MSc Economics

Monopolistic Competition, Idiosyncratic Productivity, and Information Constraints

A Master's Thesis submitted for the degree of
"Master of Science"

supervised by

Tamás K. Papp

Gabriel Ziegler

0710817

Vienna, 23. June 2014

INSTITUT FÜR HÖHERE STUDIEN
INSTITUTE FOR ADVANCED STUDIES
Vienna

TU WIEN CONTINUING EDUCATION CENTER

MSc Economics

Affidavit

I, Gabriel Ziegler

hereby declare

that I am the sole author of the present Master's Thesis,

Monopolistic Competition, Idiosyncratic Productivity, and Information Constraints

56 pages, bound, and that I have not used any source or tool other than those referenced or any other illicit aid or tool, and that I have not prior to this date submitted this Master's Thesis as an examination paper in any form in Austria or abroad.

Vienna, 23. June 2014

Signature

Table of Contents

List of Figures

Abstract

A firm might not have perfect information about its own productivity, because observing the true realization might be too costly in terms of opportunity costs of attending to other sources of uncertainty. Building on Sims (1998, 2003), I use insights from information theory to study the behavior of information constrained firms. The focus is on the informational friction on the firm side only. This allows analyzing the implications of imperfect information for individual firms. Additionally, a welfare analysis is conducted by solving a social planner's problem. Interestingly, a comparison using approximations in log-deviations is not applicable, although such an approximation is often used in the literature. In the decentralized economy, output might not be maximized under full information. Furthermore, the distortions due to monopolistic competition are increasing in the amount of information agents are allowed to process.

1 Introduction

Firm owners or managers have many decisions to make and many of these decision need different sources of information. Information acquisition is costly—most notably because of opportunity costs—and thus getting the knowledge of the exact level of a relevant variable might not be feasible. Hence, most decisions are based on imperfect information, because a good guess or a very noisy observation is the best information available. This, in turn, can affect the aggregate outcomes of the economy.

Although the economics of information was initially mostly used studying microeconomic questions, there is a literature exploring the role of imperfect information in a macroeconomic setting. For example, Phelps (1970) and Lucas (1972) study how imperfect information leads to real effects of monetary policy. Barro (1976, 1977) analyzes effects of monetary shocks, which are unanticipated due to imperfect information. All these early approaches assume exogenous information structures, which are ad hoc and hard to justify. For example, in the model of Lucas (1972), it is not clear why agents would not be able to observe the current state of monetary policy, especially, nowadays with an easy access to the internet this assumptions seems unreasonable. Thus, informational frictions in macroeconomics have received less emphasis until recently.

A more recent approach—known as sticky-information—was introduced by Mankiw and Reis (2002). In this setup only a fraction of firms update their information set. The other firms stick with their old information. The updating process—where updating gives full information—is modeled as reduced form *á la* Calvo (1983). Reis (2006) provides a micro-foundation for sticky-information models, endogenizing the updating decision. If firms need to a pay fixed cost to acquire new information then same results as in the reduced form updating can be obtained. Even in this microfounded approach firms update to full information if they choose to update. Thus, the choice is about when to update and not about what information the firm should acquire. In this sense, the information choice is not entirely endogenous.

Allowing agents to choose which information they want to obtain leads to an information structure, which is determined within the model. Sims (1998, 2003) introduces such a framework explicitly augmenting economic models with this choice dimensions and it is known as *rational inattention*. In models with rational inattention one combines economics with insights from information theory—most notably the results from Shannon (1948). A key feature of these models is that agents only have a limited capacity for processing information about random variables in the economy. The agents are rational insofar that they choose on which variables they should focus their attention. Thus, the key difference to other approaches is that no assumptions about the information structure are needed *a priori*. The information acquired by agents is determined as an equilibrium object and thus completely endogenous. Recently, models of rational inattention were used to study macroeconomic and finance questions like price inertia, monetary policy, or consumption and portfolio choice. Examples include Adam (2007), Luo (2008), Maćkowiak and Wiederholt (2009a), Nieuwerburgh and Veldkamp (2010), Mondria (2010), and Paciello and Wiederholt (2014). Veldkamp (2011) provides a more general textbook treatment of macroeconomic and finance applications of imperfect information.

Using the framework of rational inattention, I show that in a decentralized economy increasing the information capacity leads to a reduction of output. For a social planner this result turns around and more information leads to higher output. The difference of the social optimum and the decentralized economy are increasing in information capacity; this is, the distortion increases with an increase in information. Eventually, in the full information equilibrium the distortion is maximal. However, a simple proportional tax can restore the socially optimal behavior.

The model is a monopolistically competitive economy, where firms face an information constraint. Due to this information constraint, firms are not able to observe the true realization of their productivity. A similar framework is used in other models with rational inattention, for example the ones cited above. However, these paper assume that idiosyncratic shocks or errors in decision making do not have aggregate effects, in the sense that these shocks on the individual level do not effect aggregate outcomes directly. To explicitly address this aggregation issue, I assume that the only source of uncertainty comes from idiosyncratic productivity. Another interesting question—which has not been addressed by the literature yet—is what welfare implications does such an informational friction have. To study this question, I solve a social planner's problem. The objective of the social planner is to maximize the agent's utility, in my specific setup this is equivalent to maximizing output. A comparison of the social planner's problem and the decentralized economy allows the analysis of the distortions arising due to the information constraints.

In general, most models of rational inattention cannot be solved analytically. To circumvent this problem, the model is approximated with log-deviations from the non-stochastic steady state. For the decentralized economy, this is done in the first part of this paper as well. However, for the social planner's problem this approach is not applicable and thus an approximation in level deviations is used. I show that the outcome of the level- and log-deviation approximation in the decentralized economy are very similar. Therefore, the comparison with the social planner is valid and the welfare analysis with these approximation is sensible. This issue for the social planner's problem might be a reason why this welfare question was not yet addressed by the literature.

For the decentralized economy, several papers in the literature are related to the analysis of price setting provided here. As already mentioned above, in the sticky-information model of Mankiw and Reis (2002) a fraction of firms obtains full information. For other firms the information set does not change. All the firms set prices optimally given their information set and thus only a small fraction of firms sets prices given the newest information. Mankiw and Reis claim that their model generates a Phillips curve which suites data better than the usual New Keynesian Phillips curve. Woodford (2003) assumes that agents observe aggregate demand only through a noisy signal. Given the realization of the signal the firms set prices optimally. Woodford's analysis relates informational frictions to real effects of monetary policy similar to Lucas (1972). Thus, in both models there is an aggregate effect through individual imperfect information. However, both models also have the information structure exogenously assumed and this is in stark contrast to models of rational inattention, although, parts of the information acquisition choice can be endogenized in sticky-information models as shown by Reis (2006).

My analysis is more closely related to Maćkowiak and Wiederholt (2009a) and Matejka (2010). Both models incorporate rational inattention to study the behavior of firms. Maćkowiak and Wiederholt use a log-quadratic approximation to show that it is optimal for firms to focus most of their attention to idiosyncratic variables. By assumption, the idiosyncratic uncertainty does not matter in the aggregate. This point is explicitly addressed in this thesis. Matejka studies the behavior of a monopoly which faces an information constraint. In contrast to most other models of rational inattention, Matejka does not use an approximation of the objective. He shows that it might be optimal for firms to have discrete pricing—even when the underlying uncertainty stems from a continuous distribution. Aggregation is not needed in his paper as there is only one monopoly. In this sense, the analysis presented here is somewhere in-between Maćkowiak and Wiederholt (2009a) and Matejka (2010). Furthermore, to the best of my knowledge, a welfare analysis was not considered in the literature yet.

The thesis is structured as follows. In Section 2, I introduce some basic concepts from information theory. Moreover, a key result of rational inattention—the analytic solution to quadratic Gaussian problems—is presented. This result is applied to economic models in the later sections. The main contribution of the thesis is in Section 3. There I derive the optimal behavior of firms and discuss the implications on aggregate variables. The firms problem is analyzed with a log- and a level-quadratic approximation of the profit function. In this section also the social planner's problem is discussed. Furthermore, a comparison between the decentralized economy and the social planner's solution is provided. In the end of the section, numerical examples are used to check if the approximations of the firm's objective are valid. I conclude in Section 4.

2 Information Theory and Rational Inattention

The purpose of this section is to provide a self-contained summary of the concepts and results I use in the model below. Readers familiar with the technical aspects of the rational inattention framework of Sims (1998, 2003, 2006) are advised to proceed to Section 3 directly. In the first section some important definitions and key results of information theory are introduced. In the second part, the main theorem of rational inattention is presented and discussed.

2.1 Information Theory

Most of the definitions and results in this section are taken from Cover and Thomas (2006). However, note that in this thesis only continuous random variables are considered and for this reason only the definitions for these variables are presented. This section focuses on concepts and results necessary for the economic model below. The reader interested in a more in-depth coverage of information theory is referred to Cover and Thomas (2006).

Information theory was introduced by Shannon (1948). The basic idea is to quantify information to answer questions about communication. In order to do so, several well-defined concepts are needed. First of all, we need a measure of uncertainty associated with a random variable. Shannon (1948) derived such a measure from four axioms. Here, I only present the result, which is known as the *entropy* of a random variable. Let $f(x)$ denote the probability density function (pdf) of the random variable X. From nnow on, it is assumed that the pdf exists. Then entropy of X is defined as follows.[1]

Definition 1. *The* entropy $H(X)$ *of a continuous random variable* X *with pdf* $f(x)$ *is defined as*

$$H(X) = -\int_{\mathcal{X}} f(x) \log f(x) \, \mathrm{d}x,$$

where $\mathcal{X}$ *denotes the support of* X.

It is straightforward to extend this definition to random vectors and conditional entropy (Cover and Thomas, 2006, p. 229–30). Note that entropy is a property of the distribution and does not depend on the realizations. Furthermore, the logarithm is usually taken to base 2 so that the entropy of a fair coin flip is normalized to one. The unit of entropy with base 2 is called *bits*. If the natural logarithm is used in the definition, then the units care called *nats*. This latter approach is used throughout this thesis. Of course, by the properties of the logarithm the two different measures of entropy are related through multiplication by a constant.

In the main analysis most random variables are normally distributed. The entropy of a Gaussian random variable has an intuitive expression. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$H(X) = \frac{1}{2} \log \left(2\pi e \sigma^2\right).$$

For a normal distribution, entropy is an increasing function that only depends on the variance. This means that the uncertainty associated with a Gaussian

[1] Henceforth, with a slight abuse of notation, I denote the pdf associated with a random variable X by $f(x)$ only. Thus, another random variable Y has pdf $f(y)$ as well. Then, $f(x)$ and $f(y)$ might be different functions.

random variable is captured by its variance. Note that the entropy of normal distribution is invariant to the mean, which is true more generally.

As mentioned above, entropy measures the uncertainty associated with a random variable (or vector). Now, with that in hand it possible to quantify information contained in one random variable about another variable. This measure is known as *mutual information*.

Definition 2. *Let X and Y be two random variables with joint pdf $f(p, x)$ and marginal pdfs $f(x)$ and $f(p)$, respectively. Then* mutual information $\mathcal{I}[X;Y]$ *is defined as*

$$\mathcal{I}[X;Y] = H(X) - H(X|Y).$$

Here, $H(X|Y)$ denotes the entropy of X conditional on Y. Intuitively, mutual information measures the reduction of uncertainty of X due to the knowledge of Y. Mutual information of two random variables is zero if and only if the two variables are independent. Mutual information is a special case of a Kullback–Leibler distance. In particular, the Kullback–Leibler distance of the joint distribution and the product of the marginals is the same as mutual information. An immediate consequence of this result is that mutual information is always non-negative, because the Kullback–Leibler is always non-negative as well (Cover and Thomas, 2006, pp. 26–27).

In the economic model presented in the subsequent section, mutual information is constrained by an exogenous parameter. To describe the optimal behavior of firms under such a constraint it is important to know if mutual information is a convex or a concave function. This question is answered by the following theorem.

Theorem 1. *Let X and Y be random variables with joint pdf $f(x, y) = f(x)f(y|x)$.*

- *For a fixed $f(y|x)$, $\mathcal{I}[X;Y]$ is concave in $f(x)$.*
- *For a fixed $f(x)$, $\mathcal{I}[X;Y]$ is convex in $f(y|x)$.*

Proof. See Cover and Thomas (2006, p. 31). □

Another important feature of mutual information is its invariance under monotonic transformations. This feature is exploited in the economic model as well and is stated formally next.

Proposition 1 (Invariance of Mutual Information)**.** *Let X and Y be random variables. If $Z = a + bY$, where a and $b \neq 0$ are scalars, then $\mathcal{I}[X;Y] = \mathcal{I}[X;Z]$.*

Proof. This is a special case of Kraskov et al. (2004). In their appendix they show that the statement generalizes to homeomorphisms. □

This result is intuitive: a linear transformation of one random variable contains the same information about some other random variable as the original one. This is because each realization of the transformed variable can be associated with one and only one realization of the original variable.

