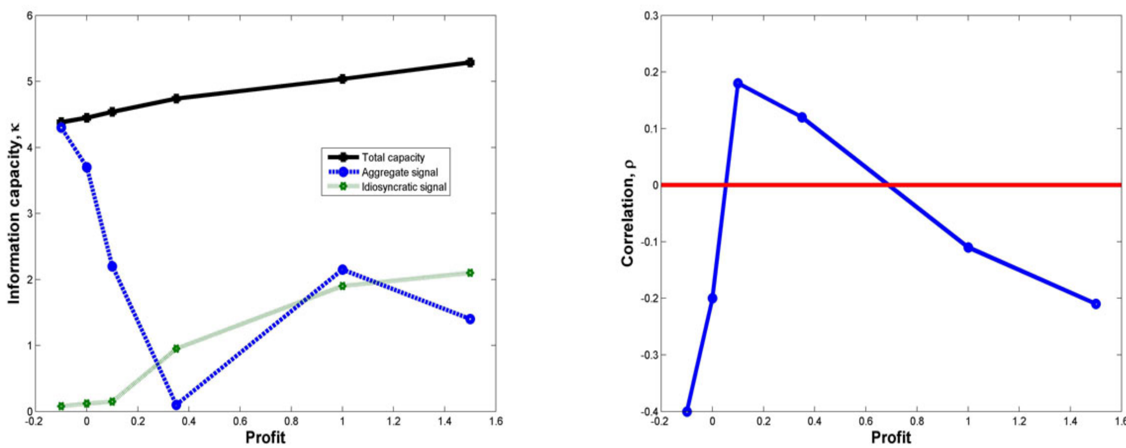

FIGURE 3. **Capacity Allocation and Correlation.**

We look at three additional sets of variables to understand changes in the signals per se as well as how the relative values of signals about the aggregate and idiosyncratic state evolve during the profit cycle of the firm. First, we look at the firm's relative capacity allocated to idiosyncratic and aggregate variables conditional on a particular value of profits, shown in the left panel of Figure 3. Second, we show the conditional cross-correlation between aggregate and idiosyncratic condition, ρ_{μ,v_i}, in the right panel of Figure 3. Finally, Figure 4 shows the conditional precision of the firm's estimate of profits and aggregate markups. The left-panel of Figure 4 plots expected profits against realized profits while the right panel shows perceived markups plotted against realized markups. In both illustrations, the black solid line is the median estimate, the dotted lines indicate 10th and 90th percentiles, while the red solid line depicts the locus where actual and expected variables coincide.

The conditional capacities illustrated in the left panel of Figure 3 show that a firm with low profits optimally chooses to spend most of its information capacity on a precise signal of the aggregate component, leaving little attention to the idiosyncratic component. The firm's decision to closely track the aggregate markup stems from the fact that the aggregate state is more persistent and, hence, a more predictable determinant of profits. Firms know the persistence of markups from the law of motion of the state variable and they realize the importance of markups when forecasting profits by observing the history of realized profits. Thus, when profits are low, and attention is relatively costly, a good signal of aggregate markup, μ, is more valuable to the firm than a good signal on the idiosyncratic state v_i. It follows that firms are far more interested in the aggregate state of the economy (Figure 3, left panel) and acquire a precise signal about μ (Figure 4, left panel) when profits are low.

The only way a firm can track the aggregate state precisely while expending overall low information capacity is to boost perception of its product, v_i. As shown in the right panel of Figure 3, this goal is achieved by setting the perceived correlation between μ and v_i to a large negative value, which implies both low overall information capacity and a high share of attention devoted to the aggregate signal. However, by boosting the perception of v_i, the firm overestimates its own profits. This upward bias in expected profits is the reason why firms decide to stay in the market even when realized profits are negative (Figure 4, right panel).

Indeed, in our model, expected profits of firm i are given by

$$E\pi_i = (E\mu)(Ev_i) - \sigma_\mu \sigma_{v_i} \rho_{\mu,v_i},$$

where $E\mu$ is the expected value of aggregate markup, Ev_i is the expected value of firm's idiosyncratic taste v_i, σ_μ and σ_{v_i} are standard deviations of μ and v_i, and ρ_{μ,v_i} is the cross-correlation between μ and v_i. Conditional on low realized profits, acquiring a precise signal on μ (with low $E\mu$ and low σ_μ) produces an overestimate of profits if the firm chooses a high perception of taste, Ev_i, and sets a large negative cross-correlation, ρ_{μ,v_i}. While the stochastic properties of v_i, captured by the prior, constrain its precision from below, the firm can simultaneously increase precision of the signal on μ and reduce overall information capacity κ by lowering ρ_{μ,v_i}.

The right panel of Figure 3 shows that, indeed, when profits are low, the perceived correlation between the idiosyncratic and aggregate state is strongly negative, whereas it is close to zero for high and medium levels of profits. Figure 4 shows that, conditional on low profits, the estimate of profits is more precise but more upwardly biased than estimates conditional on high profits. With limited capacity, an accurate signal of μ implies a noisy signal of v_i. Overestimating the idiosyncratic state to better understand the aggregate state is why an incumbent firm delays its exit decision past the point at which operating is no longer profitable.

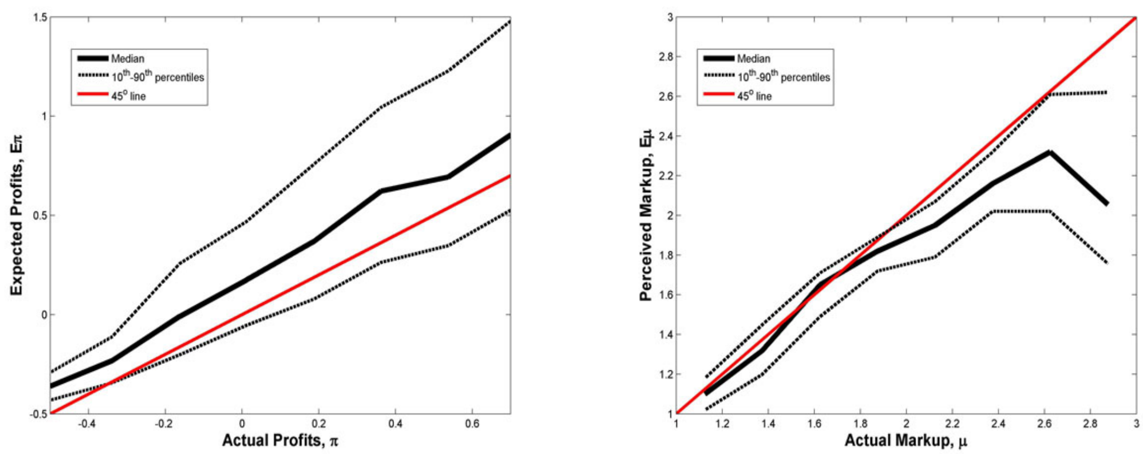

FIGURE 4. **Expectations of Markups and Profits.**

4.2.2 Aggregate attention allocation and the business cycle

Now that we have described the motives behind decisions of individual firms, we can look at implications for aggregate behavior of the model economy. Figures 5 and 6 show the aggregate behavior of the model economy over a sample 50-year (200-quarter) simulation. The top panel of Figure 5 shows profits for the firms that exit (red dots) and stay (blue dots) over the course of the simulation. The central panel shows information capacity of those staying and exiting, and finally the bottom panel shows perceived markups. The top panel of Figure 6 represents output of the simulated economy, which equals aggregate consumption. The central panel represents the behavior of the aggregate markup and the bottom panel displays the exit rate. In both figures, shaded areas indicate recessions defined as periods following declines of at least 10 percent in aggregate output.

Figures 5 and 6 show that information capacity is at its peak in the immediate aftermath of recessions and it slowly decreases in expansionary periods. This is because a recession is characterized by many firms simultaneously exiting the market. In the aftermath, with little competition left, the aggregate markup increases, and incumbents enjoy high profits. As we have seen in Section 4.2.1, when profits are high, the value of being informed is high for each incumbent, so firms increase their attention. As new firms join the market, markups and profits wind down as does the value of paying attention to economic conditions. With less overall attention to economic conditions, low-profit firms accumulate little information about idiosyncratic conditions and have an overly optimistic perception of their profits. As a result, they delay exit decisions past the point where they cease being profitable.

The asymmetry of attention allocation to economic conditions over an incumbent firm's profit cycle maps into a testable prediction: Business owners acquire and process more information, and, hence, have a more precise estimate of economic conditions, during a recession and in its immediate aftermath. The precision of their estimates declines as the economy expands. This prediction is strongly supported by the data, as shown by Coibion and Gorodnichenko(2012). Using U.S. Survey of Professional Forecasters (SPF) data to study expectation formation and quantify the informational rigidities across the business cycle, Coibion and Gorodnichenko (2012) document that economic agents process more information about macroeconomic conditions faster in the aftermath of a recession rather than at its onset. Our model provides a rationale for their findings: The cost of misjudging market conditions in the aftermath of a recession is low relative to the value of learning about market conditions. Coibion and Gorodnichenko (2012) show that their findings are robust across alternative U.S. surveys and across countries.[34] They point out that this evidence poses a challenge for both sticky-information[35] and noisy-information[36] models that fail to capture the state-dependent nature of information acquisition and revisions of expectations. In contrast, the predictions of our model agree with their evidence. The rational inattention mechanism that we propose rationalizes state-dependent information processing.

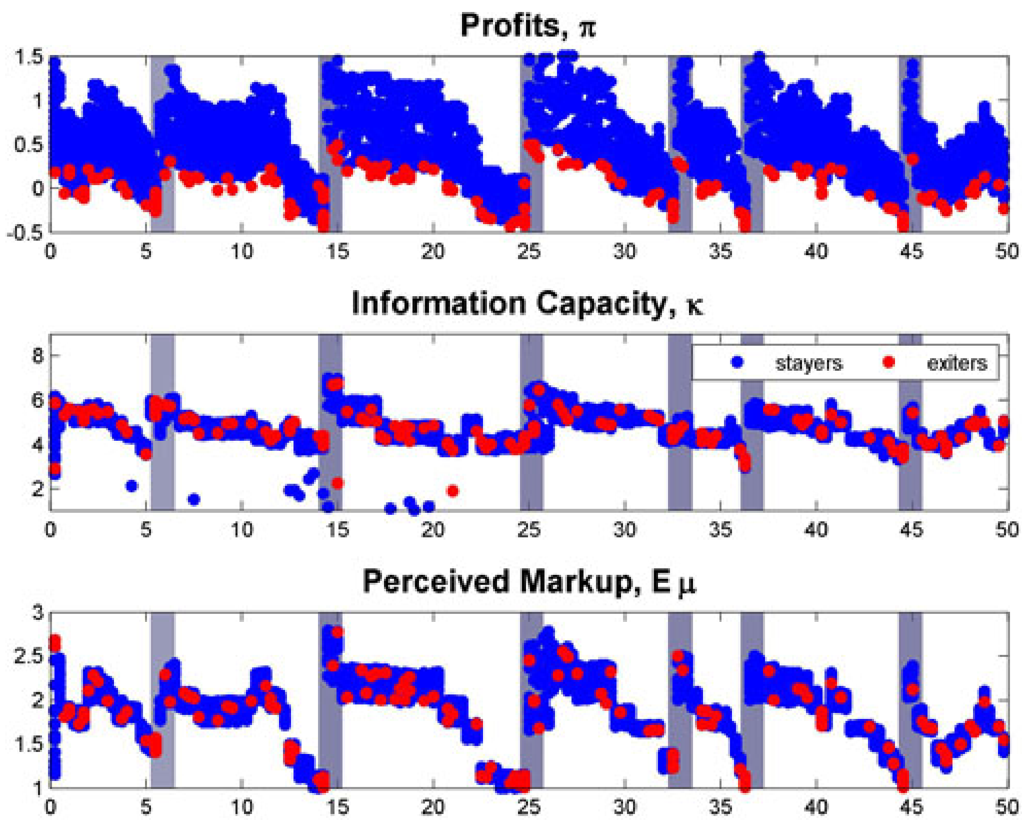

FIGURE 5. **Firm Dynamics under Inattention.**

We can see from Figure 5 that information capacity declines in expansions, and is especially low in periods immediately preceding recessions. As we learned in the

[34]For the U.S., in addition to SPF they use the Michigan Survey and the Livingston Survey. They also construct a dataset of quarterly forecasts from the international survey of professional forecasters from Consensus Economics, which includes 12 countries.

