Firm Heterogeneity and Aggregate Fluctuations

Marco Errico
IMF

Simone Pesce
CBI

Luigi Pollio*
UMBC

This Draft: November 17, 2025

Abstract

We study how firm heterogeneity shapes the transmission of aggregate shocks. The aggregate response of macroeconomic outcomes to any source of aggregate shocks depends on both the average response across firms and the covariance between firms' response and their economic weight, which determines whether heterogeneity amplifies or dampens fluctuations. Using U.S. Compustat data from 1990 to 2019 and the Generalized Random Forest estimator, we estimate firm-level responses of sales, investment, and debt issuance to business cycle fluctuations. We uncover substantial heterogeneity driven primarily by non-financial characteristics—particularly industry scope and firm size. Aggregating these responses reveals that firm heterogeneity dampens aggregate fluctuations, especially for investment and debt issuance, as larger firms tend to be less cyclical than the average firm. Our results carry over to exogenously identified shocks and to financial outcomes.

JEL Codes: D22, E32, C14
Keywords: Firm Heterogeneity, Firm Sensitivity, Aggregate Fluctuations, Machine Learning.

*Contact: Luigi Pollio, luigi.pollio@umbc.edu; Marco Errico, merrico@imf.org; Simone Pesce, simone.pesce@centralbank.ie. We thank one anonymous referee, Michele Lenza, Daniele Massacci, Livia Paranhos, and all the participants at the 2025 ECONDAT, 2024 CFE-CMStatistics Conference, 2024 SNDE-IIF Symposium, Bank of Italy, Central Bank of Ireland and ECB for their comments and suggestions. Part of the research was conducted while Simone Pesce was visiting the European Central Bank. The views expressed here do not necessarily reflect those of the IMF, the Central Bank of Ireland or the Euro-System. This paper circulated as ECB working paper with the title: "Nonlinearities and Heterogeneity in Firms Response to Aggregate Fluctuations: What Can We Learn From Machine Learning?".

1 Introduction

A large body of the corporate finance and macroeconomics literature documents substantial cross-sectional differences in how non-financial firms respond to aggregate shocks and business cycle fluctuations (Crouzet and Mehrotra, 2020; Ottonello and Winberry, 2020). Some adjust sharply to changes in GDP growth and interest rates, while others remain largely unaffected. These differences in sensitivity to aggregate shocks across firms appear to reflect differences in observable characteristics across firms—such as leverage, liquidity, and size (Ottonello and Winberry, 2020; Jeenas, 2018a; Gertler and Gilchrist, 1994). The relative importance of the underlying characteristics and, more generally, whether such heterogeneity matters at the aggregate level, however, remain open questions. Understanding these mechanisms is crucial for policy makers, as aggregate outcomes depend not only on average responses but also on how shocks propagate across firms with different balance sheet positions.[1] In this paper, we study how heterogeneity in firms' characteristics affect firms' response to aggregate fluctuations and, thus, the aggregate economy, using a nonparametric machine learning approach.

We propose a theory of aggregation that links firm-level responses to macroeconomic outcomes by weighting firms' sensitivities to shocks according to their contribution in the economy. The aggregate response of macroeconomic outcomes to any source of aggregate shocks depends not only on individual firm reactions but also on their relative weight in the economy. It follows that the aggregate response of macroeconomic outcomes is driven by two components: the average firm-level response to a given shock, and the covariance between firm sensitivities and their weight in the aggregate. The first captures how firms, on average, react to shocks, while the second reflects whether heterogeneity across firms matters for the aggregate economy. A positive covariance implies that more sensitive firms hold greater weight, amplifying aggregate fluctuations, whereas a negative covariance suggests that less sensitive firms dominate, dampening macroeconomic volatility.

