Monetary Policy, Investment and Business Cycles in the Data Economy*

Carl-Christian Groh[†] Oliver Pfäuti[‡]

First version: April 26, 2023
This version: September 2, 2023
Link to most recent version

Abstract

We propose a tractable heterogeneous firms framework to analyze how the availability of data in modern economies affects firm investment, its responsiveness to monetary policy, and its evolution along the business cycle. In our framework, data resolves a firm's uncertainty, i.e. reduces the variance of a firm's productivity and is informative about future productivity draws. Moreover, data accumulates endogenously through a data feedback loop. We establish that data-rich firms respond more strongly to monetary policy shocks and expansionary monetary policy raises market concentration. Moreover, the data feedback loop can induce firms that are more exposed to risk to hold higher levels of capital, which amplifies cyclical fluctuations driven by changes in aggregate uncertainty. By contrast, data that endows firms with knowledge about their future productivities diminishes the strength of cyclical fluctuations driven by aggregate productivity shocks.

Keywords: Data, Uncertainty, Investment, Monetary Policy, Business Cycles
JEL Codes: D21, D81, E22, E52

---

*We would like to thank Matthias Meier, Benny Moldovanu, and Volker Nocke for valuable discussions and comments. Carl-Christian Groh and Oliver Pfäuti gratefully acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224. This work was supported by the University of Mannheim's Graduate School of Economic and Social Sciences.

[†]Institute for Microeconomics, University of Bonn, carlchristian.groh@gmail.com
[‡]Department of Economics, University of Bern, pfaeuti.oliver@gmail.com

1 Introduction

Data is a key resource for businesses in the digital age and is accumulating rapidly (Brynjolfsson and McElheran, 2016b; Zolas et al., 2021; Galdon-Sanchez et al., 2022). Data generates value for firms by facilitating targeted advertising, by promoting innovation, and by aligning what consumers demand with what is supplied. The rapidly increasing prevalence of data inevitably affects the behavior of firms and the way in which they respond to changes in their environment. Understanding the role of data for the propagation of business cycles and the transmission of monetary policy is hence of first order importance, as noted by Federal Reserve Chair Jerome H. Powell (Powell, 2021).

In this paper, we set up a theoretical model to understand the key channels through which data affects firm investment as well as its responsiveness to monetary policy and cyclical fluctuations. Our model incorporates three key features of the data economy: First, data yields economic value by resolving uncertainty—superior access to data reduces the variance of a firm's productivity and enables firms to predict future realizations of their productivity.[1] Second, data accumulates endogenously through a data feedback loop—firms that sell more attain greater access to data. Third, access to data is unequally distributed: Some firms hold large advantages in terms of the data they can utilize.[2] We incorporate these features into a tractable framework which allows us to derive analytical results pertaining to the questions we pose.

We show that data amplifies the investment channel of monetary policy in several ways: Firms with access to better data adjust their capital stock more strongly in response to monetary policy shocks. Uniform increases in the availability of data increase the market shares of firms with superior access to data, which renders aggregate investment more responsive to monetary policy. However, monetary policy stimuli may have undesirable effects in the data economy: Given that data-rich firms respond comparatively strongly to changes in the interest rate, these firms become even larger such that expansionary monetary policy raises market concentration.

Whether the availability of data amplifies or dampens cyclical fluctuations within our framework is less clear. As such, the data feedback loop amplifies cyclical fluctuations driven by aggregate productivity shocks. In addition, the data feedback loop weakens the negative relationship between a firm's exposure to risk and its size. This magnifies the effects of aggregate uncertainty shocks, which are known to be a key feature of cyclical fluctuations (Bloom et al., 2018), but dampens the impact of changes in the strength of the risk-return

[1] Paine (2022), Mukerji (2022), and Wu (2023) provide empirical evidence that firms with access to superior data face lower idiosyncratic uncertainty.

[2] See Brynjolfsson and McElheran (2016a), Zolas et al. (2021), and Galdon-Sanchez et al. (2022).

relationship along the business cycle. Crucially, if access to data endows firms with exact knowledge about their future productivities, this dampens cyclical fluctuations.

Data, uncertainty, and investment. In section 4, we focus on the role of data in reducing the variance of a firm's productivity. In our model, this has economic benefits because, as in Eeckhout and Veldkamp (2022), a firm's cost of capital is increasing in the idiosyncratic uncertainty it faces. This assumption can be microfounded by explicitly modelling a risk-return relationship where the cost of capital increases with risk. We explicitly consider heterogeneity in firms' access to data, the data feedback loop, and differences in firms' exposure to risk. When there is an active data feedback loop, firms that produce more also gain better access to data. Differences in a firm's exposure to risk can be understood as heterogeneity in the slope of the risk-return relationship. Throughout the analysis, we model monetary policy shocks as exogenous changes in the firms' cost of capital.

We first establish that data-rich firms respond more strongly to monetary policy shocks, both in absolute terms and relative to their size. This is because the marginal cost of raising capital is smaller for a data-rich firm, which implies that changes in the interest rate set by the central bank have greater effects (in percentage terms) on its cost of capital, and therefore, on the firm's investment decision.[3] Conceptually, this prediction is related to the well-known result that firms facing high uncertainty are less responsive to monetary policy shocks, which has been established theoretically (Bloom et al., 2018) and empirically (Castelnuovo and Pellegrino, 2018; Pellegrino, 2021; Lakdawala and Moreland, 2022).

Thereafter, we focus on the role of heterogeneity in firms' access to data. We show that uniform increases in the availability of data raise the market share of firms with data advantages. Thus, such uniform increases in data availability raise the responsiveness of aggregate investment to monetary policy through two channels, namely by (i) increasing the responsiveness of any firm and (ii) by granting the data-rich firm higher market shares. Afterwards, we establish that heterogeneity in data access, as such, also amplifies the effects of monetary policy. Formally, we show that a mean-preserving spread in the quality of data that firms hold increases the responsiveness of aggregate investment to monetary policy shocks, both in absolute and in relative terms. This is because the responsiveness of a firm which attains better data in such a mean-preserving spread increases more strongly than the responsiveness of the other firm falls. By a similar logic, expansionary monetary policy will increase market concentration because it affects data-rich firms the most.

[3] These effects are especially pronounced at low interest rates. Thus, accounting for heterogeneity in firms' data holdings offers an explanation for the empirical finding in Kroen et al. (2021) that larger firms adjust their investment more aggressively in response to a monetary shock, especially at low interest rates.

The data feedback loop and exposure to risk. Our next set of results concerns the data feedback loop. We first show that the data feedback loop amplifies the effects of monetary policy on firm investment. This holds by the following logic: Because of the data feedback loop, firms with a higher volume of transactions acquire better data. The benefits of better data—e.g., attaining lower uncertainty —are greater, the more output a firm produces. Together, these two notions imply that the presence of a data feedback loop introduces a term that is convex in capital into the firm's optimization problem, which makes the marginal product of capital less responsive to changes in capital. In turn, this implies that the effects of monetary policy shocks on investment are amplified.

Increases in the strength of the data feedback loop weaken the negative relationship between a firm's risk exposure and its size. In the absence of the data feedback loop, firms that are more risk-sensitive have higher costs of capital and thus invest less. However, the existence of a data feedback loop means that firms can reduce their uncertainty by growing larger — the resulting incentives to attain scale weigh particularly strongly for firms that are more risk-sensitive. In fact, the relationship between a firm's risk exposure and its size becomes positive when the data feedback effect becomes large enough.

Business cycles and the data feedback loop. The presence of a data feedback loop has non-trivial implications for the propagation of cyclical fluctuations. As such, this feedback loop amplifies cyclical fluctuations driven by aggregate productivity shocks. The underlying logic is similar to the working channel in Fajgelbaum et al. (2017): Suppose that a firm anticipates a negative productivity shock and thus invests less. Because of the data feedback loop, this decision worsens the firm's access to data, which further reduces its incentives to invest and amplifies the effects of the initial shock.

The interplay of the data feedback loop and heterogeneity in firms' exposure to risk can amplify or dampen the effects of cyclical fluctuations. We established that increases in the strength of the data feedback loop weaken the negative relationship between a firm's risk exposure and its size. When there is cross-sectional heterogeneity in firm's risk exposure, the presence of a data feedback loop thus increases the market shares of firms with high levels of risk exposure. Because such firms respond relatively strongly to uncertainty, the data feedback loop amplifies the effects of aggregate uncertainty shocks, which are known to be a key feature of cyclical fluctuations (Bloom et al., 2018). The above discussions also imply that the effects of temporal changes in the strength of the risk-return relationship, which may be procylical (Alemany et al., 2023) or countercyclical (Corradi et al., 2013), will be dampened.

Data about idiosyncratic productivity. In section 5, we focus on another economic benefit of data, namely that it allows firms to predict their future productivities. We consider two different economies: In an *economy without data*, no firm has information about their future idiosyncratic productivity realizations. In an *economy with data*, all firms know the realization of their future idiosyncratic productivity when choosing their capital holdings.

We study aggregate capital and output and pin down how these aggregates are affected by monetary policy and aggregate productivity shocks in the two economies. We explicitly note that endowing firms with the aforementioned information is not equivalent to reducing the variance of their productivity: This can be seen by noting that changes in the variance of productivity do not affect firm investment or output in the economy without data.

In the economy with data, firms with above-average (below-average) productivity will invest more (less) than any firm in the economy without data. Because investment is convex in a firm's idiosyncratic productivity, aggregate capital and output are higher in the economy with data. We note that this holds true even though there is no misallocation, given that capital is supplied inelastically, and data does not favorably affect the distribution of productivity in this model. We document that the absolute effects of a monetary policy shock on aggregate investment and output are larger in the economy with data. However, the difference inbetween the strength of the effects is proportional to the difference in total output across the economies. Thus, the effects of a monetary policy shock on investment (output), relative to the initial level of investment (output), are the same in the two economies.

By contrast, the effects of aggregate productivity shocks on aggregate investment (output) are weaker in the economy with data, both in absolute terms and relative to the initial level of investment (output). Intuitively, this is because changes in aggregate productivity induce less pronounced changes in a firm's information set in the economy with data. This result suggests that the widespread availability of data in the modern economy may dampen the effects of cyclical fluctuations.

The rest of the paper is structured as follows: We discuss the related literature in section 2 and present our theoretical framework in section 3. Section 4 establishes the effects of data that changes a firm's idiosyncratic uncertainty. In section 5, we focus on the role of data in predicting firms' future productivity realizations. We conclude in section 6.

2 Related literature

To the best of our knowledge, ours is the first paper to study how data impacts the effects of monetary policy on firm investment and the propagation of cyclical fluctuations. Nevertheless, our work is related to five different strands of the literature, namely (i) the recent work studying the role of data and digitization and (ii) intangibles for macroeconomic outcomes, (iii) the contributions on the investment channel of monetary policy, (iv) the papers studying the effects of uncertainty on firm investment, and (v) the literature on rational inattention.

First, our work contributes to the rapidly growing literature on the relevance of digitization and data for macroeconomic outcomes. Veldkamp and Chung (2019) provide an overview of the role of data in the economy.[4] The firm's per-period objective function we adopt is taken from Eeckhout and Veldkamp (2022), who show that data can be a source of market power. The data feedback effect we incorporate builds on the work of Farboodi and Veldkamp (2022b), who integrate this channel into a growth model.[5] Acemoglu et al. (2022) show that data markets are not efficient in the presence of data externalities, i.e. when a user's data reveals information about others. Bergemann and Bonatti (2022) study, among others, how access to data can grant platforms market power. Our key innovation relative to these papers is that we consider the effects of monetary policy and the propagation of cyclical fluctuations in the data economy.

Within this literature, the paper that is closest to ours is Glocker and Piribauer (2021), who empirically document that increases in the amount of sales that are conducted through digital retail reduce the real effects of monetary policy. The authors rationalize this in a model with sticky prices, where prices can be adjusted more easily in online settings than in traditional markets.[6] In contrast to our work, Glocker and Piribauer (2021) do not consider the role of uncertainty or data as we define it, and there are no firm-level differences in the distribution of productivity in Glocker and Piribauer (2021).

Second, our work is related to the macroeconomic literature on research and development (R&D) and intangible assets. De Ridder (2019) and Chiavari and Goraya (2022) show that the increasing importance of intangible inputs can account for recent trends such as the rise of market power, reduced business dynamism, and lower productivity growth. Ansari (2023) documents that firms with a larger stock of customer data, which is modeled as an intangible

[4]Lashkari et al. (2018) and Brynjolfsson et al. (2023) document a positive correlation between firm size and IT intensity. Begenau et al. (2018) and Veldkamp (2023) show that firms are to a large extent valued for the data they posses and Arvai and Mann (2022) study how digitization affects consumption inequality.

[5]Wang et al. (2022), Xie and Zhang (2022), and Wu and Zhang (2022) build on the work of Farboodi and Veldkamp (2022b) and also study the role of data in growth models.

[6]Empirical evidence for lower nominal rigidities in digital markets is provided by Gorodnichenko and Talavera (2017) and Gorodnichenko et al. (2018).

asset, have lower labor shares and are differentially affected by antitrust policy.[7]

The key distinction between our paper and this literature is that data is fundamentally different from intangible assets and R&D. From a theoretical perspective, intangible capital is always modeled as a factor of production: Investment into intangible capital increases output and has a different cost structure than traditional capital. However, intangible capital does not affect the variance of productivity or grant knowledge about future productivity realizations. Empirically, measures of intangible assets encompass more than proxies for a firm's access to data and contain R&D spending, overhead, intellectual property, and software (Peters and Taylor, 2017; Döttling and Ratnovski, 2022).

Third, our paper is related to previous work that studies heterogeneity in the responsiveness of firms to monetary policy, namely with respect to a firm's size (Gertler and Gilchrist, 1994; Kroen et al., 2021), liquidity (Jeenas, 2019), default risk (Bernanke et al., 1999; Ottonello and Winberry, 2020), industry (Durante et al., 2022), price rigidities (Meier and Reinelt, 2022), and age (Cloyne et al., 2023). Heterogeneity in firms' access to data is not considered by any of these papers.

Fourth, our paper relates to the research on the role of uncertainty for firm-level investment. The seminal contribution of Bloom (2009) documents that increases of uncertainty reduce firm-level hiring and investment.[8] Bloom et al. (2018) establish that firms which face higher levels of uncertainty are less responsive to shocks such as monetary policy stimuli.[9] This insight is closely related to our result that data-rich firms respond more strongly to monetary policy shocks because they face lower idiosyncratic uncertainty. We build on this line of analysis by explicitly studying heterogeneity in firms' access to data, their exposure to risk, and by integrating a data feedback loop. Moreover, our analysis in section 5 substantially differs from previous work on the effects of uncertainty.

Our discussion of the interplay between cyclical fluctuations and the data feedback loop has parallels to Fajgelbaum et al. (2017), which builds on Veldkamp (2005) and Ordonez (2013). In these papers, there is a data feedback loop at the aggregate level, which directly amplifies business cycles. Going beyond this, we demonstrate how a firm-level data feedback loop shapes the relationship between a firm's exposure to risk and its size, and how this matters for the propagation of cyclical fluctuations.

[7] Döttling and Ratnovski (2022) and Caggese and Pérez-Orive (2022) empirically document that the investment of firms with high levels of intangible capital is less responsive to monetary policy.

[8] Using survey expectations data, Bachmann et al. (2013) show that increases in uncertainty reduce output. Kumar et al. (2022) provide causal evidence that increases in perceived uncertainty lead firms to reduce employment, investment, and sales.

[9] Empirical evidence corroborating this finding has been put forth by Castelnuovo and Pellegrino (2018), Pellegrino (2021), and Lakdawala and Moreland (2022).

Fifth, our work is related to the literature on rational inattention, which was pioneered by Sims (2003). However, there are substantial differences in focus and setup: Generally speaking, papers in this literature establish how rational inattention can account for inertia in macroeconomic outcomes. Moreover, this literature focuses on the extent to which agents choose to process different sources of information and how this matters for the propagation of shocks. These questions are not central to our analysis. In contrast to the work on rational inattention, we focus on the role of heterogeneity in firms' access to data, the data feedback loop, and the role of firms' exposure to risk. Moreover, most papers in the rational inattention literature consider models without capital.[10] In terms of setup, our analysis in section 5 is related to Charoenwong et al. (2022) and Gondhi (2023), who consider models in which firms receive signals about their idiosyncratic productivity draws.[11] However, these papers do not consider the effects of monetary policy or changes in the first moment of a firm's productivity distribution.