[35]See, e.g., Mankiw and Reis (2002).

[36]See, e.g., Lucas (1972).

previous section, when profits are low, incumbent firms allocate most of their information processing capacity to getting a precise signal of aggregate markups, but pay very little attention to idiosyncratic tastes for their own products (Figure 3, left panel). In these periods a large fraction of firms are guided by signals about a common variable: the aggregate markup. Thus, optimally chosen signals on markups serve as a market cleansing device, eliminating firms that have low profitability. This mechanism explains the pattern of exits in Figure 5 whereby many firms exit simultaneously, triggering a recession.

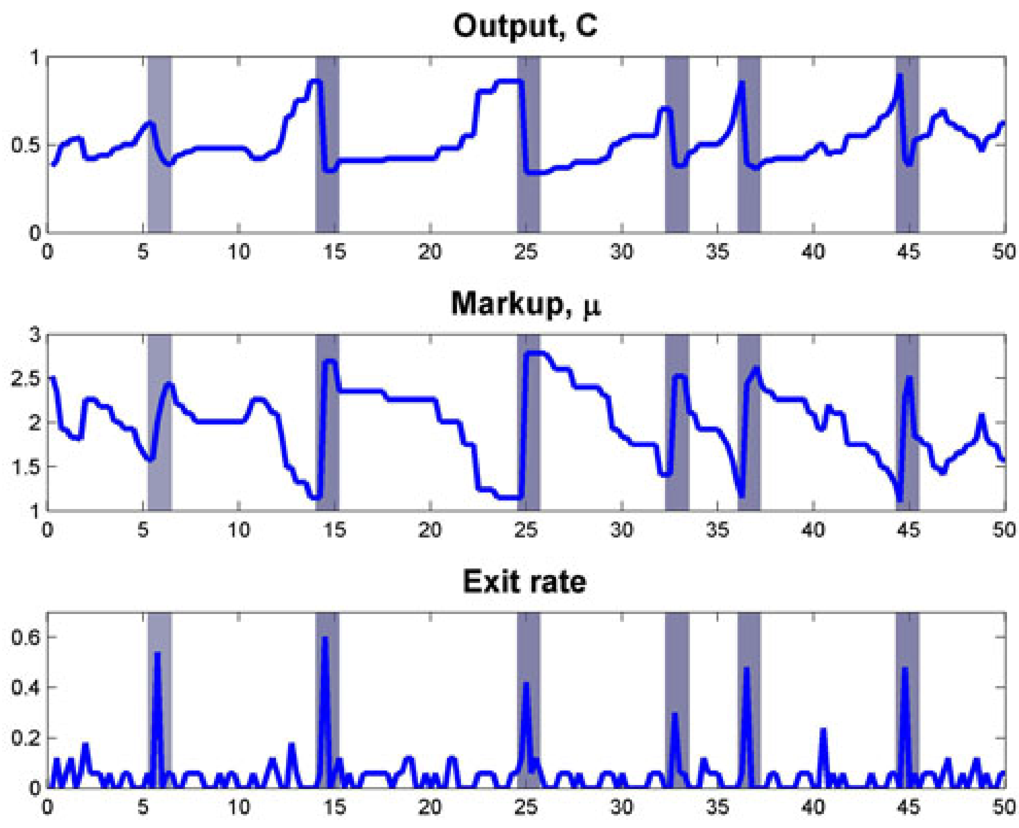

FIGURE 6. **Aggregate Economy with Inattention.**

Given that firms misjudge the value of their product when profits are low (see Section 4.2.1), they tend to stay in the market beyond the point when they are profitable. With too many firms in the market, markups keep shrinking. These events contribute further to firms' losses until they eventually exit.

The asymmetry in the allocation of attention is the driving force of the business cycle in the model economy. Firms optimally decide to monitor closely economic conditions in the aftermath of a recession, as large scale exits trigger high markups. As the economy recovers, markups and profits decrease, and firms lower the optimal amount of attention to the economy. With decreasing attention and markups, firms delay their exit up to a point where signals on deteriorating profits can no longer be ignored.

In this environment, the economy-wide markup is a lagged mirror image of output: output rises in expansions when markups, profits and firms' attention decrease; and it declines in contractions when markups, profits and firms' attention rise. Markups lag output because a decline in the number of incumbent firms (in a recession) implies a decline in production immediately, and a higher markup (with less competition) beginning in next period.

Having established the key firm-level and aggregate predictions of the model with rational inattention, we turn to comparing these predictions of with U.S. data in the next section.

4.3 Comparison of model predictions with U.S. business cycles

Our model predicts that information processing constraints lead to delays in exit decisions that result in asymmetric aggregate fluctuations in output and markups. To compare these predictions with empirical regularities characterizing the U.S. economy, we use measures such as skewness and correlation that are virtually immune to changes in the number of firms. Figure 7 shows the series for output, markups, job destruction and job creation in the manufacturing sector of the U.S. economy. It illustrates that the asymmetry present in the data is very similar to that produced by our model with inattention.[37]

Table 4 shows that our model successfully captures the qualitative properties of business cycles in the U.S. It matches the sign of the asymmetry between contractions and expansions and the asymmetry of firm exit rates - facts serving as a motivation of our paper. Another key prediction of the model: the high positive skewness of the growth rate of markups, is also overwhelmingly supported by the data. Consistent with classical evidence, our model predicts markups to be contemporaneously a-cyclical. Our model also fits well the novel empirical finding that markups lag the business cycle, and lagged markups are counter-cyclical.

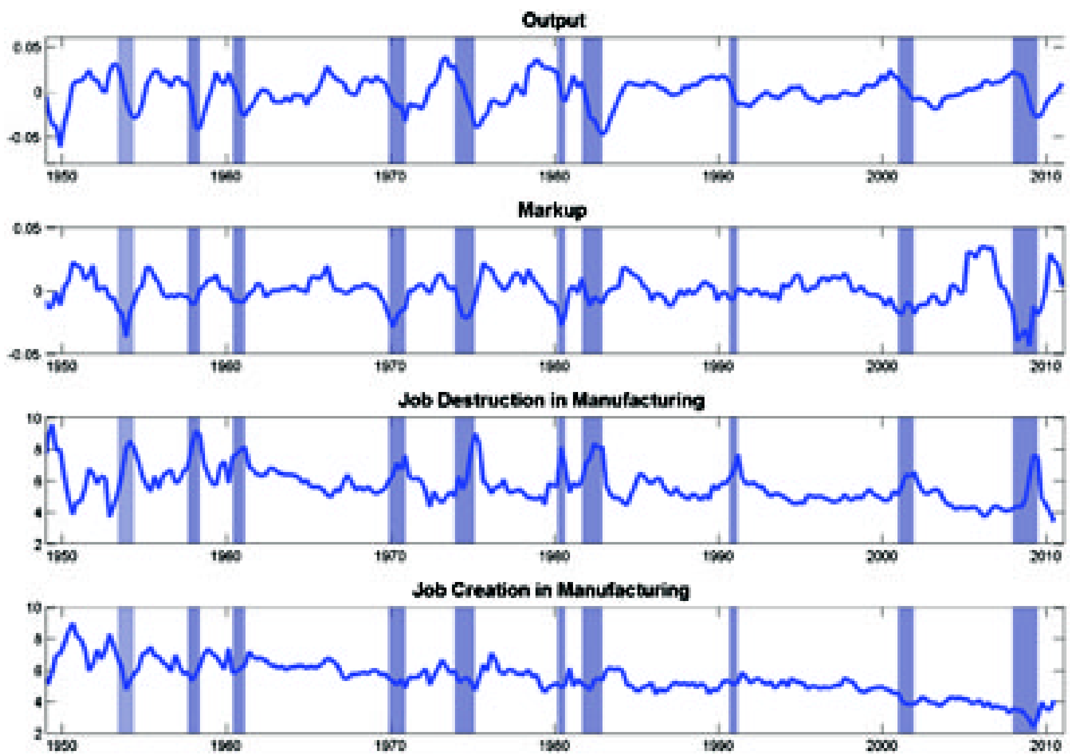

FIGURE 7. **Asymmetries in the Data**

Sources: NIPA, Nekarda and Ramey (2013), Davis et. al. (2006)

As a result, our theory of lagged counter-cyclical markups agrees with the empirical

[37]For a discussion of the use of job creation and destruction rates see Appendix 2.

	$\rho_{\Delta P,\Delta Y}$	$\rho_{X,\Delta Y}$	$\gamma_{\Delta Y}$	$\gamma_{\Delta P}$	γ_X
Data	-.24	-.49	-1.07	1.05	1.12
Model	-.92	-.65	-4.0	3.8	4.0

ρ denotes cross-corellation; γ - skewness; ΔY - growth rate of GDP;

ΔP - lagged growth rate of markups; X - exit rate.

Table 4: **Model Performance**

evidence. This theory implies that markups rise sharply in the aftermath of a recession due to a sharp decline in the number of competitors, and then fall gradually in a boom as new businesses populate the economy. This interpretation is consistent with the pattern of cross-correlations of growth rates of GDP and markups generated by our model (depicted with dashed green line) in Figure 8. For instance, the severe decline in the correlation between markups and GDP at lead one corresponds to the notion that markups peak one quarter following the trough in GDP in the model.

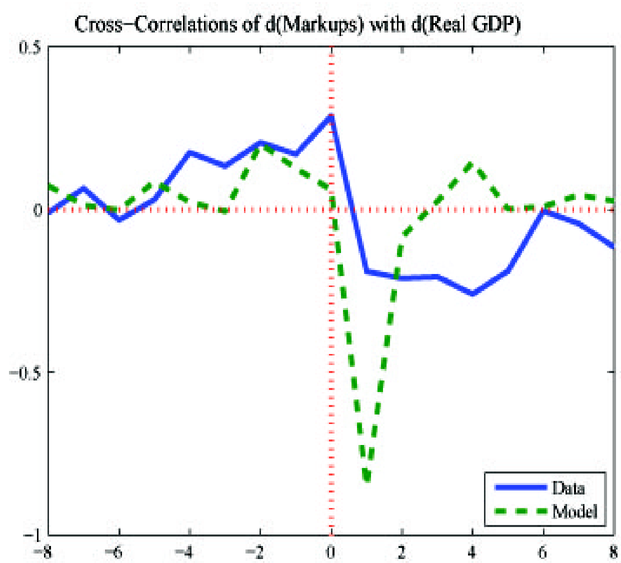

FIGURE 8. **Lead and Lag Properties of Markups.**

Sources: NIPA, Nekarda and Ramey (2013), authors' calculations

4.4 Sensitivity of model's predictions to alternative assumptions

In this section we discuss the sensitivity of our findings to simplifying assumptions that helped in solving the model and increased the transparency of the underlying mechanism. We start by describing how these assumptions affect the model's predictions and performance when compared with U.S. data . We then move on to discussing alternatives to two core modelling assumptions: (1) exogenous entry; (2) the exit rule as deterministic function of expected profits.

The first dimension of U.S. data that our model does not capture involves the length of a recession. Figure 8 indicates that in the U.S. economy it takes on average 4-6 quarters for the survivors to realize that many competitors are gone and that they can start charging higher prices. In our model, this adjustment is accomplished within a single quarter. We could introduce additional inertia into the model by either postulating a slow pass-through of changes in tastes to changes in prices or by assuming a frictional bankrupcy process that delay exits. Both sources of inertia would slow down the recessionary transition without affecting the decision-making process of incumbent firms. The resulting model would be more convoluted than the one presented but it would likely account for a large part of the current discrepancy between the lead-lag predictions of the model and the cyclical properties of the data.

The second dimension of the data that our model does not match quantitatively is the strength of the asymmetry: our model over-predicts the size of the asymmetry in output, markups and exit rates. The quantitative mismatch between these moments in Table 4 is expected, as the model is designed to abstract from many features of the real world that would dampen the asymmetry. Since cyclicality and asymmetry are crucial predictions of our model, we next explore how endogenous entry and stochastic exit would affect these predictions.