Implementing this framework empirically requires measuring how individual firms respond to aggregate shocks. While firms' aggregate shares with respect to specific outcomes are directly observable, their sensitivities to these shocks are not. Estimating these firm-level sensitivities is therefore essential to quantify how heterogeneity shapes aggregate dynamics. Machine-learning methods offer two key advantages for this task relative to a standard linear panel model. First, they do not require specifying a data-generating process ex ante, allowing the data to flexibly reveal the functional form linking firm characteristics to shock sensitivities. Second, they can accommodate a high-dimensional set of firm characteristics, capturing nonlinearities and complex interactions that standard linear models would miss. These features make machine learning particularly well suited

[1] See, e.g., Durante et al. (2020); Ottonello and Winberry (2020); Jeenas (2018b); Cloyne et al. (2018) among the others.

to estimating and assessing the aggregate implications of firm-level heterogeneity.

We apply our theoretical framework to quantify how firm heterogeneity shapes business cycle fluctuations in the United States (Crouzet and Mehrotra, 2020). We use financial information on publicly listed non-financial firms from the quarterly Compustat dataset, covering 1990Q1-2019Q4. We employ the Generalized Random Forest algorithm of Athey et al. (2019) to estimate firm-level sensitivities of sales, investment, and debt issuance to short-run output fluctuations as a function of a rich set of observable financial and non-financial characteristics widely examined in the literature, including leverage, liquidity, distance to default, share of short-term debt, size, return on assets, sales volatility, and industry scope. Using the estimated firm-level sensitivities, we document a set of results at both the micro and aggregate level.

We estimate an extremely rich cross-sectional heterogeneity in firms' sensitivity to business cycle fluctuations. We estimate that a 1 percent increase in real GDP growth is associated, on average, with a 2.1 percent increase in firms' sales, and an increase in investment and debt issuances increase by 0.9 percent and 1.2 percent, respectively, in line with economic intuition and previous estimates. However, the average sensitivity masquerades substantial cross-sectional heterogeneity, with a sizeable share of firms exhibiting either a weaker response or even an opposite sign relative to the mean, particularly for investment and debt issuance. While the average sensitivity across firms is statistically indistinguishable from the one obtained using a standard linear panel model, the higher-order moments of the distribution differ markedly. This discrepancy highlights the importance of capturing the nonlinear relationships among firm characteristics that influence the firms' responsiveness to aggregate fluctuations.

We find that non-financial characteristics—especially industry scope and firm size—are the primary drivers of the heterogeneity in firm sensitivities to the business cycle fluctuations. Following standard practice in the machine learning literature, we evaluate the importance of each firm characteristic in explaining the heterogeneity of firms' sensitivities for a given outcome variable using the proportion of splits in which it appears in the random forest. One advantage of machine learning approaches is in fact the ability to mitigate the curse of dimensionality and analyze a large set of characteristics while automatically detecting their relative importance. Our results suggest that heterogeneity is driven by a relatively small group of characteristics. We find that, taken together, non-financial characteristics account for approximately 76 percent of the heterogeneity in sales sensitivity, 87 percent in investment sensitivity, and 67 percent in debt issuance sensitivity. In particular, industry scope is the dominant driver of the heterogeneity in the sensitivity of sales, accounting for about 45 percent of the algorithm's splits; similarly, firm size accounts for 25 and 50 percent for debt and investment, respectively. These results suggests that heterogeneity is firms' sensitivity to business cycle fluctuations is primarily driven by demand-side factors, rather than changes in the cost of external financing.

We evaluate the role of individual firm characteristics in shaping firms' sensitivities to business cycle fluctuations and provide evidence of strong and complex nonlinearities. Using accumulated local effects, we show that small and medium-sized firms are less sensitive than large firms to business-cycle fluctuations in terms of sales, but they tend to increase investment and issue more debt during booms. Moreover, sales of firms with higher leverage are less cyclical than the average firm, while firms with higher liquidity, by contrast, exhibit stronger increases in investment than average. We also find that the marginal effect of each firm characteristic on firm-level sensitivities is not constant but exhibits kinks, U-shaped, or inverted U-shaped patterns. For instance, the sensitivity of investment and debt issuance is decreasing in firm size, but the marginal effect is strongly muted on the tail of the size distribution; similarly, the sensitivity of sales is decreasing in firm leverage, but a decreasing rate. Additionally, using the Friedman's H-statistic, we find that between 10% and 40% of the total effect of each characteristic on firm outcomes is mediated through its interaction with other characteristics, with firm size playing a particularly prominent role.

