Data, Business Cycles, and the Cyclical Effectiveness of Monetary Policy*

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

December 23, 2023

Abstract

We study how the availability of data in modern economies shapes the propagation of cyclical fluctuations and the effectiveness of monetary policy along the business cycle. We consider a tractable heterogeneous firms framework in which data enters investment decisions by favorably affecting a firm's productivity distribution and by allowing firms to predict their future productivity realizations. Data accumulates endogenously through a data feedback loop, i.e., firms that produce more accumulate more data. We show that increased availability of data dampens cyclical fluctuations if and only if the data feedback loop is sufficiently weak, that is, when firms accumulate only little data through production. This is because data-rich firms respond less strongly to aggregate productivity shocks when the data feedback loop is weak, while the converse holds true if the data feedback loop is strong. Given that data-rich firms respond more strongly to monetary policy, this result also makes the effectiveness of monetary policy countercyclical if the data feedback loop is weak and vice versa. Moreover, the data feedback loop weakens the negative relationship between a firm's risk sensitivity and its size, which amplifies the effects of aggregate uncertainty shocks. Our work also sheds light on the macroeconomic effects of digital markets regulation such as the EU GDPR.

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

*We would like to thank Mridula Duggal and Luigi Paciello for insightful comments. Carl-Christian Groh gratefully acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224.

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

1 Introduction

Modern economies increasingly revolve around data, i.e., digitized information which firms use to predict payoff-relevant states. Data is accumulating rapidly and more and more firms are adopting data-driven decision-making (Galdon-Sanchez et al., 2022; Brynjolfsson et al., 2023). Data yields value in many ways, for example by allowing firms to forecast future sales (Bajari et al., 2019) and the success of individual products (Munos et al., 2021), by predicting costs (Ajit, 2016), and by enabling targeted advertising (Lau, 2020). Given the various utilizations of data, the increasing availability thereof inevitably affects the way in which firms make decisions and respond to changes in their environment. Motivated by these facts, we study how data affects two central topics in macroeconomics, namely the magnitude of cyclical fluctuations and the effectiveness of monetary policy along the business cycle.

To do so, we consider a tractable model of firm investment that captures key features of the data economy: First, we integrate the empirically documented facts that firms with superior access to data have a higher expected productivity (Bajari et al., 2019; Corrado et al., 2022) and a lower variance of productivity (Paine, 2022; Mukerji, 2022; Wu, 2023). Having a lower variance of productivity is beneficial for a firm because this lowers its cost of capital. Second, having access to superior data not only favorably affects a firm's productivity distribution, but also enables the firm to predict its productivity realizations with greater precision. Formally, firms with access to more data receive more precise signals about their future productivity realizations. Third, we integrate a data feedback loop as in Farboodi and Veldkamp (2022a): Firms that produce more accumulate more data. This is a cornerstone of the data accumulation process and reflects the importance of smart devices and the algorithmic analysis of click-through rates in the generation of data. The combination of these features distinguishes data from other economic resources, which necessitates a novel approach to understand the macroeconomic impacts of data.

We show that when access to data is exogenous—i.e., when the data feedback loop is turned off—firms with superior access to data respond more strongly to monetary policy shocks, but less strongly to aggregate productivity shocks. Data therefore tends to dampen cyclical fluctuations that are driven by aggregate productivity shocks. That data-rich firms respond less strongly to aggregate productivity shocks follows from the fact that the expected productivity of these firms is higher. Aggregate productivity shocks increase the expected productivity of all firms by the same absolute amount, which translates into smaller percentage increases in the expected productivity of data-rich firms, thereby triggering a smaller response of these firms. In contrast, data-rich firms respond more strongly to monetary policy shocks. This is because data-rich firms face lower idiosyncratic risk, which implies that they have a comparatively low cost of capital since a firm's cost of capital is increasing in the idiosyncratic uncertainty it faces.[1] Thus, any changes in the aggregate component of a firm's cost of capital, which is directly affected by monetary policy, will impact a firm's total cost of capital more strongly (in percentage terms) if the firm has access to better data. Therefore, data-rich firms respond more strongly to monetary policy shocks.

Given that firms in the economy have vastly different access to data (Brynjolfsson and McElheran, 2016; Zolas et al., 2021), these findings suggest that data renders the effectiveness of monetary policy countercylical: Because firms with access to superior data respond less strongly to movements in aggregate productivity, they attain comparatively high market shares in recessions. Consequently, the effectiveness of monetary policy becomes countercyclical because data-rich firms respond more strongly to these stimuli.

The result that data-rich firms respond less to aggregate productivity shocks holds true even if data does not favorably affect a firm's productivity distribution, but only grants a firm signals about its future idiosyncratic productivity realizations. Intuitively, this holds by the following logic: Endowing a firm with a signal about its future productivity implies that a given aggregate productivity shock affects the information set of the firm to a lesser extent, thereby inducing a smaller investment response. This effect is particularly pronounced when aggregate productivity is low: In particular, raising the share of firms with access to this type of data mitigates recessions without dampening aggregate output in booms.

Once we account for the data feedback loop, however, these results may flip. In particular, firms with exogenously superior access to data respond more strongly to aggregate productivity shocks when the data feedback loop is sufficiently strong. This holds by the following logic: An increase in aggregate productivity raises a firm's investment and production, which now—through the data feedback loop—improves that firm's access to data, inducing a further increase in its capital investment. This effect is particularly strong for firms with access to superior data since these firms are larger. This means that changes in their environment trigger a large response (in absolute terms) of their investment, and thus, a relatively large improvement in their access to data. Hence, data-rich firms respond more strongly to aggregate productivity shocks if the data feedback loop is sufficiently strong.

These insights imply that the presence of a strong data feedback loop can reverse the previous predictions regarding the transmission of aggregate productivity shocks and the effectiveness of monetary policy along the business cycle. First, the presence of a data feedback loop as such amplifies the effects of shocks by creating a self-reinforcing relationship between output and the marginal benefits of investment—this idea is reminiscent of Fajgelbaum et al.

[1]This is based on the risk-return relationship at the heart of models in finance (Merton, 1973).

(2017). Second, when there is a strong data feedback loop, increases in the overall availability of data amplify the effects of aggregate productivity shocks by creating a positive relationship between a firm's access to data and its responsiveness to aggregate productivity shocks. Third, the aforementioned composition effects now have an opposite sign: Firms with superior access to data attain lower market shares in recessions driven by negative productivity shocks, which makes the effectiveness of monetary policy more procyclical.[2]

The presence of a data feedback loop not only matters for the heterogeneous responses to shocks, but also for differences in firms' size. Firms with exogenously superior access to data are larger even absent the data feedback loop. A strong data feedback loop intensifies this relationship: the differences in firm size between firms with exogenously better access to data to those with less data become larger in the presence of a strong data feedback loop. One could expect that an endogenous accumulation process for data might *level the playing field* with respect to the competitive advantages granted by exogenous data advantages. The opposite holds true, which is based on the following logic: In general, firms with superior access to data are larger. When the data feedback loop is active, their larger size magnifies the data advantages of such firms, thereby further increasing their size.

The data feedback loop weakens the negative relationship between a firm's sensitivity to risk 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 sensitivity and its size becomes positive when the data feedback effect becomes large enough. When there is cross-sectional heterogeneity in firm's risk sensitivity, the presence of a data feedback loop thus raises the market shares of firms with high levels of risk sensitivity. Because such firms respond strongly to aggregate uncertainty increases, 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).

Our insights also establish how digital markets regulation affects macroeconomic outcomes. The design of existing regulatory work in this area such as the EU GDPR, the DMA, and the UK DPA focused on consumer protection, privacy, and contestability. However, given the scope of digital markets regulation and the economic importance of these markets, it is clear that such regulation affects the entire economy. Our paper formalizes this notion and offers a conceptual framework to evaluate the macroeconomic impacts of digital markets regulation, given that different types of regulation in this area can be viewed as changes in

[2]Notably, this working channel makes the effectiveness of monetary policy procyclical even if uncertainty remains constant and the capital costs remain the same for all firms.

the strength of the data feedback loop or the exogenous availability of data to firms.

Related literature: To the best of our knowledge, ours is the first paper to study how data (with the properties discussed above) shapes cyclical fluctuations and the effectiveness of monetary policy along the business cycle. Our modelling framework is unique (up to its companion paper, namely Groh and Pfäuti (2023)) in the sense that it jointly (i) includes a data feedback loop as introduced in Farboodi and Veldkamp (2022a), (ii) captures that access to superior data raises a firm's expected productivity and reduces its cost of capital as in Eeckhout and Veldkamp (2022), and (iii) grants firms signals about their idiosyncratic productivity realizations as featured in Maćkowiak and Wiederholt (2009). In the following, we highlight how our work is related to and builds on various strands of the literature.

First, our work contributes to the rapidly growing literature on the relevance of digitization and data for macroeconomic outcomes.[3] Veldkamp and Chung (2019) provide an overview of the role of data in the economy. Eeckhout and Veldkamp (2022) show that data can be a source of market power if firms price risk. The data feedback effect we incorporate builds on the work of Farboodi and Veldkamp (2022b), who integrate this channel into a growth model.[4] 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. [5] Our key innovation relative to these papers is that we jointly incorporate the relevant features of the data economy in a macroeconomic model. Moreover, we study monetary policy and cyclical fluctuations.

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.[6] The key distinction between our paper and this literature is that data is fundamentally different from intangible assets and R&D. None of the features of data that we model can be found in the

[3]Several papers have proposed ways of estimating firms' stock of data. Examples of these are Begenau et al. (2018), Lashkari et al. (2018), Calderón and Rassier (2022), Mukerji (2022), Galdon-Sanchez et al. (2022), Corrado et al. (2022), Arvai and Mann (2022), Quan (2022), Babina et al. (2022), Demirer et al. (2022), Veldkamp (2023), Brynjolfsson et al. (2023), and Wu (2023).

[4]Wang et al. (2022), Xie and Zhang (2022), Wu and Zhang (2022), He et al. (2023), Ansari (2023), and wGomes et al. (2023) build on Farboodi and Veldkamp (2022b) and study the role of data in growth models.

[5]Glocker and Piribauer (2021) empirically document that, because prices are more easily adjustable in digital markets (Gorodnichenko and Talavera, 2017; Gorodnichenko et al., 2018), increases in the amount of sales that are conducted through digital retail reduce the real effects of monetary policy.

[6]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.

theoretical models studying the role of intangible assets. They key features of data that we model are not discussed in the work on intangible assets.

Third, 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.[7] Bloom et al. (2018) establish that firms which face higher uncertainty are less responsive to shocks such as monetary policy stimuli. This insight is 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 considering firms which are heterogenous not only in their idiosyncratic uncertainty, but also in their expected productivity and their ability to predict future outcomes. In Veldkamp (2005), Van Nieuwerburgh and Veldkamp (2006), Ordonez (2013), and Fajgelbaum et al. (2017), there is a data feedback loop at the aggregate level, which amplifies business cycles by creating countercylical movements in aggregate uncertainty. Going beyond this, we study how a firm-level data feedback loop shapes the propagation of cyclical fluctuations.