3 Framework

In this section, we present our theoretical model of firm investment and data. The model is kept deliberately stylized to focus on the main channels of how data affects the transmission of monetary policy and the strength of cyclical fluctuations.

Output, productivity, and data

There is a unit mass of infinitely-lived firms, indexed by $i \in [0, 1]$, and time is discrete and denoted by $t = 1, 2, ..., \infty$. Each firm produces according to its production function

$$Y_{i,t} = A_{i,t} K_{i,t}^{\alpha}, \tag{1}$$

where $Y_{i,t}$ denotes the output produced by firm i, $A_{i,t}$ its productivity, and $K_{i,t}$ its capital stock. Firms choose their capital before observing their productivity. That is, firm i chooses its capital stock $K_{i,t+1}$ in t before observing $A_{i,t+1}$. The parameter $\alpha \in (0, 1)$ is assumed to be identical across firms.[12]

In the aim of incorporating the manifold economic benefits generated by superior access

[10] Exceptions are Maćkowiak and Wiederholt (2015), Zorn (2020), Gondhi (2023), and Maćkowiak and Wiederholt (2023).

[11] Benhabib et al. (2019) study how financial markets interpret information sent by rationally inattentive firms.

[12] Our model can be easily augmented to include labor, provided it is hired on the spot market. Then, by plugging in the optimal labor choices, the parameter α can be understood as a combination of the parameters on the labor and capital inputs.

to data, we specify that a firm's productivity $A_{i,t}$ depends on (i) the aggregate productivity level $\bar{A}_t$, (ii) the firm's idiosyncratic productivity $\epsilon_{i,t}$, and (iii) how well the firm matches a payoff relevant state, which we call $\theta_{i,t}$. Based on Farboodi and Veldkamp (2022b), we formalize the third feature as follows: The payoff-relevant state $\theta_{i,t}$ can be understood as the optimal product variety or the ideal form of marketing in a given period, which the firm wishes to mirror by its choice of marketing/production approach ($a_{i,t}$). Reflecting these three features, a firm's productivity takes the following form:

$$A_{i,t} = \bar{A}_t - \bar{\mu} d(a_{i,t}, \theta_{i,t}) + \epsilon_{i,t}, \tag{2}$$

where $d(a_{i,t}, \theta_{i,t})$ is some distance metric and $\theta_{i,t}$ and $\epsilon_{i,t}$ are random variables that are independently drawn according to the distributions F_θ and F_ϵ. The parameter $\bar{\mu} \geq 0$ governs how important it is for firms to match the payoff-relevant state $\theta_{i,t}$.

In the absence of data, firms have no information about $\theta_{i,t}$ and $\epsilon_{i,t}$ when choosing their capital stock $K_{i,t}$. By contrast, firms with access to data receive signals about these random variables before choosing their capital stock. We assume that these signals are unbiased, and that a firm's data quality is measured by two objects, namely $\sigma_{\epsilon,i,t}$ and $\sigma_{\theta,i,t}$, where $\sigma^2_{\epsilon,i,t}$ is the variance of the signal a firm obtains about $\epsilon_{i,t}$ and $\sigma^2_{\theta,i,t}$ is the variance of the signal a firm obtains about $\theta_{i,t}$. We say that a firm has access to better data about one of these random variables if and only if this firm's signal about this random variable has lower variance.

Access to data about $\epsilon_{i,t}$ and $\theta_{i,t}$ serves different purposes. A more precise signal about $\epsilon_{i,t}$ reduces the variance of a firm's productivity and enables a firm to pinpoint the location of its productivity distribution with greater precision. However, having access to a more precise signal about $\epsilon_{i,t}$ does not increase a firm's expected productivity as such. By contrast, a firm with better data about $\theta_{i,t}$ can match this state more effectively, which increases the firm's expected productivity and reduces the variance of productivity.

To ease exposition, we study the economic effects of access to data about $\theta_{i,t}$ and $\epsilon_{i,t}$ separately. In Section 4, we consider economies in which firms receive signals about $\theta_{i,t}$, but not about $\epsilon_{i,t}$. In section 5, we study settings in which firms receive signals about $\epsilon_{i,t}$, but not about $\theta_{i,t}$.

The data feedback loop

A key feature of the way in which data accumulates is the data feedback loop as discussed by Farboodi and Veldkamp (2022a). The presence of the data feedback loop is based on the idea that data is a byproduct of production and transactions: A firm that produces more learns more about its customers' preferences, about the optimal inventory, etc. The bigger a firm, the more data it accumulates. Formally, we incorporate the data feedback loop by linking the signal quality $\sigma_{\theta,i,t}$ of a firm to its capital stock, $\sigma_{\theta,i,t} = \tilde{\sigma}(K_{i,t})$ with

$$\frac{\partial \tilde{\sigma}\left(K_{i,t}\right)}{\partial K_{i,t}} < 0. \tag{3}$$

In words, bigger firms, i.e., firms with a higher capital stock $K_{i,t}$, have access to more or better data which reduces the variance of the signal they receive about the payoff relevant state $\theta_{i,t}$.

Having access to better data therefore helps the firm to more precisely predict the relevant state $\theta_{i,t}$ and then choose $a_{i,t}$ accordingly. This increases the firm's expected productivity and reduces its uncertainty, which in turn, incentivizes the firm to grow even bigger and accumulate even more data. Figure 1 illustrates this data feedback loop graphically (see Farboodi and Veldkamp (2022a) for a similar graph).

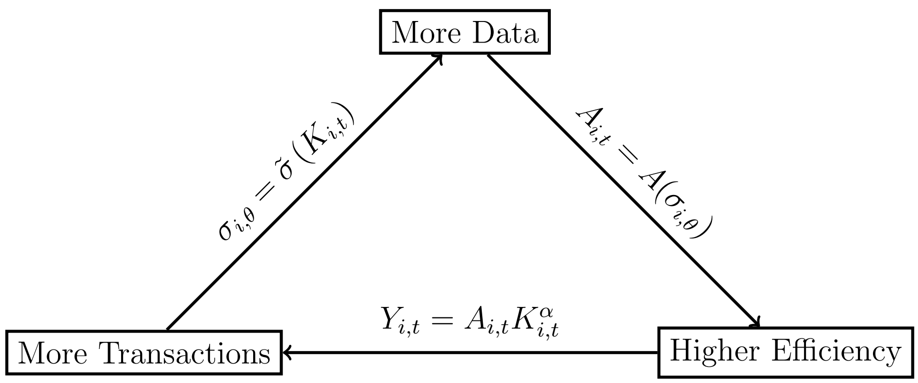

Figure 1: Data Feedback Loop

We only consider a data feedback loop that affects the information a firm obtains about the ideal variety ($\theta_{i,t}$). The ramifications of a data feedback loop pertaining to $\epsilon_{i,t}$ are likely analogous.

A firm's optimization problem

A firm's objective function in any period t is given by

$$\max_{\{K_{i,t+1+j}\}_{j=0}^{\infty}} \quad \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j \Pi_{i,t+j}, \tag{4}$$

where $\beta \in (0, 1]$ is the discount factor and the flow profits $\Pi_{i,t+j}$ are, as in Eeckhout and Veldkamp (2022), given by:

$$\Pi_{i,t+1} = \mathbb{E}_t \left[Y_{i,t+1} \right] - \frac{\rho}{2} VAR_t \left[Y_{i,t+1} \right] - r_t I_{i,t}. \tag{5}$$

We define r_t as the interest rate and assume that it is directly controlled by the monetary policy authority. A firm's investment in period t is $I_{i,t} = K_{i,t+1} - (1-\delta)K_{i,t}$, where δ is the rate of capital depreciation. Capital is inelastically supplied.

The firm's objective (5) allows for the possibility that a firm is negatively affected by the uncertainty it experiences, which holds true whenever $\rho > 0$. A natural microfoundation for this is the positive relationship between the riskiness and the return of an asset: Firms who face higher idiosyncratic uncertainty will have a higher cost of raising capital. In fact, when $\alpha = 1/2$, our setup is isomorphic to a setting in which a firm's profits are given by the difference between demand and total costs, where the total cost of capital is the sum of a risk-free rate $r_t^f I_t$ and a risk premium paid on the chosen capital stock, given by $(\rho/2)VAR[A_{i,t+1}]K_{i,t+1}$.

Discussion of key model features

Our objective in this paper is to present a model of firm investment to help us understand the mechanisms how the availability of data affects the transmission of monetary policy and the propagation of business cycles. In this endeavour, it is imperative to explicitly model the different ways in which data can yield economic value, which we have focused on. To ensure tractability, we have imposed several simplifying assumptions which we now discuss here in more detail.

We model monetary policy shocks as unexpected changes in the firms' cost of capital and abstract from how exactly changes in the short-term nominal interest rate are transmitted to the cost of capital firms face when investing. We do not model these channels explicitly as they are not the focal point of our analysis: We consider the investment channel of monetary policy and focus on how a firm's responsiveness to a change in it's cost of capital (which is induced by a monetary policy shock) is affected by the firm's access to data. Nevertheless, it is very important to analyse how access to data in capital markets affects the transmission of nominal shocks to the cost of capital and interest rate spreads. We leave the analysis of these issues to be addressed by future research.

Further, our model is set in partial equilibrium to shed the spotlight on the role of data in firms' investment decisions. Given that we focus on the investment channel of monetary policy, this specification can be motivated using recent evidence that monetary policy affects investment mainly through direct channels (Cao et al., 2023). More generally speaking, the effects we find would still be active in general equilibrium — thus, our analysis can be viewed as an initial appraisal of larger questions at hand.

4 Data and the resolution of uncertainty

In this section, we focus on how data about θ_{it} changes a firms' investment decision by increasing the expected productivity and reducing its variance. Our analysis establishes that the reduced variance granted by access to this type of data is particularly consequential. To shed the spotlight on this form of data, we assume that firms receive no signals about their idiosyncratic productivities, i.e. that $\sigma_{\epsilon,i,t} = \infty$. Throughout this section, we also impose that data enters the moments of productivity linearly.

Assumption 1 *We impose the following functional form assumptions on the relationship between data and productivity:*

$$\mathbb{E}[A_{i,t}] = \tilde{A}_t - \kappa_e \sigma_{\theta,i,t} \quad ; \quad VAR[A_{i,t}] = \tilde{V}_t + \kappa_v \sigma_{\theta,i,t} \tag{6}$$

$$\sigma_{\theta,i,t} = \underbrace{\bar{\sigma}_i - zK_{i,t}}_{:=\tilde{\sigma}(K_{i,t})}. \tag{7}$$

The parameters κ_e and κ_v capture the effects of data $\sigma_{\theta,i,t}$ on the expected productivity and its variance, respectively. When κ_v is high, relative to κ_e, this means that the primary economic value of access to better data is the associated reduction in the variance of the firm's outcomes. In order for data to be meaningful in our framework, we assume that at least one of the parameters κ_e and κ_v is always strictly positive. The parameter $z \geq 0$ captures the strength of the data feedback loop.

These linearity assumptions imply that a marginal increase in the quality of data has the same absolute impact on the decision situation of any firm. Thus, these assumptions play an important role in our analysis of firm heterogeneity, namely by enabling us to conduct meaningful comparisons between the effects of such marginal increases in data availability across firms.

Furthermore, we impose the following assumptions throughout the analysis in this section:

Assumption 2 *We assume that:*

- *The flow profit function is strictly concave in capital, i.e.* $\frac{\partial^2 \Pi_{i,t}}{\partial K_{i,t}^2} < 0$.

- *At the optimally chosen levels of $K_{i,t}$, both $E_t[A_{i,t}; \sigma_{i,t}]$ and $VAR[A_{i,t}; \sigma_{i,t}]$ remain strictly positive.*

The first assumption is necessary to ensure that a unique optimal capital choice exists for every firm. It also suffices to ensure that the basic comparative statics result with respect to the effect of an increase in r_t is such that that a firm's chosen level of capital is falling in the interest rate.

We now study the impact of monetary policy and uncertainty shocks on firm investment. A monetary policy shock is modelled as a one-time change in the interest rate r_t. An aggregate uncertainty shock is understood as a one-time increase in the variance of productivity ($\tilde{V}_t$). For notational convenience, we write $\sigma_{\theta,i,t}$ as $\sigma_{i,t}$ throughout this section and drop the subscript i when unambiguous.

4.1 Data and the optimal investment choices

We start by treating the level of a firm's data as exogenous, i.e. we assume that σ_{t+1} is fixed for every firm and that there is no data feedback effect (i.e., $z = 0$). Consider the optimization problem of a firm in period t, which chooses the level of capital K_{t+1}. The variable K_{t+1} only enters the flow profits of this firm in periods $t+1$ and $t+2$. This is because the optimal level of K_{t+2} (and any optimal level of capital chosen for periods thereafter) will not depend on the chosen level of K_{t+1}. Thus, the first-order condition for K_{t+1} reads as follows:

$$\underbrace{\alpha \mathbb{E}_t[A_{t+1}; \sigma_{t+1}] K_{t+1}^{\alpha-1}}_{\text{marg. product of capital}} = \underbrace{\rho \alpha VAR_t[A_{t+1}; \sigma_{t+1}] K_{t+1}^{2\alpha-1} + r}_{\text{marg. cost of capital}}, \tag{8}$$

where we defined $r \equiv r_t - \beta(1-\delta)\mathbb{E}_t[r_{t+1}]$. Assuming that $\alpha = 1/2$ holds, one can solve for the optimal capital stock in closed form.

$$K^*_{t+1}(\sigma_{t+1}) = \left(\frac{\mathbb{E}[A_{t+1}; \sigma_{t+1}]}{\rho VAR[A_{t+1}; \sigma_{t+1}] + 2r} \right)^2. \tag{9}$$

Examining this solution reveals that firms with access to superior data respond more strongly to monetary policy shocks (a change in r). This is formalized in the following proposition:

Proposition 1 *Suppose $\alpha = 1/2$ and $z = 0$. Firms with access to better data respond more strongly to monetary policy, both in absolute and in relative terms:*

$$\frac{\partial K^*_{t+1}}{\partial r_t} < 0 \quad ; \quad \frac{\partial^2 K^*_{t+1}}{\partial r_t \partial \sigma_{t+1}} > 0 \quad ; \quad \frac{\partial\left(\frac{\partial K^*_{t+1}/\partial r_t}{K^*_{t+1}}\right)}{\partial \sigma_{t+1}} \geq 0. \tag{10}$$

The latter inequality is strict if and only if $\kappa_v > 0$.

To understand the results regarding the absolute effects, examine the optimal capital choice of the firm, which is pinned down in equation (9): Firms with access to better data will hold higher levels of capital, because they have higher expected productivity and face lower uncertainty. The former effect raises a firm's marginal product of capital, and the latter reduces the cost of capital. Thus, both effects push up the firm's optimal capital choice.

The fact that firms with better data hold more capital implies that, in absolute terms, these firms respond more strongly to monetary policy shocks. In response to a rise of r_t, a firm responds by adjusting capital in a way that re-equates the marginal product of capital and its marginal cost. Because the flow profit function is concave in capital by assumption, the marginal product of capital is more responsive to changes in the chosen capital stock at low levels of capital. Thus, a monetary policy shock induces smaller changes in the amount of capital selected by firms with limited access to data, because their capital levels were already low ex ante.

Now let's examine the relative effect of a monetary policy shock on a firm's capital choice, which is given by:

$$\frac{\partial K^*_{t+1}/\partial r_t}{K^*_{t+1}} = \frac{-4}{[VAR[A_{t+1}]\rho + 2r]} \tag{11}$$

In relative terms, the effects of a monetary policy shock depend on a firm's access to data if and only if access to better data reduces the variance of productivity, namely $VAR[A_{t+1}]$. If access to better data does not reduce a firm's uncertainty, i.e. $\kappa_v = 0$ holds, having access to better data merely shifts up the marginal product of capital via the higher level of $\mathbb{E}[A_{t+1}]$. Because data-rich firms are larger, the effects of a monetary policy shock on firms, relative to their size, thus stay unaffected by the access to data.