We quantify the role of heterogeneity in firm sensitivity for the aggregate using the theory of aggregation and our estimates of firm-level sensitivities. We find that firm heterogeneity dampens the aggregate response to the business cycle fluctuations, but its magnitude depends on the outcome variable. On average, we find that firms increase sales, investment, and debt issuance when real GDP grows, in line with economic intuition. However, more relevant firms in the economy tend to exhibit a lower sensitivity. Firm heterogeneity has a stronger dampening effect on debt issuance, effectively offsetting the average increase in borrowing that would otherwise occur. Similarly, firm heterogeneity reduces aggregate responses in investment by 53 percent, while the effect is more muted (6 percent) for sales, indicating sales of larger firms are only modestly less sensitive.

We assess the robustness of the aggregate implications along several dimensions. Using a rolling window framework, we find that the aggregate role of firm heterogeneity has not changed for sales response to business cycle over the last 30 years. However, firms' investment has become less sensitive on average, but, at the same time, firms with different weights now respond more homogeneously. Moreover, we show that both within-sector and across- sector heterogeneity equally contribute to the aggregate role of firm heterogeneity. This indicates that sectors with larger economic shares exhibit lower sensitivities in absolute terms, but firms with larger shares within each sector also exhibit lower sensitivities relative to the sectoral average. In addition, using the estimates from a linear panel model, we show that ignoring nonlinearities in firm sensitivities tend to misestimate the economy's sensitivity to business cycle fluctuations by 0.1 and 0.3 p.p., entirely driven by a bias in the role of firm heterogeneity.

Lastly, we extend our result to other exogenous shocks and financial variables, showing that the aggregate effect of firm heterogeneity depends on the outcome variable. First, we study how heterogeneity shapes the response of sales, investment, and debt issuance to two identified exogenous shocks – monetary policy and oil price shocks (Bauer and Swanson, 2023; Känzig, 2021). Differently from business cycle fluctuations, we find that financial characteristics such as leverage and distance-to-default are especially relevant for explaining the heterogeneity in firm sensitivity to monetary policy shock. On the contrary, we find a more balanced split between financial and non-financial characteristics, and more generally, a broader set of firm characteristics contributes to the heterogeneity in the sensitivity to oil price shocks. At the aggregate level, firm heterogeneity reduces the aggregate impact of the shocks, in line with its role on the effect of business cycle fluctuations. Second, we examine the role of heterogeneity for the response of stock market value, extending the analysis to financial variables in addition to real outcomes. In this case, the heterogeneity in firm sensitivity is driven mostly by financial variables such as distance to default, leverage, liquidity; however, differently from other shocks, the covariance term is positive, indicating that firm heterogeneity amplifies the aggregate response.

The remainder of the paper is organized as follows. Section 2 proposes an aggregation theory and provides information on the methodology we use. Section 3 presents the data used in the empirical application and the key results on the heterogeneity in firm-level sensitivities and its aggregate implication. Section 4 extends our result to other exogenous shocks and outcome variables. Section 5 concludes.

Literature. Our work contributes to the literature that studies the heterogeneity in firm-level sensitivity to aggregate shocks and its determinants. Among others, Ottonello and Winberry (2020), Jeenas (2018a), Gertler and Gilchrist (1994) and Jungherr et al. (2024) study the role of leverage and distance to default, liquidity, size, and debt maturity for the response of firm investment to monetary policy shocks, respectively; Gürkaynak et al. (2022) investigate how liquidity and leverage influence the response of market value; Gertler and Gilchrist (1994) and Crouzet and Mehrotra (2020) examine respectively the role of size and industry scope for the response of sales and investments to monetary policy and business cycle fluctuations; Covas and Haan (2011) and Begenau and Salomao (2019) study debt issuance by firm size over the business cycle. Prior research typically examines a single firm characteristic at a time and imposes linearity in how firm characteristics influence responses to aggregate shocks. In contrast, we depart from the standard linear panel regression approach and apply machine learning methods to incorporate a large set of firm characteristics simultaneously.[2] We show that the heterogeneity in firm sensitivities is driven by several characteristics and is highly non-linear with strong interactions among characteristics,