Fourth, our work is related to the literature on rational inattention (pioneered by Sims (2003)). However, there are substantial differences in focus and setup: Generally speaking, papers in this literature establish how agents optimally allocate their limited attention and how this can account for inertia in macroeconomic outcomes. By contrast, we study how exogenous heterogeneity in firms' access to data and the data feedback loop shape cyclical fluctuations and the effectiveness of monetary policy along the business cycle. Moreover, most papers in the rational inattention literature consider models without capital.[8] In terms of setup, our analysis in section 4 is related to Charoenwong et al. (2022) and Gondhi (2023), who consider models in which firms receive signals about their idiosyncratic productivity draws. However, these papers do not consider the effects of monetary policy, changes in the first moment of a firm's productivity distribution, or the data feedback loop.

Outline: The paper proceeds as follows: Section 2 introduces our theoretical framework. We study how data affects outcomes by favorably affecting a firm's productivity distribution and enabling it to predict its future productivities in sections 3 and 4, respectively. We discuss the implications for digital markets regulation in section 5 and conclude thereafter.

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

[8] Exceptions are Maćkowiak and Wiederholt (2015), Zorn (2020), Gondhi (2023), and Maćkowiak and Wiederholt (2023). Benhabib et al. (2019) study how financial markets interpret signals of inattentive firms.

2 Framework

In this section, we present our theoretical model of firm investment and data. The model is kept deliberately stylized in certain parts to put the spotlight on various new features of the data economy.

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.[9] Capital is inelastically supplied.

In the aim of incorporating the manifold economic benefits generated by superior access to data, we suppose that superior access to data generates value for firms in three different ways, namely by (1) increasing the firm's expected productivity, (2) by reducing the variance of it's productivity, and (3) by granting firms more precise signals about their future productivity realizations. The first two channels thus imply that data favorably affects a firm's productivity *distribution*, while the third channel captures the idea that data allows firms to better *predict* their future productivity.

Throughout the analysis, the quality of a firm's access to data is captured by the objects $(\sigma_{i,t}, \xi_{i,t})$. The object $\sigma_{i,t}$ governs the relationship between a firm's access to data and it's productivity distribution. This relationship takes the following functional form:

$$\mathbb{E}[A_{i,t}|\sigma_{i,t}] = \bar{A}_t - \kappa_e \sigma_{i,t} \quad ; \quad VAR[A_{i,t}|\sigma_{i,t}] = \bar{V}_t + \kappa_v \sigma_{i,t} \tag{2}$$

The parameters $\kappa_e \geq 0$ and $\kappa_v \geq 0$ capture the effects of data $\sigma_{i,t}$ on the expected productivity, $\mathbb{E}[A_{i,t}|\sigma_{i,t}]$, and its variance, $VAR[A_{i,t}|\sigma_{i,t}]$, respectively. Firms with a lower $\sigma_{i,t}$ have access to superior data. When κ_v is high, relative to κ_e, the primary economic value of access to better data is the associated reduction in the variance of the firm's productivity.

We model the third channel by specifying that any firm also receives a signal about its

[9] 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.

productivity: $\hat{A}_{i,t} = A_{i,t} + e_{i,t}$, with $e_{i,t} \sim N(0, \xi_{i,t}^2)$. Reductions of $\xi_{i,t}$ or $\sigma_{i,t}$ can thus be understood as increases in the quality of data a firm has access to.

To ease exposition, we separate the analysis of the different channels. In Section 3, we focus on the effect of data via favorably affecting a firm's productivity distribution (i.e. we consider $\sigma_{i,t} \in \mathbb{R}_{\geq 0}$ and set $\xi_{i,t} \to \infty$). In section 4, we focus on the effect of data by allowing firms to predict their future productivities more accurately (i.e. we consider $\xi_{i,t} \in \mathbb{R}_{\geq 0}$).

<u>Microfoundation:</u> While we focus on the implications of data access on firm investment generally, we now provide a potential microfoundation for our assumptions on how data shapes a firm's future productivity and its expectations thereof using a particular example. Based on Farboodi and Veldkamp (2022b), 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 $\theta_{i,t}$. 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:

$$A_{i,t} = \bar{A}_t - d(a_{i,t}, \theta_{i,t}) + \epsilon_{i,t}, \tag{3}$$

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.

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 $\sigma_{i,t}^2$ is the variance of the signal a firm obtains about $\theta_{i,t}$ and $\xi_{,i,t}^2$ is the variance of the signal a firm obtains about $\epsilon_{i,t}$. 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, i.e., a higher precision.

Within this example, decreases in $\sigma_{i,t}$ (i.e. when the firm gets a more precise signal about $\theta_{i,t}$) allow a firm to more precisely match $\theta_{i,t}$. This decreases the expected value of $d(a_{i,t}, \theta_{i,t})$ and makes high values thereof less likely. This means that decreases of $\sigma_{i,t}$ raise the expected productivity and reduce the variance of productivity. By contrast, the precision of the signal about $\epsilon_{i,t}$ does not influence the underlying productivity distribution, but merely helps a firm forecast their future productivities. Notably, all firms with a given level of $\sigma_{i,t}$ will have the same productivity distribution, while the relevant productivity distribution of any firm which receives a signal about $\epsilon_{i,t}$ depends on the realization of the signal.

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. This effect is particularly pronounced in the data economy and stems from various sources.[10] Formally, we incorporate the data feedback loop by linking the signal quality of a firm to its capital stock, i.e. we suppose that $\sigma_{i,t} = \tilde{\sigma}(K_{i,t})$ and set:

$$\tilde{\sigma}(K_{i,t}) = \sigma_i - zK_{i,t}. \tag{4}$$

The parameter $z \geq 0$ governs the strength of the data feedback loop. In words, bigger firms, i.e., firms with a higher capital stock $K_{i,t}$, have access to more or better data. 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).

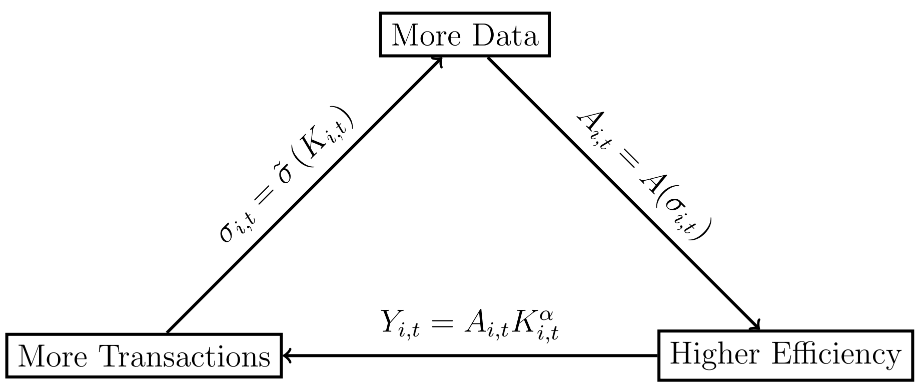


Figure 1: Data Feedback Loop

We suppose that the data feedback loop only influences the type of data which favorably affects a firm's productivity distribution $\sigma_{i,t}$, and not the precision of the signal about the productivity draws.

[10]Firstly, the creation of smart devices which send data back to their manufacturers creates a direct causal link between the output of a firm and the data it has access to. Second, an important mechanism through which data is created is through the algorithmic analysis of click-through rates online. Firms that produce more get more website visitors (which means the firm gets more data points) and can offer consumers more varieties (which means the firm can learn more about the tastes of any consumer who visits its website).

A firm's optimization problem

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

$$\max_{\{K_{i,t+1+j}\}_{j=0}^{\infty}} \quad \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j \Pi_{i,t+j}, \tag{5}$$

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

$$\Pi_{i,t+1} = A_{i,t+1}(K_{i,t+1})^{\alpha} - C_i(\sigma_{i,t})I_{i,t}, \tag{6}$$

where $I_{i,t} = K_{i,t+1} - (1-\delta)K_{i,t}$ is a firm's investment in period t and $C_i(\sigma_{i,t})$ is a firm-specific function that governs a firm's cost of investment. We specify that this is given by:

$$C_i(\sigma_{i,t}) = r_t + \rho_i VAR[A_{i,t}|\sigma_{i,t}] \tag{7}$$

We define r_t as the interest rate that is directly controlled by the monetary policy authority, which is the same for every firm. If $\rho_i > 0$, a firm's cost of capital is increasing in the uncertainty it faces. A natural microfoundation is the risk-return relationship which is at the heart of finance: There is a positive relationship between the riskiness and the return of an asset, for example due to default risk or because investors are risk averse. As a consequence, firms who face higher idiosyncratic uncertainty will have a to pay a higher cost in order to raise capital. Our results would hold analogously if firms have mean-variance preferences over output as in Eeckhout and Veldkamp (2022), and the cost of capital is fixed and controlled by the monetary policy authority.

Discussion of key model features. Our objective in this paper is to present a tractable model that allows us to study how data shapes the effects of various aggregate shocks on firm investment, namely (i) aggregate productivity shocks (changes in $\bar{A}_t$), (ii) aggregate uncertainty shocks (changes in $\bar{V}_t$), and (iii) monetary policy shocks (changes in r_t). 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 made several simplifying assumptions which we now discuss 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. For the purposes of our analysis, this is suitable, because our focus is on the the investment channel of monetary policy and how data shapes a firm's responsiveness to a change in it's cost of capital (which is induced by a monetary policy shock). 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 (partial equilibrium) channels (Cao et al., 2023). Modelling the rest of the economy and embedding our framework in a general equilibrium setup would endogenize the real interest rate r_t, but not affect how firms adjust their investment to a given change in r_t, which is what we are interested in. 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.

3 Data and the firm-level productivity distribution

In this section, we study how data affects a firm's optimization calculus by favorably affecting its productivity distribution. Formally, we thus consider $\sigma_{i,t} \in \mathbb{R}$ but shut down any other effects of data, i.e., we set $\xi_{i,t} \to \infty$. Furthermore, we impose the following assumptions throughout the analysis in this section:

Assumption 1 *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 guarantees that a firm's chosen level of capital is falling in the interest rate and rising in the expected productivity. The second assumption ensures that we can always compute firms' optimal behaviour using first order conditions.