The definitions and results presented here are enough to use them in an economic context. However, before turning to the economic model, a key theorem from rational inattention is presented in the next section.

2.2 Rational Inattention

In models with rational inattention (RI), one combines economics with insights from information theory. Rational inattention was introduced by Sims (1998, 2003) to incorporate imperfect information into economic models. A key feature of these models is that agents have a limited capacity in processing information about random variables in the economy. This is modeled as a exogenous limit on mutual information of the choice and the state variables.[2] The agents are rational insofar that they choose on which variables they should focus their attention. However, compared to other economic models with imperfect information, in RI models the information structure is *endogenously* determined within the model.

[2]It is also possible to allow agents to buy more capacity, but this approach is not used in this work.

Most models with RI are rather complicated and thus rely on numerical solutions. Currently only a very specific class of RI problems has an analytic solution.[3] This solution was first mentioned by Sims (2003), but without a formal statement. In this section a static rational inattention problem is discussed and then this specific solution is presented.

Consider an economic agents who wants to minimize the expected loss of a given loss function L.[4] The loss functions depends on a choice p and on a state x. The agent faces a constraint, which limits the mutual information of the choice and the state variable. The problem of the agent is to choose a joint distribution which minimizes the expected losses taking as given the marginal distribution of the state variable (the prior distribution) and which respects the information constraint. Formally the problem can be stated as

$$
\begin{aligned}
&\min_f \left\{ \mathbb{E}[L(X,P)] = \int_{\mathcal{X}} \int_{\mathcal{P}} L(x,p)\, f(x,p)\, \mathrm{d}p\, \mathrm{d}x \right\} \qquad (1)\\
&\text{s.t.} \quad \int_{\mathcal{P}} f(x,p)\, \mathrm{d}p = g(x), \quad \forall x \in \mathcal{X}\\
&\qquad f(x,p) \geq 0, \quad \forall (x,p) \in \mathcal{X} \times \mathcal{P}\\
&\qquad \mathcal{I}[X;P] \leq \kappa,
\end{aligned}
$$

where κ is the exogenously given maximum capacity of information flow. The objective of this minimization problem is, as mentioned above, the expected loss. The first constraint describes the available choices for the joint distribution. It describes consistency of a Bayesian agent in the sense that the posterior adds up to the prior $g(x)$. The second non-negativity constraint is necessary for a probability distribution. The most interesting constraint is the last one, the

[3] Actually, Maćkowiak and Wiederholt (2009a) derived an analytic solution for a very similar problem, but in a dynamic context. I consider only a static problem here.

[4] In economics, it is more natural to have an agent maximizing a utility function. Maximizing a quadratic utility function is equivalent to minimizing a quadratic loss function. The use of a loss function is in line with the model presented below.

information constraint. The constraint states that the reduction of uncertainty is on average less or equal to κ. If κ would be large enough, the agent could choose a distribution $f(x,p)$ which puts all the mass on the optimal p^* given x. However, usually this is not possible and the agent must choose $f(x,p)$ as to minimize expected losses.

Without any further assumptions this problem does not have an analytic solution. However, note some important features of this optimization problem. The objective is linear in the choice f and by Theorem 1 the constraint set is convex. Even with this well behaved functional forms strong assumptions need to be imposed to get an analytic solution. These special assumptions and the solution are summarized in the following theorem.

Theorem 2. *Let the loss function be quadratic, i.e.* $L(X,P) = (\alpha X - P)^2$ *for some constant* α. *Furthermore, suppose the prior is a normal distribution, i.e.* $X \sim \mathcal{N}(\mu, \sigma_x^2)$. *Then a bivariate normal distribution* $(X,P) \sim \mathcal{N}_2(\mu_x, \mu_p, \sigma_x^2, \sigma_p^2, \rho)$ *is a solution to* (1). *The parameters are given by*

$$
\begin{aligned}
\mu_p &= \alpha \mu_x \\
\sigma_p^2 &= \alpha^2 \left(1 - e^{-2\kappa}\right) \sigma_x^2 \\
\rho^2 &= 1 - e^{-2\kappa}
\end{aligned}
$$

Proof. This is a simplified and static version of Maćkowiak and Wiederholt (2009b, Technical Appendix, Section 3). □

An intuitive result is implied by this theorem. As a bivariate normal distribution is the optimal joint distribution, the behavior is the same as if the agent receives are signal about the state variable. The noise in the signal depends on the information capacity κ.

Corollary 1. *Under the assumption of Theorem 2 it is optimal to choose a signal S of the form*

$$S = X + \varepsilon,$$

where $\varepsilon \sim \mathcal{N}(0, (e^{2\kappa} - 1)^{-1}\sigma_x^2)$ and $X \perp \varepsilon$. The optimal choice P is then given by the optimal behavior conditional on this signal.

Proof. Follows from Maćkowiak and Wiederholt (2009a, Proposition 3 and 4). Also recall if X and Y follow a bivariate normal distribution $\mathcal{N}_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, then

$$X|Y \sim \mathcal{N}\left(\mu_x + \frac{\sigma_x}{\sigma_y}\rho(y - \mu_y), (1 - \rho^2)\sigma_x^2\right),$$

where ρ denotes the correlation coefficient. (See e.g. Bickel and Doksum, 2001, p. 501) . □

Although this corollary shows that the agents chooses a signal with independent Gaussian noise, it is not a standard signal extraction problem, because the signal is an endogenous result. The agents chooses this signal and thus it stems from the agents internal *system*. This information structure is not assumed *a priori*. Indeed, if the assumptions of Theorem 2 are relaxed this form of signal might not be optimal anymore. An example of this case is discussed in Subsection 3.3.

In this section basic elements of information theory were introduced. Moreover, a basic rational inattention model was presented, where the agent's mutual information of the state and his choice is constraint. Now, this machinery is used to study the behavior of monopolistic competitive firms. All of these firms are subject to such an information constraint. This is done in the next section.

3 The Economic Model

In this section the behavior of information constrained firms is studied. In addition to the individual behavior also the aggregate outcomes are discussed, in particular, the question how the information constraint on the individual level affects the aggregate economy is addressed. The focus of the following analysis is on firms. The consumer side is modeled without any friction in information processing. Of course, including an information constraint on the consumer side too raises further interesting questions, but studying such questions is left for future research.

The economy is accommodated by a representative household with a continuum of members and by N small firms, where N is a large number such that firms take aggregates as given.[5] The market structure is assumed to be monopolistically competitive to allow firms to set individual prices. Thus, this part of the model is very similar to the economy studied by Dixit and Stiglitz (1977), where the number of firms is also finite.

Moreover, in the rational inattention literature the finiteness of firms usually plays an important role. For example, in the DSGE model of Maćkowiak and Wiederholt (2011) an infinite number of firms is not possible, because consumers would need infinite information to track all the prices in the economy. However, as it is usual in the case of finite firms it is important that firms do not take into account that their actions might influence aggregate outcomes.

The representative household consumes a Dixit–Stiglitz aggregator of the N different goods and supplies labor inelastically. Total labor supply is normalized to N. This normalization guarantees that the endogenous variables do not depend on the size of the economy. I discuss this issue in more detail below. The income of the household comes from two sources. First, there is the labor income coming

[5] For the aggregate equilibrium outcomes the limit as $N \to \infty$ is used. First, this allows to use a law of large number which yields analytic expressions for the aggregate variables. Second, this limit circumvents any mathematical issues arising from a continuum of firms *and* consumers.

from the inelastic labor supply and second, firm profits are income as well. In this setup aggregate output and consumption (demand) are the same and are given by

$$Y = \left(\frac{1}{N} \sum_{i=1}^{N} Q_i^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}. \tag{2}$$

Here, $\eta > 1$ is the elasticity of substitution and Q_i is demand for good i. The representative household maximizes aggregate output taking prices and profits as given. Formally, the household solves

$$\max_{\{Q_i\}_{i=1}^{N}} \left\{ Y = \left(\frac{1}{N} \sum_{i=1}^{N} Q_i^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right\}$$
$$\text{s.t. } \mathcal{P}Y = \frac{1}{N} \sum_{i=1}^{N} P_i Q_i = \frac{1}{N} \sum_{i=1}^{N} (\Pi_i + wL_i)$$
$$\frac{1}{N} \sum_{i=1}^{N} L_i = 1,$$

where $\mathcal{P}$ denotes the aggregate price level, P_i is the price of good i. Profits and labor demand of firm i are denoted by Π_i and L_i, respectively. The wage rate w is normalized to one. Optimizing behavior of the household implies an isoelastic demand function for good i.

$$Q_i = Y \left(\frac{P_i}{\mathcal{P}} \right)^{-\eta} \tag{3}$$

Aggregate output and the budget constraint include averages over N different goods for the following reason. As N is large, one can think of this sum as an approximation of an integral over a continuum of goods with unit mass. Because of this averaging, the normalization of labor supply to N is needed.

Firms are indexed by $i \in \{1, \ldots N\}$. A linear technology transforms one unit labor into one unit of good i. The uncertainty a firm faces is only about the productivity. Idiosyncratic productivity is determined exogenously by a random variable. Let A_i denote the productivity level of firm i. It is assumed that A_i is independent of A_j for all $i \neq j$. There is no capital or any other factor needed for production. I assume that firms maximize expected profits taking the demand function into account. Under full information this is a simple optimization problem.[6]

$$\max_{P} \left\{ \Pi(Y, \mathcal{P}, X, P) = Y \left(\frac{P}{\mathcal{P}} \right)^{-\eta} (AP - 1) \right\} \tag{4}$$

Due to the linear production assumption, the labor demand is equal to the demand for good i, i.e. $L_i = Q_i$. In this baseline problem, optimal price setting is the standard result for models of monopolistic competition. Under full information the optimal behavior is to charge a mark-up over the costs, i.e. $P = \frac{\eta}{\eta - 1} \frac{1}{A}$. However, the question is how the firm behaves if it cannot fully observe its productivity. Therefore, a constraint on mutual information as described in the previous section is introduced. Note that this implies that the uncertainty for a firm comes from the uncertainty in idiosyncratic productivity only.

If the information flow is limited the firm has a prior distribution for productivity $g(A)$ and solves

$$\begin{aligned} \max_{f} \; & \int_{\mathcal{A}} \int_{\mathcal{P}} Y \left(\frac{P}{\mathcal{P}} \right)^{-\eta} (AP - 1) \, f(A, P) \, \mathrm{d}P \, \mathrm{d}A \qquad (5) \\ \text{s.t.} \; & \int_{\mathcal{P}} f(A, P) \, \mathrm{d}P = g(A), \quad \forall A \in \mathcal{A} \\ & f(A, P) \geq 0, \quad \forall (A, P) \in \mathcal{A} \times \mathcal{P} \\ & \mathcal{I}[A; P] \leq \kappa. \end{aligned}$$

This is a very similar problem to (1). As in the case for the rational inattention problem, there is no analytic solution known in the general case. The problem can be solved solved numerically, and this is done later in this thesis. However,

[6] Whenever no confusion arises I drop the firm index.

to obtain an analytic solution I follow Maćkowiak and Wiederholt (2009a) and proceed with a log-quadratic approximation around the non-stochastic steady state. This approximation reshapes the problem such that Theorem 2 applies. Before proceeding with the approximation, let me define the equilibrium in this general setup. This is useful, because the equilibrium concept is used throughout the thesis and applies to the approximation as well. The decentralized equilibrium is defined as follows.