[2]For data limitation, we do not consider additional firm characteristics such as paying dividends (Farre-Mensa and Ljungqvist, 2016) or firm age (Cloyne et al., 2018).

underscoring the importance of a comprehensive analysis with a high dimensional characteristic space.[3]

Our paper contributes to the growing literature on the macroeconomic implications of firm-level heterogeneity–including differences in size, leverage, industrial sector, idiosyncratic shock process and debt maturity structure–for aggregate fluctuations and the transmission of shocks. Prominent contributions in this area include Cooley and Quadrini (2001), Cooley and Quadrini (2006), Buera and Moll (2015), Crouzet (2018), Ottonello and Winberry (2020), Deng and Fang (2022), Jaimovich et al. (2023), and Krusell et al. (2023), among others. Our findings on the dampening effects of firm-level heterogeneity on the transmission of shocks to macroeconomic aggregates are broadly in line with previous works. However, unlike most earlier contributions, which rely on quantitative macroeconomic models, we develop an aggregation framework in the spirit of Crouzet and Mehrotra (2020) and leverage the estimated distribution of firm-level sensitivities from the random forest model to gauge the impact of firm heterogeneity. Chang et al. (2024a), Chang et al. (2024b), and Lenza and Savoia (2024) offer an alternative approach based on functional VARs and heterogeneous VARs, which integrate aggregate variables with cross-sectional distributions to study their dynamic interactions. In contrast, our approach uses machine learning techniques to estimate firm-level sensitivities to aggregate shocks, which we then aggregate in a bottom-up framework to assess macroeconomic implications.

Lastly, this paper contributes to the rapidly growing literature applying machine learning techniques to economic analysis. Machine learning offers advantages both for the estimation of conditional average treatment effects and causal inference in high-dimensional settings (Athey and Imbens, 2017; Varian, 2014; Breiman, 2001) and for predicting outcomes to improve targeting and forecasting (Mullainathan and Spiess, 2017). Our work relates to the estimation of conditional average treatment effects using machine learning; however, few studies have applied machine learning techniques to examine firm-level heterogeneous sensitivity to aggregate shocks and macroeconomic

[3]Moreover, the macroeconomic literature has devoted limited attention to heterogeneity in firm-level sensitivity to oil shocks, despite these being key drivers of macroeconomic fluctuations (Känzig, 2021). Narayan and Sharma (2011) and Tsai (2015) study how a firm's market value reacts to oil shocks depending on its size and industry scope. We expand this literature by providing novel evidence on the heterogeneity in firm-level sensitivity to oil shocks, its determinants, and its aggregate impact.

fluctuations more broadly.[4] [5] [6] The closet study to our work is Paranhos (2024), which examines the relationship between firms' default risk and the effectiveness of monetary policy transmission to investment decisions, by generalizing standard local projection methods nonparametrically. Differently from their work, we apply random forest models to study firm heterogeneity in sensitivity to business cycle fluctuations and multiple firm outcomes, incorporating a high-dimensional firm characteristic space. Our findings highlight the strong quantitative role of interactions among characteristics in shaping firm-level sensitivities, indicating that multiple characteristics jointly drive heterogeneity in firm-level outcomes.

2 Firm Heterogeneity and Aggregate Fluctuations

We first present a theory of aggregation to compute the response of any aggregate variable to aggregate shocks by aggregating firm-level responses. We then discuss how machine learning algorithms provide significant advantages for the estimation of firm-level responses.