When choosing $K_{i,t+1}$, the relevant part of the firm's objective function only contains the profits in period $t+1$ and period $t+2$. This is because profits that are further in the future and the future optimal capital choices do not depend on $K_{i,t+1}$. We define $K^*_{t+1}(\sigma_i, z, \rho_i)$ as the optimal capital choice of a firm in period t, which maximizes:

$$\mathbb{E}[A_{i,t+1}; \sigma_{i,t+1}] K^{\alpha}_{i,t+1} - \big(r_t + VAR[A_{i,t+1}; \sigma_{i,t+1}]\big)\big(K_{i,t+1} - (1-\delta)K_{i,t}\big) +$$

$$\beta\Big[\mathbb{E}[A_{i,t+2};\sigma_{i,t+2}]K_{i,t+2}^{\alpha}-\big(r_{t+1}+VAR[A_{i,t+2};\sigma_{i,t+2}]\big)\big(K_{i,t+2}-(1-\delta)K_{i,t+1}\big)\Big] \tag{8}$$

In the following, we investigate the impact of aggregate productivity shocks, uncertainty shocks, and monetary policy shocks on the optimal capital stock of firms. When evaluating which type of firms respond particularly strongly to these shocks, we work with elasticities. We define the elasticity of a firm's optimal capital choice with respect to an increase in aggregate productivity as $\varphi(\sigma_i, z, \rho_i)$. The elasticity of a firm's optimal capital choice with respect to an increase in the interest rate r_t is defined as $\gamma(\sigma_i, z, \rho_i)$. Note that:

$$\varphi(\cdot)\equiv\frac{\partial K_{t+1}^{*}}{\partial\bar{A}_{t+1}}\frac{\bar{A}_{t+1}}{K_{t+1}^{*}}>0 \quad ; \quad \gamma(\cdot)\equiv\frac{\partial K_{t+1}^{*}}{\partial r_t}\frac{r_t}{K_{t+1}^{*}}<0 \tag{9}$$

An increase in aggregate productivity will raise the optimal capital choice of any firm, and an increase in the interest rate will reduce the optimal capital choice of a firm. Thus, a firm responds more strongly to an aggregate productivity shock if $\varphi(\cdot)$ is higher. By contrast, a firm responds more strongly to an monetary policy shock if $\gamma(\cdot)$ is lower.

3.1 Benchmark results

In this subsection, we start by considering a data economy in which there is no data feedback loop and study how exogenous differences in access to data shape the responsiveness of firms to different types of shocks. In the next subsection, we will discuss how the presence of the data feedback loop affects these results. In the following, we always assume that the economy is in the steady state when the shock hits.

Proposition 1 *Suppose that the data feedback loop is inactive ($z = 0$). Then:*

- $\frac{\partial\gamma(\cdot)}{\partial\sigma_i} \geq 0$*, with strict inequality if $\kappa_v > 0$, i.e., firms with better data respond more strongly to monetary policy shocks.*
- $\frac{\partial\varphi(\cdot)}{\partial\sigma_i} \geq 0$*, with strict inequality if $\kappa_e > 0$, i.e., firms with better data respond less strongly to aggregate productivity shocks.*

If better data reduces the uncertainty a firm faces (i.e. $\kappa_v > 0$) firms with access to better data will respond more strongly to monetary policy shocks. 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 total cost of capital. As a result, changes in the interest rate affect data-rich firms to a greater extent. However, 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 $\mathbb{E}[A_{i,t+1}|\sigma_{i,t+1}]$, so any increase in the absolute responsiveness of a firm would be proportional to its size.

The second result in Proposition 1 says that firms with better data respond less strongly to aggregate productivity shocks. This is because firms with better access to data (smaller σ_i) have a higher expected productivity $\mathbb{E}_t[A_{i,t+1}; \sigma_i] = \bar{A}_{t+1} - \kappa_e \sigma_i$. Thus, any increase of $\bar{A}_{t+1}$ will trigger a smaller change of the expected productivity (in percentage terms) of firms with access to better data, which thus respond less strongly to changes in $\bar{A}_{t+1}$. In section 4, we show that this result holds true even if access to better data does not increase the expected productivity of firms, but only enables firms to predict their future productivity realizations.

These results directly imply that greater availability of data will dampen cyclical fluctuations caused by aggregate productivity shocks. Moreover, in an economy in which data is unequally distributed, firms with superior access to data will attain relatively higher market shares in recessions (that are driven by aggregate productivity declines). Because these firms respond strongly to monetary policy shocks, the effectiveness of monetary policy becomes countercyclical, ceteris paribus.

3.2 The role of the data feedback loop

From now on, 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) and establish the following insights: First, the presence of the data feedback loop strengthens the positive relationship between a firm's exogenous access to data (governed by the parameter σ_i) and its size. Second, the presence of the data feedback loop amplifies the effects of aggregate productivity shocks and monetary policy shocks by a logic similar to the one presented in Fajgelbaum et al. (2017). Third, while firms with access to better data respond less strongly to aggregate productivity shocks when the data feedback loop is inactive, the sign of this relationship flips if the data feedback loop becomes strong enough. Fourth, the presence of the data feedback loop weakens the negative relationship between a firm's risk sensitivity (measured by the parameter ρ_i) and its size.

The first issue we analyze is how the data feedback loop affects the relationship between a firm's exogenous data access (given by the parameter σ_i) and its size:

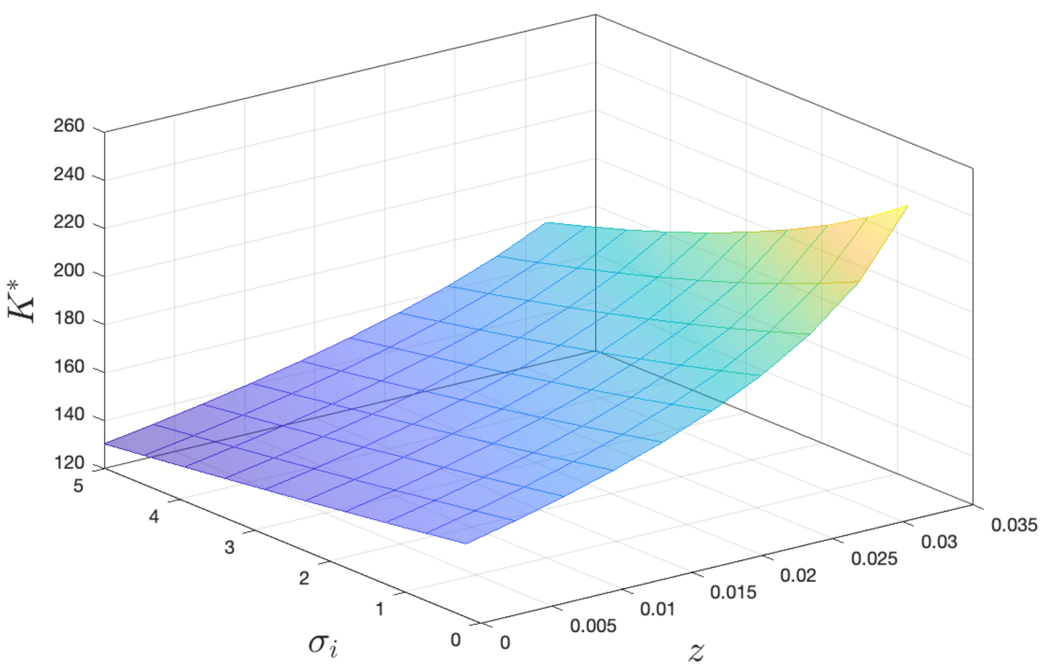
Proposition 2 *For any $z \geq 0$, firms with better access to data are larger, i.e.* $\frac{\partial K^*_{t+1}}{\partial \sigma_i} < 0$. *Moreover, the magnitude of this relationship increases in z, i.e.* $\frac{\partial^2 K^*_{t+1}}{\partial \sigma_i \partial z} < 0$ *holds, if:*

- $\kappa_e = 0$ *and* $\kappa_v > 0$ *holds true* *<u>or</u>*
- $\kappa_e > 0$ *and* $\kappa_v = 0$ *holds true.*

One might have expected that an endogenous accumulation process for data levels the playing field in the sense that the firms with exogenously better access to data hold less comparative advantages. The opposite holds true by the following intuition: In general, firms with access to better data (smaller σ_i) are larger. When the data feedback effect is active, this size difference grants such firms additional advantages in the data they can utilize, which further increases their size.

Numerical analysis suggests that these results hold true even when $\kappa_e > 0$ and $\kappa_v > 0$, but analyzing said relationship for such parameter constellations is analytically intractable.

Figure 2: Firm size, data and the data feedback loop



Note: This figure shows the optimal capital stock in steady state, K^*, for different values of exogenous access to data, σ_i, and different strengths of the data feedback loop, z. Calibration: $\alpha = 0.3$, $\kappa_e = 0.002$, $\kappa_\nu = 0.025$, $\bar{A} = 2$, $\rho_i = 0.5$, $\delta = 0.02$, $\bar{V} = 1$, $r = 0.1$, $\beta = 0.99$.

We now move on to establish how the data feedback loop shapes the effects of aggregate productivity shocks (modelled as increases in $\bar{A}_t$):

Lemma 1 *Increases in the strength of the data feedback loop amplify the relative effects of an aggregate productivity shock, i.e.* $\frac{\partial \varphi(\cdot)}{\partial z} > 0$.

This result is based on a logic that is similar to the one presented in Fajgelbaum et al. (2017), who consider a data feedback loop that runs on the aggregate level. In general, firms respond to an increase in $\bar{A}_{t+1}$ by increasing their capital input. When the data feedback loop is active, this further boosts their access to data, thereby raising their capital input even more.