Definition 3. *The tuple* $((f_i, A_i, P_i, Q_i, L_i)_{i \in \{1,\ldots N\}}, Y, \mathcal{P})$ *constitutes a* decentralized equilibrium with information capacity κ *if*

1. f_i *solves* (5) *for every firm* $i \in \{1, \ldots N\}$ *taking* Y *and* $\mathcal{P}$ *as given*
2. $(A_i, P_i) \sim f_i(A, P), \quad \forall i \in \{1, \ldots N\}$
3. *Output of good* i*,* Q_i*, is given by* (3) *for every firm* $i \in \{1, \ldots N\}$
4. *Aggregate price level is given by* $\mathcal{P} = \left(\frac{1}{N} \sum_{i=1}^{N} P_i^{1-\eta} \right)^{\frac{1}{1-\eta}}$
5. *Aggregate real output* Y *is given by* (2)
6. *The labor market clears, i.e.* $\frac{1}{N} \sum_{i=1}^{N} L_i = 1$ *and* $Q_i = L_i, \quad \forall i \in \{1, \ldots N\}$

3.1 The log-quadratic case

As the previous stated problem of the firm does not have an analytic solution, I use an log-quadratic approximation of the objective. Already Maćkowiak and Wiederholt (2009a) use this approach in a very similar setup and can be outlined as follows:

1. Find the non-stochastic steady state (NSSS)
2. Find a log-quadratic approximation around the non-stochastic steady state
3. Derive the optimal behavior with full information using the approximation of step 2

4. Find a quadratic loss function

5. Apply Theorem 2 with the quadratic loss function to find the optimal behavior

6. Derive the equilibrium (Definition 3) with the optimal choice of step 5.

The non-stochastic steady state is found by setting the variance of each random variable to zero. In the case here, the only random variable is the productivity A. I normalize NSSS productivity to one. Denote the NSSS variables with a bar, then

$$\bar{A}_i = \bar{A} = 1, \quad \forall i \in \{1, \dots N\} \tag{6}$$

$$\bar{\mathcal{P}} = \bar{P} = \bar{P}_i = \frac{\eta}{\eta - 1}, \quad \forall i \in \{1, \dots N\} \tag{7}$$

$$\bar{Y} = \bar{Q} = \bar{Q}_i = \bar{L}_i = 1, \quad \forall i \in \{1, \dots N\} \tag{8}$$

$$\bar{\Pi} = \bar{\Pi}_i = \frac{1}{\eta - 1}, \quad \forall i \in \{1, \dots N\} \tag{9}$$

Next, the log-quadratic approximation around this NSSS is derived. From now on, let small letter variables denote the deviations from the NSSS. Then the profit function can be written as

$$\pi(a, p) = \Pi(\bar{A}e^a, \bar{P}e^p).$$

The first and second partial derivatives evaluated at zero are

$$\pi_1 = \left.\frac{\partial \pi(a,p)}{\partial a}\right|_{\substack{p=0\\a=0}} = Y\left(\frac{\bar{P}}{\mathcal{P}}\right)^{-\eta}\frac{\eta}{\eta-1}$$
$$\pi_2 = \left.\frac{\partial \pi(a,p)}{\partial p}\right|_{\substack{p=0\\a=0}} = Y\left(\frac{\bar{P}}{\mathcal{P}}\right)^{-\eta}(\eta-\eta) = 0$$
$$\pi_{11} = \left.\frac{\partial^2 \pi(a,p)}{\partial a^2}\right|_{\substack{p=0\\a=0}} = \pi_1$$
$$\pi_{22} = \left.\frac{\partial^2 \pi(a,p)}{\partial p^2}\right|_{\substack{p=0\\a=0}} = -\eta Y\left(\frac{\bar{P}}{\mathcal{P}}\right)^{-\eta} < 0$$
$$\pi_{12} = \left.\frac{\partial^2 \pi(a,p)}{\partial a \partial p}\right|_{\substack{p=0\\a=0}} = \pi_{22}.$$

Denote the approximated profit function with $\tilde{\pi}$. Using the partial derivatives from above and the profits in the NSSS it follows that

$$\tilde{\pi}(a,p) = \frac{1}{\eta-1} + \pi_1 a\left(1+\frac{a}{2}\right) + \frac{\pi_{22}}{2}p^2 + \pi_{22}ap.$$

This approximation allows to solve for the optimal pricing behavior without an information constraint. The optimal choice is given by

$$p_{NC} = -a.$$

Thus, it is optimal to adjust the price one-to-one for any deviations of productivity. Intuitively, the firms charge a monopoly markup already in the NSSS. Moreover, the optimal pricing behavior in the general form is log-linear in productivity. Hence, only an one-to-one adjustment for deviations in prices is needed outside of the NSSS.[7] With this optimal behavior, the loss due to suboptimal pricing is given by

[7] Actually, due to the functional forms the optimal behavior in log-deviations with the log-quadratic approximation of the objective yields the same policy as without approximation.

$$
\begin{aligned}
L(a,p) \equiv& \tilde{\pi}(a, p_{NC}) - \tilde{\pi}(a,p) \\
=& \frac{\pi_{22}}{2}\left(p_{NC}^2 - p^2\right) + \pi_{12} a \left(p_{NC} - p\right) \\
=& \frac{\pi_{12}}{2}\left(a^2 - p^2\right) - \pi_{12} a \left(a + p\right) \\
=& -\frac{\pi_{12}}{2}\left(a + p\right)^2.
\end{aligned}
$$

This is a quadratic loss function as required by Theorem 2. The other requirement of the rational inattention theorem is a Gaussian prior. Thus, I assume that $a \sim \mathcal{N}(0, \sigma_a^2)$. Then the previous derivations together with Theorem 2 and Corollary 1 prove the following theorem.

Proposition 2. *Let $a_i \sim \mathcal{N}(0, \sigma_a^2)$ with $a_i \perp a_j$ for all $i \neq j$ and each firm solves (5) with a log-quadratic approximation around the non-stochastic steady state. Then, the optimal prices without an information constraint are given by*

$$
\begin{aligned}
p_{i,NC} &= -a_i \\
p_{i,NC} &\sim \mathcal{N}(0, \sigma_a^2).
\end{aligned}
$$

If there is an information constraint κ, then the unconditional distribution of optimal prices is given by

$$
p_i \sim \mathcal{N}(0, (1 - e^{-2\kappa})\sigma_a^2).
$$

The last statement is equivalent to first choosing a signal of the form

$$
\begin{aligned}
s_i &= a_i + \epsilon_i \\
\textit{with } \epsilon_i &\sim \mathcal{N}\left(0, \frac{\sigma_a^2}{e^{2\kappa} - 1}\right) \textit{ and } a_i \perp \epsilon_i \perp \epsilon_j, \forall i \neq j,
\end{aligned}
$$

then setting the price according to

$$p_i = -\mathbb{E}[a_i|s_i] = -\left(1 - e^{-2\kappa}\right) s_i.$$

Discussing some comparative statics of this behavior reveals some intuitive interpretation of the price setting. If $\kappa \to \infty$, then $p_i \to p_{i,NC}$, which makes sense, because an infinite amount of information capacity allows to set the optimal price, which is the price without an information constraint. On the other hand, if $\kappa \to 0$, then $p_i \to 0$: if the agents are not allowed to process any information, then they set the steady state value. The last case obtains also if $\sigma_a^2 \to 0$: if there is no variation in a, then the economy is in the NSSS.

Ultimately, the goal is to conduct welfare analysis by comparing the decentralized equilibrium with the solution of a social planner who maximizes output. The social planner chooses quantities directly without considering prices. For this reason the optimal behavior of firms in terms of quantities is needed as well. Fortunately, there is a simple relationship between optimal pricing and optimal quantities, which is summarized in the following corollary.

Corollary 2. *If firms choose quantities instead of prices, the optimal behavior is*

$$q_i = -\eta p_i.$$

This relationship holds with and without an information constraint.

Proof. The demand function (3) relates prices to quantities. Expressing Q and P in log-deviations from the NSSS yields the result without an information constraint. With an information constraint, the result follows from Proposition 1. □

Now—as the individual behavior is characterized—it is possible to aggregate and obtain the other equilibrium objects. For the aggregation I work with the optimal pricing behavior. Of course, by the relation in Corollary 2 the aggregate equilibrium objects are the same in either case. The remaining equilibrium objects are aggregate demand and the price level. With these objective we can characterized the equilibrium as follows.

Theorem 3. *Let $a_i \sim \mathcal{N}(0, \sigma_a^2)$ with $a_i \perp a_j$ for all $i \neq j$ and each firm solves (5) with a log-quadratic approximation around the non-stochastic steady state. Then the individual behavior in equilibrium as defined by Definition 3 is characterized by Proposition 2. Although, Definition 3 needs to be adjusted so that f_i solves (5) with the log-quadratic approximation of the profit function.*[8] *Furthermore, in the limit as $N \to \infty$ the equilibrium aggregate price level is given by*

$$\mathcal{P} = \bar{P} e^{\frac{1}{2}(1-\eta)\left(1-e^{-2\kappa}\right)\sigma_a^2}$$

and aggregate output is

$$Y = e^{-\frac{1}{2}\eta\left(1-e^{-2\kappa}\right)\sigma_a^2}.$$

Proof. The first part is Proposition 2. Recall that $p_i \sim \mathcal{N}(0, (1-e^{-2\kappa})\sigma_a^2)$. Using the definition of the price level allows to solve for the equilibrium price level.

$$\begin{aligned}
\mathcal{P} &= \lim_{N\to\infty} \bar{P} \left(\frac{1}{N} \sum_{i=1}^{N} e^{p_i(1-\eta)} \right)^{\frac{1}{1-\eta}} \\
&= \bar{P} \left(\mathbb{E}\left[e^{p_i(1-\eta)} \right] \right)^{\frac{1}{1-\eta}} \\
&= \bar{P} e^{\frac{1}{2}(1-\eta)\left(1-e^{-2\kappa}\right)\sigma_a^2}
\end{aligned}$$

The second equality follows form a weak law of large numbers (WLLN) as N approaches infinity. For the second equality note that the expected value in the second line is the moment generating function (MGF) of normal distribution evaluated at $1-\eta$.

[8] Again, due to the functional forms, the equilibrium under full information without approximation is the same as with approximation and full information.

The aggregate output is determined by labor market clearing and the aggregate labor constraint:

$$
\begin{aligned}
1 &= \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} L_i = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} Q_i \\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} Y \left(\frac{\bar{P}e^{p_i}}{\mathcal{P}}\right)^{-\eta} \\
&= Y e^{\frac{1}{2}\eta(1-\eta)\left(1-e^{-2\kappa}\right)\sigma_a^2}\mathbb{E}[e^{-\eta p_i}] \\
&= Y e^{\frac{1}{2}\eta\left(1-e^{-2\kappa}\right)\sigma_a^2}
\end{aligned}
$$

Again, third and fourth equality follow from a WLLN with N approaching infinity and the MGF, respectively. □

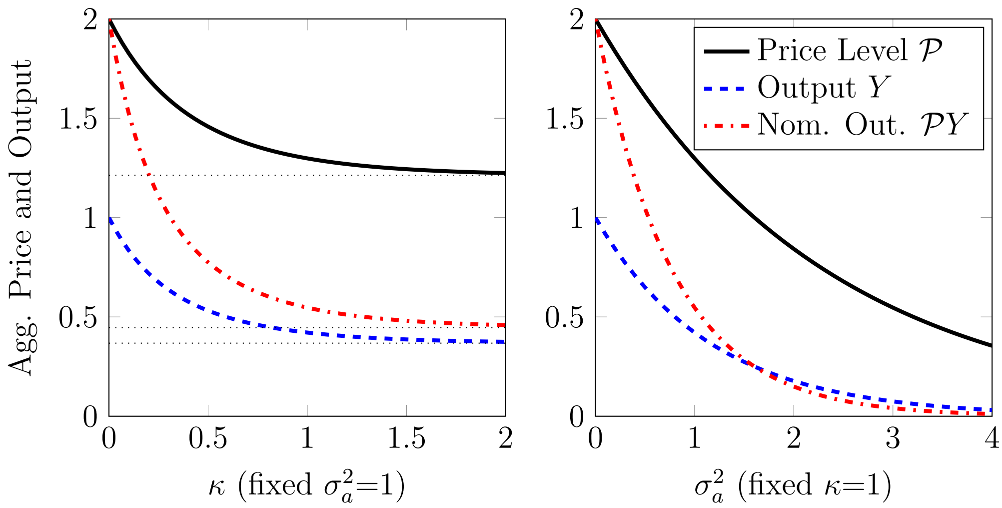

Figure 1: Comparative statics of κ and σ_a^2

To see what this equilibrium implies, Figure 1 plots the aggregate variables for different values of information capacity κ and variance of the productivity σ_a^2. For this specific case, I used $\eta = 2$. However, the qualitative results of the comparative statics are general and not specific to the choice of parameters. As for the individual behavior, zero information capacity (or zero variance of a) leads to the steady state outcome of Y and $\mathcal{P}$. If $\kappa \to \infty$, the equilibrium without an information constraint is approached. This full information equilibrium is indicated with the dotted lines in the left figure. The decrease in price level and aggregate output with an increase in κ follows the same logic as with an increase in σ_a^2. The behavior of the price level is intuitive, because a higher σ_a^2 (or κ) gives some firms the possibility to set a lower price and have a bigger market share. In turn, the aggregate price level decreases as well. To see why aggregate output drops as well, note that the Dixit–Stiglitz aggregator is symmetric in its arguments. With the aggregate labor constraint in place, the highest output is achieved in the NSSS. As soon as there is some variation, the symmetry and the labor constraint combined lead to a reduction in aggregate output. Nominal demand is the product of the price level and (real) output. Thus, the behavior of nominal demand is a combination of the other functions. It decreases faster than real output with an increase in both, σ_a^2 and κ.

In the decentralized economy allowing more information processing leads to a decrease in output. Although, this result is mainly driven by—as explained above—the symmetry of the Dixit–Stiglitz aggregator and the aggregate labor constraint. Therefore, the question arises how does the decentralized equilibrium perform compared to a social planner's solution. Is it always better to have full information for a social planner or does a similar issue arise as well? These and other questions are answered next.