2.1 Decomposition Aggregate Variables

Consider a set I_t of firms continuing to operate between two consecutive periods t and $t-1$.[7] Let G_t and $g_{i,t}$ denote the aggregate and the firm-level response of variable Y_t following an aggregate shock W_t, respectively:

$$G_t(W_t) = \frac{Y_t}{Y_{t-1}} \qquad g_{i,t}(W_t) = \frac{Y_{i,t}}{Y_{i,t-1}} \equiv \beta_{it} W_t, \tag{1}$$

[4]Estimating conditional average treatment effects using machine learning is more common on the consumers and household side rather than on the firm side. For instance, Belloni et al. (2017) estimates the effect of 401(k) eligibility and participation on accumulated assets using local quantile treatment effects. Khazra (2021) explores the heterogeneity of house price elasticity of consumption using micro panel data via GRF (Athey et al., 2019), finding that neglecting local heterogeneities in elasticity leads to overestimating the total consumption response during housing market booms and busts.

[5]The forecasting advantages of machine learning have been explored in macroeconomics in relation to inflation forecast, with Paranhos (2025) and Nakamura (2005) both using neural networks to predict future inflation.

[6]Machine learning is more widely used in finance and asset prices; for instance, Freyberger et al. (2020) and Gu et al. (2020) use machine learning techniques to predict stock market returns and asset risk premiums, respectively, accounting for non-linearities and many characteristics.

[7]This framework can be readily extended to account for firms' entry and exit. We abstract from the extensive margin given that, in our empirical application, the sample of firms is fairly balanced.

where $\beta_{it} = \frac{\partial Y_{it}}{\partial W_t}$ is the sensitivity of firm i's outcome variable at time t to aggregate shock W_t. Let $\omega_{i,t-1}$ be the share of Y_{t-1} accounted for by firm i:

$$\omega_{i,t-1} = \frac{Y_{i,t-1}}{Y_{t-1}} \qquad \text{where } Y_{t-1} = \sum_{i \in I_t} Y_{i,t-1}. \tag{2}$$

It follows that the aggregate response of variable Y_t to an aggregate shock at time t is:

$$G_t = \sum_{i \in I_t} \omega_{i,t-1} \cdot g_{i,t}. \tag{3}$$

The aggregate response G_t is a weighted average of firm-level responses $g_{i,t}$, with weights $\omega_{i,t-1}$ given by each firm's share of the aggregate in the previous period. We decompose Equation (3) to understand how firm heterogeneity contributes to the aggregate response:

$$G_t = \overline{g_t} + \mathbb{C}\mathrm{ov}(w_{i,t-1}, g_{i,t}), \tag{4}$$

where the first term is the unweighted average response across firms, $\frac{1}{|I_t|} \sum_{i \in I_t} g_{i,t}$, and the second term is the covariance between firm sensitivity and firms' relevance in the aggregate, $\sum_{i \in I_t} \left(\omega_{i,t-1} - \frac{1}{|I_t|} \right) (g_{i,t} - \overline{g_t})$. The first term captures how, on average, firms respond to aggregate fluctuations without considering their relative relevance in the economy. The second term captures how heterogeneity in firms' sensitivities interacts with the heterogeneity in their weights.

Equation (4) suggests that heterogeneity matters for the aggregate response only when there is a systematic link between firms' relevance and response. For example, suppose Firm A produces 80% of output and responds weakly to a shock ($g_{A,t} = 1.02$), while Firm B produces 20% and responds strongly ($g_{B,t} = 1.12$). The unweighted average response is 1.07; however, the aggregate response is only 1.04. The reason is that the firm contributing the most to output is less sensitive than the average firm, so variations in output shares are negatively associated with sensitivities. Conversely, if the firm contributing the most were also the more responsive one, the aggregate response would exceed the average, as output shares are positively correlated with sensitivities. Thus, a negative covariance indicates that heterogeneity dampens the aggregate response to shocks, whereas a positive covariance indicates that it amplifies the response.[8]

[8] It is trivial to observe that, in the special case of atomistic firms, the distribution of sensitivities does not matter for the aggregate, since no single firm carries weight. In that case, only the average sensitivity determines the aggregate response.