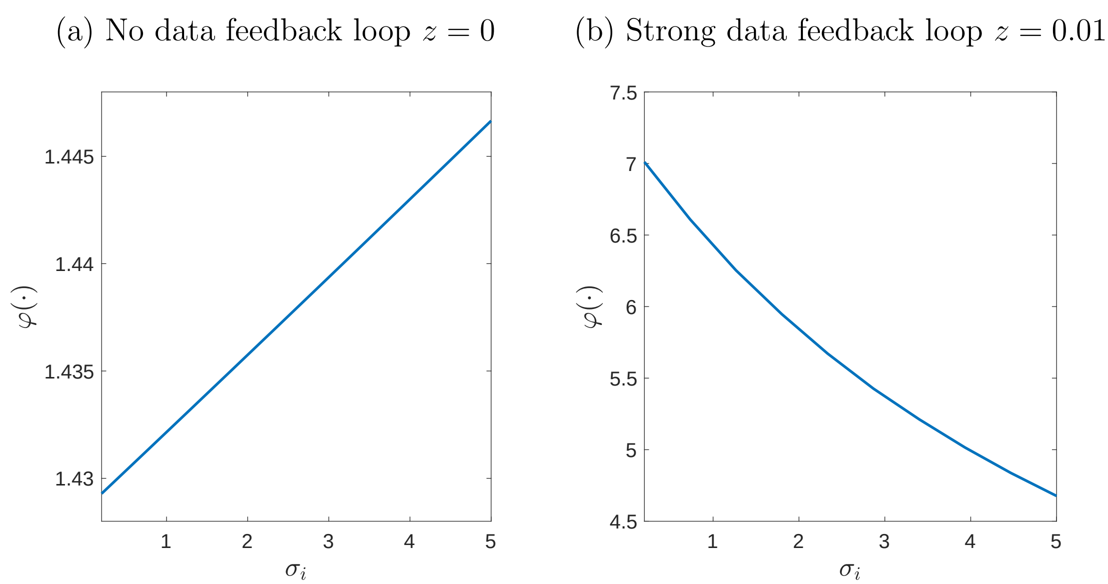
Moreover, increases in the strength of the data feedback loop also affect the relationship between a firm's exogenously given access to data (governed by the parameter σ_i) and it's responsiveness to an aggregate productivity shock:

Proposition 3 *If z is large enough, firms with better data respond more strongly to an aggregate productivity shock, i.e.:*

$$\frac{\partial \varphi(\cdot)}{\partial \sigma_i} < 0 \iff \left(\alpha(\alpha+1)\kappa_e z + 2\rho_i \kappa_v z(2-\alpha)(K^*_{t+1})^{1-\alpha}\right)\frac{\partial K^*_{t+1}}{\partial \sigma_i} < \kappa_e \alpha(\alpha-1) \tag{10}$$

To understand this result, recall the case in which $z = 0$. Then, firms with better data respond less strongly to aggregate productivity shocks. This is because firms with better access to data (smaller σ_i) have a higher expected productivity $\mathbb{E}[A_{i,t+1}] = \bar{A} - \kappa_e \sigma_i$. Thus, any increase of $\bar{A}$ will trigger a smaller change of the expected productivity (in percentage terms) of firms with access to better data, which thus respond less strongly to changes in $\bar{A}$.

Figure 3: Response to productivity shocks, data and the data feedback loop



Note: This figure shows the elasticity of capital to aggregate productivity $\bar{A}$, $\varphi(\cdot)$, for different values of data σ_i. The left panel shows this relationship for the case in which there is no data feedback loop, $z = 0$, and the right panel shows it for the case of an active data feedback loop, $z > 0$.

If the data feedback effect becomes sufficiently strong, the sign of this relationship flips:

Firms with better data will respond more strongly to aggregate productivity shocks. This is because increases in the strength of the data feedback loop (i.e. an increase of z) amplify the responsiveness of any firm to an aggregate productivity shock, and this effect is particularly strong for firms with access to better data. The latter statement holds true by the following logic: Any increase of z will add a convex term into the profit function of any firm. In percentage terms, this reduces the (local) curvature of profits the most for firms with access to better data, because these firms are larger, i.e. produce at a point where the profit function is already relatively linear. By inducing relatively large decreases (in percentage terms) in the curvature of profits for firms with better data, increases in the data feedback loop thereby amplify the responsiveness of these firms to aggregate productivity shocks.

Importantly, whether the relationship between a firm's access to data is positive or negative also depends on the parameters κ_e and κ_v. For example, if $\kappa_e = 0$, $\kappa_v > 0$, and $z > 0$, then $\frac{\partial \varphi(\cdot)}{\partial \sigma_i} < 0$ holds true.

These results imply that the strength of the data feedback loop determines whether the increasing availability of data amplifies or reduces cyclical fluctuations. In the previous subsection, we have seen that greater availability of data amplified cyclical fluctuations and made the strength of monetary policy countercyclical. These results hinged on the fact that data-rich firms respond less strongly to aggregate productivity shocks. When the data feedback loop becomes strong enough, the sign of this relationship flips, which implies that a greater availability of data amplifies cylical fluctuations.

Finally, we consider how the data feedback effect affects the relationship between firms' risk sensitivity (represented by the parameter ρ_i) and their capital holdings (and by extension, their size). In a nutshell, we show that the presence of the data feedback effect may induce firms who are more sensitive to risk 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_i. We establish the following key result:

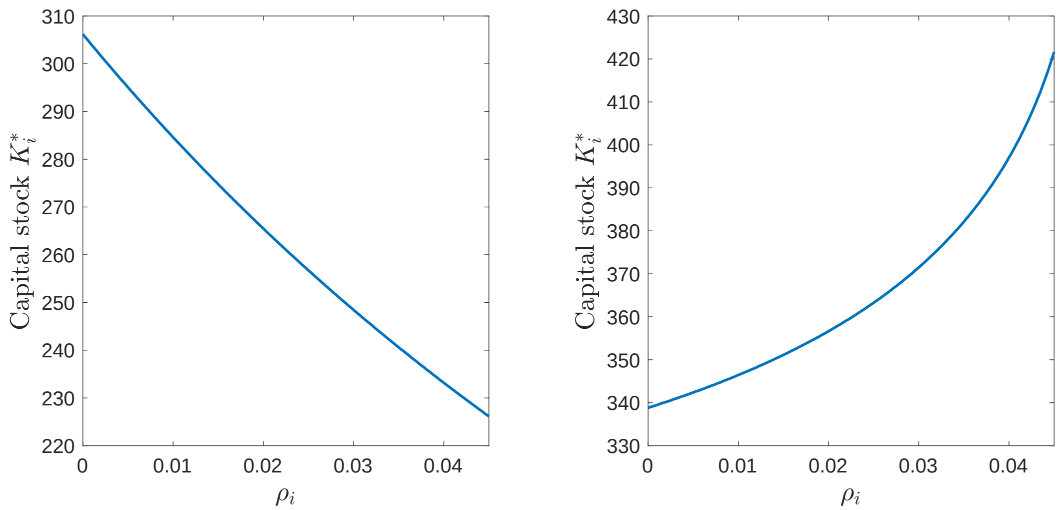
Proposition 4 *If $z = 0$, then $\frac{\partial K^*_{t+1}}{\partial \rho_i} < 0$. By contrast, $\frac{\partial K^*_{t+1}}{\partial \rho_i} > 0$ holds if:*

$$\left(\bar{V}_{t+1} + \kappa_v \sigma_i - \kappa_v z K^*_{t+1}\right) - \kappa_v z \left(K^*_{t+1} - (1-\delta) K_t\right) < 0 \tag{11}$$

The strength of the data feedback effect z thus determines whether more risk sensitive firms hold more or less capital. When $z = 0$, firms that are more exposed to risk (i.e., have a higher ρ_i) hold less capital, because increases in ρ_i 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_i) and it's capital level may flip.

Figure 4: Capital, risk-sensitivity and the data feedback loop

(a) No data feedback loop $z = 0$ (b) Strong data feedback loop $z = 0.02$



Note: This figure shows the steady state capital stock for different values of risk sensitivity ρ_i. The left panel shows this relationship for the case in which there is no data feedback loop, $z = 0$, and the right panel shows it for the case of an active data feedback loop, $z > 0$.

These results affect the magnitude of cyclical fluctuations by changing the average risk sensitivity of firms in the economy. To see this, suppose that ρ_i is heterogeneous across firms/industries but constant over time. This type of heterogeneity naturally arises for many reasons, including differences in bankruptcy risk and the cyclicality of firm's productivity (David and Zeke, 2023). By the results of proposition 4, the presence of the data feedback effect will increase the overall sensitivity of firms to risk, because firms with high risk sensitivity (high ρ_i) have stronger incentives to grow larger. This composition effect implies that the response of aggregate investment to movements in aggregate uncertainty will be amplified when the strength of the data feedback loop increases. This is because the market share of firms with high ρ_i, who are particularly susceptible to aggregate uncertainty shocks, increases when the data feedback effect becomes stronger. 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.

4 Data and idiosyncratic productivity

4.1 Framework and firm optimization

Throughout this section, we focus on the role of data in helping firms forecast their future idiosyncratic productivity realizations. The key insight of this section is the following: Even if data does not favorably affect the distribution of a firm's productivity nor its cost of capital, the presence of data dampens cyclical fluctuations and will make the effectiveness of monetary policy countercyclical (if there is no data feedback loop).

Formally, we consider the following special case of the environment outlined in section 2: There are two time periods $t \in \{1, 2\}$. In period 1, there is no production, but firms choose their capital stock for period 2. There is a unit mass of firms indexed i and the second-period productivity of a firm is given by $A_{i,2} = \bar{A} + \epsilon_i$, where $\bar{A}$ is common knowledge and can be understood as the aggregate productivity component. We model aggregate productivity shocks as changes of $\bar{A}$.

The are two different types of firms, namely *firms with data* and *firms without data*. We index the type of a firm using the indicator $j \in \{d, nd\}$, where d refers to firms with data. Firms with data receive a perfect signal about their realization of ϵ_i in period 1 (i.e. they have $\xi_{i,t} = 0$ as defined in the framework of section 2), while firms without data receive no signal about ϵ_i (i.e. $\xi_{i,t} \to \infty$ holds). To allow for the previously discussed features to be active, we allow the productivity distribution to vary across firms with and without data. We define the distribution of ϵ_i for a firm with data as G^d, with support $[\underline{\epsilon}^d, \bar{\epsilon}^d]$, and the distribution of ϵ_i for a firm without data as G^{nd}, with support $[\underline{\epsilon}^{nd}, \bar{\epsilon}^{nd}]$. The cost of acquiring one unit of capital is $r_d := r + \tilde{\rho}_d$ for firms with data and $r_{nd} := r + \tilde{\rho}_{nd}$ for firms without data. The favorable impact of data on a firm's cost of capital, which was considered in the previous section, can hence be modelled by specifying that $\tilde{\rho}_d \leq \tilde{\rho}_{nd}$. The interest rate r is controlled by the monetary authority. For tractability, we assume that the data feedback loop is inactive. We define the share of firms with data in the economy as $\omega \in (0, 1)$.