4.4.1 Endogenous entry

Motivated by evidence of mild and symmetric fluctuations of entry rates over the U.S. business cycle, we assume in the model that firms enter at an exogenously given constant rate. Alternatively, we could postulate an environment where entry is costly as in Caballero and Hammour (1996), with rationally inattentive firms deciding whether to enter. Given the assumption that new firms get the highest possible valuation v_i, each perspective entrant would base its decision on the comparison of the value of entry determined by expectations of aggregate markups and the costs of entry and information-processing. Since we assume no strategic interactions between incumbents, it is also natural to abstract from strategic interactions between entrants.

In this case, the entry rate would change only *in response* to changes in aggregate conditions, which are in turn determined by incumbent firms' decisions. The relative size of information costs and costs of entry would jointly determine the elasticity of the entry rate to variations in markups. The response of entry would be bigger if entry costs or the costs of processing information are low, and the response would be smaller with high costs. Our baseline model is equivalent to the case where both costs are high. With low costs, the entry rate follows the behavior of aggregate mark-ups: low at the onset of recessions and high in the aftermath of recessions. In this instance, the aggregate fluctuations would be dampened with respect to the baseline model as more businesses would open once there are few firms left in the market. However, since entry is purely a response to decisions of incumbents, this mechanism would only lead to a decrease in the size of the aggregate asymmetry, but would not reverse it or change its shape. Looking at U.S. data that exhibit only mild fluctuations and do not exhibit any asymmetry in entry rates, a calibration of the model would naturally set the entry costs to a high value and produce the same outcome as our baseline model.

4.4.2 Stochastic exit

In the model, we assume that information frictions restrict signals about state variables, but we place no restrictions on information flow involved in making the discrete choice of whether to exit. As a consequence, the exit rule is a deterministic function of a firm's information choice. While this assumption simplifies the computational problem, in this section we argue that allowing for stochastic exit by further constraining information flow would not affect a firm's choice of information structure. However, it might mildly dampen aggregate fluctuations, with business cycles displaying less pronounced peaks than the baseline model.

To see this, note that with stochastic exit, the new choice variable of firm i would be the joint distribution of (1) markups μ and idiosyncratic taste shocks, v_i, denoted as $S_i = \{\mu, v_i\}$; (2) the observable profits, π_i; and (3) the discrete exit choice e_i. Let this new choice variable be denoted by the joint distribution $p(S_i, \pi_i, e_i)$. Let q denote the marginal distribution of exit, i.e., $q \equiv \Pr(e_i = 1)$ and $(1-q) = \Pr(e_i = 0)$. Then, one can rewrite the firm's information structure $p(S_i, \pi_i, e_i)$ as the firm's choice of $(q, 1-q)$ and the simultaneous choice of the conditional distribution $f(S_i, \pi_i | e_i = 0)$.[38] In this setup, the problem solved by the choice of f would be identical to that solved by the choice of p in the baseline model. In addition, incumbent firms would exit by mistake with probability q. The frequency of these erroneous exits would vary over the business cycle in response to changes in expected profits.

This specification shows that the basic mechanism guiding the firm's optimal information choice is unchanged. However, since now the probability of exit is linked to markups via Shannon's channel, the exit probability would be low when profits are high, and increase as profits decline in an expansion. Firms would time exit more precisely in the aftermath of a recession when their forecasts of markups are more precise.

[38] The expectations conditional on exit, $e_i = 1$, are degenerate since in this case profits equal 0 forever.

Similarly, since firms pay little attention and misjudge their profitability when markups are low, there would be an increase in the number of firms exiting with relatively high profits during expansions. These two effects together are likely to dampen the size of peaks and troughs over the business cycle compared with the baseline model. However, they are unlikely to be strong enough to completely offset the asymmetries.

5 Conclusion

This paper shows that information-processing constraints based on rational inattention matter for business cycles analysis. In particular, modelling firms' decisions to exit the market as arising from limited information-processing capacity helps rationalize three main U.S. business cycles regularities: (1) markup growth is asymmetric and countercyclical; (2) markups lag output; (3) firms' exit rates are asymmetric and countercyclical.

The key predictions of the model hinge on the mechanism behind firms' optimal choice of information over the business cycle and the cognitive nature of the attention cost. As far as the mechanism is concerned, we show that firms optimally choose to pay more attention to economic conditions in the aftermath of a recession, when the value of being attentive is superior to the cost. They remain relatively uninformed in economic expansions when acquiring precise signals is not worth their attention.

The cognitive nature of information-processing costs that we postulate is corroborated by experimental evidence in economics, neurology and psychology. Survey evidence on attention allocation to economic conditions of entrepreneurs over the business cycles documented by Coibion and Gorodnichenko (2012) provides direct empirical support to our mechanism by confirming its predictions for the allocation of attention over the business cycle.

Our mechanism is capable of yielding starkly different policy implications compared to standard business cycle models. To smooth business cycles, a planner might be unable to acquire and process information about product demand to command firm exit. Instead, he might want to provide businesses with (low-bit) systematic information about aggregate variables to foster coordination.

Alternative policies of managing exits can smooth the cycle at the cost of slowing down long-term economic growth. This prediction is consistent with multiple central planning experiments, undertaken in different parts of the world. Our model has the potential to provide estimates of how such policies affect long-term economic growth. However, due to space limitations, we leave all these normative questions for future research.

References

[1] Abbritti, Mirko and Stephan Fahr. 2013. "Downward wage rigidity and business cycle asymmetries," Journal of Monetary Economics, 60(7): 871-886.

[2] Angeletos, George-Marios, and Alessandro Pavan. 2007. “Efficient Use of Information and Social Value of Information.” Econometrica, 75(4): 1103–1142.

[3] Basu, Susanto, 1996. "Procyclical Productivity: Increasing Returns or Cyclical Utilization?" The Quarterly Journal of Economics, 111(3): 719-51.

[4] Bilbiie, Florin, Fabio Ghironi, and Marc Melitz. 2012. “Endogenous Entry, Product Variety, and Business Cycles.” Journal of Political Economy, 120(2): 304–345.

[5] Caballero, Ricardo J, and Mohamad L Hammour. 1996. “On the Timing and Efficiency of Creative Destruction.” The Quarterly Journal of Economics, 111(3): 805–52.

[6] Canova, Fabio. 1998. "Detrending and business cycle facts," Journal of Monetary Economics, 41(3): 475-512.

[7] Chamley, Christophe, and Douglas Gale. 1994. "Information Revelation and Strategic Delay in a Model of Investment." Econometrica, 62(5): 1065–1085.

[8] Chatterjee, Satyajit, and Russell Cooper. 1993. "Entry and Exit, Product Variety and the Business Cycle." National Bureau of Economic Research, Inc NBER Working Papers 4562.

[9] Cheremukhin, Anton, Anna Popova and Antonella Tutino. 2014. "A Theory of Discrete Choice with Information Costs." mimeo.

[10] Cociuba, Simona E., Edward C. Prescott and Alexander Ueberfeldt. 2012. "U.S. Hours and Productivity Behavior Using CPS Hours Worked Data: 1947-III to 2011-IV", mimeo.

[11] Coibion, Oliver and Yuriy Gorodnichenko, 2012. "What Can Survey Forecasts Tell Us About Informational Rigidities?", Journal of Political Economy 120(1), 116-159.

[12] Cooper, Russell, John Haltiwanger, and Jonathan L. Willis. 2007. "Search frictions: Matching aggregate and establishment observations." Journal of Monetary Economics, 54(Supplement): 56–78.

[13] Davis, Steven J., R. Jason Faberman, and John Haltiwanger. 2006. "The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links." Journal of Economic Perspectives, 20(3): 3–26.

[14] Devereux, Michael B., Allen C. Head, and Beverly J. Lapham. 1996. "Aggregate fluctuations with increasing returns to specialization and scale." Journal of Economic Dynamics and Control, 20(4): 627–656.

[15] Edmond, Chris, and Laura Veldkamp. 2009. “Income dispersion and counter-cyclical markups.” Journal of Monetary Economics, 56(6): 791–804.

[16] Francois, Patrick, and Huw Lloyd-Ellis. 2003. “Animal Spirits Through Creative Destruction.” American Economic Review, 93(3): 530–550.

[17] Gabaix, Xavier. 2011. “The Granular Origins of Aggregate Fluctuations.” Econometrica, 79(3): 733–772.

[18] Hellwig, Christian, and Laura Veldkamp. 2009. “Knowing What Others Know: Coordination Motives in Information Acquisition.” Review of Economic Studies, 76(1): 223–251.

[19] Jaimovich, Nir. 2007. “Firm dynamics and markup variations: Implications for sunspot equilibria and endogenous economic fluctuations.” Journal of Economic Theory, 137(1): 300 –325.

[20] Jaimovich, Nir, and Max Floetotto. 2008. “Firm dynamics, markup variations, and the business cycle.” Journal of Monetary Economics, 55(7): 1238–1252.

[21] Jovanovic, Boyan. 2006. “Asymmetric Cycles.” Review of Economic Studies, 73(1): 145–162.

[22] Kahneman, D. 1973. "Attention and effort." Englewood Cliffs, NJ: Prentice-Hall.

[23] Lorenzoni, Guido. 2009. “A Theory of Demand Shocks.” American Economic Review, 99(5): 2050–84.

[24] Loungani, P., H. Stekler, and N. Tamirisa, 2013. "Information rigidity in growth forecasts: Some cross-country evidence," International Journal of Forecasting, 29(4): 605-621.

[25] Lucas, R.E., 1972, “Expectations and the Neutrality of Money,” Journal of Economic Theory, Vol. 4 (April), pp. 103–24.

[26] Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." American Economic Review, 99(3): 769–803.

[27] Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve." The Quarterly Journal of Economics, 117(4): 1295–1328.

[28] Matsuyama, Kiminori. 1999. "Growing Through Cycles." Econometrica, 67(2): 335–348.

[29] McKay, Alisdair and Ricardo Reis. 2008. "The brevity and violence of contractions and expansions," Journal of Monetary Economics, 55(4): 738-751.

[30] Murto, Pauli, and Juuso Valimaki. 2011. "Learning and Information Aggregation in an Exit Game." Review of Economic Studies, 78: 1426–1461.

[31] Myatt, David P, and ChrisWallace. 2012. "Endogenous Information Acquisition in Coordination Games." Review of Economic Studies, 79(1): 340-374.

[32] Nekarda, Christopher J, and Valery A Ramey. 2013. "The Cyclical Behavior of the Price-Cost Markup."National Bureau of Economic Research, Inc NBER Working Papers 19099.

[33] Psaradakis, Zacharias and Martin Sola. 2003. "On detrending and cyclical asymmetry," Journal of Applied Econometrics, 18(3): 271-289.

[34] Rotemberg, Julio J, and Garth Saloner. 1986. "A Supergame-Theoretic Model of Price Wars during Booms." American Economic Review, 76(3): 390–407.

[35] Rotemberg, Julio J., and Michael Woodford. 1999. "The cyclical behavior of prices and costs."1: 1051–1135.

[36] Shleifer, Andrei. 1986. "Implementation Cycles." Journal of Political Economy, 94(6): 1163–90.

[37] Sims, Christopher A. 2003. “Implications of rational inattention.”Journal of Monetary Economics, 50(3): 665–690.

[38] Sims, Christopher A. 2006. “Rational Inattention: Beyond the Linear-Quadratic Case.”American Economic Review, 96(2): 158–163.

[39] Sperling, George, and Barbara A. Dosher. 1986. "Strategy and optimization in human information processing." In K. Boff, L. Kaufman, & J. Thomas (Eds.), Handbook of perception and human performance: Vol. 1. Sensory processes and perception. New York: Wiley. pp. 2-1 to 2-65.

[40] Tutino, Antonella. 2013. "Rationally inattentive consumption choices," Review of Economic Dynamics, 16(3): 421-439.

[41] Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2006. “Learning asymmetries in real business cycles.”Journal of Monetary Economics, 53(4): 753–772.