However, if better data reduces the uncertainty a firm faces (i.e. $\kappa_v > 0$) firms with access to better data will also respond more strongly to monetary policy shocks in relative terms. This is because the variance of future productivity enters the cost of capital. When this variance is small (i.e. when a firm has access to high-quality data), changes in the interest rate r_t imply large (in percentage terms) changes in a firm's cost of capital. As a result, changes in the interest rate affect data-rich firms to a greater extent.

Before moving forward, we highlight that our model matches several empirically documented facts on the relationship between uncertainty, monetary policy, and investment. Firms that face higher uncertainty invest less in our model (consistent with Bachmann et al. (2013); Kumar et al. (2022)) and firms with access to better data grow larger (consistent with Brynjolfsson et al. (2023)). Slightly rephrased, proposition 1 states that firms which face higher uncertainty respond less to monetary policy shocks. This is consistent with the insights of Bloom et al. (2018) and the empirical findings of Castelnuovo and Pellegrino (2018), Pellegrino (2021), and Lakdawala and Moreland (2022).

We further find that larger firms (i.e., firms with access to better data and lower idiosyncratic uncertainty) respond more strongly to monetary policy shocks, especially when interest rates are low, consistent with the empirical findings in Kroen et al. (2021). This holds because the derivative of $\frac{\partial K^*_{t+1}/\partial r_t}{K^*_{t+1}}$ with respect to the variance term $VAR[A_{t+1}]$ is positive, but moves closer towards zero as the interest rate increases.

Numerical insights. In Appendix B, we show that our analytical insights from Proposition 1 extend beyond the case of $\alpha = 0.5$. Firms with more data respond more strongly to monetary policy shocks, both in absolute and in relative terms. We further find that this relationship is concave in σ_{t+1}, especially for higher values of κ_v. This follows because with a higher κ_v, more data reduces the firm's uncertainty more, which induces the firm to respond more strongly to the interest rate change.

4.2 Heterogeneity in firms' data availability

Firms are vastly heterogeneous in their access to data and some firms have attained significant advantages in terms of the data they can utilize (Brynjolfsson and McElheran, 2016a; Zolas et al., 2021; Galdon-Sanchez et al., 2022). We therefore now turn to the role of heterogeneity in firms' access to data for the overall effectiveness of monetary policy and the role of the interest rate environment. We begin by characterizing the effects of a uniform increase in the firms' ability to utilize data, and how this relationship is shaped by the interest rate.

Proposition 2 *Suppose that $\alpha = 1/2$, $z = 0$, $\kappa_e = 0$, and consider an economy with two firms $i \in \{l, h\}$, whose data quality levels are given by $\sigma_l = \bar{\sigma}_l - b$ and $\sigma_h = \bar{\sigma}_h - b$, with $\bar{\sigma}_l < \bar{\sigma}_h$. An increase of b raises the market share of the firm with access to superior data.*

This result establishes that uniform increases in the availability of data in the economy (i.e. increases of the parameter b) can be causally responsible for the upward trends in market concentration that have been witnessed over the last decade. The intuition underlying the proposition is as follows: A uniform increase in the availability of data reduces the marginal costs of all firms. Because the marginal costs of a firm with access to high-quality data is low ex ante, the marginal costs of this firm are affected strongly (in percentage terms). Thus, an increase of b has particularly strong positive effects on the output of a firm with access to superior data, thus granting this firm a higher market share.

To see how this relationship is shaped by the prevailing interest rate, consider an example of the aforementioned two-firm environment in which $\bar{\sigma}_h = 2$, $\bar{\sigma}_l = 1$, $\tilde{A} = 1$, $\tilde{V} = 1$, $\kappa_v = 0.5$, $\rho = 0.25$, and $\alpha = 0.5$. In the following graph, we plot the market shares of the firm with access to superior data (on the y-axis) for different levels of b (plotted on the x-axis) at three different interest rates, namely $r \in \{0.01, 0.05, 0.2\}$.

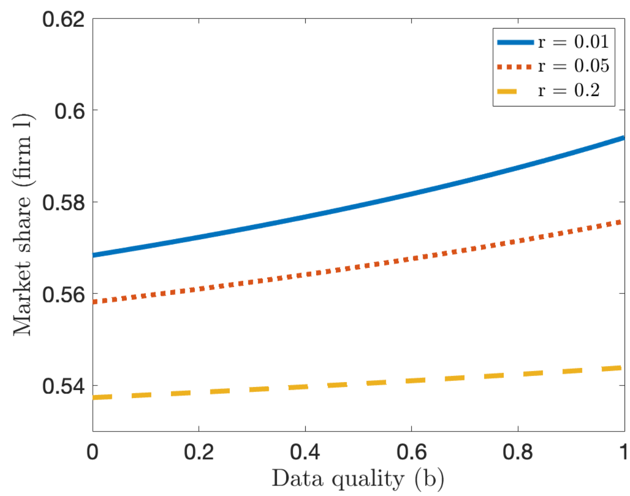

Figure 2: Data quality and market concentration

This numerical analysis reveals that the effects of uniform increases in the availability of data on market concentration are particularly pronounced in low-interest rate environments such as those we have experienced in the last decades before the Covid-19 pandemic. This is because as interest rates rise, the relative differences in the marginal costs between the two firms (which are driven by differences in the availability of data) become smaller. Thus, the effects of any shock on the choices of these firms converge when the interest rate increases.

The results of proposition 2 have an important corollary: A uniform increase in the availability of data amplifies the responsiveness of aggregate investment to a monetary policy shock in two ways, namely by (i) increasing the responsiveness of each firm and (ii) by raising the market share of more responsive firms. We formalize this in the following corollary:

Corollary 1 *Suppose that $\alpha = 1/2$, $z = 0$, $\kappa_e = 0$, and consider an economy with two firms $i \in \{l, h\}$, whose data quality levels are given by $\sigma_l = \bar{\sigma}_l - b$ and $\sigma_h = \bar{\sigma}_h - b$, with $\bar{\sigma}_l < \bar{\sigma}_h$.*

A rise of b increases the relative effect of monetary policy on aggregate investment, i.e.:

$$\frac{\partial\left(\frac{\frac{\partial(K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h))}{\partial r_t}}{K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h)}\right)}{\partial b} < 0. \tag{12}$$

Defining $C_i(\sigma_l, \sigma_h) = \frac{K^*_{t+1}(\sigma_i)}{\sum_j K^*_{t+1}(\sigma_j)}$ and $\gamma(\sigma_i) = \frac{\partial K^*_{t+1}(\sigma_i)/\partial r}{K^*_{t+1}(\sigma_i)}$, the effect of a uniform increase in the availability of data on the effectiveness of monetary policy can be decomposed as follows:

$$\frac{\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)}}{\partial b} = \underbrace{C_l(\sigma_l, \sigma_h)\frac{\partial\gamma(\sigma_l)}{\partial b} + C_h(\sigma_l, \sigma_h)\frac{\partial\gamma(\sigma_h)}{\partial b}}_{\text{direct effects}} + \underbrace{\frac{\partial C_l}{\partial b}\big(\gamma(\sigma_l) - \gamma(\sigma_h)\big)}_{\text{composition effect}} \tag{13}$$

All terms are negative. The increase in the direct effects captures that the uniform increase in the availability of data raises the responsiveness of individual firms to a monetary policy shock. The composition effect that the increase of b raises the relative size of the firm with access to superior data, which increases the overall impact of the shock because this firm is more responsive to monetary policy than its rival.

The previous discussion foreshadows that expansionary monetary policy will raise the market shares of firms with access to superior data. We formalize this result in the following proposition:

Proposition 3 *Suppose that $\alpha = 1/2$, $z = 0$, $\kappa_e = 0$, and consider an economy with two firms $i \in \{l, h\}$, whose data quality levels are given by $\sigma_l < \sigma_h$. The market share of the firm with access to superior data increases as the interest rate falls.*

Whenever access to data is unequally distributed, expansionary monetary policy increases market concentration as measured by the Herfindahl-Hirschman index, ceteris paribus. To understand this, note that firm l is a firm with superior access to data, which generally has a higher market share than its rival. An expansionary monetary policy shock will further increase this firm's market share. This is because changes in the interest rate have comparatively greater effects on the output levels of firms with superior access to data, given that their cost of capital is affected more strongly (in percentage terms) by changes in the interest rate.

Finally, we characterize how the effects of a monetary policy shock are shaped by the firms' unequal access to data itself:

Proposition 4 *Suppose that $\alpha = 1/2$, $z = 0$, $\kappa_e = 0$, and consider an economy with two firms $i \in \{l, h\}$ with data quality levels given by $\sigma_l = \sigma_m - y$ and $\sigma_h = \sigma_m + y$, with*

$y > 0$. *Mean-preserving spreads of* (σ_l, σ_h) *amplify the effects of monetary policy on aggregate investment (in absolute and relative terms), i.e.*

$$\frac{\partial\left(\frac{\partial(K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h))}{\partial r_t}\right)}{\partial y} < 0 \quad ; \quad \frac{\partial\left(\frac{\frac{\partial(K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h))}{\partial r_t}}{K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h)}\right)}{\partial y} < 0. \tag{14}$$

The responsiveness of the firm that attains better data in the considered mean-preserving spread increases more strongly than the responsiveness of the other firm falls. Given that the data disparities will likely grow further over time, the result in the above proposition indicates that the effects of monetary policy on firm investment are likely to become stronger.

4.3 The data feedback loop

In this subsection, we focus on the role of the data feedback loop, which we abstracted from up to now. Formally, we consider arbitrary levels of $z > 0$ in this subsection (under the constraint that assumption 1 is still satisfied). The key insight of this section is that the presence of the data feedback loop amplifies the effects of monetary policy as such. This notion is formalized in the following proposition.

Proposition 5 *Suppose that* $\alpha = 1/2$ *and that* $\kappa_e = 0$. *The data feedback loop amplifies the effect of a monetary policy shock, i.e.:*

$$\frac{\partial^2 K^*_{t+1}}{\partial r_t \partial z} < 0 \quad ; \quad \frac{\partial\left(\frac{\partial K^*_{t+1}/\partial r_t}{K^*_{t+1}}\right)}{\partial z} < 0. \tag{15}$$

To understand the result, note that a firm's optimal capital choice must solve the following first-order condition:

$$\left(\bar{A} - \kappa_e \sigma_{t+1}\right) \alpha K_{t+1}^{\alpha-1} - \rho\left(\bar{V} + \kappa_v \sigma_{t+1}\right) \alpha K_{t+1}^{2\alpha-1} + \underbrace{\kappa_e z K_{t+1}^{\alpha} + \frac{\rho}{2}(\kappa_v z) K_{t+1}^{2\alpha}}_{\text{data feedback loop}} = r \tag{16}$$

This first order condition shows that the data feedback effect makes the flow profits of the firm more convex in capital. This is based on the following logic: A firm can increase its data access by raising its capital stock through the data feedback loop. The benefits of attaining access to better data are greater, the more output the firm produces. Taken together, these two considerations imply that the data feedback loop thus introduces convex terms into any firm's objective function.

Thus, increases in z (which reflect a stronger data feedback loop) make the marginal product of capital less responsive to changes in capital. This, in turn, directly implies that the effects of a monetary policy shock are amplified when z increases (i.e., when the data feedback loop becomes stronger).

Numerical insights. In Appendix B, we show that our main insights from Proposition 5 continue to hold for $\alpha \neq 0.5$. The data feedback loop especially tends to amplify the investment response to monetary policy shocks when κ_v or κ_e are high, that is when more data reduces the firm's uncertainty more effectively or increases its productivity more strongly.

4.4 Firms' risk exposure & the data feedback effect

In this section, we consider how the data feedback effect affects the relationship between firms' exposure to risk (represented by the parameter ρ) and their capital holdings. In a nutshell, we show that the presence of the data feedback effect may induce firms who are more sensitive to risk (higher ρ) to hold higher levels of capital. This is particularly relevant when considering the impact of cyclical fluctuations, which coincide with movements in aggregate uncertainty and the slope of the risk-return relationship, which directly relates to the parameter ρ. We establish the following key result:

Proposition 6 *Suppose $\alpha = 1/2$. If $z = 0$, then $\frac{\partial K^*_{t+1}}{\partial \rho} < 0$. By contrast, $\frac{\partial K^*_{t+1}}{\partial \rho} > 0$ holds if z is high enough, i.e., when:*

$$\bar{V} + \kappa_v \bar{\sigma} - 2\kappa_v z K^*_{t+1} < 0. \tag{17}$$

The strength of the data feedback effect z thus determines whether more risk sensitive firms (higher ρ) hold more or less capital. When $z = 0$, firms that are more exposed to risk (i.e., have a higher ρ) hold less capital, because increases in ρ go along with higher costs of capital. When $z > 0$, there is an opposing effect: Attaining scale by increasing K_{t+1} allows a firm to reduce the idiosyncratic risk it faces through the data feedback loop. The economic benefits of this channel are particularly high for firms that are very sensitive to risk. If this channel becomes strong enough, which happens if z becomes large, then the sign of the relationship between a firm's level of risk aversion (ρ) and it's capital level may flip.

Figure 3 illustrates this graphically.[13] We plot a firm's optimal capital stock (vertical axis) as a function of its exposure to risk, ρ (horizontal axis). We do this for different strengths of the data feedback effect, z. We see that when z is relatively low (the two left panels), an increase in ρ decreases a firm's optimal capital stock. However, this negative

[13] We set $\delta = 0.1$, $\alpha = 0.3$, $\bar{\sigma} = 0.5$, $\kappa_v = 0.05$, $\kappa_e = 0.005$, $\tilde{A} = 2$, and $r = 0.1$.

relationship becomes weaker as the strength of the data feedback effect increases. Moreover, if z is large enough, firms that with higher levels of ρ hold higher levels of capital because size reduces uncertainty through the data feedback loop:

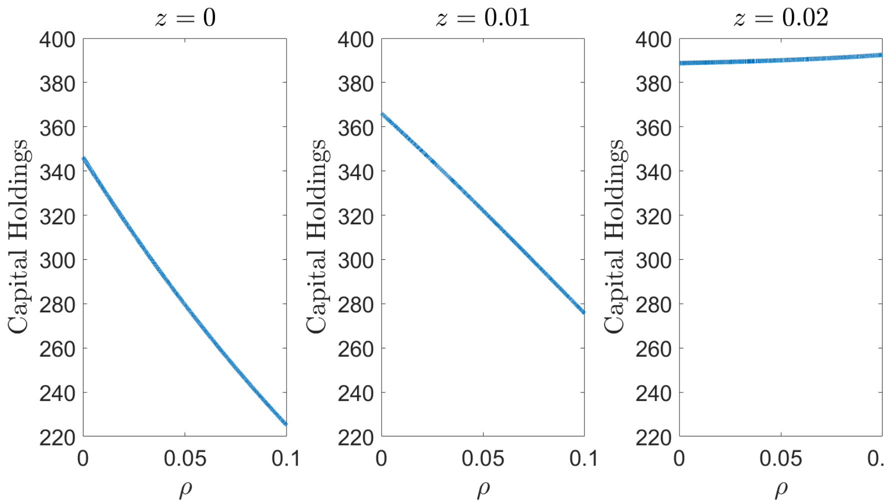

Figure 3: Risk exposure & optimal capital levels

Figure 4 shows how this channel exacerbates the effects of uncertainty shocks. In particular, we plot the responsiveness of a firm's capital choice (on the vertical axis) to an increase in $\tilde{V}$ (from 1 to 1.1) as a function of ρ (on the horizontal axis) for the different levels of z we have considered. In all cases, the uncertainty shock induces a stronger investment response when firms are more exposed to risk. These effects are amplified through the data feedback effect. In fact, we see how the relationship between the firm's responsiveness to uncertainty and its risk aversion ρ becomes convex at higher levels of z (see right panel).