2.2 Estimating Firm Sensitivities

While firm's share in the aggregate, $\omega_{i,t-1}$, can be readily measured in the data, firm-level responses $g_{i,t}$ depends on the underlying firm-level sensitivities β to aggregate shocks, which are not directly observed and must be estimated. Assessing the importance of firm heterogeneity for the aggregate hinges on carefully measuring the heterogeneity in β. We turn our attention to machine learning to flexibly estimating firm-level sensitivities.

We use the Generalized Random Forest (GRF henceforth) algorithm by Athey et al. (2019) to estimate firm-level sensitivities as a function of firm characteristics. Consider an empirical setting where we observe outcome variables and characteristics of a set of firms, indexed by i, over multiple periods, indexed by t. The outcome variable of interest, $Y_{i,t}$, represents firm-level outcomes such as sales growth, investment, or other key indicators. Let W_t denote the source of aggregate fluctuation that is common to all firms. Firm-level characteristics, $X_{i,t-1}$, can influence the sensitivity of $Y_{i,t}$ to W_t. GRF allows us to estimate the heterogeneous response of firm outcomes $(Y_{i,t})$ to an aggregate shock (W_t), conditional on a set of firm-level characteristics $(X_{i,t-1})$, in a nonparametric way:

$$Y_{i,t} \;=\; b\big(X_{i,t-1}\big) \,\cdot W_t \;+\; \varepsilon_{i,t} \quad , \quad \beta(x) \;=\; \mathbb{E}\big[b\big(X_{i,t-1}\big)\,\big|\, X_{i,t-1} = x\big], \tag{5}$$

where $\epsilon_{i,t}$ is an i.i.d. error term, b is a flexible function of firms' characteristics, and the sensitivity $\beta(x)$ is the conditional average effect of the aggregate shock W_t on the outcome $Y_{i,t}$ for firms with characteristics equal to x. The latter, which is our object of interest, is determined as follows in the GRF algorithm:

$$\hat{\beta}(x) = \frac{\sum_{i=1}^{n} \alpha_i(x)\left(W_i - \bar{W}_\alpha\right)\left(Y_i - \bar{Y}_\alpha\right)}{\sum_{i=1}^{n} \alpha_i(x)\left(W_i - \bar{W}_\alpha\right)}, \tag{6}$$

where, $\alpha_i(x)$ is a weight determined by the causal forest, $\bar{W}_\alpha = \sum_{i=1}^{n} \alpha_i(x)\, W_i$ is a weighted average of the shock, and $\bar{Y}_\alpha = \sum_{i=1}^{n} \alpha_i(x)\, Y_i$ is a weighted average of the outcome.

The GRF algorithm estimates $\beta(x)$ in two steps: first, it constructs a forest of decision trees designed to partition the data in a way that maximizes the heterogeneity in firms' responses to aggregate shocks; second, it estimates the conditional average effect using a locally weighted regression approach. In the first stage, GRF builds a collection of honest and adaptive decision trees that recursively split the data based on firm characteristics $X_{i,t-1}$. Unlike standard regression trees, which minimize prediction errors, each tree is constructed by selecting a random subsample of the data, and splits are determined by optimizing a criterion that prioritizes variation in the estimated conditional average effects. The algorithm is considered "honest" because it uses one subsample to determine optimal splits and a separate subsample to estimate conditional average effects, thereby mitigating overfitting. In the second stage, GRF estimates the conditional average effect $\beta(x)$ by aggregating information across trees. For a given firm with characteristics $X_{i,t-1} = x$, the algorithm identifies neighbouring firms that frequently appear in the same leaf across multiple trees. Each observation is then assigned a weight $\alpha_i(x)$ based on how often it appears in the same leaf as the hypothetical firm x. Using these weights, GRF estimates $\beta(x)$ via a weighted regression of firm outcomes $Y_{i,t}$ on aggregate shocks W_t, ensuring that identification relies on variation in W_t within locally homogeneous subgroups. The way weights are constructed, which are determined by the structure of the causal forest, ensures that $\hat{\beta}(x)$ is locally smoothed and not overly sensitive to a single partition.[9]

Advantages over Linear Panel Model. The standard econometric framework for estimating heterogeneity in firms' sensitivity to aggregate shocks is a linear panel model, which poses linearity and suffer from dimensionality constraints. By contrast, the GRF agnostic approach to the function $b\big(X_{i,t-1}\big)$ allows it to account for non-linear, flexible relationships in the marginal effects of shocks, accommodating a complex, high-dimensional firm characteristic space.