[42] Woodford, Michael. 2013. "Macroeconomic Analysis Without the Rational Expectations Hypothesis," Annual Review of Economics, 5(1): 303-346.

[43] Zeira, Joseph. 1994. “Informational Cycles.”Review of Economic Studies, 61(1): 31–44.

Appendix NOT FOR PUBLICATION

Appendix 1: Data Sources

Data	Frequency, Source
1. Average Hours Worked	at annual rate, SA, Q:1947:III-2011:IV, Cociuba, Prescott and Ueberfeldt (2010) sites.google.com/site/simonacociuba/research
2. Labor share	nonfin. corp. business, SA, Q:1947:I-2011:IV, FRED2: PRS88003173
3. Price-cost Markup	Total Private, SA, Q: 1948:1-2010:4, Nekarda and Ramey (2013)
4. Price-cost Markup	Manufacturing, SA, Q: 1948:1-2010:4 Nekarda and Ramey (2013)
5. Real GDP	Index w.r.t. 2005, SA, Q: 1947:1-2010:4, FRED2: GDPC1
6. Industrial Production	Index w.r.t. 2007, SA, Q: 1947:1-2010:4, FRED2: INDPRO
7. Gross Job Gains	Manufacturing, SA, Q:1992:1-2010:4, BLS BDS0000000000100030110001RQ5
8. Gross Job Losses	Manufacturing, SA, Q:1992:3-2010:4, BLS BDS0000000000100030110004RQ5
9. Gross Job Gains	Total Private, SA, Q:1992:3-2010:4, BLS BDS0000000000000000110001RQ5
10. Gross Job Losses	Total Private, SA, Q:1992:3-2010:4 BDS0000000000000000110004RQ5
11. Job Gains	Manuf., SA, Q:1947:1-2005:1, Davis, Faberman and Haltiwanger (2006)

Data	Frequency, Source
12. Job Losses	Manuf., SA, Q:1947:1-2005:1,www.nber. .org/data/job-creation-and-destruction.html
13. Opening Establishments	Total Private, Q:1992:3-2010:4 BDS0000000000000000120003RQ5
14. Closing Establishments	Total Private, Q:1992:3-2010:4 BDS0000000000000000120006RQ5
15. Business Failures	A: 1940:2010, Dunn&Bradstreet, Statistical Abstract of the United States
16. New Incorporations	A: 1948:2010, Dunn&Bradstreet, www.census.gov/prod/www/abs/statab.html

The first two measures of markups come directly from Nekarda and Ramey (2013). To compute the inverse share of labor we took $-\log(\text{Labor Share})$. To adjust for overhead labor we took $-\log(\text{Labor Share}) - \frac{0.28}{1-0.28}\log(\text{Hours})$. We extend the series for job creation and job destruction in manufacturing from Davis, Faberman and Haltiwanger (2006) past 2005 by merging them with job destruction and creation rates in manufacturing from the BLS. We hp-filter business failures and new incorporations data with parameter 100 to remove long-run trends.

Appendix 2: Robustness of Empirical Evidence

Robustness of Fact 1

In the main text, we reported results for Granger causality tests for GDP and markups in Table 1. In Tables 5-7 we show that our findings are robust to alternative detrending procedures, such as no detrending, a linear or quadratic trend, or an hp-filter trend. Our results are very robust independent of the specification that one might prefer.

Number of lags	1	2	3	4	5	6
Change in logs:	GDP Granger causes Markups					
F-statistic	10.8***	8.6***	8.7***	11.4***	8.2***	6.2***
	Markups do not Granger cause GDP					
F-statistic	0.2	2.5*	1.4	1.4	0.9	2.1*
Levels, no detrending:	GDP Granger causes markups					
F-statistic	7.0**	8.2***	6.3***	6.2***	9.6***	7.4***
	Markups do not Granger cause GDP					
F-statistic	7.6**	0.9	0.9	1.2	2.5**	1.9
Levels, log-linear trend:	GDP Granger causes markups					
F-statistic	28.6***	13.4***	9.3***	7.8***	9.2***	7.2***
	Markups do not Granger cause GDP					
F-statistic	7.2**	0.0	1.1	0.6	0.7	0.6
Levels, log-quadratic trend:	GDP Granger causes markups					
F-statistic	28.3***	13.8***	8.3***	6.4***	7.3***	6.0***
	Markups do not Granger cause GDP					
F-statistic	14.1***	1.1	0.8	0.7	0.9	0.8
Levels, HP(100k)-trend:	GDP Granger causes markups					
F-statistic	36.4***	17.2***	10.1***	7.2***	7.5***	6.2***
	Markups do not Granger cause GDP					
F-statistic	17.7***	1.8	0.7	0.7	0.8	0.8

Frequency: Quarterly 1948:1-2010:4. Observations: 251. *, ** and *** show significance at 10%, 5% and 1%.

Sources: NIPA; Nekarda and Ramey (2013).Markups are measured for the whole economy .

Table 5: **Cyclical Properties of Markups (T)**

Number of lags	1	2	3	4	5	6
Change in logs:	GDP Granger causes Markups					
F-statistic	0.4	4.7**	7.9***	9.1***	7.6***	6.7***
	Markups do not Granger cause GDP					
F-statistic	0.8	0.3	1.0	1.1	1.4	1.5
Levels, no detrending:	GDP Granger causes markups					
F-statistic	5.7**	3.4**	4.3**	7.6***	7.8***	7.4***
	Markups do not Granger cause GDP					
F-statistic	1.4	1.4	1.0	2.2	2.9**	3.4**
Levels, log-linear trend:	GDP Granger causes markups					
F-statistic	16.9***	10.1***	8.1***	8.0***	8.4***	7.2***
	Markups do not Granger cause GDP					
F-statistic	0.1	2.4*	2.2*	1.9	1.7	1.9*
Levels, log-quadratic trend:	GDP Granger causes markups					
F-statistic	10.0***	7.4***	5.8***	6.0***	6.7***	5.9***
	Markups do not Granger cause GDP					
F-statistic	0.6	0.9	0.7	0.9	1.0	1.4
Levels, HP(100k)-trend:	GDP Granger causes markups					
F-statistic	12.7***	11.4***	8.0***	6.3***	6.4***	5.5***
	Markups do not Granger cause GDP					
F-statistic	4.0*	1.1	0.6	0.7	0.7	1.2

Frequency: Quarterly 1948:1-2010:4. Observations: 251. *, ** and *** show significance at 10%, 5% and 1%.

Sources: NIPA; Nekarda and Ramey (2013).Markups are measured for the manufacturing sector.

Table 6: **Cyclical Properties of Markups (M)**

Number of lags	1	2	3	4	5	6
Change in logs:	GDP Granger causes Markups					
F-statistic	9.8***	8.5***	7.5***	6.6***	5.0***	4.6***
	Markups do not Granger cause GDP					
F-statistic	0.2	2.8*	1.9	2.2	1.1	1.0
Levels, no detrending:	GDP Granger causes markups					
F-statistic	5.0**	8.2***	6.9***	5.7***	4.9***	4.2***
	Markups do not Granger cause GDP					
F-statistic	1.4	0.3	1.0	0.5	0.9	0.7
Levels, log-linear trend:	GDP Granger causes markups					
F-statistic	30.0***	15.7***	10.5***	8.4***	7.4***	6.7***
	Markups do not Granger cause GDP					
F-statistic	7.0**	0.4	1.5	1.2	1.6	1.0
Levels, log-quadratic trend:	GDP Granger causes markups					
F-statistic	25.3***	13.4***	8.8***	7.1***	6.1***	5.9***
	Markups do not Granger cause GDP					
F-statistic	14.9***	2.4*	1.7	1.3	1.5	0.9
Levels, HP(100k)-trend:	GDP Granger causes markups					
F-statistic	20.9***	11.1***	7.3***	5.9***	4.9***	4.9***
	Markups do not Granger cause GDP					
F-statistic	15.9***	3.3**	1.8	1.4	1.4	1.0

Frequency: Quarterly 1959:1-2010:4. Observations: 207. *, ** and *** show significance at 10%, 5% and 1%.

Sources: NIPA. Markups are measured as inverse labor share.

Table 7: **Cyclical Properties of Markups (S)**

Robustness of Fact 2.

In Table 8 we demonstrate the robustness of our results regarding the skewness of GDP and markups. First, we find that our results are robust to removing a quadratic trend from growth rates, before or after taking changes in logs. Second, we show that the magnitude of skeweness depends on the measure of markup being used (markups for whole economy and manufacturing from Nekarda and Ramey (2013), inverse labor share). However, the sign of skewness is robust.

Third, looking at time series going further back in time (compared with Table 2) it is easy to spot a break point in the estimates of skewness. We believe that this is driven by the apparent reduction in the volatility of these series post-1980. We find that for the 1947-1980 period the log changes in GDP and markups are more volatile and the magnitude of skewness is lower. This reduction in volatility after 1980 might be due to the so-called Great Moderation, or simply to improvements in measurement. Either way, the estimates for the post-1980 period have the same magnitude and sign as those reported in Table 2. We also find that excluding the Great Recession has a negligible effect on the estimates.

The finding of high positive skewness of changes in markups sheds new light on the controversial behavior of markups over the business cycle. While Rotemberg and Woodford (1999) find markups to be counter-cyclical, Nekarda and Ramey (2013) recently argued that markups are virtually a-cyclical.

These findings rely on computing correlations of de-trended time series, which makes them particularly sensitive to the statistical model for the mean used to de-trend markups and GDP. The two main facts we document in this paper are robust to this concern because many of them can be established purely based on properties of first-differenced data. This intuition has been discussed in different contexts, among others, by Canova (1998), Psaradakis and Sola (2003), Veldkamp and Niewenburg (2006). In addition, we have demonstrated that they are immune to the choice

Series	GDP	μ^T	μ^M	μ^S
Filtering method		1947-2010		
1	-.08	+.68***	+.39**	+.77***
2	-.25*	+.65***	+.32**	+.76***
3	-.29*	+.64***	+.25*	+.75***
		1947-1980		
1	-.08	+.26	+.19	+.92***
2	-.11	+.31	+.19	+.87***
3	-.09	+.30	+.19	+.84***
		1981-2007		
1	-.88***	+1.06***	+.41*	+.67***
2	-1.01***	+1.02***	+.29	+.67***
3	-.53**	+.70***	+.11	+.71***

μ^T,μ^M,μ^S,represent measures of markups for the whole economy (T), manufacturing (M), inverse labor share (S).

Filtering methods include (1) change in non-detrended log, (2) change in detrended log, (3) detrended change in log, where detrending removes quadratic trend. Frequency: Quarterly, *, ** and *** show significance at 10%, 5% and 1%.

Sources: NIPA; Nekarda and Ramey (2013).

Table 8: **Robustness for Skewness of Markups in the U.S.**

of de-trending procedure.

Additional evidence on Markups, Entry and Exit

We explore the robustness of our empirical findings along four additional dimensions. First, in order to extend the analysis of entry and exit to a longer time period, we use job destruction and creation rates in manufacturing since 1948 from Davis et al. (2006) as a proxy for exit and entry rates in manufacturing.

In our theoretical model, there is no distinction between firms and establishments, so firm entry and exit is synonymous to an opening or closing of an establishment.

	d(GDP)	d(Mrk(+3))	Exit	Entry	Job Des	Job Cre
Volatility	1.0%	0.82%	5.3%	4.0%	5.1%	3.1%
Skewness	-1.07***	1.05***	1.12***	-0.09	1.06***	-0.70*
			Cross-corellations			
d(Mrk(+3))	$-.24^{**}$					
Exit rate	$-.49^{***}$	$+.19^{*}$				
Entry rate	$+.23^{**}$	$-.21^{**}$	$-.42^{***}$			
Job Des	$-.51^{***}$	$+.35^{***}$	$+.71^{***}$	$-.13$		
Job Cre	$+.49^{***}$	$-.27^{**}$	$-.26^{**}$	$+.65^{***}$	$-.35^{***}$	

Frequency: Quarterly 1992:3-2010:4. Observations: 71. *, **, *** show significance at 10%, 5%, 1%.

d(GDP) and d(Mrk(+3)) denote percentage changes in GDP and 3-quarter lead of markups respectively.