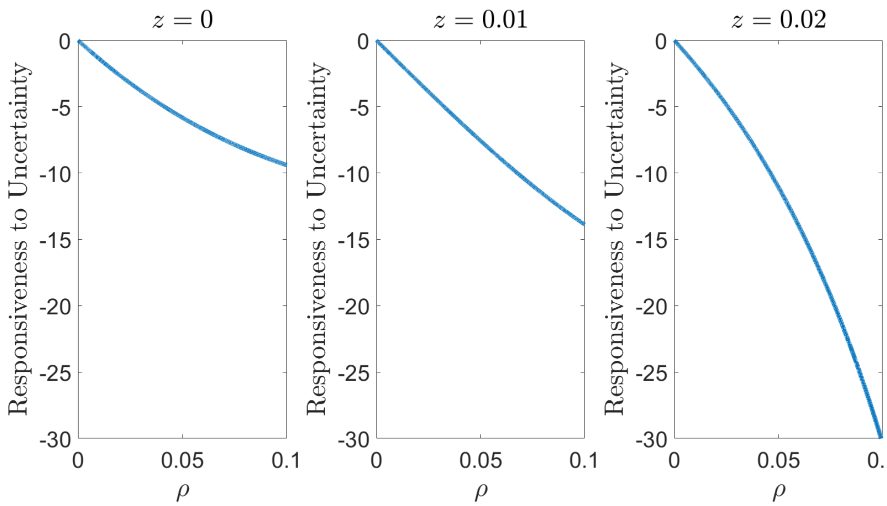

Figure 4: Risk exposure & responsiveness to an uncertainty shock

A simple back-of-the-envelope quantification of these effects shows that they may be quite substantial. Consider an aggregate uncertainty shock (during which $\tilde{V}$ increases from 1 to 1.1) in partial equilibrium. Assume there is a unit mass of firms and the risk preference of firms in the economy is distributed according to $\rho \sim U[0, 0.1]$. The absolute effects of the uncertainty shock on aggregate investment are -5.4 when $z = 0$ and -12.4 when $z = 0.02$, respectively. The relative effects of an uncertainty shock (i.e. the change in aggregate investment, divided by the ex-ante value of aggregate investment) are -1.9% when $z = 0$ and -3.2% when $z = 0.02$, respectively. The presence of the data feedback effect may thus substantially amplify the effects of aggregate uncertainty shocks.

4.5 Business cycles and the data feedback effect

The data feedback loop may alter the magnitude of business cycles through various channels. First, the presence of the data feedback effect will directly amplify business cycles by a logic analogous to the one discussed in Fajgelbaum et al. (2017): Suppose the economy is hit by a negative productivity shock, which directly induces firms to produce less. Via the data feedback effect, this decrease in production reduces the data quality of firms, which further lowers their output, thereby exacerbating the effects of such a productivity shock.

Second, the data feedback loop also affects the propagation of cyclical fluctuations by shaping the relationship between a firm's exposure to risk (ρ) and its capital holdings. Whether or not this channel amplifies or mitigates cyclical fluctuations depends on the nature of the variation in ρ, i.e. whether this varies across firms/industries and whether it is pro- or countercylical. To understand the role of firms' exposure to risk in shaping cyclical fluctuations, note that (i) uncertainty is strongly countercyclical (Bloom, 2009; Bloom et al., 2018) and (ii) there is evidence that the slope of the risk-return relationship varies along the business cycle (Corradi et al., 2013; Alemany et al., 2023). In our framework, both these features shape the propagation of cyclical fluctuations.

Suppose ρ is heterogeneous across firms/industries but constant over time. By the results of proposition 6, the presence of the data feedback effect will increase the overall exposure of firms to risk, because firms with high exposure to risk have stronger incentives to grow larger. This composition effect implies that the response of aggregate investment to movements in aggregate uncertainty will be amplified. Given that countercyclical movements of aggregate uncertainty are a key feature of business cycles (Bloom et al., 2018), this channel will amplify the effects of business cycles.

Suppose instead that ρ does not vary across firms/industries, but only over time. The empirical evidence on the evolution of this parameter along the business cycle is inconclusive: There is evidence that the strength of the risk-return relationship is countercylical (Corradi et al., 2013), i.e. that ρ is high in recessions and small in booms, but also evidence for the contrary (Alemany et al., 2023). To fix ideas, suppose that ρ is countercyclical, which means that these movements exacerbate cyclical fluctuations. When there is no data feedback effect, the increased risk premium in recessions reduces firm investment. However, when the data feedback effect is strong, this relationship gets weaker. In particular, the relationship between ρ and optimal capital holdings can become positive when the data feedback effect is sufficiently strong, as documented by the numerical example in the previous section. Via this channel, a strong data feedback effect thus mitigates the effects of countercyclical changes in ρ along the business cycle.

5 Data and idiosyncratic productivity

5.1 Setup and firm optimization

Throughout this section, we focus on the role of data in helping firms predict the location of their productivity distribution. For tractability, we fully abstract from the the role of data in predicting optimal varieties and assume that $\bar{\mu} = 0$, i.e. that such considerations are irrelevant for productivity. Moreover, we also assume that firms do not price risk in this section, i.e. that $\rho = 0$, that there are only two periods $t \in \{1, 2\}$, and that $\epsilon_{i,2} \sim U[\epsilon^{min}, \epsilon^{max}]$.

Given these assumptions, the economic environment is as follows: In period 1, there is no production, and firms choose their capital stock to produce in period 2. Firm profits in period 2 are given by:

$$\Pi_{i,2} = A_{i,2}(K_{i,2})^{\alpha} - rK_{i,2}, \tag{18}$$

where the cost of capital is r and a firm's productivity $A_{i,2}$ is given by:

$$A_{i,2} = \bar{A}_2 + \epsilon_{i,2}. \tag{19}$$

To understand the role of data, we consider two stylized economies that share the above characteristics but differ in the degree to which firms have access to data. In one economy, which we call *the economy with data*, all firms have access to a perfect signal about their draw of $\epsilon_{i,2}$, i.e. $\sigma_{i,\epsilon} = 0$ holds true for all firms. In the other economy, which we call the *economy without data*, all firms receive a completely uninformative signal about $\epsilon_{i,2}$, i.e. $\sigma_{i,\epsilon} = \infty$ holds true for all firms. Crucially, access to better data does not affect the aggregate distribution of productivity, but only informs individual firms about their productivity draws. We suppose that capital supply is perfectly inelastic, so aggregate capital is equal to aggregate capital demand. We compare aggregate capital and output, their responses to a monetary policy shock, and the magnitude of business cycles across the two economies.

Consider first the economy with data, in which the optimization problem of a firm is standard. In period 1, the firm knows its realization of $A_{i,2} = \bar{A} + \epsilon_{i,2}$ and will optimally choose its future capital stock based on this information. We define $K_2^d(A_{i,2})$ as the optimal capital stock of a firm in the economy with data, which is a function of the firm's productivity $A_{i,2}$, and is given by

$$K_2^d(A_2) = \arg\max_{K_{i,2}} \left[A_{i,2}(K_{i,2})^{\alpha} - rK_{i,2}\right].$$

Now consider any firm in the economy without data. Given that no firm in this economy knows its future productivity, they all face the same optimization problem. We define K_2^{nd} as the optimally chosen capital stock of any firm in the economy without data. This optimal capital stock solves:

$$K_2^{nd} = \arg\max_{K_2} \left[\int_0^1 A_{i,2}(K_2)^{\alpha} dA_{i,2} - rK_2\right]$$

Solving the aforementioned optimization problems allows us to pin down the aggregate capital stocks and output in the two economies. We define $\bar{K}_2^{nd}$ and $\bar{K}_2^d$ as aggregate (expected) capital in the economy without and with data, respectively. Further, we define $\bar{Y}_2^{nd}$ and $\bar{Y}_2^d$ as aggregate (expected) output in the economy without and with data, respectively.

Lemma 1 *The levels of aggregate capital in the economies with and without data are:*

$$\bar{K}_2^{nd} = (\alpha)^{\frac{1}{1-\alpha}} (r)^{\frac{1}{\alpha-1}} \left(\mathbb{E}[A_{i,2}]\right)^{\frac{1}{1-\alpha}} \quad ; \quad \bar{K}_2^{d} = (\alpha)^{\frac{1}{1-\alpha}} (r)^{\frac{1}{\alpha-1}} \mathbb{E}\left[\left(A_{i,2}\right)^{\frac{1}{1-\alpha}}\right]. \tag{20}$$

The levels of aggregate output in the economies with and without data are:

$$\bar{Y}_2^{nd} = (\alpha)^{\frac{\alpha}{1-\alpha}} (r)^{\frac{\alpha}{\alpha-1}} \left(\mathbb{E}[A_{i,2}]\right)^{\frac{1}{1-\alpha}} \quad ; \quad \bar{Y}_2^{d} = (\alpha)^{\frac{\alpha}{1-\alpha}} (r)^{\frac{\alpha}{\alpha-1}} \mathbb{E}\left[\left(A_{i,2}\right)^{\frac{1}{1-\alpha}}\right]. \tag{21}$$

Note that:

$$\mathbb{E}[A_{i,2}] = \int_{\epsilon^{\min}}^{\epsilon^{\max}} \left(\bar{A} + \epsilon\right) dF(\epsilon) \quad ; \quad \mathbb{E}\left[\left(A_{i,2}\right)^{\frac{1}{1-\alpha}}\right] = \int_{\epsilon^{\min}}^{\epsilon^{\max}} \left(\bar{A} + \epsilon\right)^{\frac{1}{1-\alpha}} dF(\epsilon) \tag{22}$$

Both aggregate capital and aggregate output are strictly higher in the economy with data than in the economy without data.

This lemma underscores an important feature of the data economy. Even when it does not favorably affect the aggregate distribution of productivity, the presence of data will increase aggregate output. This holds because compared to the economy without data, firms with below-average (above-average) productivity draws will produce less (more) in the economy with data. Because the optimal capital demand of a firm is convex in its productivity (which follows from the concavity of the firm's production function), the latter effect dominates and total output is increased by the availability of data. Crucially, these results hold true even though there is no misallocation, given that capital is supplied inelastically.

We can relate these insights to another important issue in digital markets, namely market tipping. Suppose that there is an active data feedback effect as described in section 3 and consider two different firms, one without access to data and another with access to a signal about its future productivity realizations. In expectation, the firm with data will produce higher levels of output, which allows it to obtain even more access to data. This feedback loop will propagate over time, enabling the firm that initially had a data advantage to obtain higher and higher market shares.

5.2 The impact of monetary policy and business cycles

The aforementioned results allow us to establish how the effect of a monetary policy shock differs in the two economies:

Proposition 7 *The absolute effects of a monetary policy shock on capital and output are* *higher in the economy with data, but the relative effects are the same in both economies, i.e.:*

$$\frac{\partial \bar{K}_2^d}{\partial r} < \frac{\partial \bar{K}_2^{nd}}{\partial r} \qquad ; \qquad \frac{\partial \bar{K}_2^d/\partial r}{\bar{K}_2^d} = \frac{\partial \bar{K}_2^{nd}/\partial r}{\bar{K}_2^{nd}} \tag{23}$$

$$\frac{\partial \bar{Y}_2^d}{\partial r} < \frac{\partial \bar{Y}_2^{nd}}{\partial r} \qquad ; \qquad \frac{\partial \bar{Y}_2^d/\partial r}{\bar{Y}_2^d} = \frac{\partial \bar{Y}_2^{nd}/\partial r}{\bar{Y}_2^{nd}} \tag{24}$$

This follows because in the economy with data, the response of any firm to a monetary policy shock is proportional to its productivity. Because firms with higher productivity produce more output, their response to a shock will be comparatively large, i.e. the magnitude of their response is larger than that of a firm with lower productivity. Thus, the absolute effect of a monetary policy shock is larger in the economy with data. However, this proportionality also implies that the relative effects of a monetary policy shock are identical in the two economies, which can be seen directly when studying the closed-form solutions for aggregate capital and output given in lemma 1.

Next, we establish how the magnitude of business cycles driven by productivity shocks differs across the two economies:

Proposition 8 *Suppose $\alpha \leq 1/2$. Then, the absolute and relative effects of an aggregate productivity shock are smaller in the economy with data, i.e.:*

$$\frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} \geq \frac{\partial \bar{Y}_2^d}{\partial \bar{A}} \qquad ; \qquad \frac{\partial \bar{Y}_2^{nd}/\partial \bar{A}}{\bar{Y}_2^{nd}} > \frac{\partial \bar{Y}_2^d/\partial \bar{A}}{\bar{Y}_2^d}. \tag{25}$$

Given that estimates for the parameter α are commonly in the range $[0.3, 0.5]$, this result indicates that cyclical fluctuations may be dampened when firms have data about their future idiosyncratic productivities.

These insights are visualized in the following figure, in which we consider the dynamics of output in response to anticipated cyclical fluctuations in $\bar{A}$. We suppose that $\alpha = 0.4$, $r = 0.2$, $\epsilon^{min} = 0$, and $\epsilon^{max} = 5$. In any graph, we depict the time periods on the x-axis. A hypothetical cyclical path for $\bar{A}$ is plotted in the left graph and the implied values of aggregate output under data and no data are visualized in the middle figure. Finally, the relative deviations of output from its value at the expected productivity are plotted in the right figure.

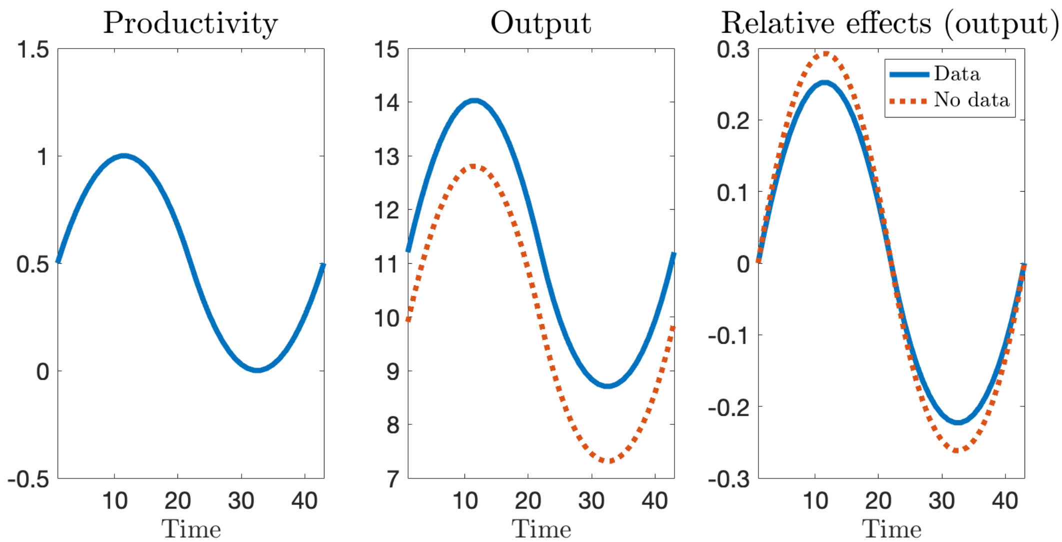

Figure 5: Data and business cycles

5.3 The role of firm bankruptcy

The impact of data on output and its responsiveness to shocks is closely connected to the issue of firm bankruptcy, from which we have abstracted sofar. Suppose that producing a non-zero amount incurs fixed cost $\tau \geq 0$ to any firm and that firms can, after observing their productivity draw in a period, choose whether or not to produce this period. These firms will choose not to produce, i.e. declare bankruptcy, whenever their variable profits are below the fixed costs. Moreover, firms are forward-looking: When deciding their capital stock in the first period, they take into account their optimal decisions in the period thereafter.

In an economy with data, the incidence of bankruptcy will be lower. This can be seen most easily when considering the case in which fixed costs are zero, i.e. $\tau = 0$. Then, no firm will ever go bankrupt in the economy with data, because positive profits can always be attained by choosing the statically optimal capital stock for a given level of productivity. In the economy without data, firms that draw very low productivities will have an excessively high capital stock and will go bankrupt. This result extends to more general levels of $\tau > 0$. This is because the profits of a firm in the economy with data will always be higher, given that this firm holds a capital stock that maximizes profits for its given productivity draw.