3.1.1 Social Planner

How would a social planner allocate the production if he faces similar information constraints as the individual firms do in the last section? The objective of the social planner is to maximize the utility of the representative household. However, he takes the effects of different productivity levels and the aggregate labor constraint into account. This allows me to relate the decentralized economy with the social planner's solution. Two possible differences might appear. On the one hand, there is the direct effect of inefficient individual's actions. On the other hand, the individual behavior might lead to further distortions on the aggregate level. To study the social optimal choice, I need to proceed in a similar manner as in the decentralized economy.

In general the social planner solves

$$
\begin{aligned}
\max_{f} & \left\{ \mathbb{E}\left[\Psi(\boldsymbol{A},\boldsymbol{Q})\right] = \int_{\boldsymbol{\mathcal{A}}}\int_{\boldsymbol{\mathcal{Q}}} \left(\frac{1}{N}\sum_{i=1}^{N} A_i Q_i^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} f(\boldsymbol{A},\boldsymbol{Q})\,\mathrm{d}\boldsymbol{Q}\,\mathrm{d}\boldsymbol{A} \right\} \qquad (10)\\
\text{s.t. } & \int_{\boldsymbol{\mathcal{Q}}} f(\boldsymbol{A},\boldsymbol{Q})\,\mathrm{d}\boldsymbol{Q} = g(\boldsymbol{A}), \quad \forall \boldsymbol{A} \in \boldsymbol{\mathcal{A}}\\
& f(\boldsymbol{A},\boldsymbol{Q}) \geq 0, \quad \forall (\boldsymbol{A},\boldsymbol{Q}) \in \boldsymbol{\mathcal{A}} \times \boldsymbol{\mathcal{Q}}\\
& \mathcal{I}[A_i;Q_i] \leq \kappa, \quad \forall i \in \{1,\dots,N\}\\
& \frac{1}{N}\sum_{i=1}^{N} Q_i = 1,
\end{aligned}
$$

where $\boldsymbol{Q} = (Q_i)_{i=1}^{N}$ and $\boldsymbol{A} = (A_i)_{i=1}^{N}$. As in the previous section, I use a log-quadratic approximation around the NSSS. Because of the aggregate labor constraint the NSSS is the same as in the decentralized economy (see equations (6) to (9)). Note that there is no inefficiency due to monopolistic competition, because real output is pinned down by the labor constraint. In the decentralized economy the price level adjusts so that these equations constitute an equilibrium. The social planner chooses the quantities directly, but of course has to respect the aggregate labor constraint.

The next step is to take a log-quadratic approximation around this NSSS. However, it is useful to have a closer look at the Dixit–Stiglitz aggregator first. By construction (and assumption $\eta > 1$), the aggregator is concave in every argument (Dixit and Stiglitz, 1977). This is not the case anymore if one considers the log-deviations from any positive point $\tilde{Q}$ at the point of zero deviations. Formally, this means

$$
\frac{\partial^2}{\partial q_i^2}\,\tilde{Q}\left(\frac{1}{N}\sum_{i=1}^{N} A_i e^{q_i \frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\Bigg|_{\substack{\forall A_i=1\\ \forall q_i=0}} = \tilde{Q}\frac{1}{N\eta}\left(\frac{1}{N}+\eta-1\right) > 0.
$$

This derivative is positive everywhere, not only if $q_i = 0$ and $A_i = 1$ for all $i \in \{1, \ldots, N\}$. Hence, the Dixit–Stiglitz aggregator is convex in each argument for log-deviations. For illustration purposes, it is instructive to have a closer look at the aggregator for the two good case at steady state productivity, i.e. $N = 2$ and $A_i = 1$. The aggregator for level arguments is just the standard two goods CES utility

$$U(Q_1, Q_2) = \left(\frac{1}{2} Q_1^{\frac{\eta-1}{\eta}} + \frac{1}{2} Q_2^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} . \tag{11}$$

The log-deviations aggregator reads as

$$U(q_1, q_2) = \tilde{Q} \left(\frac{1}{2} e^{q_1 \frac{\eta-1}{\eta}} + \frac{1}{2} e^{q_2 \frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} . \tag{12}$$

Graphically, this is illustrated in Figure 2 for the case of $N = 2$ and $\eta = 3$. In the left panel the level sets for the first case (11) is plotted. It has the usual convex *upper* contour sets. On the contrary, the graph on the right panel has convex *lower* contour sets. This graph corresponds to the case of log-deviations as in (12).

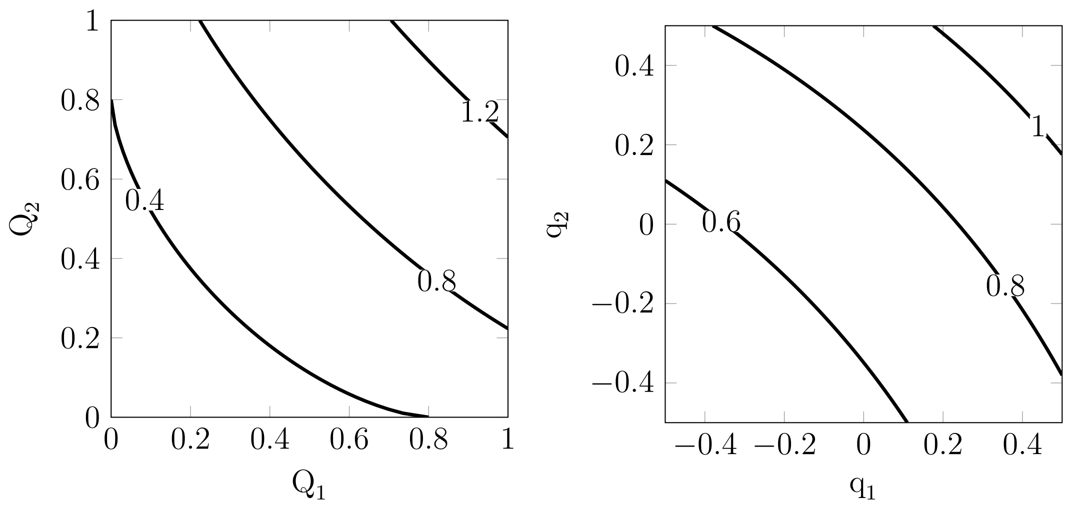

Figure 2: Level sets of the Dixit–Stiglitz aggregator

This observation is important for the further analysis. The goal is to derive a quadratic loss function to apply Theorem 2. If the objective is convex, a corner solution is the optimal choice. Hence, a quadratic loss function cannot be derived. This is a fundamental deficiency of the log-approximation, which changes the nature of the problem so that the usual technique is not applicable. Maybe this is the reason why no one conducted a social planner's analysis for such a problem before. However, I pursue a different way to get an analytic solution to the social planner's problem and use a quadratic approximation in levels. Expressing the Dixit–Stiglitz aggregator in level deviations from a particular value preserves the concavity. However, this social planner's solution is not comparable to the previously presented decentralized equilibrium. Hence, I characterize the decentralized equilibrium with a quadratic approximation in levels first.

3.2 The level-quadratic case

In the previous section I derived the optimal behavior of firms facing an information constraint and discussed the implications for the decentralized economy. I found that a comparison with a social planner's solution is not possible, because the Dixit–Stiglitz aggregator is convex in its arguments expressed in log-deviations. For this reason a level approximation is needed. To get a valid comparison, the decentralized economy with level-deviations is presented first. The steps are very similar to the log-quadratic case, also note that the NSSS is, of course, the same.

I start out with expressing the objective of firm, the profits Π, in level deviations from the NSSS.

$$\pi(a,q) = \Pi\left(\bar{A}(1+a), \bar{Q}(1+q)\right)$$

Note that I work with the quantity choice right away. This gives the relevant behavior which is comparable to the social planner's behavior. For the second order Taylor approximation, the partial derivatives evaluated at zero are needed. They are

$$\begin{aligned}
\pi_1 &= \left.\frac{\partial \pi(a,q)}{\partial a}\right|_{\substack{q=0\\a=0}} = Y^{\frac{1}{\eta}}\mathcal{P} \\
\pi_2 &= \left.\frac{\partial \pi(a,q)}{\partial q}\right|_{\substack{q=0\\a=0}} = \frac{\eta-1}{\eta}Y^{\frac{1}{\eta}}\mathcal{P}\bar{A}\bar{Q}^{\frac{\eta-1}{\eta}} - \bar{Q} = 0 \\
\pi_{11} &= \left.\frac{\partial^2 \pi(a,q)}{\partial a^2}\right|_{\substack{q=0\\a=0}} = 0 \\
\pi_{22} &= \left.\frac{\partial^2 \pi(a,q)}{\partial q^2}\right|_{\substack{q=0\\a=0}} = -\frac{1}{\eta}\frac{\eta-1}{\eta}Y^{\frac{1}{\eta}}\mathcal{P}\bar{A}\bar{Q}^{\frac{\eta-1}{\eta}} < 0 \\
\pi_{12} &= \left.\frac{\partial^2 \pi(a,q)}{\partial a \partial q}\right|_{\substack{q=0\\a=0}} = \frac{\eta-1}{\eta}Y^{\frac{1}{\eta}}\mathcal{P}\bar{A}\bar{Q}^{\frac{\eta-1}{\eta}} = -\eta\pi_{22} > 0.
\end{aligned}$$

This implies—and is easily verified—that the optimal quantity with full information is

$$q_{NC} = \eta a.$$

Note that this is the same functional form as in the log-quadratic case (see Corollary 2). Deviations from this optimal behavior yield a quadratic loss function:

$$L(x,q) = \frac{\pi_{12}}{2}\left(\eta a - q\right)^2.$$

As $\pi_{12} > 0$, the loss is positive as it should be. This result allows to apply Theorem 2 and the results are summarized in the following theorem. The distribution of the deviations is adjusted to match the first two moments of the log-normal distribution of the choice as derived in the previous section. With this adjustment the problem is still parametrized by the variance of the log-deviations σ_a^2.

Proposition 3. *Let $a_i \sim \mathcal{N}\left(\frac{1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right), e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right)$ with $a_i \perp a_j$ for all $i \neq j$ and each firm solves (5) with a level-quadratic approximation around the non-stochastic steady state. Then, the optimal quantities without an information* *constraint are given by*

$$
\begin{aligned}
q_{i,NC} &= \eta a_i \\
q_{i,NC} &\sim \mathcal{N}\left(e^{\frac{1}{2}\sigma_a^2} - 1, \eta^2 e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right).
\end{aligned}
$$

If there is an information constraint κ, then the unconditional distribution of optimal quantities is given by

$$
q_i \sim \mathcal{N}\left(e^{\frac{1}{2}\sigma_a^2} - 1, \eta^2(1 - e^{-2\kappa})e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right).
$$

The last statement is equivalent to first choosing a signal of the form

$$
\begin{aligned}
s_i &= a_i + \epsilon_i \\
\text{with } \epsilon_i &\sim \mathcal{N}\left(0, \frac{e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)}{e^{2\kappa} - 1}\right) \text{ and } a_i \perp \epsilon_i \perp \epsilon_j, \forall i \neq j,
\end{aligned}
$$

then setting the price according to

$$
q_i = \eta \mathbb{E}[a_i | s_i] = \eta\left(1 - e^{-2\kappa}\right) s_i.
$$

Proof. Follows from the previous discussion, Theorem 2, and Corollary 1. □

The intuition and the comparative statics are very similar to the log-quadratic case. A higher information capacity brings the behavior closer to the behavior under full information. Asymptotically, the information constraint behavior converges to the full information case. Whenever $\kappa \to 0$, firms behave as if the economy is in the non-stochastic steady state.

The fact that the behavior of a level approximation is the same as in the log-approximation gives hope that two models are comparable. To see if this is indeed the case, I look at the other aggregate variables which determine the equilibrium.