Job Des and Job Cre denote job destruction and creation rates. Sources: NIPA; Nekarda and Ramey (2013); Business Employment Dynamics.

Table 9: **Markups, Exit and Entry rates in the U.S.**

To check the robustness of these properties we build on the findings of Jaimovich and Floetotto (2008) who show that closing and opening establishments account for a large fraction of cyclical variations in job destruction and job creation rates. As we document in Table 9 (an extension of Table 2), the correlation between exit and job destruction rates is 0.71, and the correlation between entry and job creation rates is 0.65. This suggests that job destruction and creation rates are informative proxies for the numbers of entering and exiting establishments in the U.S. economy. The properties we document in the main text are consistent with the fact that volatility of job destruction rates has been more than 50% higher that the volatility of job creation rates. The correlation of the establishment entry rate with GDP has been at 0.23 with skewness at -0.09. Overall, job destruction and creation rates show similar patterns to exit and entry rates, with the behavior of job creation slightly more cyclical than entry.

	d(Ind.Prod)	d(Markup(+3))	Job Destr	Job Cre
Volatility			16.0%	9.2%
Skewness	-0.20^{*}	$+0.39^{***}$	$+1.11^{***}$	$+0.01$
		Cross-corellations		
d(Markup(+3))	-0.18^{***}			
Job Destruction	-0.73^{***}	$+0.28^{***}$		
Job Creation	$+0.75^{***}$	-0.16^{***}	-0.59^{***}	

Frequency: Quarterly 1948:1-2010:4. Observations: 251. *, **, *** show significance at 10%, 5%, 1%.

Source: Board of Governors; Nekarda and Ramey (2013); Davis, Faberman and Haltiwanger (2006).

Table 10: **Markups, Exit and Entry rates in U.S. Manufacturing**

	d(GDP)	d(Markup(+1))	Bus. Fail.	New Inc.
Skewness	-0.38^{*}	$+0.16$	$+0.94^{***}$	$+0.03$
		Cross-corellations		
d(Markup(+1))	-0.56^{***}			
Business Failures	-0.21^{**}	$+0.23^{**}$		
New Incorporations	$+0.25^{**}$	-0.22^{**}	-0.02	

Frequency: Annual 1948-2010. Observations: 61. *, ** and *** show significance at 10%, 5% and 1%.

Sources: NIPA; Nekarda and Ramey (2013); Dunn&Bradstreet.

Table 11: **Markups, Business Failures and New Incorporations**

We combine these series with data on markups in manufacturing from Nekarda and Ramey (2013) and with the index of industrial production as a proxy for manufacturing output. Table 10 shows that these data are consistent with our facts. Lagged markups are counter-cyclical and exhibit a positive growth asymmetry. Job destruction is much more volatile than job creation. Job destruction is counter-cyclical and highly positively skewed, while job creation is pro-cyclical and symmetric.

Second, we use annual data on the number of business failures and new incorporations in the U.S. economy from the Statistical Abstract of the United States to construct measures of exit and entry for a longer time period. To eliminate trends in these series we use the standard hp-filter. Table 11 shows that business failures are counter-cyclical and highly positively skewed at business cycle frequencies, while new incorporations are pro-cyclical and symmetric.

Third, we have looked at data from the Business Dynamics Statistics (BDS) database constructed by the Census Bureau. This database contains numbers of firms and establishments that enter and exit, and the numbers of jobs that these firms/establishments create and destroy. The main limitation of this source is that the data are collected on an annual basis, which constrains the number of points for each series to just 34 (since 1977), and makes it impossible to evaluate correlations at business-cycle (quarterly) frequencies.

We computed analogs of Table 9 for both the whole economy and the manufacturing sector. We report properties of the number of entering/exiting firms/establishments, and of job creation/destruction by entering/exiting/continuing establishments. We report the skewness of these series as well as their correlations with changes in logs of GDP and markups in Tables 12-13. Few of the results are statistically significant due to the short length of the series. The only conclusion one can make is that the results are qualitatively consistent with our story (the signs are largely the same). It is largely impossible to make inferences about the magnitudes or the differences between the behavior of establishments and firms.

Fourth, we consider two alternative measures of markups, the inverse of the labor share in the nonfinancial corporate sector used by Rotemberg and Woodford (1999)[39], and the labor share adjusted for overhead labor as proposed by Basu (1996). We compare their statistical properties to the two measures of markups provided by Nekarda and Ramey (2013), for the whole economy and for the manufacturing sector.

[39] Rotemberg and Woodford (1999) argue that the labor share in the corporate sector is the preferred measure because the labor share in the whole economy includes government services which are not sold in the market. The inclusion of the financial sector does not make a quantitative difference.

	d(GDP)	d(Mrk(+1))	Firm Entry	Est. Entry	JC: Entrng	JC: Cont
			Whole Economy			
Skewness	-.87**	+.07	-1.15**	-.15	-.12	-.83*
d(GDP)		-.48***	+.49***	+.22	+.14	+.69***
d(Mrk(+1))			-.11	-.31*	-.23	-.45**
			Manufacturing Sector			
Skewness	-.87**	+.68	-1.49***	-1.08**	+.70*	-.09
d(GDP)		-.38**	+.57***	+.29*	+.04	+.70***
d(Mrk(+1))			+.05	-.19	+.09	-.39**

Frequency: Annual 1977-2010. Observations: 34. *, **, *** show significance at 10%, 5%, 1%.

Source: NIPA, Census Bureau; Nekarda and Ramey (2013).

Table 12: **Properties of Entry and Job Creation using BDS data.**

	d(GDP)	d(Mrk(+1))	Firm Exit	Est. Exit	JD: Extng	JD: Cont
			Whole Economy			
Skewness	-.87**	+.07	+.66	+.82*	+.43	+.83*
d(GDP)		-.48***	-.14	-.23	+.04	-.59***
d(Mrk(+1))			+.05	+.01	-.00	+.45**
			Manufacturing Sector			
Skewness	-.87**	+.68	-0.55	+.77*	+.84**	+.80*
d(GDP)		-.38**	-.07	-.22	-.14	-.62***
d(Mrk(+1))			+.36**	+.29*	+.19	+.45**

Frequency: Annual 1977-2010. Observations: 34. *, **, *** show significance at 10%, 5%, 1%.

Source: NIPA, Census Bureau; Nekarda and Ramey (2013).

Table 13: **Properties of Exit and Job Destruction using BDS data.**

	d($\mu^T(+3)$)	d($\mu^M(+3)$)	d($\mu^S(+3)$)	d($\mu^H(+3)$)
Skewness	1.05***	0.28**	0.39*	0.30*
	Cross-corellations			
d($\mu^M(+3)$)	+.65***			
d($\mu^S(+3)$)	+.75***	+.67***		
d($\mu^H(+3)$)	+.72***	+.63***	+.92***	
d(Ind.prod.)	−.24**	−.21**	−.17*	−.28**
Exit rate	+.19*	+.12*	+.06	+.14*
Entry rate	−.21**	−.19*	−.22**	−.22**
Job Destruction	+.35***	+.22**	+.22**	+.27**
Job Creation	−.27**	−.15*	−.30***	−.29**

μ^T,μ^M,μ^S,μ^H,represent measures of markups for the whole economy (T), manufacturing (M), inverse labor share (S), including the adjustment for overhead hours (H).

Frequency: Quarterly, 1992:3-2010:4. Observations: 71. *, ** and *** show significance at 10%, 5% and 1%. Sources: NIPA; Nekarda and Ramey (2013); Business Employment Dynamics.

Table 14: **Alternative Measures of Markups in the U.S.**

Table 13 shows that all four measures of markups have positively skewed growth rates, are positively correlated with exit rates, and negatively correlated with entry rates. Thus, the facts we document are robust, whether one just considers the inverse of the labor share, or additionally takes into account overhead labor, or adjusts for the difference between marginal and average costs following Nekarda and Ramey (2013). Correcting the markup measure for additional margins of adjustment, as described by Rotemberg and Woodford (1999), is known to amplify the counter-cyclicality of markups, thereby strengthening our results.

Appendix 3: Substitution between products

Here we describe a generalization of the model to the case of non-perfect substitution. The representative household trades off leisure for consumption, maximizing a standard utility function:

$$\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\gamma} - 1}{1-\gamma} - E_t \right), \tag{23}$$

with respect to the supply of labor, E_t, and a Dixit-Stiglitz consumption aggregator, C_t, which weights differentiated products, q_{it}, by their tastes, v_{it}:

$$C_t = \left(\sum_{i=1}^{K_t} v_{it}^{\frac{1}{\sigma}} q_{it}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{24}$$

The household owns all the firms in the economy. It spends wage and profit income on contemporaneous consumption, maximizing utility subject to a budget constraint:

$$\sum_{i=1}^{K_t} p_{it} q_{it} = w_t E_t + \sum_{i=1}^{K_t} \pi_{it}. \tag{25}$$

Maximization yields the following first-order condition, which in a Walrasian equilibrium determines the price for each good indexed by i:

$$p_{it} = w_t C_t^{-\gamma} \left(\frac{C_t}{q_{it}} \right)^{\frac{1}{\sigma}} v_{it}^{\frac{1}{\sigma}}, \tag{26}$$

which is driven by variations in idiosyncratic tastes v_{it}. Variations in tastes are the only source of uncertainty in the economy.

We define a consumption price index as follows:

$$P_t = \frac{1}{C_t} \sum_{i=1}^{K_t} p_{it} q_{it} = w_t C_t^{-\gamma} \tag{27}$$

The economy is populated by K_t firms, which profit from producing and selling differentiated products q_{it} at price p_{it}. In addition to wages, a firm pays a fixed cost of operating the technology, f:

$$\pi_{it} = p_{it}q_{it} - w_t l_{it} - f \tag{28}$$

Firms use identical production functions, which are linear in labor inputs, l_{it} :

$$q_{it} = Al_{it} \tag{29}$$

Within each period firms maximize profits (28) with respect to output, $q_{i,t}$, and labor input, $l_{i,t}$, subject to the production function (29) and given the individual demand curve (26). The first order condition of the firm pins down the optimal level of output as a function of the idiosyncratic shock, $v_{i,t}$:

$$\frac{q_{it}}{A} = \frac{\sigma - 1}{\sigma} C_t^{\frac{1}{\sigma} - \gamma} v_{it}^{\frac{1}{\sigma}} q_{it}^{1 - \frac{1}{\sigma}} \tag{30}$$

We substitute output as a function of taste from (30) into (24) to show how tastes determine consumption and prices:

$$P_t = C_t^{-\gamma} = \frac{1}{A} \frac{\sigma}{\sigma - 1} \left(\sum_{i=1}^{K_t} v_{it} \right)^{-\frac{1}{\sigma - 1}} \tag{31}$$

Without loss of generality normalize operating cost, f, to one and let the wage, w_t, be the numeraire. The expression for profits then simplifies to:

$$\pi_{it} = \mu_{t+1} v_{it} - 1. \tag{32}$$

where the aggregate markup, μ_{t+1}, is each firm's sufficient statistic, which characterizes the aggregate state of the economy:

$$\mu_{t+1} = A^{\gamma - 1} \frac{(\sigma - 1)^{\gamma - 1}}{\sigma^{\gamma}} \left(\sum_{i=1}^{K_t} v_{i,t} \right)^{\frac{\gamma - \sigma}{\sigma - 1}} \tag{33}$$

This functional form is equivalent to the one presented in the text, except that γ is substituted by $\frac{\gamma - \sigma}{1 - \sigma}$ as long as $\gamma \approx 1$, which we assume in the calibration. Depending on the degree of substitution, $\sigma > 0, \sigma \neq 1$, this expression determines the behavior of the economy instead of γ. The rest of the simulation procedure remains intact. In the limit, as $\sigma \to \infty$, products become perfect substitutes, and we arrive at a model similar to that presented in the text. We work with the simpler version for the purpose of transparency.