The possibility of firm bankruptcy can amplify the effects of a monetary policy shock in the economy without data, compared to the economy with data. To see this, consider the effects of a contractionary monetary policy shock when $\tau = 0$. In the economy with data, such a shock will not affect the share of firms going bankrupt. In the economy without data, a contractionary monetary policy shock will increase the share of firms going bankrupt, which further amplifies the negative effects of the shock on output.

These notions are visualized in the following figure. We fix $\alpha = 0.5$, $\bar{A} = 0$, $\epsilon^{min} = 0$, and $\epsilon^{max} = 1$ and set $\tau = 1$. On the x-axis, we plot different values of $r \in \{0.06, ..., 0.1\}$. In the left graph, we plot the associated levels of output in two different economies, namely (i) an economy with data (red) and (ii) an economy without data (green). The absolute and relative effects of a monetary policy shock in the two different economies are plotted in the middle and right graphs.

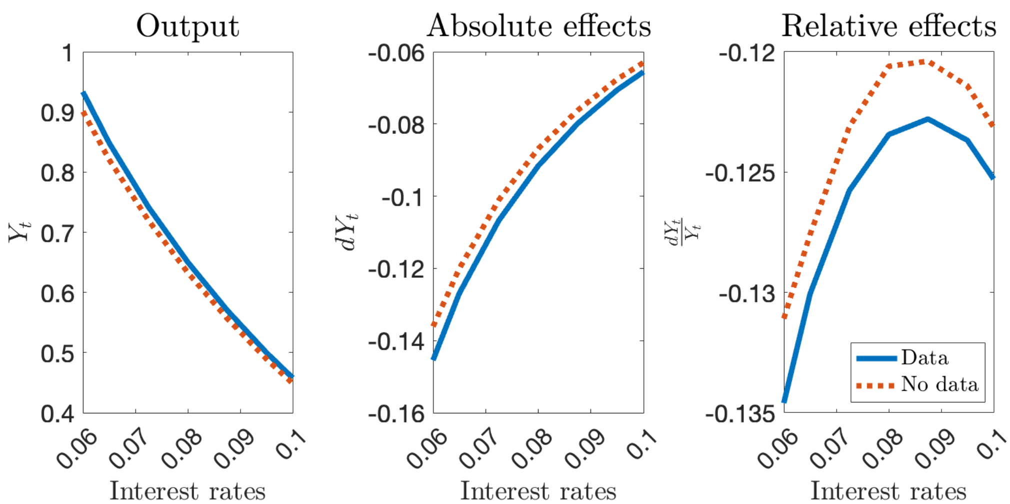

Figure 6: Data, bankruptcy, and monetary policy

6 Conclusion

Modern economies increasingly revolve around data, which yields value to firms by facilitating targeted advertising, by guiding innovation, and by allowing firms to tailor their products to the desires of consumers. In this paper, we propose a novel framework that allows us to study how the widespread availability of data to firms affects the transmission of monetary policy and the magnitude of cyclical fluctuations. In our model, data yields value by resolving the uncertainty a firm faces. We also incorporate a data feedback loop, which captures the idea that firms who generate a lot of sales receive high-quality data as a byproduct.

We find that the investment of data-rich firms responds comparatively strongly to monetary policy shocks, both in absolute and relative terms. This follows from the fact that data rich firms are both larger and face lower idiosyncratic uncertainty. Moreover, we show that heterogeneous access to data can increase the effects of a monetary policy shock on aggregate investment. However, expansionary monetary policy increases market concentration when access to data is unequally distributed.

The interaction between data and the strength of cyclical fluctuations is non-trivial. As such, the data feedback loop amplifies the strength of cyclical fluctuations. Moreover, the data feedback loop can incentivize firms which are sensitive to uncertainty to attain greater scale, thus raising the vulnerability of the economy to aggregate uncertainty shocks. However, the availability of data which allows firms to accurately predict their future idiosyncratic productivities dampens cyclical fluctuations driven by aggregate productivity shocks.

Our framework is deliberately kept stylized to highlight how key features of data affect the transmission of monetary policy and cyclical fluctuations. Quite naturally, there is a wide range of other frictions or channels that we abstract from—for example, price setting, general equilibrium, competition or financial frictions. We leave these topics for future work.

A Mathematical proofs

A.1 Proof of proposition 1:

Part 1: Calculating the effects of a contractionary MP shock:

Suppose firstly that there is no data feedback effect (i.e. $z = 0$) and that $\alpha = 1/2$. Then, the corresponding first order condition becomes:

$$0.5K_{t+1}^{-0.5}\left(\bar{A} - \kappa_e \sigma_{t+1}\right) - 0.5\rho\left(\bar{V} + \kappa_v \sigma_{t+1}\right) K_{t+1}^{0} - r_t + \beta(1-\delta)\mathbb{E}_t r_{t+1} = 0$$

Thus, the optimal capital stock will satisfy:

$$K_{t+1}^{-0.5}\mathbb{E}[A_{t+1};\sigma_{t+1}] = 2\underbrace{\left(r_t - \beta(1-\delta)\mathbb{E}_t r_{t+1}\right)}_{:=r} + \rho VAR[A_{t+1};\sigma_{t+1}] \iff$$

$$K_{t+1}^{*} = \left(\frac{\mathbb{E}[A_{t+1};\sigma_{t+1}]}{2r + \rho VAR[A_{t+1};\sigma_{t+1}]}\right)^2$$

Thus, we have:

$$\frac{\partial K_{t+1}^{*}}{\partial \sigma_{t+1}} = 2\left(\frac{\mathbb{E}[A_{t+1}]}{VAR[A_{t+1}]\rho + 2r}\right)\left(\frac{[VAR[A_{t+1};\sigma_{t+1}]\rho + 2r]\frac{\partial \mathbb{E}[A_{t+1}]}{\partial \sigma} - \mathbb{E}[A_{t+1}]\frac{\partial VAR[A_{t+1}]}{\partial \sigma_{t+1}}\rho}{[VAR[A_{t+1}]\rho + 2r]^2}\right) \implies$$

$$\frac{\partial K_{t+1}^{*}}{\partial \sigma_{t+1}} = 2\left(\frac{\mathbb{E}[A_{t+1}]}{[VAR[A_{t+1}]\rho + 2r]^3}\right)\left([VAR[A_{t+1}]\rho + 2r]\frac{\partial \mathbb{E}[A_{t+1}]}{\partial \sigma_{t+1}} - \mathbb{E}[A_{t+1}]\frac{\partial VAR[A_{t+1}]}{\partial \sigma_{t+1}}\rho\right) < 0 \tag{26}$$

Now let's examine the effect of a monetary policy shock. The relevant derivative is:

$$\frac{\partial K_{t+1}^{*}}{\partial r_t} = \frac{-4[\mathbb{E}[A_{t+1}]]^2[VAR[A_{t+1}]\rho + 2r]}{[VAR[A_{t+1}]\rho + 2r]^4} \implies \frac{\partial K_{t+1}^{*}}{\partial r_t} = \frac{-4[\mathbb{E}[A_{t+1}]]^2}{[VAR[A_{t+1}]\rho + 2r]^3} \tag{27}$$

Note: These derivatives are derived using the implicit function theorem. All cross-derivative from now on cannot be derived this way, but have to be manually differentiated. When evaluating the effects of parameter changes, we also have to take into account how these changes affect the optimal K_{t+1}.

Now let's examine how this function depends on σ_{t+1}. To see this, note that:

$$\frac{\partial^2 K^*_{t+1}}{\partial r_t \partial \sigma_{t+1}} = \frac{[VAR[A_{t+1}]\rho + 2r]^3\Big(- 8\mathbb{E}[A_{t+1}]\frac{\partial \mathbb{E}[A_{t+1}]}{\partial \sigma_{t+1}}\Big) + 4[\mathbb{E}[A_{t+1}]]^2\Big(3[VAR[A_{t+1}]\rho + 2r]^2\rho\frac{\partial VAR[A_{t+1}]}{\partial \sigma_{t+1}}\Big)}{[VAR[A_{t+1}]\rho + 2r]^6}$$

$$\Longrightarrow$$

$$\frac{\partial^2 K^*_{t+1}}{\partial r_t \partial \sigma_{t+1}} = \frac{[VAR[A_2]\rho + 2r]^3\big(8\mathbb{E}[A_{t+1}]\kappa_e\big) + 4[\mathbb{E}[A_{t+1}]]^2\big(3[VAR[A_{t+1}]\rho + 2r]^2\rho\kappa_v\big)}{[VAR[A_{t+1}]\rho + 2r]^6} > 0 \tag{28}$$

This also implies that:

$$\frac{\partial^2 K^*_{t+1}}{\partial r \partial \sigma_{t+1}} = \frac{8\mathbb{E}[A_{t+1}]\kappa_e}{[VAR[A_{t+1}]\rho + 2r]^3} + \frac{12[\mathbb{E}[A_{t+1}]]^2\rho\kappa_v}{[VAR[A_{t+1}]\rho + 2r]^4}$$

In turn, this means that:

$$\frac{\partial^3 K^*_{t+1}}{\partial r_t \partial \sigma^2_{t+1}} = \frac{[VAR[A_{t+1}]\rho + 2r]^3\big(- 8\kappa_e^2\big) - 8\mathbb{E}[A_{t+1}]\kappa_e\big(3[VAR[A_{t+1}]\rho + 2r]^2\kappa_v\rho\big)}{[VAR[A_{t+1}]\rho + 2r]^6} +$$

$$\frac{[VAR[A_{t+1}]\rho + 2r]^4\big(- 24\mathbb{E}[A_{t+1}]\kappa_e\rho\kappa_v\big) - 12[\mathbb{E}[A_{t+1}]]^2\rho\kappa_v\big(4[VAR[A_{t+1}]\rho + 2r]^3\kappa_v\rho\big)}{[VAR[A_{t+1}]\rho + 2r]^8}$$

$$\Longrightarrow \frac{\partial^3 K^*_{t+1}}{\partial r_t \partial \sigma^2_{t+1}} < 0 \tag{29}$$

<u>Part 2:</u> The effects in relative terms

Let's also examine the changes in the relative effects of monetary policy. These are given by:

$$\gamma(\sigma_{t+1}) := \frac{\frac{\partial K^*_{t+1}}{\partial r_t}}{K^*_{t+1}} = \frac{\frac{-4[\mathbb{E}[A_{t+1}]]^2}{[VAR[A_{t+1}]\rho + 2r]^3}}{\frac{[\mathbb{E}[A_{t+1}]]^2}{[VAR[A_{t+1}]\rho + 2r]^2}} = \frac{-4}{[VAR[A_{t+1}]\rho + 2r]}$$

We can analyse whether these relative effects are larger for high-data firms by investigating:

$$\frac{\partial \gamma(\sigma_{t+1})}{\partial \sigma_{t+1}} = 4\kappa_v\rho[VAR[A_{t+1}]\rho + 2r]^{-2} > 0 \tag{30}$$

$$\frac{\partial^2\gamma(\sigma_{t+1})}{\partial\sigma_{t+1}^2} = -8(\kappa_v\rho)^2[VAR[A_{t+1}]\rho + 2r]^{-3} < 0 \tag{31}$$

A.2 Proof of proposition 2

Consider the two firms $j \in \{l, h\}$ with $\sigma_l = \bar{\sigma}_l - b$, $\sigma_h = \bar{\sigma}_h - b$, and $\bar{\sigma}_l < \bar{\sigma}_h$. We have:

$$VAR[A_j; \sigma_j^2] = \tilde{V} + \kappa_v\sigma_j = \tilde{V} + \kappa_v\bar{\sigma}_j - \kappa_v b$$

Consider the effects of an increase in b (i.e. a general improvement in the availability of data) on the market share of a firm with access to superior data.

One can show that:

$$M_l = \frac{1}{1 + \frac{A_h}{A_l}\left(\frac{K_h^*}{K_l^*}\right)^\alpha} \implies \frac{\partial M_l}{\partial b} = \frac{-\frac{\alpha A_h}{A_l}\left(\frac{K_h^*}{K_l^*}\right)^{\alpha-1}\left[\frac{\partial}{\partial b}\left(\frac{K_h^*}{K_l^*}\right)\right]}{\left(1 + \frac{A_h}{A_l}\left(\frac{K_h^*}{K_l^*}\right)^\alpha\right)^2}$$

Note that:

$$\frac{\partial}{\partial b}\left[\frac{K_h^*}{K_l^*}\right] = \frac{K_l^*\frac{\partial K_h^*}{\partial b} - K_h^*\frac{\partial K_l^*}{\partial b}}{(K_l^*)^2} < 0$$

To see this, note that:

$$\frac{\partial K_j^*}{\partial b} = \frac{-2\rho[\mathbb{E}[A_j]]^2}{[VAR[A_j; \sigma_j]\rho + 2r]^3}\Bigg(\underbrace{\frac{\partial VAR[A_j; \sigma_j^2]}{\partial b}}_{=-\kappa_v}\Bigg)$$

This implies that:

$$K_l^*\frac{\partial K_h^*}{\partial b} = \frac{[\mathbb{E}[A]]^2}{[VAR[A_l; \sigma_l]\rho + 2r]^2}\left[\frac{2\rho\kappa_v[\mathbb{E}[A]]^2}{[VAR[A_h; \sigma_h]\rho + 2r]^3}\right]$$

$$-K_h^*\frac{\partial K_l^*}{\partial b} = -\frac{[\mathbb{E}[A]]^2}{[VAR[A_h; \sigma_h]\rho + 2r]^2}\left[\frac{2\rho\kappa_v[\mathbb{E}[A]]^2}{[VAR[A_l; \sigma_l]\rho + 2r]^3}\right]$$

It follows that:

$$K_l^*\frac{\partial K_h^*}{\partial b} - K_h^*\frac{\partial K_l^*}{\partial b} < 0$$

$$\iff$$

$$\frac{[\mathbb{E}[A]]^2}{[VAR[A_l; \sigma_l]\rho + 2r]^2}\left[\frac{2\rho\kappa_v[\mathbb{E}[A]]^2}{[VAR[A_h; \sigma_h]\rho + 2r]^3}\right] - \frac{[\mathbb{E}[A]]^2}{[VAR[A_h; \sigma_h]\rho + 2r]^2}\left[\frac{2\rho\kappa_v[\mathbb{E}[A]]^2}{[VAR[A_l; \sigma_l]\rho + 2r]^3}\right] < 0$$

$$\iff$$

$$\frac{1}{[VAR[A_l;\sigma_l]\rho+2r]^2[VAR[A_h;\sigma_h]\rho+2r]^3} < \frac{1}{[VAR[A_h;\sigma_h]\rho+2r]^2[VAR[A_l;\sigma_l]\rho+2r]^3}$$

$$\Longleftrightarrow$$

$$VAR[A_l;\sigma_l]\rho < VAR[A_h;\sigma_h]\rho$$

A.3 Proof of corollary 1

Suppose that $\alpha = 1/2$, $z = 0$, $\kappa_e = 0$, and consider an economy with two firms $i \in \{l, h\}$, whose data quality levels are given by $\sigma_l = \bar{\sigma}_1 - b$ and $\sigma_h = \bar{\sigma}_h - b$, with $\bar{\sigma}_l < \bar{\sigma}_h$. We aim to show that an rise of b increases the relative effect of monetary policy on aggregate investment, i.e.:

$$\frac{\partial\left(\frac{\frac{\partial(K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h))}{\partial r_t}}{K^*_{t+1}(\sigma_l)+K^*_{t+1}(\sigma_h)}\right)}{\partial b} < 0. \tag{32}$$

Based on previous results, we can write:

$$\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)} = \frac{-\frac{4}{[VAR[A_{t+1};\sigma_l]\rho+2r]^3} - \frac{4}{[VAR[A_{t+1};\sigma_h]\rho+2r]^3}}{\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2}}$$

Thus, the relative effect of a monetary policy shock can be written as follows:

$$\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)} = \left(\frac{\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^2}}{\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2}}\right)\left(-\frac{4}{[VAR[A_{t+1};\sigma_l]\rho+2r]}\right)+$$

$$\left(\frac{\frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2}}{\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2}}\right)\left(-\frac{4}{[VAR[A_{t+1};\sigma_h]\rho+2r]}\right) \tag{33}$$