Labor market clearing, the aggregate labor constraint, and the limit as $N \to \infty$ pin down Q:

$$\begin{aligned}1 &= \lim_{N\to\infty} Q\frac{1}{N}\sum_{i=1}^{N}(1+q_i)\\ &= Q\mathbb{E}[1+q_i] = Qe^{\frac{1}{2}\sigma_a^2}.\end{aligned}$$

Now we can characterize aggregate output. It is also possible to derive the price level. However, the price level does not have a nice algebraic expression and does not give much insight. Thus I present and discuss only output of the economy.

$$Y = \lim_{N\to\infty} Q\left(\frac{1}{N}\sum_{i=1}^{N}(1+q_i)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} = Q\left[\mathbb{E}\left[(1+q_i)^{\frac{\eta-1}{\eta}}\right]\right]^{\frac{\eta}{\eta-1}}$$

Unfortunately, these equations involve non-integer moments because $\eta > 1$. However, note that if $\sigma_a^2 = 0$, the output in the NSSS is obtained. The same is obtained in the case of $\sigma_a^2 > 0$ and $\kappa = 0$. In the latter case firms just behave as in the NSSS. For the intermediate case the non-integer moment can be calculated numerically. However, to get an analytical expression, I present a second order Taylor approximation of the expected value as well.

$$\begin{aligned}Y &\approx Q\mathbb{E}\left[1+\frac{\eta-1}{\eta}q - \frac{\eta-1}{2\eta^2}q^2\right]^{\frac{\eta}{\eta-1}}\\ &= Q\left[1+\frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2}-1\right) - \frac{\eta-1}{2\eta^2}\left(\mathbb{V}[q]+\mathbb{E}[q]^2\right)\right]\\ &= Q\Bigg[1+\frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2}-1\right) - \\ &\qquad \frac{\eta-1}{2\eta^2}\left(\eta^2(1-e^{-2\kappa})e^{\sigma_a^2}\left(e^{\sigma_a^2}-1\right)+\left(e^{\frac{1}{2}\sigma_a^2}-1\right)^2\right)\Bigg]\end{aligned}$$

An analytical expression is useful for the comparison with the social planner solution. Also note that σ_a^2 needs to be sufficiently small for a valid comparison of the log-quadratic and level-quadratic case. To see how close the different values are, Figure 3 plots a comparison of output for the log- and the level-approximation. The black line represents the output in the log-quadratic case. Blue and red lines are the exact and the approximated output for the level case, respectively.

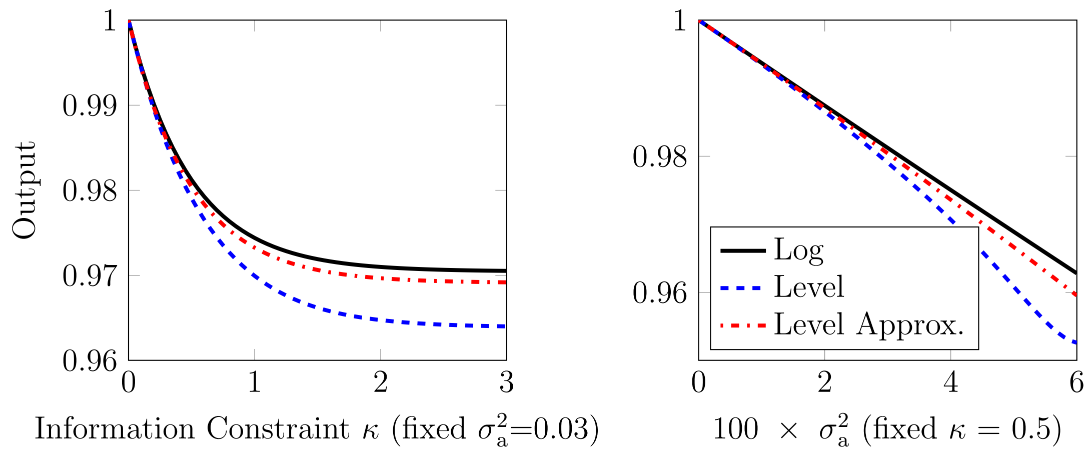

Figure 3: Model comparison of log- and level-deviations

The values are the same for the case with $\kappa = 0$ or $\sigma_a^2 = 0$. This is not surprising, because this is how the models were constructed. With an increase in either κ or σ_a^2 the difference increases, but is still reasonable small. An important issue is that the overall shape is preserved as well. Furthermore, note that the approximated expected value in the level case is closer to the log case than the exact output for the level case. This relates also to the non-monotonic behavior of output at higher variances for the exact level case (see right graph). To avoid complex numbers, I approximate the normal distribution for the level-deviation by a folded normal around zero. Thus, a greater variance puts more mass below zero and this changes the distribution of the folded normal. In turn, this approximation might not be accurate for large σ_a^2.

Although, the depicted variances seem small, they are not in economic terms. To see why, consider the case of $\sigma_a^2 = 0.05$. This implies a standard deviation of about 0.51 for the normal distribution of the level-deviations of the quantities. Economically, this means that about 32% of the quantity realizations are 51%

above or below the NSSS. Even more drastically, about 6.5% of the realizations are 102% below the NSSS. These realizations imply a negative quantity, thus they hit the lower bound of zero.

Overall, the log-quadratic and the level-quadratic equilibrium seem to be comparable for small productivity variances. The small variances are not necessary small in an economic context. This conclusion justifies the comparison of a social planner's solution with a level-quadratic approximation to the decentralized economy.

3.2.1 Social Planner

Previously, it was shown that a social planner's problem in the log-quadratic case leads to a convex objective in a maximization problem, which implies a corner solutions. For an analytic solution a quadratic loss function is needed (see Theorem 2). With a corner solution it is not possible to derive a quadratic loss function. The problem of convexity does not arise with an approximation in level-deviations. For the decentralized economy, it does not make a big difference if a log- or level-deviations approach is used given that the variance is sufficiently small.

As discussed before, the social planner solves (10). However, now I consider a level-quadratic approximation of the objective. For this reason express the objective in level-deviations from the NSSS.

$$\psi(\boldsymbol{a}, \boldsymbol{q}) = \Psi(\bar{A}\,(1+\boldsymbol{a})\,, \bar{Q}\,(1+\boldsymbol{q})) = \bar{A}\bar{Q}\left(\frac{1}{N}\sum_{i=1}^{N}\left[(1+a_i)(1+q_i)\right]^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \tag{13}$$

Note that even if $\bar{Q} \neq 1 \neq \bar{A}$, it would not influence the analysis, because it is just a positive constant multiplying the other term. As before, I use a second order Taylor approximation around the NSSS.

The partial derivatives evaluated at the origin are

$$\begin{aligned}
\psi_1 &= \frac{\partial\psi(\boldsymbol{a},\boldsymbol{q})}{\partial a_i}\bigg|_{\substack{\forall a_i=0\\ \forall q_i=0}} = \frac{1}{N}\bar{A}\bar{Q}\\
\psi_2 &= \frac{\partial\psi(\boldsymbol{a},\boldsymbol{q})}{\partial q_i}\bigg|_{\substack{\forall a_i=0\\ \forall q_i=0}} = \psi_1\\
\psi_{11} &= \frac{\partial^2\psi(\boldsymbol{a},\boldsymbol{q})}{\partial a_i^2}\bigg|_{\substack{\forall a_i=0\\ \forall q_i=0}} = \frac{1}{N}\frac{\bar{A}\bar{Q}}{\eta}\left(\frac{1}{N}-1\right) \approx -\frac{1}{N}\frac{\bar{A}\bar{Q}}{\eta} < 0\\
\psi_{22} &= \frac{\partial^2\psi(\boldsymbol{a},\boldsymbol{q})}{\partial q_i^2}\bigg|_{\substack{\forall a_i=0\\ \forall q_i=0}} = \psi_{11} < 0\\
\psi_{12} &= \frac{\partial^2\psi(\boldsymbol{a},\boldsymbol{q})}{\partial a_i\partial q_i}\bigg|_{\substack{\forall a_i=0\\ \forall q_i=0}} = \frac{1}{N}\frac{\frac{1}{N}-1+\eta}{\eta}\bar{A}\bar{Q} \approx \frac{1}{N}\frac{\eta-1}{\eta}\bar{A}\bar{Q}.
\end{aligned}$$

The approximations hold for sufficiently large values of N. Here, we also see that the concavity is preserved as $\psi_{22} < 0$. One might wonder why ψ_2, the first derivative with respect to the choice variable, is not zero. In the firms problem this derivative is zero, which follows directly from the optimality condition in the NSSS. The firm does not have to take into account the aggregate labor constraint. The social planner, on the other hand, has to satisfy this constraint. In this case the partial derivative is not zero, but is the shadow price (the Lagrange multiplier) of the aggregate labor constraint.

Using these derivatives for a second order Taylor approximation and dropping terms independent of the choice, the social planner with full information solves

$$\begin{aligned}
&\max_{\boldsymbol{q}} \sum_{i=1}^N q_i - \frac{1}{2\eta}\sum_{i=1}^N q_i^2 + \frac{\eta-1}{\eta}\sum_{i=1}^N a_i q_i\\
&\text{s.t.} \sum_{i=1}^N q_i = 0.
\end{aligned}$$

Solving for the optimal behavior yields $q_{i,SP,NC} = (\eta-1)a_i$. Note that the coefficient is smaller than the coefficient in the decentralized solution η. That is, under perfect information the individual firms respond too strongly to a change in their productivity. The social planner adjusts the quantity not as radically as an individual firm given a change in productivity. In the decentralized economy a simple tax on quantity adjustments can restore the social optimal easily.

Proposition 4. *A tax rate $\tau = {}^{1}\!/_{\eta}$ on quantity adjustments restores the optimal policy of a social planner.*

Proof. We are looking for a τ such that

$$(1-\tau)q_{NC} = q_{SP,NC}.$$

Plugging in yields the desired result:

$$\begin{aligned}(1-\tau)\eta a &= (\eta-1)a \\ \tau &= \frac{1}{\eta}.\end{aligned}$$

□

This result is very intuitive. The elasticity of substitution η determines the monopoly mark-up charged by firms. A high value of η implies only a small mark-up, i.e. only small monopoly power. In the limit, $\eta \to \infty$, the monopoly power vanishes and the economy approaches perfect competition. In this case the tax rate τ approaches zero as there is no distortion due to monopolistic competition. If η decreases, the monopoly power increases and therefore the tax rate increases as well. Note that the tax rate result carries through to the case with limited information capacity. To see this, the behavior with restricted information flow is needed first. The optimal behavior under full information implies a quadratic loss function due to suboptimal behavior.

$$
\begin{aligned}
L(\boldsymbol{a},\boldsymbol{q}) =& \psi_2 \sum_{i=1}^{N} \left(q_{i,SP,NC} - q_i\right) + \frac{1}{2}\psi_{22} \sum_{i=1}^{N} \left(q_{i,SP,NC}^2 - q_i^2\right) \\
&+ \psi_{12} \sum_{i=1}^{N} \left(q_{i,SP,NC} a_i - q_i a_i\right) \\
\triangleq& \sum_{i=1}^{N} q_i + \frac{1}{2\eta} \sum_{i=1}^{N} \left[(\eta-1)^2 a_i^2 - 2(\eta-1) q_i a_i + q_i^2\right] \\
=& \frac{1}{2\eta} \sum_{i=1}^{N} \left[(\eta-1)\, a_i - q_i\right]^2
\end{aligned}
\tag{14}
$$

In the third line, I drop the terms independent of the choice q_i. They do not affect the behavior anyway. The last equality holds due to the aggregate labor constraint. With the assumption that a_i is independent of a_j for all $i \neq j$ and because the information constraint is separately given for each good,[9] it is possible to consider each element of the sum separately and apply Theorem 2. However, also the aggregate labor constraint needs to be fulfilled. This is $\frac{1}{N}\sum_{i=1}^{N} q_i = 0$. Given that N is sufficiently large, this does not cause any problems as a law of large numbers applies and imposing $\mathbb{E}[q_i] = 0$ makes sure that the labor constraint is met in the limit.[10] The previous discussion, together with Theorem 2 proves most of the following theorem.

Theorem 4. *Let* $a_i \sim \mathcal{N}\left(\frac{1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right), e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right)$, $\forall i$, $a_i \perp a_j$, $\forall i \neq j$, *and the social planner solves* (10) *with a level-quadratic approximation around the non-stochastic steady state. Then, the output maximizing quantities without*

[9]Later in Subsection 3.3 I show that having one overall information constraint leads to the same outcome.

[10]As in the decentralized economy, the moments is matched with the log-normal distribution of the log-deviation case. In order to get rid of the level effect of the log-normal distribution, I relax the aggregate labor constraint. This means that I allow $\mathbb{E}[q_i] = \frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right)$. This value is very close to zero for the variances under consideration.

an information constraint are given by

$$
\begin{aligned}
q_{i,SP,NC} &= (\eta - 1)a_i \\
q_{i,SP,NC} &\sim \mathcal{N}\left(\frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right), (\eta-1)^2 e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right).
\end{aligned}
$$

If there is an information constraint κ for each good, then the optimal quantities are given by

$$
\begin{aligned}
q_{i,SP} &\sim \mathcal{N}\left(\frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right), \left(1-e^{-2\kappa}\right)(\eta-1)^2 e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right) \\
q_{i,SP} &= (\eta-1)\mathbb{E}[a_i|s_i] = (\eta-1)\left(1-e^{-2\kappa}\right) s_i \\
\text{where } s_i &= a_i + \epsilon_i \\
\epsilon_i &\sim \mathcal{N}\left(0, \left(e^{2\kappa}-1\right)^{-1} e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\right), \text{ with } \epsilon_i \perp \epsilon_j, \forall i \neq j \text{ and } \epsilon_i \perp a_i.
\end{aligned}
$$

In the latter case, output is given by

$$
\Psi = \mathbb{E}\left[(1 + a_i + q_i + a_i q_i)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}
$$

Proof. Everything but Ψ follows from the previous discussion and Corollary 1. Output follows from (13) and a law of large numbers as N tends to infinity. □

Comparing the behavior with an information constraint using the signal reveals that the result about the tax rate applies here as well. In general, the behavior under full information carries through to the case with an information constraint. The social planner makes smaller adjustments to productivity deviations than an individual firm does.