Appendix 4: Properties of Skewness

Note that skewness of a sum of independent random variables can be expressed as follows:

$$\gamma_{\Sigma X_i} = \frac{E\left[(\Sigma X_i - EX_i)^3\right]}{\left(\sigma_{\Sigma X_i}\right)^3} = \frac{\Sigma E(X_i - EX_i)^3}{\left(\sigma_{\Sigma X_i}\right)^3} = \frac{\Sigma_i\left[(VarX_i)^{\frac{3}{2}} SkewX_i\right]}{(\Sigma_i VarX_i)^{\frac{3}{2}}}$$

Let output in the economy be a sum of i.i.d. outputs of individual firms: $Y_t = \Sigma S_i$. Then, aggregate GDP growth follows

$$\Delta y = \frac{\Delta Y}{Y} = \Sigma \frac{S_i}{Y} \sigma_i \varepsilon_{it} = \Sigma s_i \varepsilon_{it}$$

Let for simplicity firms have the same volatility σ and skewness γ, then skewness of growth is:

$$\gamma_{\Delta y} = \gamma_{\Sigma s_i \varepsilon_{it}} = \frac{\Sigma_i\left[(VarX_i)^{\frac{3}{2}} SkewX_i\right]}{(\Sigma_i VarX_i)^{\frac{3}{2}}} = \frac{\Sigma_i\left[\left(s_i^2\sigma^2\right)^{\frac{3}{2}} Skew(s_i\varepsilon_i)\right]}{\left(\sigma^2\Sigma_i s_i^2\right)^{\frac{3}{2}}} = \frac{\Sigma_i s_i^3 \gamma_i}{\left(\Sigma_i s_i^2\right)^{\frac{3}{2}}}$$

Using the properties of a power distribution:

$$\begin{aligned}\Sigma_i s_i^2 &\sim \tfrac{1}{N^2}\Sigma\left(\tfrac{i}{N}\right)^{-\frac{2}{\zeta}} = N^{\frac{2}{\zeta}-2}\left(\Sigma i^{-\frac{2}{\zeta}}\right)\\ \Sigma_i s_i^3 &\sim \tfrac{1}{N^3}\Sigma\left(\tfrac{i}{N}\right)^{-\frac{3}{\zeta}} = N^{\frac{3}{\zeta}-3}\left(\Sigma i^{-\frac{3}{\zeta}}\right)\end{aligned}$$

Hence, we can express the skewness of GDP as the product of skewness of processes for individual firms multiplied by a ratio of finite sums:

$$\gamma_{\Delta y} = \gamma \frac{\Sigma_i s_i^3}{\left(\Sigma_i s_i^2\right)^{\frac{3}{2}}} \sim \gamma \frac{\left(\frac{1}{N}\right)^3 N^{\frac{3}{\zeta}}}{\left(\frac{1}{N}\right)^3 N^{\frac{3}{\zeta}}} \frac{\left(\Sigma i^{-\frac{3}{\zeta}}\right)^{\frac{1}{2}}}{\left(\Sigma i^{-\frac{2}{\zeta}}\right)^{\frac{3}{2}}} = \gamma\left(\frac{\Sigma i^{-\frac{3}{\zeta}}}{\left(\Sigma i^{-\frac{2}{\zeta}}\right)^3}\right)^{\frac{1}{2}}$$

If Zipf's law holds, $\zeta = 1$ (Pareto distribution), then $\gamma_{\Delta y} \sim 0.52\gamma$. If we adopt the standard estimate for the US economy, $\zeta = 1.055$, the result is little changed: $\gamma_{\Delta y} \sim 0.48\gamma$. However, if we assume the other extreme $\zeta = 2$ (no fat tails), then $\gamma_{\Delta y} \sim 0.06\gamma$.

The key message from this derivation is that the number of firms has no effect on the asymptotic skewness of GDP growth. A power law sized distribution reduces skewness of aggregate fluctuations by a constant, independent of the exact number of firms. A similar argument can be made for cross-correlations.

Appendix 5: Comparison with the full-information model

To compare the behavior of the information constrained economy with the full information economy, we first plot the paths of output and markups, entry and exit rates for a sample of 50 year periods for each of the two models. Here we adopt a simple definition of a recession - an event when the number of firms drops by at least 20 percent in a single period.[40] The top panels of Figures 9A and 9B below show the simulated paths with shaded recessions.

The bottom panels of Figures 9A and 9B compare the laws of motion of the aggregate markup for the two models. Figure 9B demonstrates, that in an inattentive economy the aggregate markup has two potential behaviors. It either drifts down slowly, or jumps up sharply. Outliers above the diagonal are more common than below the diagonal. In the full information economy, off-diagonal behavior is much less pronounced.

[40] This roughly corresponds to a 3-5% drop in GDP in a world with 200 firms.

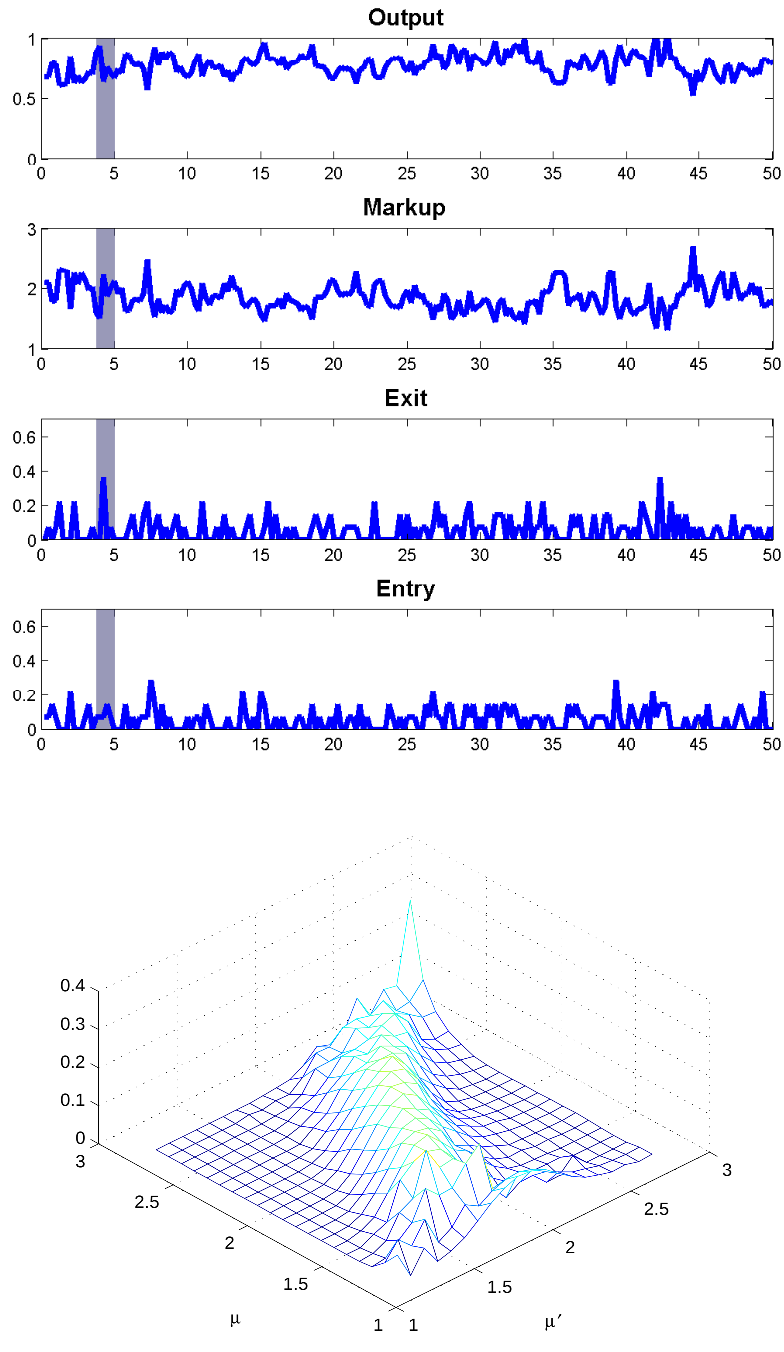

Figure 9A. **Model with Full Information.**

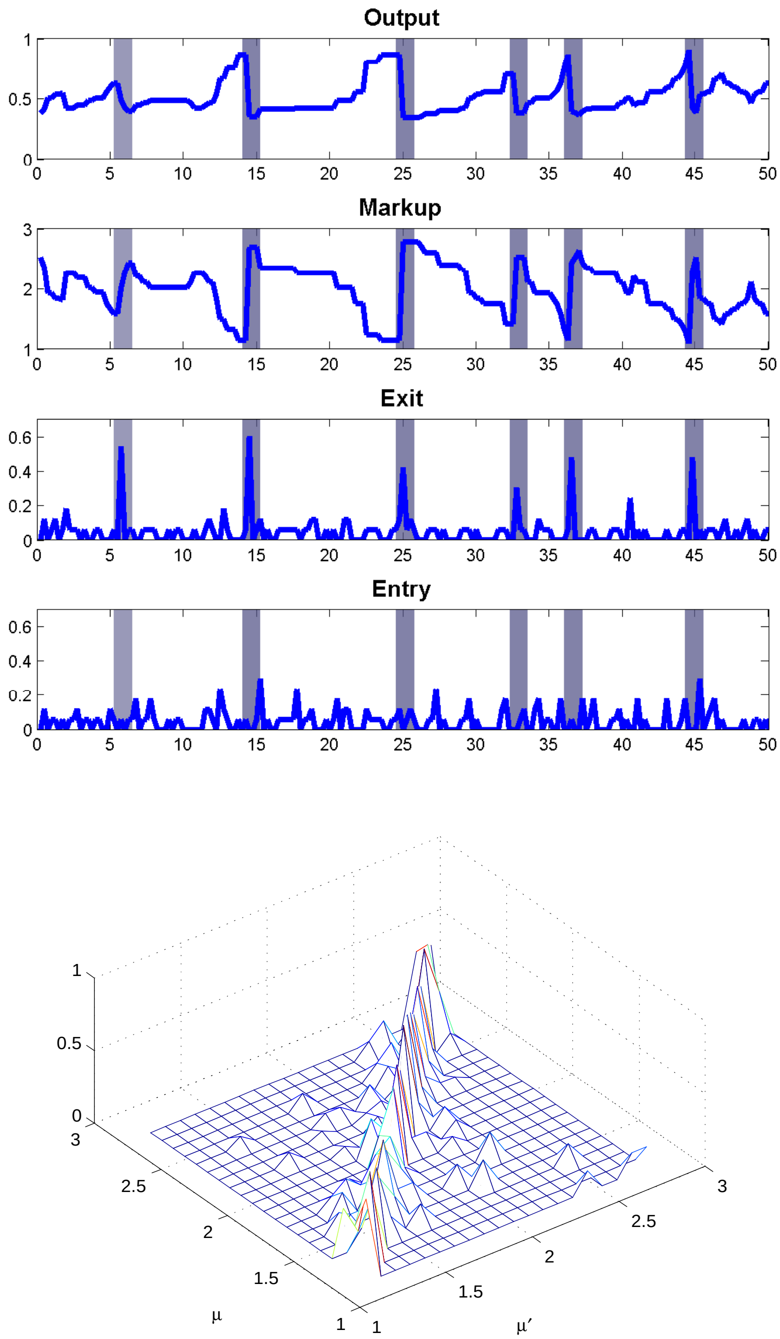

Figure 9B. **Model with Inattention.**

Appendix 6: Sensitivity to Key Parameters

First, we explore the sensitivity of our results to variations in two important parameters: the entry rate, s, and the probability of innovation, φ. Table 14 displays the behavior of the two models under the benchmark calibration ($s = 0.05, \varphi = 0.8$) and compares it to alternative calibrations ($s = 0.03$ and $\varphi = 0.2$).

When full information is available, the decrease in the entry rate, s, makes cycles slightly more asymmetric, while persistence of individual histories, $1 - \varphi$, increases aggregate persistence without affecting much the length and asymmetry of the cycle. When capacity of processing information is limited, the cycles are much bigger and much more asymmetric compared to the full information case.

In this context, both a decrease in the entry rate s and an increase in persistence of individual histories (reduction of φ) help alleviate uncertainty, reducing the amount of information that needs to be processed. A lower entry rate makes the aggregate component more predictable, while persistence of individual histories makes the idiosyncratic component more predictable. As a result, less effort is required to process information, and better coordination is achieved.