Defining the amount of capital a firm utilizes as $K_{i,2}$, the second-period profits of a firm are thus given by:

$$A_{i,2}(K_{i,2})^{\alpha} - r_j K_{i,2}. \tag{12}$$

We now pin down the optimal capital choices of firms, beginning with any firm with data. 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 with data, which conditions on the firm's productivity $A_{i,2}$, and is given by

$$K_2^d(A_{i,2}) = \arg\max_{K_2} \big[A_{i,2}(K_2)^\alpha - r_d K_2\big]. \tag{13}$$

Now consider any firm without data. Given that any such firm does not know its future productivity, they all face the same optimization problem. We define K_2^{nd} as the optimally chosen capital stock of any firm without data. This optimal capital stock solves:

$$K_2^{nd} = \arg\max_{K_2} \left[\int_{\underline{\epsilon}^{nd}}^{\bar{\epsilon}^{nd}} (\bar{A} + \epsilon_i)(K_2)^\alpha dG(\epsilon_i) - r_{nd} K_2 \right] \tag{14}$$

Solving the aforementioned optimization problems allows us to derive the expected capital stock and output of the different types of firms. We define $\bar{K}_2^{nd}$ and $\bar{K}_2^d$ as the cross-sectional expectation of the capital stock of a firm without and with data respectively. Further, we define $\bar{Y}_2^{nd}$ and $\bar{Y}_2^d$ as cross-sectional expectation of the output of a firm without and with data, respectively. Finally, we define the following moments that characterize the expected capital inputs and outputs of the two different types of firms, with $j \in \{d, nd\}$:

$$\mathbb{E}^j[A_{i,2}] = \int_{\underline{\epsilon}^j}^{\bar{\epsilon}^j} (\bar{A} + \epsilon_i) dG^j(\epsilon_i) \quad ; \quad \mathbb{E}^j[(A_{i,2})^{\frac{1}{1-\alpha}}] = \int_{\underline{\epsilon}^j}^{\bar{\epsilon}^j} (\bar{A} + \epsilon_i)^{\frac{1}{1-\alpha}} dG^j(\epsilon_i) \tag{15}$$

Throughout the following analysis, we place particular emphasis on the case in which $\tilde{\rho}_{nd} = \tilde{\rho}_d$ and $G^{nd} = G^d$, i.e. in which superior access to data only yields value by providing signals about future productivities, but not by affecting the distribution of productivity nor the capital costs of firms. This analysis allows us to establish that the results of the previous section extend even if data does not favorably affect a firm's productivity distribution, but only enables firm to forecast their future productivity realizations.

In the following lemma, we characterize the expected capital stocks and outputs of the different types of firms:

Lemma 2 *The expected capital stock and expected output of a firm with data are given by:*

$$\bar{K}_2^d = (\alpha)^{\frac{1}{1-\alpha}} (r + \tilde{\rho}_d)^{\frac{1}{\alpha-1}} \mathbb{E}^d\Big[\big(A_{i,2}\big)^{\frac{1}{1-\alpha}}\Big] \quad ; \quad \bar{Y}_2^d = (\alpha)^{\frac{\alpha}{1-\alpha}} (r + \tilde{\rho}_d)^{\frac{\alpha}{\alpha-1}} \mathbb{E}^d\Big[\big(A_{i,2}\big)^{\frac{1}{1-\alpha}}\Big] \tag{16}$$

The expected capital stock and expected output of a firm without data are given by:

$$\bar{K}_2^{nd} = (\alpha)^{\frac{1}{1-\alpha}} (r + \tilde{\rho}_{nd})^{\frac{1}{\alpha-1}} \big(\mathbb{E}^{nd}[A_{i,2}]\big)^{\frac{1}{1-\alpha}} \quad ; \quad \bar{Y}_2^{nd} = (\alpha)^{\frac{\alpha}{1-\alpha}} (r + \tilde{\rho}_{nd})^{\frac{\alpha}{\alpha-1}} \big(\mathbb{E}^{nd}[A_{i,2}]\big)^{\frac{1}{1-\alpha}} \tag{17}$$

If $\tilde{\rho}_d = \tilde{\rho}_{nd}$ and $G^d = G^{nd}$, 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 or a firm's cost of capital, 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, 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.

To study how access to the aforementioned type of data shapes the effects of aggregate productivity and monetary policy shocks, we once again consider the elasticities of capital with respect to changes in r and $\bar{A}$. We say that a firm responds more strongly to a given shock if its elasticity is larger (in absolute terms).

4.2 The impact of monetary policy and business cycles

In this subsection, we establish how the magnitude of cyclical fluctuations and the effectiveness of monetary policy along the cycle are shaped by firm's access to signals about their future productivities. We begin by establishing that firms with and without data respond to monetary policy shocks differentially only if their costs of capital differ:

Proposition 5 *If $\tilde{\rho}_d = \tilde{\rho}_{nd}$, the effect of a monetary policy shock on the expected output (capital) of firms with data equals its effect on the output (capital) of firms without data:*

$$\frac{\partial \bar{K}_2^d / \partial r}{\bar{K}_2^d} = \frac{\partial \bar{K}_2^{nd} / \partial r}{\bar{K}_2^{nd}} \qquad ; \qquad \frac{\partial \bar{Y}_2^d / \partial r}{\bar{Y}_2^d} = \frac{\partial \bar{Y}_2^{nd} / \partial r}{\bar{Y}_2^{nd}} \tag{18}$$

This result holds by the following logic. The magnitude of a firm's response to a monetary policy shock is proportional to the firm's productivity if the firm has access to data. Importantly, the size of such a firm is also proportional to its productivity. This proportionality implies that the relative effects of a monetary policy shock on the two types of firms are identical, which can be seen directly when studying the closed-form solutions for aggregate capital and output given in lemma 2.

Next, we establish how access to data shapes the responsiveness of a firm to an aggregate productivity shock:

Proposition 6 *Suppose $\alpha < 0.5$. The effects of an aggregate productivity shock on the expected output (capital) of firms without data are strictly larger:*

$$\frac{\partial \bar{K}_2^{nd}/\partial \bar{A}}{\bar{K}_2^{nd}} > \frac{\partial \bar{K}_2^{d}/\partial \bar{A}}{\bar{K}_2^{d}} \qquad ; \qquad \frac{\partial \bar{Y}_2^{nd}/\partial \bar{A}}{\bar{Y}_2^{nd}} > \frac{\partial \bar{Y}_2^{d}/\partial \bar{A}}{\bar{Y}_2^{d}}. \tag{19}$$

Moreover, the difference $\bar{Y}_2^d - \bar{Y}_2^{nd}$ decreases as $\bar{A}$ increases.

Given that estimates for the parameter α are commonly in the range $[0.3, 0.5]$, this result indicates that cyclical fluctuations will be dampened when more firms acquire data that allows them to predict their future productivities. The intuition underlying this result is as follows: When firms have access to data about their idiosyncratic productivities, changes in aggregate productivity will induce smaller changes in their information sets, thereby eliciting a smaller response.

A corollary of the previous results is that the dissemination of data not only dampens cyclical fluctuations in general, but mitigates recessions in particular:

Corollary 1 *Suppose $\alpha < 1/2$. Increasing the share of firms with access to data (ω) dampens recessions more than it amplifies booms, i.e.:*

$$\frac{\partial}{\partial \bar{A}}\left[\frac{\partial Y_2}{\partial \omega}\right] < 0 \tag{20}$$

Moreover, the aforementioned dynamics imply that the market shares of firms with access to data will be countercyclical. This renders the effectiveness of monetary policy countercyclical, which is formalized in the following corollary.

Corollary 2 *Suppose $\alpha < 1/2$ and that $\tilde{\rho}_d < \tilde{\rho}_{nd}$. Then, the effects of a monetary policy shock are countercylical, i.e.:*

$$\frac{\partial}{\partial \bar{A}}\left[\frac{\partial Y_2/\partial r}{Y_2}\right] > 0 \tag{21}$$

The underlying logic is as before: Firms with access to data respond comparatively weakly to aggregate productivity shocks, which implies that they attain relatively large market shares in recessions. These firms respond comparatively strongly to monetary policy if their costs of capital are lower (i.e. $\tilde{\rho}_d < \tilde{\rho}_{nd}$), which thus implies that monetary policy becomes relatively more effective in recessions.

5 The macroeconomic effects of digital markets regulation

Our work also establishes how digital markets regulation such as the EU GDPR, the DMA, and the UK DPA affects macroeconomic outcomes. This is because this type of regulation is centered around the governance of data and our work directly speaks to the macroeconomic role of data. It is inevitable that digital markets regulation affects the economy as a whole, given the scope of this type of regulation and the relevance of digital markets in modern economies. Thus, understanding the macroeconomic effects thereof is of first-order importance, especially because the macroeconomic perspective has essentially been absent from the policy debate surrounding these pieces of legislation.

Our work offers a conceptual framework to make progress in this endeavour. Different policy measures in the area can be understood as changes in the strength of the data feedback loop or the exogenous availability of data to firms. To reinforce this point, we consider two cornerstones of the existing regulation on digital markets, namely (1) the establishment of a right to data portability (as codified in the EU GDPR and the DMA) and (2) the prohibition of data transfer within firms, which the EU commission imposed on Google as part of its merger with Fitbit (European Commission, 2020).

Consider first the implementation of a right to data portability, as defined in the EDU GDPR and reinforced in the DMA. This right allows consumers to transfer all the data a given firm has about them to its competitor. To see how to model this within our framework, suppose that there are two firms $j \in \{A, B\}$ with different exogenous access to data. The establishment of a right to data portability allows for the transfer of data between firms and would thus induce a reduction of σ_j (i.e. an improvement in the exogenous access to data) for both firms. Moreover, the improvement in data access would likely be particularly pronounced for the firm with ex ante worse access to data. Such increases in the access to data raise the responsiveness of both firms to monetary policy, while the impact on the effect of aggregate productivity shocks depends on the strength of the data feedback loop.

Now consider the second type of legislation, which prohibits the transfer of data within a firm. Essentially, this limits the ability of a firm to use data on the behaviour of consumers in other branches of the firm (e.g. in advertising or the forecasting of individual product success). Within our framework, this can be viewed as a reduction in the strength of the data feedback loop. Through various channels, this reduces the magnitude of cyclical fluctuations caused by aggregate productivity or uncertainty shocks. It also reduces the competitive advantages generated by exogenous heterogeneity in firms' access to data.

6 Conclusion

The importance of data in modern economies cannot be understated. Data yields value to firms in many ways, namely by predicting sales, forecasting costs, and by enabling targeted advertising. In this paper, we study how data shapes two classic issues in Macroeconomics, namely the propagation of cyclical fluctuations and the effectiveness of monetary policy along the business cycle.

We demonstrate that firms with superior access to data respond less to aggregate productivity shocks. This is because access to better data increases a firm's expected productivity and enables firms to predict their idiosyncratic productivity realizations. Through both these channels, superior access to data makes fluctuations in aggregate productivity less consequential, thereby eliciting a smaller response. Moreover, firms with superior access to data respond more strongly to monetary policy shocks. This is because these firms face lower idiosyncratic uncertainty, which lowers their cost of capital. Thus, a given change in the interest rate induced by the central banks changes (in percentage terms) a firm's cost of capital more substantially if it has data, thereby eliciting a greater response. These insights imply that the increased availability of data in modern economies amplifies cyclical fluctuations and makes the effectiveness of monetary policy countercyclical.

In the presence of smart devices and algorithms to analyse click-through rates, data accumulates endogenously through a data feedback loop: Firms which produce more get access to larger amounts of data. The presence of a data feedback loop has substantial implications for the questions at hand. First, we show that the positive relationship between a firm's exogenously given access to data and its size is reinforced in the presence of a data feedback loop. Second, when the data feedback loop becomes strong enough, data-rich firms respond more strongly to aggregate productivity shocks. This has substantial implications for the propagation of aggregate productivity shocks and the relative effectiveness of monetary policy along the business cycle. Thirdly, we show that the data feedback loop amplifies cyclical fluctuations by weakening the negative relationship between a firm's exposure to risk and its size, which amplifies the effects of uncertainty shocks.