Here, the expression for output not only involves a non-integer moment, but also random variables which are not normally distributed. Thus, only some special cases can be considered analytically. The behavior with an information constraint asymptotically approaches the full information case. Therefore, also output converges to the full information case. If $\sigma_a^2 \to 0$, the output converges to the NSSS.

If the social planner is not allowed to process any information ($\kappa = 0$), then $q_i = \frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right)$ and output is

$$\left[1 + \frac{\eta-1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right)\right] \mathbb{E}\left[(1+a_i)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}. \tag{15}$$

This still involves a non-integer moment, but now the random variable is Gaussian which helps investigating the behavior numerically. However, an analytical expression, also for $\kappa > 0$, might help to understand output for the social planner's choice. In the decentralized economy, it turned out that a second order Taylor approximation of the term in the expectation was very accurate. Using this approximation in the social planner's context yields an approximate output of

$$\begin{aligned}
\Psi \approx & \mathbb{E}\left[1 + \frac{\eta-1}{\eta}(a_i + q_i) - \frac{\eta-1}{2\eta^2}\left(a_i^2 + q_i^2\right) + \left(\frac{\eta-1}{\eta}\right)^2 a_i q_i\right]^{\frac{\eta}{\eta-1}} \\
= & \left\{1 + \frac{\eta}{\eta-1}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right)\right. \\
& - \frac{\eta-1}{2\eta^2}\left[e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right)\left(1 + \left(1 - e^{-2\kappa}\right)(\eta-1)^2\right) + \left(e^{\frac{1}{2}\sigma_a^2} - 1\right)^2 \frac{1 + (\eta-1)^2}{\eta^2}\right] \\
& \left. + \frac{(\eta-1)^3}{\eta^2}\left(1 - e^{-2\kappa}\right)\left[e^{\sigma_a^2}\left(e^{\sigma_a^2} - 1\right) + \frac{1}{\eta^2}\left(e^{\frac{1}{2}\sigma_a^2} - 1\right)^2\right]\right\}^{\frac{\eta}{\eta-1}}.
\end{aligned}$$

Figure 4 adds the approximated social planner's output Ψ to the left panel of Figure 3. The graph reveals that the difference between the social planner's solution and the decentralized economy increases with κ. With a high value of κ the full information is approached and the difference is the usual inefficiency due to monopolistic competition. If less information processing is allowed the gap narrows. If $\kappa \to 0$, the social planner and the firms set the NSSS values. The gap arises because $\sigma_a \neq 0$. Thus, there is a distortions, because firms never take the aggregate labor constraint explicitly into account. The labor market clears because aggregate demand adjusts downward accordingly. The social planner always takes the labor constraint explicitly into account. Therefore, the social planner is able to use movements in productivity explicitly to increase output.[11] The dashed line indicates Ψ for $\kappa = 0$ using (15). This line indicates that the approximation is close to the true value - at least at $\kappa = 0$.

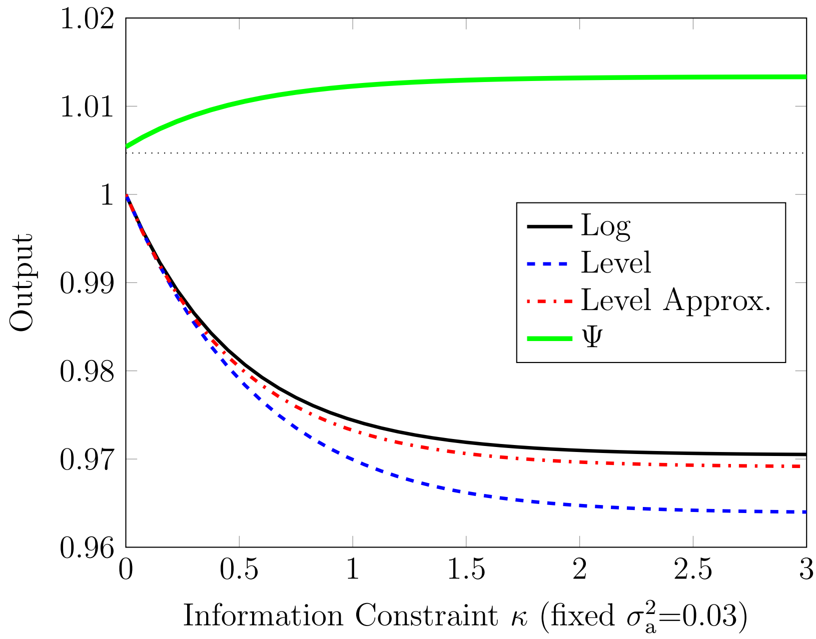

Figure 4: Output of social planner and decentralized economy

Figure 4 shows the output for a specific value of σ_a^2 only. However, the result is more general. For this reason, Figure 5 plots the output ratio (social planner's value divided by the decentralized output, level-approximated value) for different values of κ and σ_a^2. In the NSSS, i.e. $\sigma_a^2 = 0$, the difference is zero. This is as expected: by the aggregate labor constraint the NSSS is the same in both cases. An increase in σ_a^2 with $\kappa = 0$ leads to the gap as described above. More important is that for any value of σ_a^2 the difference increases with an increase in κ. It is also intuitive that at a lower value of σ_a^2 the full information difference is reached with lower values of κ. If there is not much variation, then having a lot information does not improve the decisions significantly.

To sum up, even with the strict assumption that the NSSS is the same for the social planner as in the decentralized economy, there emerge substantial distortions

[11]A small level shift is also due to the distributional assumptions. Also see Footnote 10. However, this effect is very small and does lead to a different conclusion.

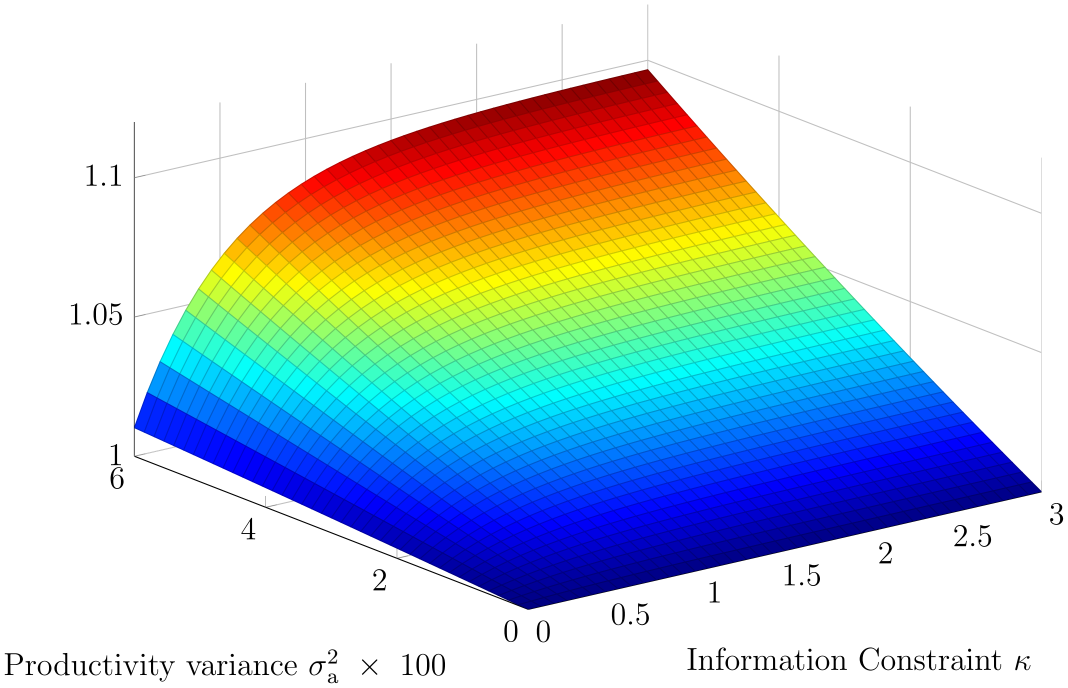

Figure 5: Output ratio of social planner and decentralized economy

outside the NSSS. At the firm level, the adjustments to changes in productivity are too large compared to a social optimal behavior—a social planner would make smaller adjustment if productivity levels change. On the aggregate this implies that the biggest difference of social planner to decentralized economy in terms of output arises in a full information equilibrium. This is independent of the productivity variance. Intuitively, in a full information equilibrium the usual distortions due to monopolistic competition arise. If the agents (social planner and firms) have finite information capacity their behavior is closer to the behavior in the NSSS. By assumption, in the NSSS the aggregate variables are the same for the social planner and for the decentralized economy.

This concludes the main analysis. Even in this simple model interesting results emerge due to limited information processing. An interesting further step would be to include an information friction on the consumer side as well. This might lead to interesting feedback effects *á la* Maćkowiak and Wiederholt (2011).

3.3 Extensions

The main analysis in the previous section revealed some interesting findings even with a quadratic approximation of the objective functions. I discuss two extensions in this section. The one is about the approximation. Do firms which solve the general problem stated in Equation 5, behave close to the behavior predicted by the quadratic approximation? The other is about the social planner. How does the social planner choose if he faces only one overall information constraint and not one constraint for each firm or good?

3.3.1 The general firm problem

In this section, I discuss the behavior of a firm which solves the general non-quadratic problem stated in (5). As the profit function is not quadratic, Theorem 2 does not apply. An analytic solution to such a problem is not known. Hence, I rely on numerical results in this section. Matejka (2010) showed that the quadratic Gaussian case is very special and a Gaussian posterior is not optimal in general. Often, this leads to discrete behavior.

In this thesis I consider four different possible setups. The first is the same setup as in the log-quadratic case. This case is presented mostly for illustration of how such problems can be analyzed. Furthermore, it is a good reference point as an analytic solution is known. The second case is that log-deviations are normally distributed, but the objective is not approximated. In the third example I consider normally distributed level-deviations, but again without approximating the objective. The last example shows the behavior if productivity is uniformly distributed. Especially the last one is in spirit of Matejka (2010). I do not consider any general solution to the social planner's problem, as such a problem is far more complex. A firms problem is relatively simple as it requires to find a joint distribution of two variables. This means that an optimization over a two dimensional grid needs to be done. In the social planner's problem the objective of interest is a joint distribution of $2 \times N$ variables, which is an optimization problem on a $2N$ dimensional grid.

For the firms problem I use the following numerical procedure. First, 25 grid points on each dimension discretize the 2 dimensional space. Second, invariance of the solution to the size of the choice space determines the size of the space under consideration. The size of the space and the 625 grid points allow to use a standard optimization routine to solve (5). The result is a joint distribution over the 2 dimensional discretized space.

The log-quadratic case. If firms solve (5) with a log-quadratic approximation of the objective an analytic solution is available. The result is stated in Proposition 2. Recall that the optimal behavior with full information is $p_{NC} = -a$. Thus, with full information all the mass of the joint distribution $f(a, p)$ would be on the diagonal. With an information constraint,[12] this cannot be achieved, but the optimal behavior puts most of the mass close to the diagonal. Figure 6 plots the optimal joint distribution $f(a, p)$ for the same economic parameters as in the examples in the previous section. The information capacity is set to $\kappa = 1$ and the prior variance is $\sigma_a^2 = 0.01$. The graph is exactly as just described. Most of the mass sits at the diagonal. Due to the information constraint not all the mass is exactly on the diagonal. The theory predicts that the posterior distribution is Gaussian as well. At first glance, this seems to be fulfilled as well. Graphically, this is confirmed by looking at the conditional distributions $f(p|a)$. Figure 7 plots these marginal distributions for five different values of a. Note that the small asymmetry at the end points is due to the approximation, because there is no grid point exactly at the mode.

Figure 8 plots the marginal distributions $f(p)$ for different values of κ. For all κ, the marginal is a normal distribution. Furthermore, the variance decreases with a decrease in κ. This is also intuitive and in line with theory. The lower κ the

[12] Note that in all the numerical examples, a finite κ would correspond to full information, because the distributions is discretized.