Second, since the main difference between the two models we consider is the increase in the cost of processing information, θ, from 0 to 20 percent of average profits, it is instructive to explore the effect of a further increase in the cost of information processing. Surprisingly, an increase by a factor of five in the cost of processing information, which now accounts for about 80 percent of average firm profits, leads to a decrease in both the asymmetry and persistence of cycles. The reason for this is that in this case the costs of processing information become so high, that firms give up on getting a precise signal even about aggregate conditions, and base their exit decisions only on profits in the previous period. This behavior leads to almost uniformly distributed random exits of firms which are not necessarily the ones producing the most outdated products.

Table 15 shows how the degree of asymmetry depends on various other changes

	σ_Y	σ_P	ρ_Y	$\tau_{5-10\%}$	$\tau_{10-20\%}$	$\tau_{>20\%}$	$\gamma_{\Delta Y}$	$\gamma_{\Delta P}$	γ_X	γ_R
Full Info	0.13	0.13	0.57	1.8	2.6	15	-.18	.26	1.6	1.1
$s = 0.03$	0.12	0.13	0.65	2.2	5.1	62.5	-.39	.29	2.6	1.9
$\varphi = 0.2$	0.13	0.13	0.67	2.2	2.8	7.8	-.42	.58	2.3	1.0
Inattention	0.26	0.21	0.79	9.4	10	8.3	-4.0	3.8	4.0	1.6
$s = 0.03$	0.15	0.13	0.84	1.8	6.3	9.4	-1.7	0.6	3.6	1.6
$\varphi = 0.2$	0.13	0.11	0.49	1.8	2.4	11	-.08	.31	1.6	1.3
$\theta = 0.05$	0.13	0.12	0.52	1.5	24	19	-.01	.03	1.1	1.2

σ denotes standard deviation; ρ - autocorellation; $\tau_{\%}$ - average periods between contractions;

γ - skewness; ΔY - growth rate of GDP; ΔP - growth rate of markups; X - exit rate; R - entry rate.

Table 15: **Sensitivity to variations in parameters**

s	θ	φ	$\gamma_{\Delta Y}$	$\gamma_{\Delta P}$	γ_X	γ_R
5%	1%	0.8	-4.0	3.8	4.0	1.6
3%	1%	0.8	-1.7	0.6	3.6	1.6
3%	1%	**0.65**	-1.1	1.0	2.0	1.3
5%	1%	**0.2**	-.08	.31	1.6	1.3
5%	**5%**	0.8	-.01	.03	1.1	1.2
5%	**0.5%**	0.8	-.41	.34	1.3	1.1
3%	**2%**	0.8	-1.2	0.95	2.0	1.3

s denotes entry rate; θ - cost of processing information; φ - probability of innovations; γ - skewness;

ΔY - growth rate of GDP; ΔP - growth rate of markups; X - exit rate; R - entry rate.

Table 16: **Sensitivity of skewness to variations in parameters**

in parameters. Results reported in Table 10 confirms the hump-shaped response of skewness to variations in the cost of information, as well as the reduction in skewness associated with increased predictability of (reduction in uncertainty about) idiosyncratic shocks and entry rates.

Appendix 7: Sensitivity to Number of Firms

Table 16 shows what happens when we gradually increase the number of firms in the full-information case. Second moments of fluctuations gradually decrease to values, which are similar to those observed in developed countries. The length of cycles increases to values, which are much closer to the average lengths of cycles in developed countries.

Note from Table 14 that the Poisson entry process does play a role in generating the asymmetry in aggregate fluctuations. Table 16 shows that an increase in the number of firms decreases the asymmetry in the entry rate and as a result the asymmetry in the exit rate in the full-information economy. (Tables 17-19 give more details on the simulations.) However, we know from the empirical evidence that cyclicality of entry is not a source of asymmetric fluctuations in the data. Hence, in order to explain the observed asymmetric fluctuations for a given number of firms, the asymmetry in the exit margin must drive the asymmetry in the business cycles. This is precisely the case in the information-constrained economy but it is not the case in the full-information model.

A power law sized distribution reduces skewness of aggregate fluctuations by a constant on the order of 0.5, independent of the exact number of firms. Therefore, the asymptotic behavior of skewness of GDP growth in an economy with a million firms will be a fraction on the order of one half of skewness of GDP growth in our information-constrained economy with fifteen firms. Thus, we can be confident that the key business cycle predictions of our model will be preserved if the number of

	σ_C	σ_P	ρ_C	$\tau_{5-10\%}$	$\tau_{10-20\%}$	$\tau_{>20\%}$	$\gamma_{\Delta Y}$	$\gamma_{\Delta P}$	γ_X	γ_R
K=15	0.13	0.13	0.57	1.8	2.6	15	-.18	.26	1.6	1.1
K=35	0.09	0.09	0.59	2.9	8.1	∞	-.13	.15	1.1	0.7
K=70	0.07	0.07	0.64	5.3	∞	∞	-.05	.08	0.5	0.6
K=100	0.05	0.05	0.62	7.1	∞	∞	-.14	.17	0.5	0.4

σ denotes standard deviation; ρ - autocorellation; $\tau_{\%}$ - average periods between contractions;

γ - skewness; ΔY - growth rate of GDP; ΔP - growth rate of markups; X - exit rate; R - entry rate.

Table 17: **Sensitivity to the number of firms, K**

	σ_C	σ_P	ρ_C	$\tau_{5-10\%}$	$\tau_{10-20\%}$	$\tau_{>20\%}$	$\gamma_{\Delta C}$	$\gamma_{\Delta P}$	γ_X	γ_R
Baseline	0.13	0.13	0.57	1.8	2.6	15	-.18	.26	1.6	1.1
$\gamma = 0.5$	0.23	0.10	0.83	1.7	4.3	20.8	-.47	.53	1.9	1.1
$\gamma = 1.5$	0.11	0.16	0.51	1.9	2.3	15.6	-.24	.23	1.5	1.2
$\gamma = 2.5$	0.10	0.27	0.39	2.1	2.5	8.9	-.35	.44	2.0	1.4
$\varphi = 0.5$	0.12	0.12	0.62	1.8	2.5	20.8	.03	.19	2.0	1.0
$\varphi = 0.2$	0.13	0.13	0.67	2.2	2.8	7.8	-.42	.58	2.3	1.0
$s = 0.025$	0.12	0.13	0.65	2.2	5.1	62.5	-.39	.29	2.6	1.9

Table 18: **Full Information, K=15**

	σ_C	σ_P	ρ_C	$\tau_{5-10\%}$	$\tau_{10-20\%}$	$\tau_{>20\%}$	$\gamma_{\Delta C}$	$\gamma_{\Delta P}$	γ_X	γ_R
Baseline	0.09	0.09	0.59	2.9	8.1	∞	-.13	.15	1.1	0.7
$\gamma = 0.5$	0.19	0.08	0.81	3.6	31.5	∞	-.14	.16	0.5	0.7
$\gamma = 1.5$	0.08	0.11	0.60	2.3	15.6	∞	-.01	.08	0.8	0.7
$\gamma = 2.5$	0.06	0.16	0.35	2.3	5.4	∞	-.02	.28	1.1	0.7
$\varphi = 0.5$	0.08	0.08	0.69	2.6	10.4	∞	-.14	.16	1.0	0.7
$\varphi = 0.2$	0.08	0.08	0.74	3.2	10.4	∞	-.06	.02	1.1	0.7

Table 19: **Full Information, K=35**

	σ_C	σ_P	ρ_C	$\tau_{5-10\%}$	$\tau_{10-20\%}$	$\tau_{>20\%}$	$\gamma_{\Delta C}$	$\gamma_{\Delta P}$	γ_X	γ_R
Baseline	0.07	0.07	0.64	5.3	∞	∞	-.05	.08	0.5	0.6
$\gamma = 0.5$	0.10	0.05	0.81	5.1	∞	∞	-.06	.06	1.5	0.6
$\gamma = 1.5$	0.05	0.07	0.32	2.8	62.5	∞	-.14	.17	1.3	0.6
$\varphi = 0.2$	0.09	0.09	0.87	4.1	∞	∞	-.09	.13	0.8	0.6

Table 20: **Full Information, K=70**

firms was substantially increased: the size of fluctuations would be smaller, but the asymmetry would remain.

We conclude that even though it is hard to infer properties of second moments from our simplified model, this model has strong predictions for skewness and asymmetric behavior which is virtually immune to aggregation and variations in the number of firms.

Appendix 8: Bellman Recursion

Concavity of Mutual information in the Belief State.

For a given $p(\pi|\, s)$, Mutual Information is concave in $g(S)$

Proof. Let Z be the binary random variable with $P(Z=0)=\lambda$ and let $S=S_1$ if $Z=0$ and $S=S_2$ if $Z=1$. Let the set of all profits be $\pi_i \in A = \{\pi_1, .., \pi_n\}$Consider

$$\begin{aligned} I(S,Z;\pi) &= I(S;A) + I(Z;A|\, S) \\ &= I(S;A|\, Z) + I(Z;A) \end{aligned}$$

Conditional on S, A and Z are independent, $I(A;Z|\, S)=0$. Thus,

$$\begin{aligned} I(S;A) &\geq I(S;A|\, Z) \\ &= \lambda(I(S;A|\, Z=0)) + (1-\lambda)(I(S;A|\, Z=1)) \\ &= \lambda(I(S_1;A)) + (1-\lambda)(I(S_2;A)) \end{aligned}$$

Q.E.D. ■

Lemma 1 *For a given $p(\pi|\, s)$, the expression (17) is concave in $g(s)$*

The Bellman Recursion is a Contraction Mapping.

Proposition 1. *For the discrete Rational Inattention firm's problem, value recursion H and two given functions V and U, it holds that*

$$||HV - HU|| \leq \beta\, ||V-U||\,,$$

with $0 \leq \beta < 1$ and $||.||$ the supreme norm. That is, the value recursion H is a contraction mapping.

Proof. The H mapping displays:

$$HV(g) = \max_p \left[H^p V(g)\right]_+\,,$$

with

$$H^p V(g) = \left[\sum_{s \in S} \left(\sum_{\pi \in A} \pi p(\pi|s) \right) g(s) - \theta\kappa + \beta \sum_{s \in S} \sum_{\pi \in A} \left(V\left(g'_\pi(\cdot)\right)\right) p(\pi|s)\, g(s) \right].$$

Suppose that $||HV - HU||$ is the maximum at point g. Let p_1 denote the optimal control for HV under g and p_2 the optimal one for HU

$$HV(g) = \left[H^{p_1} V(g)\right]_+,$$
$$HU(g) = \left[H^{p_2} U(g)\right]_+.$$

$$\Longrightarrow ||HV(g) - HU(g)|| = \left[H^{p_1} V(g)\right]_+ - \left[H^{p_2} U(g)\right]_+.$$

Suppose (without loss of generality) that $HV(g) \leq HU(g)$. Since p_1 maximizes HV at g , it follows that

$$\left[H^{p_2} V(g)\right]_+ \leq \left[H^{p_1} V(g)\right]_+.$$

Hence,

$$\begin{aligned}
||HV - HU|| &= \\
||HV(g) - HU(g)|| &= \\
\left[H^{p_1} V(g)\right]_+ - \left[H^{p_2} U(g)\right]_+ &\leq \\
\left[H^{p_2} V(g)\right]_+ - \left[H^{p_2} U(g)\right]_+ &\leq \\
\beta \sum_{w \in W} \sum_{a \in A} \left[\left(V^{p_2}\left(g'_a(\cdot)\right)\right) - \left(U^{p_2}\left(g'_a(\cdot)\right)\right) \right] p_2 g(w) &\leq \\
\beta \sum_{w \in W} \sum_{a \in A} \left(||V - U|| \right) p_2 g(w) &\leq \\
\beta\, ||V - U||.
\end{aligned}$$

In the derivation above we can open the positive brackets because the profit function is evaluated at the same point p_2. Then either both brackets do not bind, or only the second binds, or both bind, which implies a $\leq$ sign. Otherwise, the whole expression is exactly equal to zero, which also implies a contraction mapping. Recalling that $0 \leq \beta < 1$ completes the proof. ■