A Proofs

A.1 Proof of proposition 1

Part 1: The effects of monetary policy shocks

When choosing K_{t+1}, the relevant part of a firm's objective function is:

$$\mathbb{E}[A_{t+1};\sigma_{i,t+1}]K_{t+1}^{\alpha} - \big(r_t + \rho_i VAR[A_{t+1};\sigma_{i,t+1}]\big)\big(K_{t+1} - (1-\delta)K_t\big) +$$

$$\beta\Big[\mathbb{E}[A_{t+2};\sigma_{i,t+1}]K_{t+2}^{\alpha} - \big(r_{t+1} + \rho_i VAR[A_{t+2};\sigma_{i,t+1}]\big)\big(K_{t+2} - (1-\delta)K_{t+1}\big)\Big]$$

When $z = 0$, the first-order condition the optimal capital stock has to satisfy reads:

$$\alpha(K_{t+1})^{\alpha-1}\mathbb{E}[A;\sigma_i] - \big(1-\beta(1-\delta)\big)\big(r_t + \rho_i VAR[A;\sigma_i]\big) = 0$$

Thus, the optimal capital stock satisfies:

$$K_{t+1}^{*} = \left[\frac{\alpha\mathbb{E}[A;\sigma_i]}{\big(1-\beta(1-\delta)\big)\big(r_t + \rho_i VAR[A;\sigma_i]\big)}\right]^{1/(1-\alpha)}$$

This implies that:

$$\frac{\partial K_{t+1}^{*}}{\partial r_t} = \frac{1}{1-\alpha}\left[\frac{\alpha\mathbb{E}[A;\sigma_i]}{\big(1-\beta(1-\delta)\big)\big(r_t + \rho_i VAR[A;\sigma_i]\big)}\right]^{1/(1-\alpha)-1}\left[\frac{-\alpha\mathbb{E}[A;\sigma_i]\big(1-\beta(1-\delta)\big)}{\big[\big(1-\beta(1-\delta)\big)\big(r_t + \rho_i VAR[A;\sigma_i]\big)\big]^2}\right]$$

$$=$$

$$K_{t+1}^{*}\left[\frac{-1/(1-\alpha)}{\big[r_t + \rho_i VAR[A;\sigma_i]\big]}\right]$$

The elasticity with respect to a monetary policy shock is given by

$$\gamma(\cdot) = \frac{\partial K_{t+1}^{*}}{\partial r_t}\frac{r_t}{K_{t+1}^{*}} = \frac{-r_t/(1-\alpha)}{\big[r_t + \rho_i VAR[A;\sigma_i]\big]} < 0$$

Thus, we have:

$$\frac{\partial\gamma(\cdot)}{\partial\sigma_i} = \frac{r_t\rho_i\kappa_v/(1-\alpha)}{\big[r_{t+1} + \rho_i VAR[A;\sigma_i]\big]^2} \geq 0$$

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

Part 2: The effects of aggregate productivity shocks

When $z = 0$, we have:

$$\varphi(\cdot) = \frac{\partial K_{t+1}}{\partial \bar{A}} \frac{\bar{A}}{K^*_{t+1}} = \frac{\bar{A}}{(1-\alpha)\mathbb{E}[A;\sigma_i]} \implies \frac{\partial \varphi(\cdot)}{\partial \sigma_i} = \frac{-1(-\kappa_e)\bar{A}}{(1-\alpha)\big[\mathbb{E}[A;\sigma_i]\big]^2} > 0$$

A.2 Proof of proposition 2

Consider any $z \geq 0$. The first-order condition which the optimal capital stock has to satisfy reads:

$$T(\cdot) := \alpha K^{\alpha-1}_{t+1} \underbrace{(\mathbb{E}[A_{t+1};K_{t+1}])}_{=\bar{A}-\kappa_e\sigma_i+\kappa_e z K_{t+1}} + \underbrace{\frac{\partial \mathbb{E}[A_{t+1}]}{\partial K_{t+1}}}_{=\kappa_e z} K^{\alpha}_{t+1} - \big(r_t + \rho_i \underbrace{VAR[A_{t+1};K_{t+1}]}_{=\bar{V}+\kappa_v\sigma_i-\kappa_v z K_{t+1}}\big)$$

$$-\rho_i \underbrace{\frac{\partial VAR[A_{t+1}]}{\partial K_{t+1}}}_{=-\kappa_v z} \big(K_{t+1} - (1-\delta)K_t\big) + \beta(1-\delta)\big(r_{t+1} + \rho_i VAR[A_{t+2};K_{t+2}]\big) = 0$$

One can show that $\frac{\partial K^*_{t+1}}{\partial z} > 0$. This holds because:

$$\frac{\partial T}{\partial z} = \alpha\kappa_e K^{\alpha}_{t+1} + \kappa_e K^{\alpha}_{t+1} + \rho_i\kappa_v\big(K_{t+1} - \beta(1-\delta)K_{t+2}\big) + \rho_i\kappa_v\big(K_{t+1} - (1-\delta)K_t\big) > 0$$

$$\frac{\partial T}{\partial K_{t+1}} = \alpha(\alpha-1)(K_{t+1})^{\alpha-2}\big(\mathbb{E}[A_{t+1}]\big) + 2\alpha K^{\alpha-1}_{t+1}\kappa_e z + 2\rho_i\kappa_v z$$

The first inequality holds because the optimal capital stock of any firm is constant over time in the setting we consider.

Note further that:

$$\frac{\partial T}{\partial \sigma_i} = -\kappa_e \alpha K^{\alpha-1}_{t+1} - \rho_i\kappa_v\big(1 - \beta(1-\delta)\big) < 0$$

It follows that:

$$\frac{\partial K^*_{t+1}}{\partial \sigma_i} = \frac{\kappa_e\alpha K^{\alpha-1}_{t+1} + \rho_i\kappa_v\big(1-\beta(1-\delta)\big)}{\alpha(\alpha-1)(K_{t+1})^{\alpha-2}\big(\mathbb{E}[A_{t+1}]\big) + 2\alpha K^{\alpha-1}_{t+1}\kappa_e z + 2\rho_i\kappa_v z} < 0$$

If $\kappa_e = 0$, it is relatively easy to evaluate how this object is shaped by increases in the strength of the data feedback loop. Then, we have:

$$\frac{\partial K^*_{t+1}}{\partial \sigma_i} = \frac{\rho_i \kappa_v \big(1 - \beta(1-\delta)\big)}{\alpha(\alpha-1)(K_{t+1})^{\alpha-2}\big(\mathbb{E}[A_{t+1}]\big) + 2\rho_i \kappa_v z}$$

It follows that:

$$\frac{\partial^2 K^*_{t+1}}{\partial \sigma_i \partial z} = \frac{-\rho_i \kappa_v \big(1 - \beta(1-\delta)\big)}{\big[\alpha(\alpha-1)\big(\mathbb{E}[A_{t+1}]\big)(K_{t+1})^{\alpha-2} + 2\rho_i \kappa_v z\big]^2} \left[\alpha(\alpha-1)(\alpha-2)(\mathbb{E}[A_{t+1}])(K_{t+1})^{\alpha-3} \frac{\partial K_{t+1}}{\partial z} + 2\rho_i \kappa_v\right] < 0$$

In other words, increases in the strength of the data feedback loop exacerbate the relationship between a firm's exogenous access to data and its size.

Now let's consider a general $\kappa_e > 0$ and set $\kappa_v = 0$. In that case, we have that:

$$\frac{\partial K^*_{t+1}}{\partial \sigma_i} = \frac{\kappa_e \alpha}{\alpha(\alpha-1)(K_{t+1})^{-1}\big(\bar{A} - \kappa_e \sigma_i + \kappa_e z K_{t+1}\big) + 2\alpha \kappa_e z} =$$

$$\frac{\kappa_e \alpha}{\alpha(\alpha-1)(K_{t+1})^{-1}\big(\bar{A} - \kappa_e \sigma_i\big) + \alpha(\alpha+1)\kappa_e z}$$

Then, it follows that:

$$\frac{\partial^2 K^*_{t+1}}{\partial \sigma_i \partial z} = \frac{-\kappa_e \alpha\big[-\alpha(\alpha-1)(K_{t+1})^{-2}\big(\bar{A} - \kappa_e \sigma_i\big)\frac{\partial K_{t+1}}{\partial z} + \alpha(\alpha+1)\kappa_e\big]}{\big[\alpha(\alpha-1)(K_{t+1})^{-1}\big(\bar{A} - \kappa_e \sigma_i\big) + \alpha(\alpha+1)\kappa_e z\big]^2} < 0$$

A.3 Proof of lemma 1

Recall that, for any $z \geq 0$, the first-order condition the optimal capital stock has to satisfy reads:

$$T(\cdot) := \alpha K^{\alpha-1}_{t+1} \underbrace{\big(\mathbb{E}[A_{t+1}; K_{t+1}]\big)}_{=\bar{A} - \kappa_e \sigma_i + \kappa_e z K_{t+1}} + \underbrace{\frac{\partial \mathbb{E}[A_{t+1}]}{\partial K_{t+1}}}_{=\kappa_e z} K^{\alpha}_{t+1} - \big(r_t + \rho_i VAR[A_{t+1}; K_{t+1}]\big)$$

$$-\rho_i \underbrace{\frac{\partial VAR[A_{t+1}]}{\partial K_{t+1}}}_{=-\kappa_v z} \big(K_{t+1} - (1-\delta)K_t\big) + \beta(1-\delta)\big(r_{t+1} + \rho_i VAR[A_{t+2}; K_{t+2}]\big) = 0$$

Before moving forward, note that:

$$\frac{\partial T}{\partial \bar{A}} = \alpha(K_{t+1})^{\alpha-1} \quad ; \quad \frac{\partial T}{\partial \sigma_i} = -\kappa_e \alpha K_{t+1}^{\alpha-1} - \rho_i \kappa_v \big(1-\beta(1-\delta)\big)$$

$$\frac{\partial T}{\partial K_{t+1}} = \alpha(\alpha-1)\big(\mathbb{E}[A_{t+1}]\big)(K_{t+1})^{\alpha-2} + 2\alpha K_{t+1}^{\alpha-1}\kappa_e z + 2\rho_i \kappa_v z$$

Based on this, we can directly conclude:

$$\frac{\partial K_{t+1}}{\partial \bar{A}} = \frac{-\alpha(K_{t+1})^{\alpha-1}}{\alpha(\alpha-1)\big(\mathbb{E}[A_{t+1}]\big)(K_{t+1})^{\alpha-2} + 2\alpha K_{t+1}^{\alpha-1}\kappa_e z + 2\rho_i \kappa_v z} > 0$$