Working with the definition of capital shares $C_i(\sigma_l, \sigma_h) = \frac{K^*(\sigma_i)}{K^*(\sigma_l)+K^*(\sigma_h)}$, we can write:

$$\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)} = C_l(\sigma_l,\sigma_h)\gamma(\sigma_l) + C_h(\sigma_l,\sigma_h)\gamma(\sigma_h) \tag{34}$$

Now we can study how this is shaped by an increase of b. We can write the following decomposition:

$$\frac{\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)}}{\partial b} = C_l(\sigma_l,\sigma_h)\frac{\partial\gamma(\sigma_l)}{\partial b} + C_h(\sigma_l,\sigma_h)\frac{\partial\gamma(\sigma_h)}{\partial b} + \frac{\partial C_l}{\partial b}\gamma(\sigma_l) + \frac{\partial C_h}{\partial b}\gamma(\sigma_h) \tag{35}$$

This implies that:

$$\frac{\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)}}{\partial b} = C_l(\sigma_l,\sigma_h)\frac{\partial\gamma(\sigma_l)}{\partial b} + C_h(\sigma_l,\sigma_h)\frac{\partial\gamma(\sigma_h)}{\partial b} + \frac{\partial C_l}{\partial b}\big(\gamma(\sigma_l)-\gamma(\sigma_h)\big) \tag{36}$$

The last term is strictly negative since $\frac{\partial C_l}{\partial b} > 0$ (this holds by the previously made argument that $\frac{K_h^*}{K_l^*}$ is falling in b) and the fact that:

$$\gamma(\sigma_l)-\gamma(\sigma_h) = -\frac{4}{[VAR[A_{t+1};\sigma_l]\rho+2r]} + \frac{4}{[VAR[A_{t+1};\sigma_h]\rho+2r]} < 0 \iff$$

$$\frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]} < \frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]} \iff VAR[A_{t+1};\sigma_l] < VAR[A_{t+1};\sigma_h]$$

A.4 Proof of proposition 3

Under our assumptions, we have obtained the following closed-form solution for the chosen capital stock (and the derivative of it w.r.t. the interest rate):

$$K_{t+1}^* = \frac{[\mathbb{E}[A_{t+1}]]^2}{[VAR[A_{t+1};\sigma_{t+1}]\rho+2r]^2} \quad ; \quad \frac{\partial K_{t+1}^*}{\partial r} = \frac{-4[\mathbb{E}[A_{t+1}]]^2}{[VAR[A_{t+1};\sigma_{t+1}]\rho+2r]^3} \tag{37}$$

Define A_j and K_j as the realized productivity and the chosen capital stock of firm j. The market share of the firm with access to superior data can be written as:

$$M_l = \frac{A_l(K_l^*)^\alpha}{A_l(K_l^*)^\alpha + A_h(K_h^*)^\alpha} = \frac{1}{1+\frac{A_h}{A_l}\big(\frac{K_h^*}{K_l^*}\big)^\alpha}$$

We will show that M_l is falling in r, i.e. that an expansiory MP shock raises the market share of the firm with access to superior data.

Note that the sign of the derivative of the market share with respect to r has the opposite sign as $\frac{\partial}{\partial r}\big(\frac{K_h^*}{K_l^*}\big)$, because:

$$\frac{\partial M_l}{\partial r} = \frac{-\frac{\alpha A_h}{A_l}\big(\frac{K_h^*}{K_l^*}\big)^{\alpha-1}\big[\frac{\partial}{\partial r}\big(\frac{K_h^*}{K_l^*}\big)\big]}{\big(1+\frac{A_h}{A_l}\big(\frac{K_h^*}{K_l^*}\big)^\alpha\big)^2}$$

Note that:

$$\frac{\partial}{\partial r}\left[\frac{K_h^*}{K_l^*}\right] = \frac{K_l^*\frac{\partial K_h^*}{\partial r} - K_h^*\frac{\partial K_l^*}{\partial r}}{(K_l^*)^2} > 0$$

We can show that the numerator is strictly positive. To see this, note that:

$$K_l^* \frac{\partial K_h^*}{\partial r} = \frac{[\mathbb{E}[A]]^2}{[VAR[A_l;\sigma_l]\rho + 2r]^2} \left[\frac{-4[\mathbb{E}[A]]^2}{[VAR[A_h;\sigma_h]\rho + 2r]^3} \right]$$

$$-K_h^* \frac{\partial K_l^*}{\partial r} = -\frac{[\mathbb{E}[A]]^2}{[VAR[A_h;\sigma_h]\rho + 2r]^2} \left[\frac{-4[\mathbb{E}[A]]^2}{[VAR[A_l;\sigma_l]\rho + 2r]^3} \right]$$

It follows that:

$$K_l^* \frac{\partial K_h^*}{\partial r} - K_h^* \frac{\partial K_l^*}{\partial r} > 0$$

$$\iff$$

$$\frac{[\mathbb{E}[A]]^2}{[VAR[A_l;\sigma_l]\rho + 2r]^2} \left[\frac{-4[\mathbb{E}[A]]^2}{[VAR[A_h;\sigma_h]\rho + 2r]^3} \right] + \frac{[\mathbb{E}[A]]^2}{[VAR[A_h;\sigma_h]\rho + 2r]^2} \left[\frac{4[\mathbb{E}[A]]^2}{[VAR[A_l;\sigma_l]\rho + 2r]^3} \right] > 0$$

$$\iff$$

$$-\frac{1}{[VAR[A_l;\sigma_l]\rho + 2r]^2[VAR[A_h;\sigma_h]\rho + 2r]^3} + \frac{1}{[VAR[A_h;\sigma_h]\rho + 2r]^2[VAR[A_l;\sigma_l]\rho + 2r]^3} > 0$$

$$\iff$$

$$[VAR[A_l;\sigma_l]\rho + 2r]^2[VAR[A_h;\sigma_h]\rho + 2r]^3 > [VAR[A_h;\sigma_h]\rho + 2r]^2[VAR[A_l;\sigma_l]\rho + 2r]^3$$

$$\iff$$

$$VAR[A_h;\sigma_h]\rho + 2r > VAR[A_l;\sigma_l]\rho + 2r$$

This holds true by construction.

The increased market share of the firm with access to superior data will lead to higher market concentration if both firms had the same productivity draw ex ante, as defined by the Herfindahl-Hirschman index. To see this, note that this index is defined as:

$$H = (M_l)^2 + (1 - M_l)^2 \implies \frac{\partial H}{\partial M_l} = 2M_l - 2(1 - M_l) > 0$$

The latter holds because $M_l > 0.5$.

A.5 Proof of proposition 4

Part 1: Absolute effects

Now we investigate the effect of an MP shock on two different economies: one where data is unequally distributed (in the sense of a mean preserving spread) and one where all firms have the same level of σ. There are two different firms with data levels given by σ_l and σ_h.

We define that $\sigma_h = \sigma_m + y$, while $\sigma_l = \sigma_m - y$. If $y = 0$, both firms have access to the same level of data. Increases of y constitute mean-preserving spreads in the access to data.

Total investment is $K^*_{t+1}(\sigma_l) + K^*_{t+1}(\sigma_h)$. The effect of a monetary policy shock on total investment is given by:

$$\frac{\partial K^*_{t+1}(\sigma_h)}{\partial r_t} + \frac{\partial K^*_{t+1}(\sigma_l)}{\partial r_t} \tag{38}$$

We know that the function $\frac{\partial K^*_{t+1}(\sigma)}{\partial r_t}$ is a strictly concave function in σ. We can write the σ_m of the firm in the economy without data dispersion as $\sigma_m = 0.5\sigma_l + 0.5\sigma_h$. Thus, the effect of an MP shock on total investment in this economy is given by:

$$2\frac{\partial K^*_{t+1}(\sigma_m)}{\partial r_t} \tag{39}$$

By strict concavity, we have:

$$\frac{\partial K^*_{t+1}(\sigma_m)}{\partial r_t} \geq 0.5\frac{\partial K^*_{t+1}(\sigma_l)}{\partial r_t} + 0.5\frac{\partial K^*_{t+1}(\sigma_h)}{\partial r_t} \tag{40}$$

Thus, if there is a mean-preserving spread in the access to data, the total effect of a monetary policy shock becomes more negative. In words, MP is more effective in the economy where data is dispersed.

<u>Part 2:</u> Relative effects.

Consider the effect of an MP shock on investment, relative to the initial level of investment. This is given by:

$$\frac{\sum(\partial K^*_{t+1}(\sigma)/\partial r_t)}{\sum K^*_{t+1}(\sigma)}$$

Let's focus on the simple case $\kappa_e = 0$ again. We have shown that:

$$\frac{\partial K^*_{t+1}}{\partial r_t} = \frac{-4[\mathbb{E}[A_{t+1}]]^2}{[VAR[A_{t+1}]\rho + 2r]^3} = \frac{-4[\bar{A}]^2}{[VAR[A_{t+1}]\rho + 2r]^3}$$

Moreover, we have:

$$K^*_{t+1} = \left(\frac{\bar{A}}{2(r_t + \beta(1-\delta)\mathbb{E}_t r_{t+1}) + \rho VAR[A_{t+1}; \sigma_{t+1}]}\right)^2$$

Thus, we can write:

$$\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)} = \frac{-\frac{4}{[VAR[A_{t+1};\sigma_l]\rho+2r]^3} - \frac{4}{[VAR[A_{t+1};\sigma_h]\rho+2r]^3}}{\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2}} \approx \frac{-\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^3} - \frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^3}}{\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2}}$$

A mean-preserving spread of σ is a equivalent to an increase of y. Thus, we can derive the effects of a mean preserving spread in σ on the relative effects of monetary policy on aggregate investment by studying the derivative of $\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)}$ with respect to y.

In the following, we show that this derivative is negative, i.e. that mean preserving spreads of σ exacerbate the effects of monetary policy on aggregate investment. To do this, we define:

$$H(y) := -\frac{1}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^3} - \frac{1}{[VAR[A_{t+1};\sigma_h]\rho + 2r]^3} \tag{41}$$

$$L(y) := \frac{1}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho + 2r]^2} \tag{42}$$

Recall that $\frac{VAR[A_{t+1}]}{\partial\sigma_{t+1}} = \kappa_v$. Thus, one can compute the derivative of this expression:

$$\frac{\partial L}{\partial y} = +\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^3} - \frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho + 2r]^3} > 0$$

$$\frac{\partial H}{\partial y} = -\frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^4} + \frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho + 2r]^4} < 0$$

We can write:

$$\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)} = \frac{H(y)}{L(y)} \implies \frac{\partial\left(\frac{\sum(\partial K^*(\sigma)/\partial r)}{\sum K^*(\sigma)}\right)}{\partial y} = \frac{L(y)H'(y) - H(y)L'(y)}{[L(y)]^2}$$

Thus, we show our desired result by proving that $L(y)H'(y) - H(y)L'(y) < 0$. We have:

$$L\frac{\partial H}{\partial y} =$$

$$\left(\frac{1}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^2} + \frac{1}{[VAR[A_{t+1};\sigma_h]\rho + 2r]^2}\right)\left(\frac{-3\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^4} + \frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho + 2r]^4}\right)$$

$$=$$

$$\frac{-3\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^6} + \frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho + 2r]^2[VAR[A_{t+1};\sigma_h]\rho + 2r]^4}$$

$$-\frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^2[VAR[A_{t+1};\sigma_l]\rho+2r]^4}+\frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^6}$$
$$=$$
$$\frac{-3\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho+2r]^6}+\frac{3\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^6}$$
$$+\frac{3\rho\kappa_v\big([VAR[A_{t+1};\sigma_l]\rho+2r]^2-[VAR[A_{t+1};\sigma_h]\rho+2r]^2\big)}{[VAR[A_{t+1};\sigma_h]\rho+2r]^4[VAR[A_{t+1};\sigma_l]\rho+2r]^4}$$

Moreover, we have:

$$-H\frac{\partial L}{\partial y}=$$
$$\left(\frac{1}{[VAR[A_{t+1};\sigma_l]\rho+2r]^3}+\frac{1}{[VAR[A_{t+1};\sigma_h]\rho+2r]^3}\right)\left(\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho+2r]^3}-\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^3}\right)$$
$$=$$
$$\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho+2r]^6}-\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho+2r]^3[VAR[A_{t+1};\sigma_h]\rho+2r]^3}$$
$$+\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^3[VAR[A_{t+1};\sigma_l]\rho+2r]^3}-\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^6}$$
$$=$$
$$\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho+2r]^6}-\frac{2\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^6}$$

Thus, we can calculate:

$$L\frac{\partial H}{\partial y}-H\frac{\partial L}{\partial y}=$$
$$\frac{-\rho\kappa_v}{[VAR[A_{t+1};\sigma_l]\rho+2r]^6}+\frac{\rho\kappa_v}{[VAR[A_{t+1};\sigma_h]\rho+2r]^6}+$$
$$\frac{3\rho\kappa_v\big([VAR[A_{t+1};\sigma_l]\rho+2r]^2-[VAR[A_{t+1};\sigma_h]\rho+2r]^2\big)}{[VAR[A_{t+1};\sigma_h]\rho+2r]^4[VAR[A_{t+1};\sigma_l]\rho+2r]^4}<0$$

This expression is negative because $VAR[A_{t+1};\sigma_l] < [VAR[A_{t+1};\sigma_h]$.

A.6 Proof of proposition 5

Part 1: Absolute effects:

Recall that we defined: $\sigma_{i,t}=\bar{\sigma}_i-zK_{i,t}$, which means that we can write:

$$VAR[A_{t+1};\sigma_{t+1}]=\tilde{V}+\kappa_v\sigma_j=\tilde{V}+\kappa_v\bar{\sigma}-\kappa_v zK_{t+1}$$

Generally speaking, we also have:

$$\mathbb{E}[A_{t+1}; \sigma_{t+1}] = \tilde{A} - \kappa_e \sigma_{t+1} \implies \frac{\partial \mathbb{E}[A_{t+1}]}{\partial K_{t+1}} > 0$$

When $\kappa_e = 0$, this derivative becomes zero.

Our first-order condition for the optimal capital stock reads:

$$\alpha K_{t+1}^{\alpha-1} \left(\bar{A} - \kappa_e \sigma_t\right) + \frac{\partial \mathbb{E}[A_{t+1}]}{\partial K_{t+1}} K_{t+1}^{\alpha} - \rho \left(\bar{V} + \kappa_v \sigma_{t+1}\right) \alpha K_{t+1}^{2\alpha-1} - \frac{\rho}{2}(-\kappa_v z) K_{t+1}^{2\alpha} - r_t + \beta(1-\delta)\mathbb{E}_t r_{t+1} = 0$$

For $\alpha = 1/2$ and $\kappa_e = 0$, the first-order condition that the optimal capital stock has to satisfy reads:

$$0.5 K_{t+1}^{-0.5} \left(\bar{A}\right) - 0.5\rho \left(\bar{V} + \kappa_v \bar{\sigma} - \kappa_v z K_{t+1}\right) - 0.5\rho(-\kappa_v z) K_{t+1} - r_t + \beta(1-\delta)\mathbb{E}_t r_{t+1} = 0$$

The effect of a monetary policy shock is given by:

$$\frac{\partial K_{t+1}}{\partial r_t} = \frac{1}{-0.25 \bar{A} K_{t+1}^{-1.5} + \rho \kappa_v z} < 0 \tag{43}$$

Similarly, one can show that:

$$\frac{\partial K_{t+1}}{\partial z} = \frac{-\rho \kappa_v K_{t+1}}{-0.25 \bar{A} K_{t+1}^{-1.5} + \rho \kappa_v z} > 0 \tag{44}$$

Thus: if the data feedback effect is stronger, firms hold more capital. This makes perfect sense, as a stronger data feedback effect shifts up the marginal product of capital.