The Bellman Recursion is an Isotonic Mapping

Corollary *For the discrete Rational Inattention firm's problem recursion H and two given functions V and U, it holds that $V \leq U \Longrightarrow HV \leq HU$, that is the value recursion H is an isotonic mapping.*

Proof. Let p_1 denote the optimal control for HV under g and p_2 the optimal one for HU

$$HV(g) = H^{p_1}V(g),$$
$$HU(g) = H^{p_2}U(g).$$

By definition,

$$H^{p_1}U(g) \leq H^{p_2}U(g).$$

From a given g, it is possible to compute $g'_\pi(\cdot)|_{p_1}$ for an arbitrary c and then the following will hold $V \leq U \Longrightarrow$

$\forall (g(s), \pi),$

$$V\left(g'_\pi(\cdot)|_{p_1}\right) \leq U\left(g'_\pi(\cdot)|_{p_1}\right) \Longrightarrow$$
$$\sum_{\pi\in A} V\left(g'_\pi(\cdot)|_{p_1}\right) \cdot p_1 g \leq \sum_{\pi\in A} U\left(g'_\pi(\cdot)|_{p_1}\right) \cdot p_1 g \Longrightarrow$$
$$\sum_{s\in S}\left(\sum_{\pi\in A}\pi g(s) + \beta\sum_{\pi\in A} V\left(g'_\pi(\cdot)|_{p_1}\right) \cdot p_1 g\right)$$
$$\leq \sum_{s\in S}\left(\sum_{\pi\in A}\pi g(s) + \beta\sum_{\pi\in A} U\left(g'_\pi(\cdot)|_{p_1}\right) \cdot p_1 g\right) \Longrightarrow$$
$$H^{p_1}V(g) \leq H^{p_1}U(g) \Longrightarrow$$
$$\left[H^{p_1}V(g)\right]_+ \leq \left[H^{p_1}U(g)\right]_+ \Longrightarrow$$
$$\left[H^{p_1}V(g)\right]_+ \leq \left[H^{p_2}U(g)\right]_+ \Longrightarrow$$
$$HV(g) \leq HU(g) \Longrightarrow \; HV \leq HU.$$

Note that g was chosen arbitrarily and, from it, $g'_\pi(\cdot)|_{p_1}$ completes the argument that the value function is isotone. ■

The Optimal Value Function is Piecewise Linear

Proposition 2. *If the profit function is weakly quasi-convex and if* $\Pr(\pi_j, S_i)$ satisfies (18) and (20)-(22), *then the optimal* $n-$ *step value function* $V_n(g)$ *can be expressed as:*

$$V_n(g) = \max_{\{\alpha_n^i\}_i} \sum_i \alpha_n(S_i)\, g(S_i)$$

where the $\alpha-$*vectors,* $\alpha : S \to R$*, are* $|S|-$*dimensional hyperplanes.*

Proof. The proof is done via induction. We assume that all the operations are well-defined in their corresponding spaces. Let Γ be the set that contains constraints (18),(20)-(22) .For planning horizon $n = 0$, we have only to take into account the immediate expected rewards and thus:

$$V_0(g) = \max_{p\in\Gamma} \left[\sum_{s\in S} \left(\sum_{\pi\in A} \pi(s)\, p \right) g(S) \right] \tag{34}$$

and therefore if I define the vectors

$$\left\{\alpha_0^i(S)\right\}_i \equiv \left(\sum_{\pi\in A} \pi(s)\, p \right)_{p\in\Gamma} \tag{35}$$

We have the desired

$$V_0(g) = \max_{\left\{\alpha_0^i(S)\right\}_i} \left\langle \alpha_0^i, g \right\rangle \tag{36}$$

where $\langle .,. \rangle$ denotes the inner product $\langle \alpha_0^i, g \rangle \equiv \sum_{s\in S} \alpha_0^i(s), g(s)$. For the general case, using equations (15)-(16):

$$V_n(g) = \max_{p\in\Gamma} \left[\begin{array}{c} \sum_{s\in S} \left(\sum_{\pi\in A} \pi(s)\, p(\pi|S) \right) g(S) + \\ +\beta \sum_{s\in S} \sum_{\pi\in A} \left(V_{n-1}\left(g'_\pi(\cdot)_{p\text{ß}} \right) \right) p(\pi|w)\, g(s) \end{array} \right] \tag{37}$$

by the induction hypothesis

$$V_{n-1}\left(g(\cdot)|_\pi \right) = \max_{\left\{\alpha_{n-1}^i\right\}_i} \left\langle \alpha_{n-1}^i, g'_\pi(\cdot) \right\rangle \tag{38}$$

Plugging into the above equation (18) and by definition of $\langle .,. \rangle$,

$$V_{n-1}\left(g'_{\pi}\left(\cdot\right)\right)=\max_{\left\{\alpha_{n-1}^{i}\right\}_{i}}\sum_{s'\in S}\alpha_{n-1}^{i}\left(\sum_{s\in S}\sum_{\pi\in A}T\left(\cdot;s,\pi\right)\frac{\Pr\left(s,\pi\right)}{\Pr\left(\pi\right)}\right)\tag{39}$$

With the above:

$$\begin{aligned}V_{n}\left(g\right)&=\max_{p\in\Gamma}\left[\begin{array}{c}\sum_{s\in S}\left(\sum_{\pi\in A}\pi\left(s\right)p\right)g\left(s\right)+\\+\beta\max_{\left\{\alpha_{n-1}^{i}\right\}_{i}}\sum_{s'\in S}\alpha_{n-1}^{i}\left(s'\right)\left(\sum_{s\in S}\left(\sum_{\pi\in A}\frac{T(\cdot;s,\pi)}{\Pr(\pi)}\cdot p\right)g\left(s\right)\right)\end{array}\right]\\&=\max_{p\in\Gamma}\left[\left\langle\pi\cdot p,\ g\left(s\right)\right\rangle+\beta\sum_{s\in S}\frac{1}{\Pr\left(\pi\right)}\max_{\left\{\alpha_{n-1}^{i}\right\}_{i}}\left\langle\sum_{s'\in S}\alpha_{n-1}^{i}\left(s'\right)T\left(\cdot;s,\pi\right)\cdot p,\ g\right\rangle\right]\end{aligned}\tag{40}$$

At this point, it is possible to define

$$\alpha_{p,\pi}^{j}\left(s\right)=\sum_{s'\in S}\alpha_{n-1}^{i}\left(s'\right)T\left(\cdot:s,\pi\right)\cdot p.\tag{41}$$

Note that these hyper-planes are independent on the prior g for which I am computing V_n. Thus, the value function amounts to

$$V_{n}\left(g\right)=\max_{p\in\Gamma}\left[\left\langle\pi\cdot p,\ g\right\rangle+\beta\sum_{\pi\in A}\frac{1}{\Pr\left(\pi\right)}\max_{\left\{\alpha_{p,\pi}^{j}\right\}_{j}}\left\langle\alpha_{p,\pi}^{j},g\right\rangle\right],\tag{42}$$

and define:

$$\alpha_{p,\pi,g}=\arg\max_{\left\{\alpha_{p,\pi}^{j}\right\}_{j}}\left\langle\alpha_{p,\pi}^{j},g\right\rangle.\tag{43}$$

Note that $\alpha_{p,\pi,g}$ is a subset of $\alpha_{p,\pi}^{j}$ and using this subset results into

$$\begin{aligned}V_{n}\left(g\right)&=\max_{p\in\Gamma}\left[\left\langle\pi\left(s\right)\cdot p,\ g\right\rangle+\beta\sum_{\pi\in A}\frac{1}{\Pr\left(\pi\right)}\left\langle\alpha_{p,\pi,g},g\right\rangle\right]\\&=\max_{p\in\Gamma}\left\langle\pi\cdot+\beta\sum_{\pi\in A}\frac{1}{\Pr\left(\pi\right)}\alpha_{p,\pi,g},\ g\right\rangle.\end{aligned}\tag{44}$$

Now

$$\left\{\alpha_{n}^{i}\right\}_{i}=\bigcup_{\forall g}\left\{\pi\cdot p+\beta\sum_{\pi\in A}\frac{1}{\Pr\left(\pi\right)}\alpha_{p,\pi,g}\right\}_{p\in\Gamma}\tag{45}$$

is a finite set of linear function parameterized in the action set. Note that a maximum of a piecewise linear convex function and a zero (a constant function) is also piecewise linear and convex. ■

.. and Convex (PCWL)

Proposition 3. *Assuming weak quasi-convexity of the profit function and the conditions of Proposition 1, let V_0 be an initial value function that is piecewise linear and convex. Then the i^{th} value function obtained after a finite number of update steps for a rational inattention consumption-saving problem is also finite, piecewise linear and convex (PCWL).*

Proof. The first task is to prove that $\{\alpha_n^i\}_i$ sets are discrete for all n. The proof proceeds via induction. Assuming quasi-convex profit function and since the optimal policy belongs to Γ, it is straightforward to see that through (35), the set of vectors $\{\alpha_0^i\}_i$,

$$\left\{\alpha_0^i\right\}_i \equiv \left(\sum_{s\in S}\left(\sum_{\pi\in A}\left(\pi\left(\mu, v_i\right)\right) p\left(\pi|s\right)\right) g\left(s\right)\right)_{p\in\Gamma}$$

is discrete. For the general case, observe that for discrete controls and assuming $M = \left|\left\{\alpha_{n-1}^j\right\}\right|$, the sets $\left\{\alpha_{p,\pi}^j\right\}$ are discrete, for a given action p and profits π, I can only generate $\alpha_{p,\pi}^j-$vectors. Now, fixing p it is possible to select one of the M $\alpha_{p,\pi}^j-$vectors for each one of the observed consumption π and, thus, $\{\alpha_n^j\}_i$ is a discrete set. The previous proposition, shows the value function to be convex. The *piecewise-linear* component of the properties comes from the fact that $\{\alpha_n^j\}_i$ set is of finite cardinality. It follows that V_n is defined as a finite set of linear functions. ■

Note also, that the profit function $\pi\left(s\right) = \mu v - 1$ is strictly quasi-convex in its arguments. Therefore, the existence and uniqueness of the solution of the value iteration problem of the firm follows from the contraction mapping theorem.

Appendix 9: Pseudocode

Let θ be the shadow cost associated with $\kappa_t = I_t(B_t, D_t)$.

- Step 1: Build the transition matrix $T(\cdot; b_t, d_t)$ convoluting the stochastic properties of the random variables (B, D) on an equispaced grid.
- Approximate T by a first-order Markov process.
- Step 2: Build the simplex - an equispaced grid to approximate each $g(B_t)$- a simplex point.
- Step 3: For each simplex point, define $p(b_t, d_t)$and initialize $V\left(g'_{\pi_j}(\cdot)\right) = 0$.
- Step 4: For each simplex point, find $p^*(b, d)$ which solves

$$V_0(g(b_t))|_{p^*(b_t,d_t)} = \max_{p(b_t,sd)} \left\{ \sum_{b_t \in \Omega_w} \sum_{d_t \in \Omega_c} (\pi_t(b,d)) \, p^*(b_t, d_t) - \theta [I_t(B_t, D_t)] \right\}.$$

- Step 5: For each simplex point, compute $g'_{\pi_j}(\cdot) = \sum_{b_t \in \Omega_b} T(\cdot; b_t, d_t) \, p^*(b_t|d_t)$. Use a kernel regression to interpolate $V_0(g(b_t))$ into $g'_{\pi_j}(\cdot)$.
- Step 6: Optimize using csminwel and iterate on the value function to convergence.
- Step 7. For each model, draw from the ergodic $p^*(b, d)$, samples of (b_t, d_t) and use the consumer's F.O.C. to simulate the time series of consumption, prices, markups and expected markups, profits and exit decisions.
- Step 8. Compute the model-simulated empirical distribution of consumption, prices, markups and the idiosyncratic shocks to generate the empirical transition matrix and go back to Step 1.
- Step 9. Iterate until convergence.

Observations

1. Firms' value function takes about 20 iterations to converge.

2. Global equilibrium (law of motion) takes up to 7 iterations to converge.