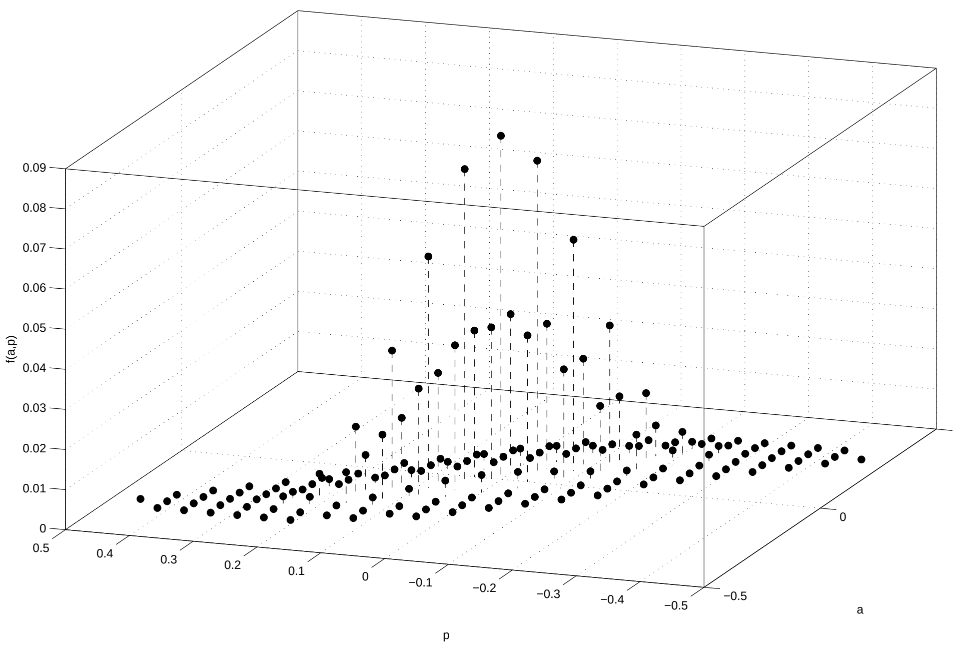

Figure 6: Optimal joint distribution $f(a,p)$ in the log-quadratic case

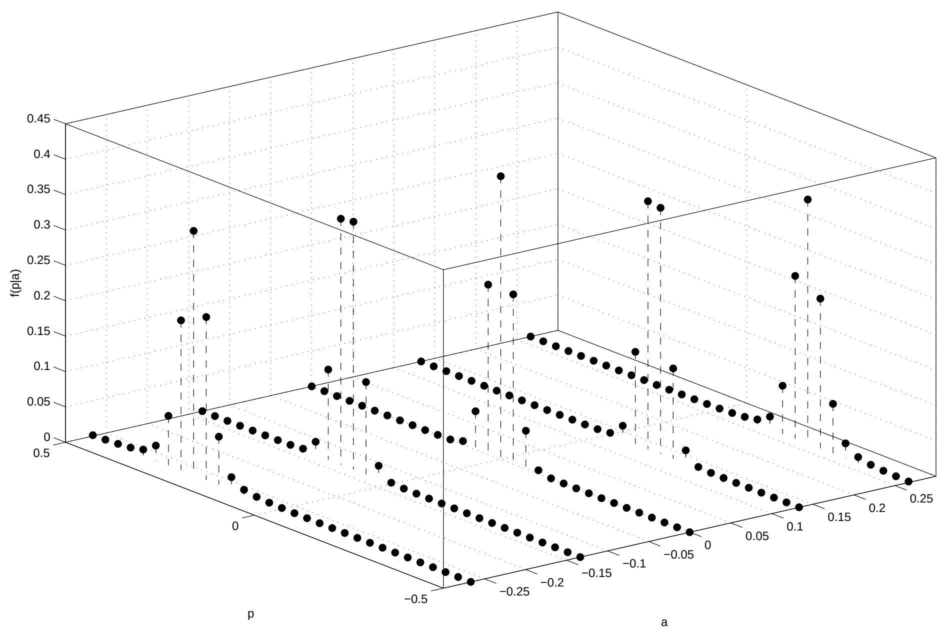

Figure 7: Selection of optimal conditional distributions $f(p|a)$ in the log-quadratic case

closer the behavior to the NSSS behavior. With $\kappa = 0.05$ almost all mass is on zero, the NSSS value. For $\kappa = 1$ and $\kappa = 2$ the distributions are almost the same. This happens because these values of κ are already close to the full information case.

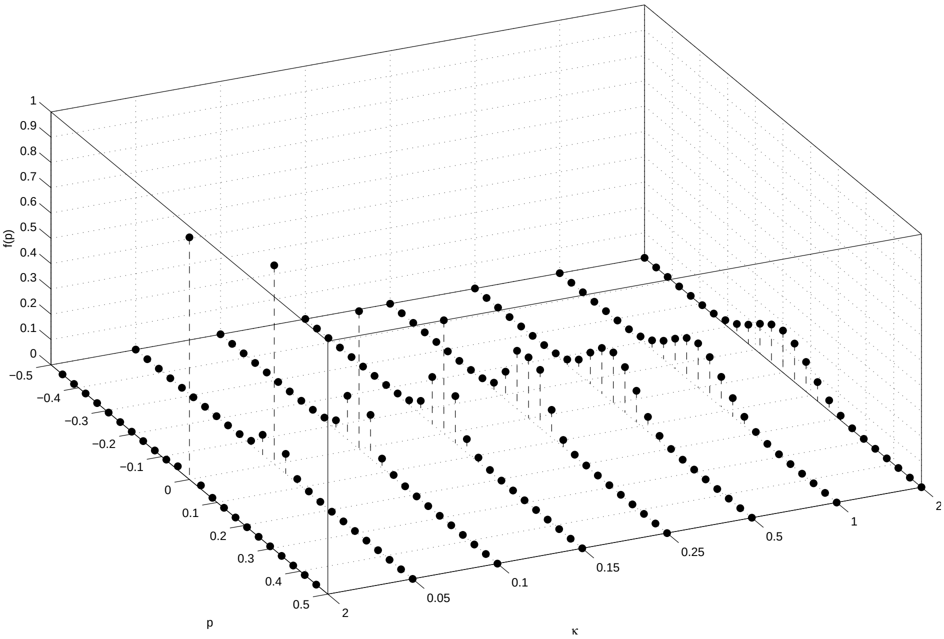

Figure 8: Optimal marginal distributions $f(p)$ in the log-quadratic case for different κ

The log-deviation case. The previous example showed that with a quadratic objective the numerical results are as predicted by theory. However, there was not much new insight in this exercise. For this reason, I consider the non-quadratic objective with log-deviations now. I keep the assumption that $a \sim \mathcal{N}(0, \sigma_a^2)$. The objective is

$$Y \left(\frac{\bar{P}}{\mathcal{P}}\right)^{-\eta} e^{-\eta p} \left(\bar{A} e^a \bar{P} e^p - 1\right),$$

where the firm takes all the upper case variables as given. Now, Theorem 2 does not apply anymore and a Gaussian posterior distribution is not guaranteed. Looking at the marginal distributions $f(p)$ in Figure 9 reveals that the distributions have a slight asymmetry for some values of κ. For this example all the parameters are the same as before. Note that with (almost) full information ($\kappa \geq 0.5$) the marginal distribution is a normal distribution. With full information it is still optimal to track the productivity change one-to-one and thus a normal distribution is optimal. With a lower information constraint the marginal distributions are not too different from a normal distribution, but have some significant asymmetries. However, the log-quadratic case seems to be a good approximation to this non-linear problem. The gain of getting an analytic solution might outweigh the approximation loss.

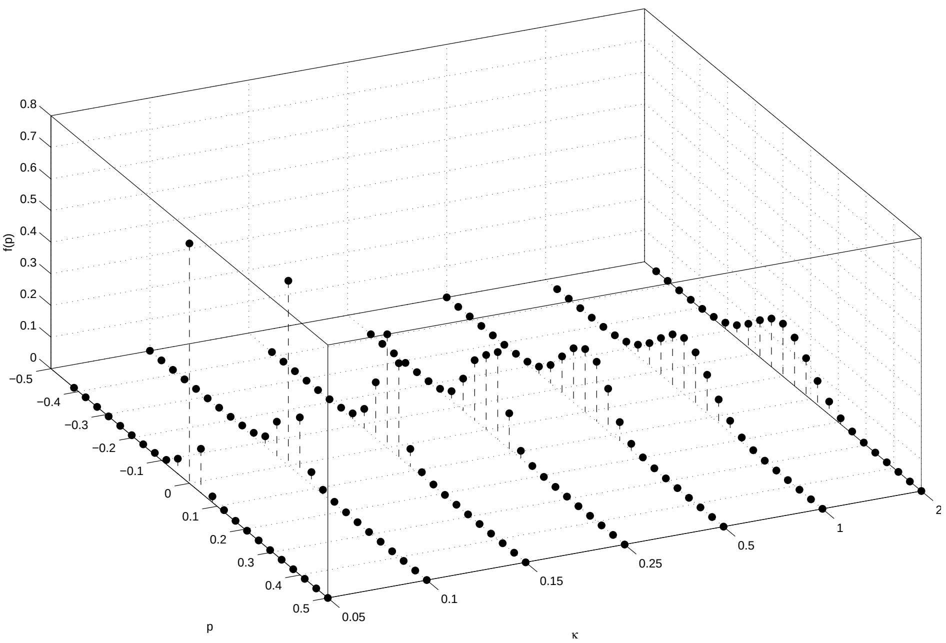

Figure 9: Optimal marginal distributions $f(p)$ in the log-deviation case for different κ

The level-deviation case. In the main analysis, I used the level-quadratic approximation of the firms objective for a comparison with the social planner's solution. For the same reason as in the analytic part I assume that the level-deviations are distributed according to $a \sim \mathcal{N}\left(\frac{1}{\eta}\left(e^{\frac{1}{2}\sigma_a^2}-1\right), e^{\sigma_a^2}\left(e^{\sigma_a^2}-1\right)\right)$.

Now, (13) is the objective and no approximation is used. The choice is the joint distribution of quantity and productivity deviations, i.e. $f(a, q)$. Figure 10 shows the optimal marginal distribution $f(q)$. In this case the linear-quadratic approximation seems to be very accurate. As in the approximation, the variance of $f(q)$ is about η^2 higher than the variance of $f(a)$ with full information. Note that the range of the q-axis is doubled compared to the previous graphs.

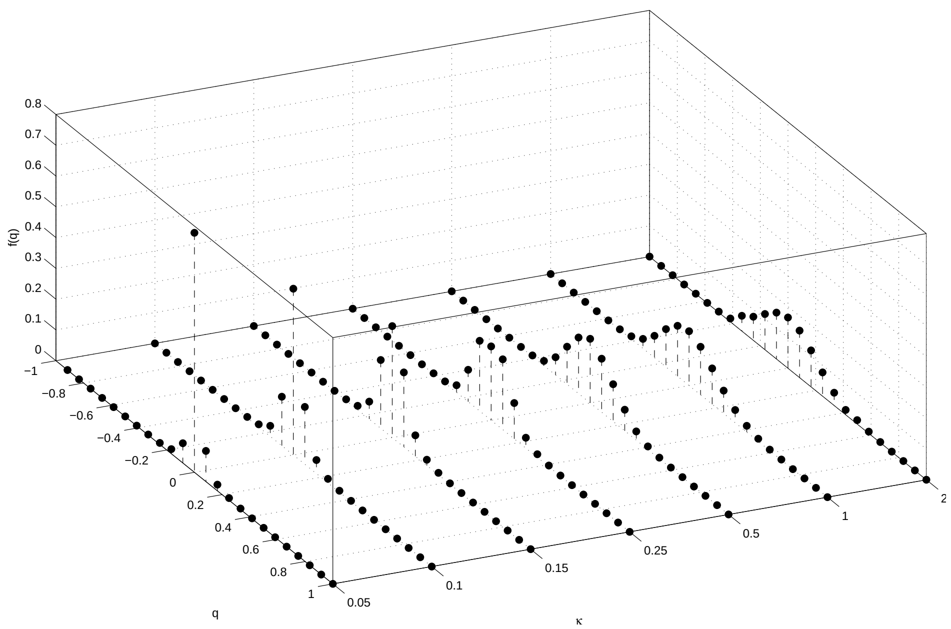

Figure 10: Optimal marginal distributions $f(q)$ in the level-deviation case for different κ

The uniform case. In this last experiment, I drop the Gaussian assumption. The firm uses the general objective (4). Now the difference is that the productivity, not the deviations, are uniformly distributed, i.e. $A \sim \mathcal{U}[0.825, 1.175]$. This distribution has mean one too, so that the NSSS is the same as before. Furthermore, the variance is similar to the variances of the normal distribution considered before. Matejka and Sims (2011) show that such problems might give rise to discrete behavior.

From Figure 11 it is not clear if discrete behavior is optimal, especially, for small values of κ. Here the objective is not quadratic and therefore the proposition about discreteness of Matejka (2010) is not applicable. However, for higher values of κ it seems that there are discrete masspoints, which is in line with the numerical results of Matejka (2010). He finds that even without an quadratic objective discrete behavior is optimal in many cases.