Let's examine the relative effects of the aggregate productivity shock, which is given by:

$$\varphi(\cdot) = \frac{\bar{A}\frac{\partial K^*_{t+1}}{\partial \bar{A}}}{K^*_{t+1}} = \frac{-\alpha \bar{A}(K_{t+1})^{\alpha-1}}{\alpha(\alpha-1)\big(\mathbb{E}[A_{t+1}]\big)(K_{t+1})^{\alpha-1} + 2\alpha\kappa_e z K_{t+1}^{\alpha} + 2\rho_i \kappa_v z K_{t+1}} =$$

$$\frac{-\alpha \bar{A}}{\alpha(\alpha-1)\big(\bar{A} - \kappa_e \sigma_i + \kappa_e z K_{t+1}\big) + 2\alpha \kappa_e z K_{t+1} + 2\rho_i \kappa_v z (K_{t+1})^{2-\alpha}} =$$

$$\frac{-\alpha \bar{A}}{\alpha(\alpha-1)\big(\bar{A} - \kappa_e \sigma_i\big) + \alpha\big(\alpha+1\big)\kappa_e z K_{t+1} + 2\rho_i \kappa_v z (K_{t+1})^{2-\alpha}} > 0$$

It follows that:

$$\frac{\partial \varphi(\cdot)}{\partial z} = \frac{\alpha \bar{A}}{\big[\alpha(\alpha-1)\big(\bar{A} - \kappa_e \sigma_i\big) + \alpha\big(\alpha+1\big)\kappa_e z K_{t+1} + 2\rho_i \kappa_v z (K_{t+1})^{2-\alpha}\big]^2}$$

$$\left[\alpha(\alpha+1)\kappa_e \frac{\partial [z K_{t+1}]}{\partial z} + 2\rho_i \kappa_v \frac{\partial [z (K_{t+1})^{2-\alpha}]}{\partial z}\right] > 0$$

This expression is strictly positive because K_{t+1} is increasing in z.

A.4 Proof of proposition 3

When $z > 0$, we have that:

$$\varphi(\cdot) = \frac{\partial K_{t+1}}{\partial \bar{A}} \frac{\bar{A}}{K_{t+1}} = \frac{-\alpha \bar{A}}{\alpha(\alpha-1)\big(\bar{A} - \kappa_e \sigma_i\big) + \alpha\big(\alpha+1\big)\kappa_e z K_{t+1} + 2\rho_i \kappa_v z (K_{t+1})^{2-\alpha}}$$

This implies that:

$$\frac{\partial \varphi(\cdot)}{\partial \sigma_i} = \frac{\alpha \bar{A}}{\big[\alpha(\alpha-1)\big(\bar{A} - \kappa_e \sigma_i\big) + \alpha\big(\alpha+1\big)\kappa_e z K_{t+1} + 2\rho_i \kappa_v z (K_{t+1})^{2-\alpha}\big]^2}$$

$$\Big[-\alpha(\alpha-1)\kappa_e + \big(\alpha(\alpha+1)\kappa_e z + 2\rho_i\kappa_v z(2-\alpha)(K_{t+1})^{1-\alpha}\big)\frac{\partial K_{t+1}}{\partial \sigma_i}\Big]$$

If $z = 0$, this expression is strictly positive. If $z > 0$, this expression is strictly negative if and only if:

$$-\kappa_e\alpha(\alpha-1) + \big(\alpha(\alpha+1)\kappa_e z + 2\rho_i\kappa_v z(2-\alpha)(K_{t+1})^{1-\alpha}\big)\frac{\partial K_{t+1}}{\partial \sigma_i} < 0$$

$$\Longleftrightarrow$$

$$\big(\alpha(\alpha+1)\kappa_e z + 2\rho_i\kappa_v z(2-\alpha)(K_{t+1})^{1-\alpha}\big)\frac{\partial K_{t+1}}{\partial \sigma_i} < \kappa_e\alpha(\alpha-1)$$

This holds true, for example, if $\kappa_e = 0$ and $\kappa_v > 0$.

A.5 Proof of proposition 4

As before, the optimal capital stock must solve the following first-order condition:

$$T(K_{t+1}, r_t) := \alpha\mathbb{E}[A_{t+1}; K_{t+1}]K_{t+1}^{\alpha-1} + \frac{\partial\mathbb{E}[A_{t+1}; K_{t+1}]}{\partial K_{t+1}}K_{t+1}^{\alpha} - \big(r_t + \rho_i VAR[A_{t+1}; K_{t+1}]\big)$$

$$-\rho_i\frac{\partial VAR[A_{t+1}; K_{t+1}]}{\partial K_{t+1}}\big(K_{t+1} - (1-\delta)K_t\big) + \beta\big(1-\delta\big)\big(r_{t+1} + \rho_i VAR[A_{t+2}; K_{t+2}]\big) = 0$$

Let's examine the derivative of the capital stock w.r.t ρ_i. To calculate this, note that:

$$\frac{\partial T}{\partial \rho_i} = -VAR[A_{t+1}; K_{t+1}] - \frac{\partial VAR[A_{t+1}; K_{t+1}]}{\partial K_{t+1}}\big(K_{t+1} - (1-\delta)K_t\big) + \beta(1-\delta)VAR[A_{t+2}; K_{t+2}]$$

Thus, the desired relationship is given by:

$$\frac{\partial K^*_{t+1}}{\partial \rho_i} = \frac{VAR[A_{t+1}; K_{t+1}] + \frac{\partial VAR[A_{t+1}; K_{t+1}]}{\partial K_{t+1}}\big(K_{t+1} - (1-\delta)K_t\big) - \beta(1-\delta)VAR[A_{t+2}; K_{t+2}]}{\frac{\partial \Pi}{\partial K_{t+1}}}$$

The denominator is negative by assumption. Thus, the entire term is negative if the numerator is positive. If $z = 0$, this holds, because the numerator becomes:

$$VAR[A_{t+1}; K_{t+1}] + \underbrace{\frac{\partial VAR[A_{t+1}; K_{t+1}]}{\partial K_{t+1}}}_{=0}\big(K_{t+1} - (1-\delta)K_t\big) - \beta(1-\delta)VAR[A_{t+2}; K_{t+2}] =$$

$$\big(1 - \beta(1-\delta)\big)VAR[A; \sigma] > 0$$

By contrast, the numerator becomes negative if z is large enough. This is satisfied if:

$$\big(\bar{V} + \kappa_v \sigma_i - \kappa_v z K_{t+1}\big) - \kappa_v z\big(K_{t+1} - (1-\delta)K_t\big) < 0$$

A.6 Proof of lemma 2

Firms without data do not know the realizations of A_2 when making their capital choices. Firms with data know their realizations of A_2 when making their capital choices. The cost of capital of any firm with data (any firm without data) is given by $r_d := r + \tilde{\rho}_d$ ($r_{nd} := r + \tilde{\rho}_{nd}$), respectively.

The optimization problem of a firm with data is just:

$$K_2^d(A_2) = arg\max_{K_2} A_2(K_2)^\alpha - r_d K_2 \iff K_2^d(A_2) = \big(\alpha A_2\big)^{1/(1-\alpha)}(r_d)^{1/(\alpha-1)}$$

The expected capital of a firm with data is

$$\bar{K}_2^d = \int_{\underline{\epsilon}^d}^{\bar{\epsilon}^d} K_2^d(\bar{A}+\epsilon)dG^d(\epsilon)$$

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

$$K_2^{nd} = arg\max_{K_2}\left[\int_{\underline{\epsilon}^{nd}}^{\bar{\epsilon}^{nd}} (\bar{A}+\epsilon)(K_2)^\alpha dG^{nd}(\epsilon)\right] - r_{nd}K_2 \iff \alpha\mathbb{E}^{nd}[A_2](K_2)^{\alpha-1} - r_{nd} = 0$$

Taking all this together, the expected capital stocks of firms with data and without data are:

$$\bar{K}_2^{nd} = \big(\mathbb{E}^{nd}[A_2]\big)^{1/(1-\alpha)}\big(\alpha\big)^{1/(1-\alpha)}(r_{nd})^{1/(\alpha-1)}$$

$$\bar{K}_2^d = \mathbb{E}^d\big[(A_2)^{1/(1-\alpha)}\big]\big(\alpha\big)^{1/(1-\alpha)}(r_d)^{1/(\alpha-1)}$$

Now we determine the levels of output, beginning with the expected output of firms with data, which is:

$$\bar{Y}_2^d = \int_{\underline{\epsilon}^d}^{\bar{\epsilon}^d} (\bar{A}+\epsilon)\bigg(\big(\alpha A_2\big)^{1/(1-\alpha)}(r_d)^{1/(\alpha-1)}\bigg)^\alpha dG^d(\epsilon) =$$

$$(\alpha)^{\alpha/(1-\alpha)}(r_d)^{\alpha/(\alpha-1)}\int_{\underline{\epsilon}^d}^{\bar{\epsilon}^d} (\bar{A}+\epsilon)\big(A_2\big)^{\alpha/(1-\alpha)}dG^d(\epsilon) = (\alpha)^{\alpha/(1-\alpha)}(r_d)^{\alpha/(\alpha-1)}\mathbb{E}^d\big[(A_2)^{\frac{1}{1-\alpha}}\big]$$

Finally, the expected output of a firm without data is:

$$\bar{Y}_2^{nd} = \int_{\underline{\epsilon}^{nd}}^{\bar{\epsilon}^{nd}} (\bar{A}+\epsilon)\Big((\alpha)^{1/(1-\alpha)}\big(\mathbb{E}[A_2]\big)^{1/(1-\alpha)}(r_{nd})^{1/(\alpha-1)}\Big)^{\alpha} dG^{nd}(\epsilon) =$$

$$(\alpha)^{\alpha/(1-\alpha)}(r_{nd})^{\alpha/(\alpha-1)} \int_{\underline{\epsilon}^{nd}}^{\bar{\epsilon}^{nd}} (\bar{A}+\epsilon)\big(\mathbb{E}[A_2]\big)^{\alpha/(1-\alpha)} dG^{nd}(\epsilon) = (\alpha)^{\alpha/(1-\alpha)}(r_{nd})^{\alpha/(\alpha-1)}\big(\mathbb{E}^{nd}[A_2]\big)^{1/(1-\alpha)}$$

A.7 Proof of proposition 5

The relative effect of a monetary policy shock on the expected output of firms with type $j \in \{d, nd\}$:

$$\frac{\partial \bar{Y}_2^j/\partial r}{\bar{Y}_2^j} = \frac{\alpha}{\alpha-1}\frac{1}{r+\tilde{\rho}^j}$$

The result follows directly. Similar arguments imply the desired results for the expected capital stocks.