In a linear panel model (LPM henceforth), the heterogeneous response of firms' outcomes ($Y_{i,t}$) to an aggregate shock (W_t), conditional on a set of firm-level characteristics ($X_{i,t-1}$) is estimated as follow:

$$Y_{i,t} = \alpha + \beta_0 \cdot W_t + \beta_1 \cdot X_{i,t-1} + \beta_2' \left(W_t \cdot X_{i,t-1}\right) + \epsilon_{i,t}, \tag{7}$$

where $\epsilon_{i,t}$ is an i.i.d. error term, and firm characteristics are predetermined at $t-1$. The parameter vector of interest, β_2, captures how firms' sensitivity to aggregate shocks varies with firm characteristics. The marginal effect of the aggregate shock W_t on firm outcomes is given by $\beta_0 + \beta_2' X_{i,t-1}$, which depends linearly on $X_{i,t-1}$.

The GRF algorithm in Equation (5) offers two advantages over the standard linear panel regression in Equation (7), making it particularly well-suited for estimating the heterogeneous effects of aggregate shocks. First, the linear panel regression model assumes that firms' characteristics linearly influence firm sensitivity to aggregate shocks. However, this linearity assumption may be restrictive and could lead to misspecification if the nonlinear component of heterogeneity is significant. While the linear panel regression model can incorporate more complex forms of heterogeneity by including polynomial terms in firm characteristics, it remains a parametric approach that requires the econometrician to take a stance on the unknown forms of non-linearities, making the LPM vulnerable to errors from model misspecification. In contrast, GRF explores the covariate space non-parametrically, adaptively detecting intricate relationships without requiring a pre-specified form. Second, GRF can efficiently handle a high-dimensional characteristics space, automatically putting more weight on the most important covariates. This feature of the GRF mitigates the curse of dimensionality inherent in models with large sets of covariates and interactions. Enumerating all possible pairwise (or higher-order) interactions in a linear model quickly leads to

[9] Additional technical details on the GRF algorithm are in Appendix A.1.

over-parameterization and multicollinearity, while GRF adaptively partitions the data, freeing the researcher from having to manually specify functional forms or interactions. Thus, the linearity and parametric features of the linear panel regression model become more restrictive in the presence of high-dimensional characteristics space. We illustrate the relative performance of the two models in the Monte Carlo simulation exercise in Appendix A.2.[10]

3 Application to U.S. Firms

We estimate the sensitivity of firm sales, investment, and debt issuance to business cycle fluctuations, conditional on eight firm-level balance-sheet characteristics, applying the GRF algorithm on firm-level data from the U.S. over the period 1990–2019. We begin by describing the data, followed by an analysis of the estimated firm-level sensitivities and implications for aggregate response.

3.1 Data and measurement

Our primary data source is the quarterly Compustat dataset, which provides comprehensive financial statement information for publicly listed companies in the U.S. All variables are deflated using the implied price index of gross value added in the U.S. non-farm business sector. Outcome variables are constructed as one-year percentage changes, thus we lag the firm characteristics by four periods in the empirical application. The final dataset includes 220,259 firm-quarter observations spanning from 1990 Q1 to 2019 Q4.