Now we can analyse how the data feedback effect will come into play:

$$\frac{\partial K_{t+1}}{\partial r_t \partial z} = \frac{-\left[(3/8)\bar{A} K_{t+1}^{-2.5} \frac{\partial K_{t+1}}{\partial z} + \rho \kappa_v\right]}{\left[-0.25(\bar{A}) K_{t+1}^{-1.5} + \rho \kappa_v z\right]^2} < 0 \tag{45}$$

<u>Part 2:</u> Relative effects

We begin by defining the relative effect of an interest rate increase:

$$\phi(\sigma_{t+1}) := \frac{\frac{\partial K_{t+1}}{\partial r_t}}{K_{t+1}} \tag{46}$$

When $\alpha = 1/2$ and $\kappa_e = 0$, the first-order condition the optimal capital stock has to satisfy reads:

$$0.5K_{t+1}^{-0.5}\left(\bar{A}\right) - 0.5\rho\left(\bar{V} + \kappa_v\bar{\sigma} - \kappa_v z K_{t+1}\right) - 0.5\rho(-\kappa_v z)K_{t+1} - r_t + \beta(1-\delta)\mathbb{E}_t r_{t+1} = 0$$

Thus, we have:

$$\phi(\sigma_{t+1}) = \frac{1}{(-0.25\bar{A}K_{t+1}^{-1.5} + \rho\kappa_v z)K_{t+1}}$$

Thus, we can write:

$$\frac{\partial\phi(\sigma_{t+1})}{\partial z} = \frac{-1}{[(-0.25\bar{A}K_{t+1}^{-1.5} + \rho\kappa_v z)K_{t+1}]^2}$$

$$\left[(-0.25\bar{A}K_{t+1}^{-1.5} + \rho\kappa_v z)\frac{\partial K_{t+1}}{\partial z} + \left((3/8)\bar{A}K_{t+1}^{-2.5}\frac{\partial K_{t+1}}{\partial z} + \rho\kappa_v\right)K_{t+1}\right]$$

$$=$$

$$\frac{-1}{[-0.25\bar{A}K_{t+1}^{-1.5} + \rho\kappa_v z)K_{t+1}]^2}\left[(1/8)\bar{A}K_{t+1}^{-1.5}\frac{\partial K_{t+1}}{\partial z} + \rho\kappa_v z\frac{\partial K_{t+1}}{\partial z} + \rho\kappa_v K_{t+1}\right] < 0$$

A.7 Proof of proposition 6

Examine first whether firms who are more sensitive to risk will hold more or less capital. We begin, once again, by examining the FOC when $\alpha = 1/2$:

$$0.5K_{t+1}^{-0.5}\left(\bar{A} - \kappa_e\sigma_{t+1}\right) + \frac{\partial\mathbb{E}[A_{t+1}]}{\partial K_{t+1}}K_{t+1}^{0.5} - \rho\left(\bar{V} + \kappa_v\sigma_{t+1}\right)0.5K_{t+1}^{0} - 0.5\rho(-\kappa_v z)K_{t+1} - r = 0$$

Let's investigate the following derivative:

$$\frac{\partial K_{t+1}}{\partial\rho} = \frac{0.5\left(\bar{V} + \kappa_v\bar{\sigma} - \kappa_v z K_{t+1}\right) - 0.5\kappa_v z K_{t+1}}{\partial^2\Pi_{t+1}/\partial K_{t+1}^2} \tag{47}$$

If $z = 0$, this derivative is clearly negative, because the denominator is negative. If z is high enough such that the numerator becomes negative, the result flips.

A.8 Proof of lemma 1

In an economy without data, firms do not know the realizations of A_{t+1} when making their capital choices. In an economy with data, all firms know A_{t+1} when choosing K_{t+1}.

The optimization problem of a firm with data is just:

$$K^d_{t+1}(A_{t+1}) = arg\max_{K_{t+1}} A_{t+1}(K_{t+1})^\alpha - rK_{t+1} \iff K^d_{t+1}(A_{t+1}) = \big(\alpha A_{t+1}\big)^{1/(1-\alpha)} r^{1/(\alpha-1)}$$

Aggregate capital in the economy with data is:

$$\bar{K}^d_{t+1} = \int_{A^{min}}^{A^{max}} K^d_{t+1}(A_{t+1}) dA_{t+1}$$

Consider any firm in the economy without data. This firm maximizes the following profit function:

$$K^{nd}_{t+1} = arg\max_{K_{t+1}} \left[\int_{A^{min}}^{A^{max}} A_{t+1}(K_{t+1})^\alpha dA_{t+1}\right] - rK_{t+1} \iff \alpha\mathbb{E}[A_{t+1}](K_{t+1})^{\alpha-1} - r = 0$$

Taking all this together, the aggregate capital stocks under data and no data are:

$$\bar{K}^{nd}_{t+1} = \big(\alpha\mathbb{E}[A_{t+1}]\big)^{1/(1-\alpha)}(r)^{1/(\alpha-1)}$$

$$\bar{K}^d_{t+1} = \mathbb{E}\big[(\alpha A_{t+1})^{1/(1-\alpha)}\big](r)^{1/(\alpha-1)}$$

Now we determine the levels of output, beginning with output in the economy with data, which is:

$$\bar{Y}^d_{t+1} = \int_{A^{min}}^{A^{max}} A_{t+1}\Big(\big(\alpha A_{t+1}\big)^{1/(1-\alpha)} r^{1/(\alpha-1)}\Big)^\alpha dA_{t+1} =$$

$$(\alpha)^{\alpha/(1-\alpha)}(r)^{\alpha/(\alpha-1)}\int_{A^{min}}^{A^{max}} A_{t+1}\big(A_{t+1}\big)^{\alpha/(1-\alpha)} dA_{t+1} = (\alpha)^{\alpha/(1-\alpha)}(r)^{\alpha/(\alpha-1)}\mathbb{E}\big[(A_{t+1})^{\frac{1}{1-\alpha}}\big]$$

Finally, output in the economy without data is:

$$\bar{Y}^{nd}_{t+1} = \int_{A^{min}}^{A^{max}} A_{t+1}\Big((\alpha)^{1/(1-\alpha)}\big(\mathbb{E}[A_{t+1}]\big)^{1/(1-\alpha)} r^{1/(\alpha-1)}\Big)^\alpha dA_{t+1} =$$

$$(\alpha)^{\alpha/(1-\alpha)}(r)^{\alpha/(\alpha-1)}\int_{A^{min}}^{A^{max}} A_{t+1}\big(\mathbb{E}[A_{t+1}]\big)^{\alpha/(1-\alpha)} dA_{t+1} = (\alpha)^{\alpha/(1-\alpha)}(r)^{\alpha/(\alpha-1)}\big(\mathbb{E}[A_{t+1}]\big)^{1/(1-\alpha)}$$

A.9 Proof of proposition 7

We begin by showing that the absolute effects of a monetary policy shock are larger in the economy with data. To see this, note that:

$$\frac{\partial \bar{Y}_2^d}{\partial r} = \left(\frac{\alpha}{\alpha-1}\right)(\alpha)^{\alpha/(1-\alpha)}(r)^{1/(\alpha-1)}\mathbb{E}\big[(A_2)^{\frac{1}{1-\alpha}}\big]$$

$$\frac{\partial \bar{Y}_2^{nd}}{\partial r} = \left(\frac{\alpha}{\alpha-1}\right)(\alpha)^{\alpha/(1-\alpha)}(r)^{1/(\alpha-1)}\big[\mathbb{E}(A_2)\big]^{\frac{1}{1-\alpha}}$$

Thus, we have:

$$\frac{\partial \bar{Y}_2^d}{\partial r} < \frac{\partial \bar{Y}_2^{nd}}{\partial r} \iff \left(\frac{\alpha}{\alpha-1}\right)(\alpha)^{\alpha/(1-\alpha)}\mathbb{E}\big[(A_2)^{\frac{1}{1-\alpha}}\big] < \left(\frac{\alpha}{\alpha-1}\right)(\alpha)^{\alpha/(1-\alpha)}\big[\mathbb{E}(A_2)\big]^{\frac{1}{1-\alpha}}$$

$$\iff$$

$$\mathbb{E}\big[(A_2)^{\frac{1}{1-\alpha}}\big] > \big[\mathbb{E}(A_2)\big]^{\frac{1}{1-\alpha}}$$

The last inequality holds by Jensen's inequality and since $\frac{1}{1-\alpha} > 1$. Now we show that the relative effects are the same. To see this, note that:

$$\frac{\frac{\partial \bar{Y}_2^d}{\partial r}}{\bar{Y}_2^d} = \frac{\left(\frac{\alpha}{\alpha-1}\right)(\alpha)^{\alpha/(1-\alpha)}(r)^{1/(\alpha-1)}\mathbb{E}\big[(A_2)^{\frac{1}{1-\alpha}}\big]}{(\alpha)^{\alpha/(1-\alpha)}(r)^{\alpha/(\alpha-1)}\mathbb{E}\big[(A_2)^{\frac{1}{1-\alpha}}\big]} = \frac{\alpha}{\alpha-1}r^{-1}$$

$$\frac{\frac{\partial \bar{Y}_2^{nd}}{\partial r}}{\bar{Y}_2^{nd}} = \frac{\left(\frac{\alpha}{\alpha-1}\right)(\alpha)^{\alpha/(1-\alpha)}(r)^{1/(\alpha-1)}\big[\mathbb{E}(A_2)\big]^{\frac{1}{1-\alpha}}}{(\alpha)^{\alpha/(1-\alpha)}(r)^{\alpha/(\alpha-1)}\big[\mathbb{E}(A_2)\big]^{\frac{1}{1-\alpha}}} = \frac{\alpha}{\alpha-1}r^{-1}$$

A.10 Proof of proposition 8

Note that $A_2 = \bar{A} + \epsilon_2$. Thus, the derivatives of output with respect to $\bar{A}$ in the two economies are:

$$\frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} = (\alpha)^{\frac{\alpha}{1-\alpha}}(r)^{\frac{\alpha}{\alpha-1}}\left(\frac{1}{1-\alpha}\right)\big(\mathbb{E}[\bar{A}+\epsilon]\big)^{\frac{\alpha}{1-\alpha}}$$

$$\frac{\partial \bar{Y}_2^d}{\partial \bar{A}} = (\alpha)^{\frac{\alpha}{1-\alpha}}(r)^{\frac{\alpha}{\alpha-1}}\left(\frac{1}{1-\alpha}\right)\mathbb{E}\big[(\bar{A}+\epsilon)^{\frac{\alpha}{1-\alpha}}\big]$$

Note first that:

$$\frac{\alpha}{1-\alpha} \leq 1 \iff \alpha \leq 1-\alpha \iff \alpha \leq 1/2$$

We have that:

$$\frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} \geq \frac{\partial \bar{Y}_2^{d}}{\partial \bar{A}} \iff \left(\mathbb{E}[\bar{A}+\epsilon]\right)^{\frac{\alpha}{1-\alpha}} \geq \mathbb{E}\left[\left(\bar{A}+\epsilon\right)^{\frac{\alpha}{1-\alpha}}\right] \iff \frac{\alpha}{1-\alpha} \geq 1 \iff \alpha \geq 1/2$$

This inequality is reversed when $\alpha > 1/2$. This proves the result with respect to the absolute effects.

Now we move to relative effects. When $\alpha \leq 1/2$, the relative effects of business cycles will be stronger in the economy without data. This is because we have:

$$\frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} \geq \frac{\partial \bar{Y}_2^{d}}{\partial \bar{A}}$$

In general, we have that $\bar{Y}_2^{nd} < \bar{Y}_2^{d}$. This implies that:

$$\frac{1}{\bar{Y}_2^{nd}} > \frac{1}{\bar{Y}_2^{d}}$$

This, in turn, implies the desired result regarding relative effects:

$$\frac{\frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}}}{\bar{Y}_2^{nd}} \geq \frac{\frac{\partial \bar{Y}_2^{d}}{\partial \bar{A}}}{\bar{Y}_2^{d}}$$

This directly proves that the relative effects are also weaker in the economy with data.

B Numerical Results

In figure 7, we show that the analytical insights from Proposition 1 extend beyond the analytically tractable case of $\alpha = 1/2$. We visualize the effect of a contractionary monetary policy shock (during which the interest rate rises from $r' = 0.1$ to $r'' = 0.11$) on the optimal capital choice for different levels of σ (horizontal axis). We do this for varying κ_v while we set $\kappa_e = 0$.[14] For the different κ_v, we first plot the difference $K^*_{t+1}(r'') - K^*_{t+1}(r')$, which can be considered the rough equivalent of the absolute effects. Moreover, we plot the ratio $\frac{K^*_{t+1}(r'') - K^*_{t+1}(r')}{K^*_{t+1}(r')}$, which can be considered the rough equivalent of the relative effects.

We see from Figure 7 that our analytical results from Proposition 1 continue to hold. Firms with more data (lower σ) respond more strongly to monetary policy shocks, both in absolute and in relative terms. We further see that this relationship is concave in σ, especially for higher values of κ_v. This follows from the fact that with a higher κ_v, more

[14]We use the following parameters: $\rho = 0.5$, $\delta = 0.1$, $\alpha = 0.4$, $\bar{A} = 2$, and $\bar{V} = 1$.

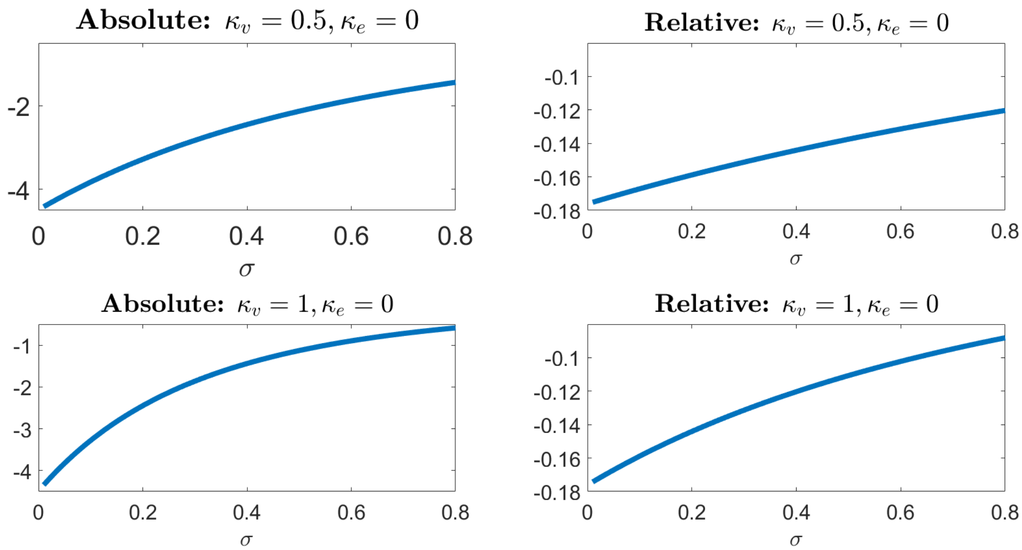

Figure 7: Effects of a contractionary MP shock

data reduces the firm's uncertainty more, which induces the firm to respond more forcefully to the interest rate change.

We now turn to the role of the data feedback loop and illustrate the effects of the data feedback effect on the response of investment to monetary policy graphically. Figure 8 shows the absolute and relative effects of a change in r_t from 0.1 to 0.11 on K_{t+1} for varying degrees of the data feedback effect, z (horizontal axis). In all cases, we set $\kappa_e = 0$ and consider different κ_v.[15]

[15] In all cases, we keep $\bar{A} = 2$, $\bar{V} = 1$, $\alpha = 0.3$ and $\rho = 0.1$, $\delta = 0.2$ and $\bar{\sigma} = 0.5$.