A.8 Proof of proposition 6

The expected outputs of firms without data and firms with data are given by:

$$\bar{Y}_2^{nd} = (\alpha)^{\frac{\alpha}{1-\alpha}}(r+\tilde{\rho}_{nd})^{\frac{\alpha}{\alpha-1}}\big(\mathbb{E}[A_{i,2}]\big)^{\frac{1}{1-\alpha}} \quad ; \quad \bar{Y}_2^{d} = (\alpha)^{\frac{\alpha}{1-\alpha}}(r+\tilde{\rho}_{d})^{\frac{\alpha}{\alpha-1}}\mathbb{E}\Big[\big(A_{i,2}\big)^{\frac{1}{1-\alpha}}\Big]. \tag{22}$$

For $\alpha < 1/2$, firms without data respond more strongly to aggregate productivity shocks, i.e.:

$$\frac{\partial \bar{Y}_2^{nd}/\partial \bar{A}}{\bar{Y}_2^{nd}} - \frac{\partial \bar{Y}_2^{d}/\partial \bar{A}}{\bar{Y}_2^{d}} > 0$$

To see this, note that:

$$\frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} = \left(\frac{1}{1-\alpha}\right)(\alpha)^{\frac{\alpha}{1-\alpha}}(r_{nd})^{\frac{\alpha}{\alpha-1}}\big(\mathbb{E}[A_{i,2}]\big)^{\frac{\alpha}{1-\alpha}} \implies \frac{\partial \bar{Y}^{nd}/\partial \bar{A}}{\bar{Y}^{nd}} = \left(\frac{1}{1-\alpha}\right)\big(\mathbb{E}[A_{i,2}]\big)^{-1}$$

Note further that:

$$\frac{\partial \bar{Y}_2^{d}}{\partial \bar{A}} = \left(\frac{1}{1-\alpha}\right)(\alpha)^{\frac{\alpha}{1-\alpha}}(r_{d})^{\frac{\alpha}{\alpha-1}}\mathbb{E}\big[(A_{i,2})^{\alpha/(1-\alpha)}\big] \implies \frac{\partial \bar{Y}^{d}/\partial \bar{A}}{\bar{Y}^{d}} = \left(\frac{1}{1-\alpha}\right)\frac{\mathbb{E}\big[(A_{i,2})^{\alpha/(1-\alpha)}\big]}{\mathbb{E}\big[(A_{i,2})^{1/(1-\alpha)}\big]}$$

By our assumption that $\alpha < 1/2$, one can establish that the relative effect on firms without data will be larger. This is because:

$$\mathbb{E}\big[(A_{i,2})^{\alpha/(1-\alpha)}\big] < \big[\mathbb{E}\big(A_{i,2}\big)\big]^{\alpha/(1-\alpha)} \quad ; \quad \mathbb{E}\big[(A_{i,2})^{1/(1-\alpha)}\big] > \big[\mathbb{E}\big(A_{i,2}\big)\big]^{1/(1-\alpha)}$$

This implies that:

$$\frac{\mathbb{E}\big[(A_{i,2})^{\alpha/(1-\alpha)}\big]}{\mathbb{E}\big[(A_{i,2})^{1/(1-\alpha)}\big]} < \frac{\big[\mathbb{E}\big(A_{i,2}\big)\big]^{\alpha/(1-\alpha)}}{\big[\mathbb{E}\big(A_{i,2}\big)\big]^{1/(1-\alpha)}} = \big[\mathbb{E}\big(A_{i,2}\big)\big]^{-1}$$

A.9 Proof of corollary 1

To see this result, note that:

$$Y_2 = \omega \bar{Y}_2^d + (1-\omega)\bar{Y}_2^{nd}$$

We have that:

$$\frac{\partial Y_2}{\partial \omega} = \bar{Y}_2^d - \bar{Y}_2^{nd} > 0$$

Thus, previous arguments directly imply that:

$$\frac{\partial}{\partial \bar{A}}\left[\frac{\partial Y_2}{\partial \omega}\right] = \frac{\partial \bar{Y}_2^d}{\partial \bar{A}} - \frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} < 0$$

A.10 Proof of corollary 2

Consider the market share of firms with data. This is given by:

$$M^d = \frac{\bar{Y}_2^d}{\bar{Y}_2^d + \bar{Y}_2^{nd}} = \frac{1}{1 + \frac{\bar{Y}_2^{nd}}{\bar{Y}_2^d}}$$

We want to show that the market share of firms with data is falling in $\bar{A}$ (i.e. is comparatively low in booms and relatively high in recessions):

$$\frac{\partial M^d}{\partial \bar{A}} < 0 \iff \frac{\partial\big[\bar{Y}_2^{nd}/\bar{Y}_2^d\big]}{\partial \bar{A}} > 0$$

$$\iff$$

$$\frac{\bar{Y}_2^d \frac{\partial \bar{Y}_2^{nd}}{\partial \bar{A}} - \bar{Y}_2^{nd}\frac{\partial \bar{Y}_2^d}{\partial \bar{A}}}{[\bar{Y}_2^d]^2} > 0 \iff \frac{\partial \bar{Y}_2^{nd}/\partial \bar{A}}{\bar{Y}_2^{nd}} > \frac{\partial \bar{Y}_2^d/\partial \bar{A}}{\bar{Y}_2^d}$$

Because firms with data get a higher market share in recessions and are more responsive to monetary policy shocks, this makes the effectiveness of monetary policy countercyclical.

To see this, note that the effect of a change in r on aggregate output are given by:

$$\frac{\frac{\partial[\omega\bar{Y}_2^d+(1-\omega)\bar{Y}_2^{nd}]}{\partial r}}{[\omega\bar{Y}_2^d+(1-\omega)\bar{Y}_2^{nd}]} = \frac{\omega}{\omega\bar{Y}_2^d+(1-\omega)\bar{Y}_2^{nd}}\frac{\partial\bar{Y}_2^d}{\partial r} + \frac{1-\omega}{\omega\bar{Y}_2^d+(1-\omega)\bar{Y}_2^{nd}}\frac{\partial\bar{Y}_2^{nd}}{\partial r} =$$

$$M^d\frac{\partial\bar{Y}_2^d/\partial r}{\bar{Y}_2^d} + \left(1-M^d\right)\frac{\partial\bar{Y}_2^{nd}/\partial r}{\bar{Y}_2^{nd}}$$

This depends on aggregate productivity in the following way:

$$\frac{\partial}{\partial\bar{A}}\left[\frac{\partial Y/\partial r}{Y}\right] = \underbrace{\frac{\partial M^d}{\partial\bar{A}}}_{<0}\underbrace{\left(\frac{\partial\bar{Y}_2^d/\partial r}{\bar{Y}_2^d} - \frac{\partial\bar{Y}_2^{nd}/\partial r}{\bar{Y}_2^{nd}}\right)}_{<0} + M^d\frac{\partial}{\partial\bar{A}}\left[\frac{\partial\bar{Y}_2^d/\partial r}{\bar{Y}_2^d}\right] + M^{nd}\frac{\partial}{\partial\bar{A}}\left[\frac{\partial\bar{Y}_2^{nd}/\partial r}{\bar{Y}_2^{nd}}\right] > 0$$

To see that this is strictly positive (the negative effect of an increase in r_t is less pronounced if $\bar{A}$ is higher, since the market share of firms in the economy with data is higher), let's evaluate the effect of a monetary policy shock on the expected output of firms with type $j \in \{d, nd\}$:

$$\frac{\partial\bar{Y}_2^j/\partial r}{\bar{Y}_2^j} = \frac{\alpha}{\alpha-1}\frac{1}{r+\tilde{\rho}^j}$$

This is independent of $\bar{A}$ and the relative effect on firms with data are more substantial, i.e.:

$$\frac{\partial\bar{Y}_2^d/\partial r}{\bar{Y}_2^d} < \frac{\partial\bar{Y}_2^{nd}/\partial r}{\bar{Y}_2^{nd}} \iff \frac{\alpha}{\alpha-1}\frac{1}{r+\tilde{\rho}_d} < \frac{\alpha}{\alpha-1}\frac{1}{r+\tilde{\rho}_{nd}} \iff \tilde{\rho}_{nd} > \tilde{\rho}_d$$

The latter holds by assumption.

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.

P. Ajit. Prediction of employee turnover in organizations using machine learning algorithms. algorithms, 4(5):C5, 2016.

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.

T. Babina, A. Fedyk, A. He, and J. Hodson. Firm investments in artificial intelligence technologies and changes in workforce composition. 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.

P. Bajari, V. Chernozhukov, A. Hortaçsu, and J. Suzuki. The impact of big data on firm performance: An empirical investigation. In AEA papers and proceedings, volume 109, pages 33–37. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2019.

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.

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), 2016.

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. B. S. Calderón and D. G. Rassier. Valuing the us data economy using machine learning and online job postings. NBER Chapters, 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.

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.

C. Corrado, J. Haskel, M. Iommi, C. Jona-Lasinio, and F. Bontadini. Data, intangible capital, and productivity. NBER/CRIW Conference on Technology, Productivity and Economic Growth, 2022.

J. M. David and D. Zeke. Risk-taking, capital allocation and monetary policy. 2023.

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

M. Demirer, D. J. Hernández, D. Li, and S. Peng. Data, privacy laws, and firm production: Evidence from gdpr. Work in progress, 2022.

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

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

European Commission. Mergers: Commission clears acquisition of fitbit by google, subject to conditions. 2020.

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.

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

O. Gomes, R. Mihet, and K. Rishabh. Growth and innovation in the modern data economy. 2023.

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.

C.-C. Groh and O. Pfäuti. Monetary policy and uncertainty in the data economy. 2023.

Y. He, H. Jiang, et al. The life cycle of firm in the economics of data. 2023.

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

D. Lashkari, A. Bauer, and J. Boussard. Information technology and returns to scale. 2018.

Y. Lau. A brief primer on the economics of targeted advertising. Economic Issues, 2020.

B. Maćkowiak and M. Wiederholt. Optimal sticky prices under rational inattention. American Economic Review, 99(3):769–803, 2009.

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.

R. C. Merton. An intertemporal capital asset pricing model. Econometrica: Journal of the Econometric Society, pages 867–887, 1973.

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

B. Munos, J. Niederreiter, and M. Riccaboni. Improving the prediction of clinical success using machine learning. medRxiv, pages 2021–02, 2021.

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

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

J. Quan. Tracing out international data flow: The value of data and privacy. 2022.

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

S. Van Nieuwerburgh and L. Veldkamp. Learning asymmetries in real business cycles. Journal of monetary Economics, 53(4):753–772, 2006.

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. 2022.

M. Wu and L. Zhang. Endogenous growth and human capital accumulation in a data economy. 2022.

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

D. Xie and L. Zhang. Endogenous growth with data generated during production. 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.