Some similarities with the examples before are worth mentioning. With a low κ the optimal choice approaches the NSSS behavior. An increase in κ leads to a higher variance until the full information case is reached. There the prior distribution can be tracked perfectly. For very small values of κ a marginal distribution similar to a Gaussian seems to be optimal. For higher values of κ a bimodal distribution arises where some points have a lot of mass. Also note that the mean of the distribution stays the same at $\bar{P} = \frac{\eta}{\eta-1}$.

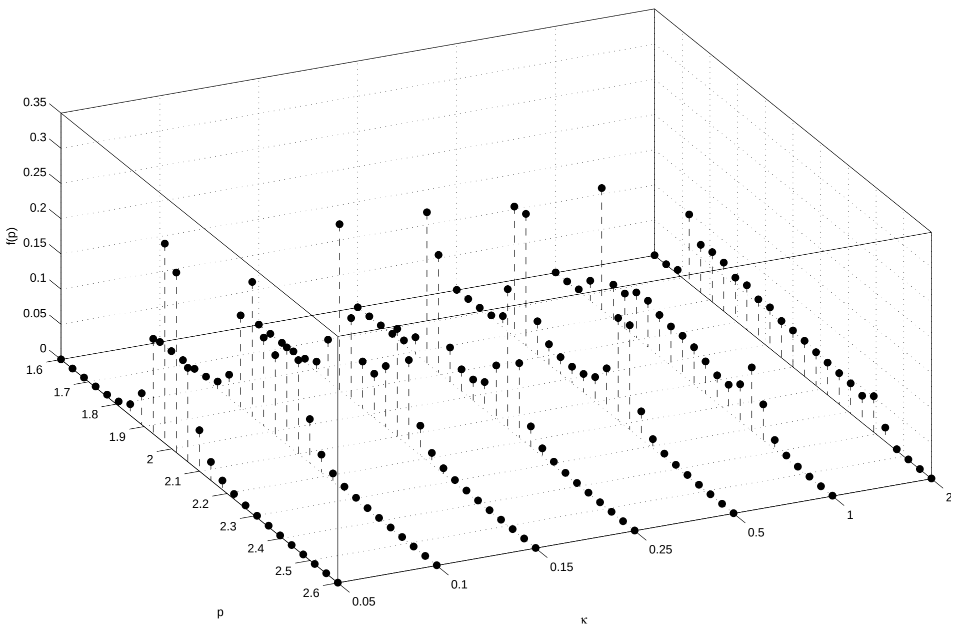

Figure 11: Optimal marginal distributions $f(p)$ in the uniform case for different κ

To sum up, the log-quadratic approximation seems to be good for relatively high values of κ. The level-quadratic approximation is very accurate for all values of

κ. In this setup even a non-Gaussian prior leads to an distribution which is close to a normal distribution for small values of κ. The clear discreteness result is not confirmed here, but this is not contradicting the formal proposition about discreteness in Matejka (2010), because the objective is not quadratic. However, a bimodal distribution with discrete masspoints arises for medium values of κ, which has some spirit of the discreteness result of Matejka (2010).

3.3.2 The Social Planner with one Information Constraint

In the main part I analyzed a planning problem, where the social planner has a separate information constraint for each good. This simplified the analysis insofar that each good could be considered independent of the others. Only the aggregate labor constraint linked the decisions about the individual goods together. One might wonder what happens if the social planner is allowed to shift the information allocation from one good to another. In other words, what are the implications of a social planner with one overall information constraint only. For comparison reason his total capacity is kept at the same level. Formally the social planner solves (10), but N information constraints are replaced by

$$\mathcal{I}[\boldsymbol{A};\boldsymbol{Q}] \leq N\kappa.$$

I assume that $a_i \sim \mathcal{N}\left(\mu,\sigma^2\right), \forall i,\ x_i \perp x_j, \forall i \neq j$. This covers the social planner's case in the previous section, but simplifies the notation. However, with this multivariate information constraint Theorem 2 does not apply directly. Fortunately, by Proposition 3 and 4 in Maćkowiak and Wiederholt (2009a) it is optimal to choose for each good a signal of the form true productivity plus independent Gaussian white noise, i.e. $s_i = a_i + \epsilon_i$ for all i with $a_i \perp \epsilon_i$ and $\epsilon_i \perp \epsilon_j$ for all

$i \neq j$. Thus, the information constraint can be rewritten as

$$\kappa_i \equiv \frac{1}{2} \log \left(\frac{\sigma^2}{\sigma^2_{\epsilon_i}} + 1 \right) \tag{16}$$
$$\sum_{i=1}^{N} \kappa_i \leq N\kappa.$$

κ_i can be thought of as the attention devoted to good i. Using the second order Taylor approximation from above, the optimal choice given the signal s_i is $q_i = (\eta - 1)\left(1 - e^{-2\kappa_i}\right) s_i$. Plugging this in the expectation of the quadratic loss function (14) gives the objective in terms of κ_i's.

$$\begin{aligned}
\mathbb{E}\left[L(\boldsymbol{a}, \boldsymbol{q})\right] &= \mathbb{E}\left[\frac{1}{2\eta} \sum_{i=1}^{N} \left((\eta - 1)a_i - q_i\right)^2\right] \\
&= \mathbb{E}\left[\frac{1}{2\eta} \sum_{i=1}^{N} \left((\eta - 1)a_i - (\eta - 1)\left(1 - e^{-2\kappa_i}\right)(a_i + \epsilon_i)\right)^2\right] \\
&= \frac{(\eta - 1)^2}{2\eta} \sum_{i=1}^{N} \mathbb{E}_i\left[e^{-4\kappa_i} a_i^2 + \left(1 - e^{-2\kappa_i}\right)^2 \epsilon_i^2\right] \\
&= \frac{(\eta - 1)^2}{2\eta} \sum_{i=1}^{N} \sigma^2 e^{-2\kappa_i},
\end{aligned}$$

where the third equality follows from the independence of a_i and ϵ_i. The last equality uses (16). Note that the objective is convex in every κ_i. This allows to characterize the optimal behavior by the first order conditions of the following problem.

$$\min_{\{\kappa_i\}_{i=1}^{N}} \frac{1}{2} \sum_{i=1}^{N} e^{-2\kappa_i}$$
$$\text{s.t.} \sum_{i=1}^{N} \kappa_i \leq N\kappa$$

The first order condition imply that $\kappa_i = \kappa_j$ for all $i, j \in \{1, \ldots, N\}$. This, in turn, means that $\kappa_i^* = \kappa$ for all i. This shows that a social planner with one overall information constraint behaves the same as if he had a separate information constraint for each good. This is intuitive for two reasons. The problem is symmetric. Therefore, if it is better to focus attention on more than one good, then the best is to split it equally on all goods. As the problem is convex, a corner solution is not optimal. This shows that it does not make a difference if the social planner has a separate information constraint for each good or one overall constraint. Either specification gives the same result.

4 Conclusion

This paper studied the behavior of firms which face an information constraint. Due to the information constraint firms are not able to observe the true realization of their productivity. The market structure is assumed to be monopolistic competitive to allow for individual price setting. The optimal behavior cannot be characterized by an analytic solution in general, however, the rational inattention literature provides an analytic solution for a special case of such a problem. This special case holds for quadratic loss functions. In order to apply this to the firms problem, a second order Taylor approximation of the profit function is needed.

In addition to the firms problem, a social planner's problem was considered to analyze the implications of the information constraint on the efficiency of the decentralized economy. The social planner's problem faces the same issue as in the firms problem. Thus, a quadratic approximation of the objective is needed in this case as well. The standard approach is to use an approximation in log-deviations. This was done, for example, also with the firm's problem. As it turned out, the Dixit–Stiglitz aggregator is not concave in log deviations, which does not allow to derive a quadratic loss function. For this reason a level-quadratic approximation was used.

The main findings are the following. The optimal behavior of a firm—where it does not matter if the price or quantity is chosen—is characterized by a simple function which is linear in productivity and decreasing in the information capacity. This optimal behavior has some important implications for aggregate outcomes. Most importantly, output is not maximal in a full information case. Output in the decentralized economy is maximized either if agents are not allowed to process any information or if the economy is the non-stochastic steady state. This result is driven by the symmetry of the Dixit–Stiglitz aggregator and the aggregate labor constraint. In equilibrium aggregate output needs to adjust to clear the labor market.

It was shown, that the optimal policy using a level-quadratic approximation of the firms objective leads to the same functional form as in the log-quadratic case. For this reason, the behaviors in both cases are comparable. However, this only holds for small productivity variances. Due to the normality assumption of the deviations, a high variance would put too much mass on negative realizations, which would not fit into the economic framework.

For the social planner a log-quadratic approximation is not possible, because the objective function is not concave. This is needed to get an an analytic solution. To cope with this problem, a level-quadratic approximation is used as well. The level-approximation preserves the concavity of the original objective. The optimal behavior of the social planner is very similar to the behavior in the decentralized economy. However, the reaction to deviations in productivity are less strong than in the decentralized economy. Introducing a simple tax in the decentralized economy could restore the social optimal behavior.

The social planner maximizes output, which is equivalent to maximizing the utility of a representative household in the model under consideration. By assumption, the non-stochastic steady state output yields the same for the social planner and the decentralized economy. Under full information the usual inefficiency due to monopolistic competition is obtained. This implies that intermediate cases have less distortions than a full information setup. Or in other words, distortions are increasing in the amount of information agents are allowed to process.

Furthermore, it was shown that the results are robust to the modeling choice of the information constraint in the social planner's problem. It does not matter whether the social planner has a separate information constraint for each good or one overall information constraint. With one information constraint the social planner would still allocate the same amount of attention to each good.

In the last part of this paper four numerical exercises were conducted to see if the approximations to the firm problems are valid. If the firm wants to maximize profits and log-deviations of productivity are normally distributed then the optimal behavior is close to the approximated one. For the level-approximation the approximation is even better. In the last experiment the Gaussian assumption was dropped. Instead, I assumed that productivity has a uniform distribution. In this case, the optimal behavior is characterized by an almost Gaussian distribution of prices for low information capacity cases. For intermediate cases a bimodal distribution is optimal. This, in turn, might have important implications for the aggregate variables as well. Further research in this direction could be fruitful.

An interesting avenue for future research could be to model an information friction on the consumer side as well. Including this feature might lead to interesting feedback effects. Because the consumer side has also implications for the firms and for aggregate variables, such as aggregate output and the price level, as well. Solving such a model might be challenging, but it is far from clear that the results presented here are the same (or turn around?) if one includes an information constraint on the consumer side as well.

References

Adam, K. (2007): "Optimal Monetary Policy with Imperfect Common Knowledge," *Journal of Monetary Economics*, 54, 267–301.

Barro, R. J. (1976): "Rational Expectations and the Role of Monetary Policy," *Journal of Monetary Economics*, 2, 1–32.

——— (1977): "Unanticipated Money Growth and Unemployment in the United States," *American Economic Review*, 67, 101–15.

Bickel, P. and K. Doksum (2001): *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day Series in Probability and Statistics, Upper Saddle River: Prentice Hall, 2nd revised ed.

Calvo, G. A. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12, 383–398.

Cover, T. M. and J. A. Thomas (2006): *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, New York: Wiley-Interscience.

Dixit, A. K. and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67, 297–308.

Kraskov, A., H. Stögbauer, and P. Grassberger (2004): "Estimating Mutual Information," *Physical Review E*, 69, 066138.

Lucas, R. J. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4, 103–124.

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

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

——— (2009b): "Optimal Sticky Prices under Rational Inattention," Working Paper Series 1009, European Central Bank.

——— (2011): "Business Cycle Dynamics under Rational Inattention," Working Paper Series 1331, European Central Bank.

Mankiw, N. G. and R. Reis (2002): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 117, 1295–1328.

Matejka, F. (2010): "Rationally Inattentive Seller: Sales and Discrete Pricing," CERGE-EI Working Papers wp408, The Center for Economic Research and Graduate Education - Economic Institute, Prague.

Matejka, F. and C. A. Sims (2011): "Discrete Actions in Information-Constrained Tracking Problems," CERGE-EI Working Papers wp441, The Center for Economic Research and Graduate Education - Economic Institute, Prague.

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

Nieuwerburgh, S. V. and L. Veldkamp (2010): "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77, 779–805.

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

Phelps, E. (1970): *Microeconomic Foundations of Employment and Inflation Theory*, New York: W.W. Norton.

Reis, R. (2006): "Inattentive Producers," *The Review of Economic Studies*, 73, 793–821.

Shannon, C. E. (1948): "A Mathematical Theory of Communication," *The Bell System Technical Journal*, 27, 379–423, 623–656.

Sims, C. A. (1998): "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, 49, 317–356.

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

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

Veldkamp, L. (2011): *Information Choice in Macroeconomics and Finance*, Princeton: Princeton University Press.

Woodford, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, Princeton: Princeton University Press, 25–58.