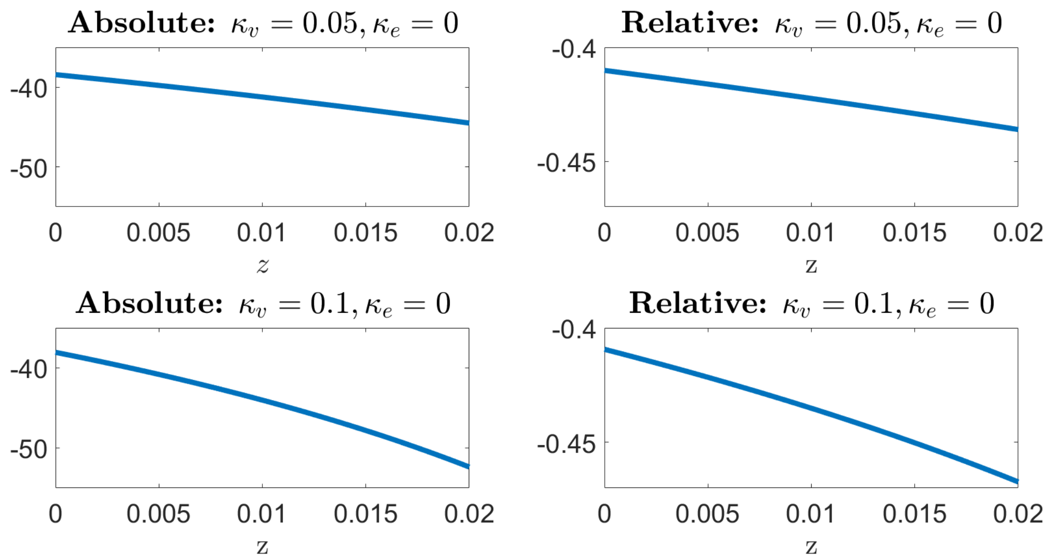

Figure 8: The data feedback effect & the effects of a MP shock ($\kappa_e = 0$)

We see that the results that we derived analytically in Proposition 5 continue to hold. A stronger data feedback effect, i.e., a higher z, amplifies the investment effects of monetary policy. This is especially the case at higher values of κ_v (see bottom panel in Figure 8). When κ_v is higher, more data reduces the firm's uncertainty more effectively, and hence, the effects are further amplified.

Figure 9 turns to the role of κ_e. We set $\kappa_v = 0$ and vary κ_e. The rest of the calibration is the same as in Figure 8. Again, and consistent with Proposition 5, the data feedback effect amplifies the investment response to a monetary policy shock. We further see that increasing κ_e has similar effects than increasing κ_v. A higher κ_e implies that data increases the firm's expected productivity more strongly, and hence, the overall effects increase.

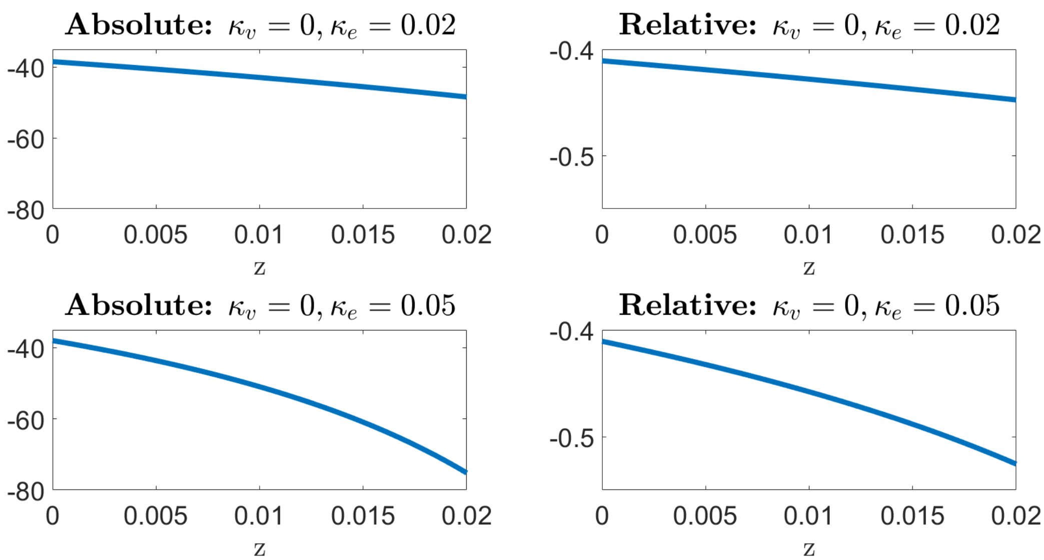

Figure 9: The data feedback effect & the effects of a MP shock ($\kappa_v = 0$)

C Omitted results

C.1 Section 5: Firm bankruptcy

We have considered the following timing: In the initial period, firms choose their capital stock in period t. At the beginning of period $t+1$, their productivity A_{t+1} realizes. All firms whose profits in $t+1$, given their capital stock and the realized productivity, are below some fixed cost, exit the market. Then, production realizes. Thus, a firm will be active in period $t+1$ if and only if the combination of productivity and capital stock satisfies:

$$A_t(K_{t+1})^\alpha - r_t K_{t+1} + \tau \geq 0$$

A firm that sets K_{t+1} will go bankrupt if and only if:

$$A_1(K_{t+1})^\alpha - rK_{t+1} + \tau < 0 \iff A_1 < \tilde{A}_1 := \frac{rK_{t+1} - \tau}{(K_{t+1})^\alpha}$$

Optimization problem in the economy without data:

Consider a firm in the economy without data. This firm maximizes the following:

$$\Pi^{nd}(K_{t+1}) = \int_{A^{min}}^{A^{max}} \max\{A_{t+1}(K_{t+1})^\alpha - rK_{t+1} + \tau, 0\}\big(1/(A^{max} - A^{min})\big)dA_{t+1} =$$

Setting $\tau = 0$, this profit function becomes:

$$\Pi^{nd}(K_{t+1}) = \int_{r(K_{t+1})^{1-\alpha}}^{A^{max}} \big(A_{t+1}(K_{t+1})^{\alpha} - rK_{t+1}\big)\big(1/(A^{max} - A^{min})\big)dA_{t+1} =$$

$$\Big[0.5(A_{t+1})^2(K_{t+1})^{\alpha} - A_{t+1}rK_{t+1}\Big]_{r(K_{t+1})^{1-\alpha}}^{A^{max}}$$

$$=$$

$$\Big[0.5(A^{max})^2(K_{t+1})^{\alpha} - A^{max}rK_{t+1}\Big] - \Big[0.5r^2(K_{t+1})^{2-2\alpha}(K_{t+1})^{\alpha} - r^2(K_{t+1})^{2-\alpha}\Big]$$

$$\implies \Pi^{nd}(K_{t+1}) = \Big[0.5(A^{max})^2(K_{t+1})^{\alpha} - A^{max}rK_{t+1}\Big] + 0.5r^2(K_{t+1})^{2-\alpha}$$

Here, one can see that the introduction of firm bankruptcy introduces a term that is convex in capital into the firm's objective function.

Now let's consider aggregate output in the economy without data. Define $\bar{K}^{nd}$ as the capital stock chosen by all firms in the economy without data. Given this, the aggregate output in this economy is given by:

$$\bar{Y}^{nd} = \int_{A^{min}}^{A^{max}} \mathbb{1}[A_{t+1}(\bar{K}^{nd})^{\alpha} - r\bar{K}^{nd} + \tau > 0]\big(A_{t+1}(\bar{K}^{nd})^{\alpha}\big)\big(1/(A^{max} - A^{min})\big)dA_{t+1}$$

<u>Optimization problem in the economy with data:</u>

Consider a firm in the economy without data. This firm maximizes the following:

$$\Pi^{nd}(K_{t+1}; A_{t+1}) = \max\{A_{t+1}(K_{t+1})^{\alpha} - rK_{t+1} - \tau, 0\}$$

The optimal choice for this firm is to set the statically optimal capital stock is this attains positive profits and to exit the market and produce nothing if the statically optimal capital stock yields negative profits. Define this statically optimal capital stock as $K^d_{t+1}(A_{t+1})$.

Given this, the aggregate output in the economy with data is given by:

$$\bar{Y}^{d} = \int_{A^{min}}^{A^{max}} \mathbb{1}[A_{t+1}(K^d_{t+1}(A_{t+1}))^{\alpha} - rK^d_{t+1}(A_{t+1}) + \tau > 0]\big(A_{t+1}(K^d_{t+1}(A_{t+1}))^{\alpha}\big)dF(A_{t+1})$$

References

D. Acemoglu, A. Makhdoumi, A. Malekian, and A. Ozdaglar. Too much data: Prices and inefficiencies in data markets. American Economic Journal: Microeconomics, 14(4):218–256, 2022.

N. Alemany, V. Aragó, and E. Salvador. The time-varying risk–return trade-off and its explanatory and predictive factors. The North American Journal of Economics and Finance, 68:101953, 2023.

M. Ansari. Essays in Macroeconomics and Firm Dynamics. PhD thesis, University of Minnesota, 2023.

K. Arvai and K. Mann. Consumption inequality in the digital age. 2022.

R. Bachmann, S. Elstner, and E. R. Sims. Uncertainty and economic activity: Evidence from business survey data. American Economic Journal: Macroeconomics, 5(2):217–249, 2013.

J. Begenau, M. Farboodi, and L. Veldkamp. Big data in finance and the growth of large firms. Journal of Monetary Economics, 97:71–87, 2018.

J. Benhabib, X. Liu, and P. Wang. Financial markets, the real economy, and self-fulfilling uncertainties. The Journal of Finance, 74(3):1503–1557, 2019.

D. Bergemann and A. Bonatti. Data, competition, and digital platforms. 2022.

B. S. Bernanke, M. Gertler, and S. Gilchrist. The financial accelerator in a quantitative business cycle framework. Handbook of macroeconomics, 1:1341–1393, 1999.

N. Bloom. The impact of uncertainty shocks. econometrica, 77(3):623–685, 2009.

N. Bloom, M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry. Really uncertain business cycles. Econometrica, 86(3):1031–1065, 2018.

E. Brynjolfsson and K. McElheran. Data in action: Data-driven decision making in us manufacturing. US Census Bureau Center for Economic Studies Paper No. CES-WP-16-06, Rotman School of Management Working Paper, (2722502), 2016a.

E. Brynjolfsson and K. McElheran. The rapid adoption of data-driven decision-making. American Economic Review, 106(5):133–139, 2016b.

E. Brynjolfsson, W. Jin, and X. Wang. Information technology, firm size, and industrial concentration. Technical report, National Bureau of Economic Research, 2023.

A. Caggese and A. Pérez-Orive. How stimulative are low real interest rates for intangible capital? European Economic Review, 142:103987, 2022.

J. Cao, T. Hegna, M. B. Holm, R. Juelsrud, T. König, and M. Riiser. The investment channel of monetary policy: Evidence from norway. 2023.

E. Castelnuovo and G. Pellegrino. Uncertainty-dependent effects of monetary policy shocks: A new-keynesian interpretation. Journal of Economic Dynamics and Control, 93:277–296, 2018.

B. Charoenwong, Y. Kimura, A. Kwan, and E. Tan. Capital budgeting, uncertainty, and misallocation. Uncertainty, and Misallocation (April 2022), 2022.

A. Chiavari and S. Goraya. The rise of intangible capital and the macroeconomic implications. Technical report, Working paper, 2022.

J. Cloyne, C. Ferreira, M. Froemel, and P. Surico. Monetary policy, corporate finance and investment. Journal of the European Economic Association, page jvad009, 2023.

V. Corradi, W. Distaso, and A. Mele. Macroeconomic determinants of stock volatility and volatility premiums. Journal of Monetary Economics, 60(2):203–220, 2013.

M. De Ridder. Market power and innovation in the intangible economy. 2019.

R. Döttling and L. Ratnovski. Monetary policy and intangible investment. Journal of Monetary Economics, 2022.

E. Durante, A. Ferrando, and P. Vermeulen. Monetary policy, investment and firm heterogeneity. European Economic Review, 148:104251, 2022.

J. Eeckhout and L. Veldkamp. Data and market power. Technical report, National Bureau of Economic Research, 2022.

P. D. Fajgelbaum, E. Schaal, and M. Taschereau-Dumouchel. Uncertainty traps. The Quarterly Journal of Economics, 132(4):1641–1692, 2017.

M. Farboodi and L. Veldkamp. Data and markets. Annual Review of Economics, 15, 2022a.

M. Farboodi and L. Veldkamp. A model of the data economy. 2022b.

J. E. Galdon-Sanchez, R. Gil, and G. Uriz-Uharte. The value of information in competitive markets: The impact of big data on small and medium enterprises. 2022.

M. Gertler and S. Gilchrist. Monetary policy, business cycles, and the behavior of small manufacturing firms. The Quarterly Journal of Economics, 109(2):309–340, 1994.

C. Glocker and P. Piribauer. Digitalization, retail trade and monetary policy. Journal of International Money and Finance, 112:102340, 2021.

N. Gondhi. Rational inattention, misallocation, and the aggregate economy. Journal of Monetary Economics, 136:50–75, 2023.

Y. Gorodnichenko and O. Talavera. Price setting in online markets: Basic facts, international comparisons, and cross-border integration. American Economic Review, 107(1):249–282, 2017.

Y. Gorodnichenko, V. Sheremirov, and O. Talavera. Price setting in online markets: Does it click? Journal of the European Economic Association, 16(6):1764–1811, 2018.

P. Jeenas. Firm balance sheet liquidity, monetary policy shocks, and investment dynamics. 2019.

T. Kroen, E. Liu, A. R. Mian, and A. Sufi. Falling rates and rising superstars. 2021.

S. Kumar, Y. Gorodnichenko, and O. Coibion. The effect of macroeconomic uncertainty on firm decisions. Technical report, National Bureau of Economic Research, 2022.

A. Lakdawala and T. Moreland. Firm-level uncertainty and the transmission of monetary policy. Available at SSRN 4255980, 2022.

D. Lashkari, A. Bauer, and J. Boussard. Information technology and returns to scale. Available at SSRN 3458604, 2018.

B. Maćkowiak and M. Wiederholt. Business cycle dynamics under rational inattention. The Review of Economic Studies, 82(4):1502–1532, 2015.

B. Maćkowiak and M. Wiederholt. Rational inattention and the business cycle effects of productivity and news shocks. 2023.

M. Meier and T. Reinelt. Monetary policy, markup dispersion, and aggregate tfp. Review of Economics and Statistics, pages 1–45, 2022.

A. Mukerji. Essays in Technological Innovation & Financial Economics. Stanford University, 2022.

G. Ordonez. The asymmetric effects of financial frictions. Journal of Political Economy, 121 (5):844–895, 2013.

P. Ottonello and T. Winberry. Financial heterogeneity and the investment channel of monetary policy. Econometrica, 88(6):2473–2502, 2020.

F. Paine. Big data and firm risk. 2022.

G. Pellegrino. Uncertainty and monetary policy in the us: A journey into nonlinear territory. Economic Inquiry, 59(3):1106–1128, 2021.

R. H. Peters and L. A. Taylor. Intangible capital and the investment-q relation. Journal of Financial Economics, 123(2):251–272, 2017.

J. Powell. Opening remarks by chair powell at the new york innovation center event. 2021.

C. A. Sims. Implications of rational inattention. Journal of monetary Economics, 50(3): 665–690, 2003.

L. Veldkamp. Valuing data as an asset. Review of Finance, 2023.

L. Veldkamp and C. Chung. Data and the aggregate economy. Journal of Economic Literature, 2019.

L. L. Veldkamp. Slow boom, sudden crash. Journal of Economic theory, 124(2):230–257, 2005.

J. Wang, Y. Dou, and L. Huang. Regulating the collection of data as a factor of production: An economic analysis. Available at SSRN 4215107, 2022.

M. Wu and L. Zhang. Endogenous growth and human capital accumulation in a data economy. Available at SSRN 4248918, 2022.

X. Wu. Mobile app, firm risk, and growth. 2023.

D. Xie and L. Zhang. Endogenous growth with data generated during production. Available at SSRN 4033576, 2022.

N. Zolas, Z. Kroff, E. Brynjolfsson, K. McElheran, D. N. Beede, C. Buffington, N. Goldschlag, L. Foster, and E. Dinlersoz. Advanced technologies adoption and use by us firms: Evidence from the annual business survey. Technical report, National Bureau of Economic Research, 2021.

P. Zorn. Investment under rational inattention: Evidence from us sectoral data. 2020.