Measurement. We investigate the heterogeneous sensitivity of firm outcomes to business cycle fluctuations. We proxy business cycle fluctuations by the annual percentage change in real GDP following Crouzet and Mehrotra (2020).[11] The set of firm outcome variables includes: annual real sales growth, debt issuance (measured by the one-year percentage change in short- and long-term debt), and the investment rate (measured as the one-year percentage change in capital stock using the perpetual inventory method). To account for heterogeneity in firm sensitivities, we consider the following set of eight balance-sheet characteristics, grouped into financial and non-financial variables. Non-financial characteristics include firm size (measured by the logarithm of total assets), industry scope (captured by NAICS 5-digit industry codes), ten-years sales volatility,

[10] GRF's flexibility comes with trade-offs, such as the potential loss of precision in smaller samples and reliance on careful hyperparameter tuning. When the true relationship between covariates and the conditional effect of shocks is linear–or can be sufficiently well captured by a modest set of polynomial terms– a linear panel regression model perform comparably to GRF. Appendix A.2 provides additional details.

[11] To account for potential changes in trend, we divide real GDP by the population trend before computing the growth rate.

and firm profitability (measured by return on assets, ROA). Financial characteristics include the liquidity ratio (cash-to-total assets), the leverage ratio (total debt-to-total assets), distance to default (Merton, 1974), and debt liquidity (measured by the proportion of short-term debt to total debt). These firm-level balance-sheet characteristics have been widely used in the literature to study the heterogeneity in the transmission of aggregate shocks onto firm outcomes.[12] Appendix B presents summary statistics of all firm-level variables used in the empirical analysis. Notably, Table 3 in Appendix B reports the pairwise correlation between all independent variables, showing that, although some correlation exists among firm-level characteristics, they provide distinct information along different dimensions.

Estimation details. In the GRF algorithm, we set the number of trees in the forest to 2,000 with equal weighting. We use honest splitting for sub-sample partitioning, allocating 50% of the data to build each tree and ensuring a minimum of five observations per tree leaf. Observations are clustered at the firm level with equal weight, so firms with more observations receive greater weight, thereby reducing the influence of entry and exit. Splitting is allowed across all characteristics, with the tuning parameter controlling the maximum imbalance of a split set at 0.05. We do not include time fixed effects, as our objective is to estimate the average unconditional effects of aggregate shocks on firm outcomes. In the case macroeconomic confounding factors are a concern, macroeconomic variables can be partialled out before estimation. We also omit firm fixed effects because one of our primary interest is in assessing the role of industry scope – which is constant at the firm level – in driving and explaining the response of outcome variables and their heterogeneity across firms. Following best practices, the data are "centered" before GRF estimation takes place: this step involves differencing out the effect of the firm-level characteristics on the outcome variables. This is done to ensure that the GRF model captures the effect of the aggregate shocks on the outcome variables, conditional on the firm-level characteristics, rather than the effect of the firm-level characteristics themselves. We effectively estimate the random forest on centered variables $\tilde{Y}_i = Y_i - \hat{y}_i^{(-i)}(X_i)$ and $\tilde{W}_i = W_i - \hat{w}_i^{(-i)}(X_i)$, where $\hat{y}_i^{(-i)}(X_i)$ and $\hat{w}_i^{(-i)}(X_i)$ are leave-one-out estimates of marginal expectations, computed without the i-th observation.[13] We also estimate the LPM counterpart to the GRF specification using OLS. In this case, we include the interaction between the aggregate shock, W_t, and industry scope, while absorbing the level of industry scope

[12] For instance, Ottonello and Winberry (2020), Cloyne et al. (2018), and Jeenas (2018a) study the role that distance to default, leverage and liquidity play in the transmission of monetary policy shocks to investment, respectively. Similarly, Alfaro et al. (2024) studies the effects of uncertainty on firms' financial variables such as liquidity and leverage, while Crouzet and Mehrotra (2020) focuses on how size and industry scope impact the response to business cycle fluctuations.

[13] Athey et al. (2019) note that the performance of the forests improves with this procedure, and that the estimator $\hat{\beta}(x)$ is more robust to confounding effects. Chernozhukov et al. (2018) also apply a similar orthogonalization procedure.

to reduce computational burden.

Notice that our baseline analysis using GDP growth does not establish causality as firm sensitivities reflect co-movement with the cycle, not the causal effect of real output fluctuations. The results should thus be read as a descriptive characterization of heterogeneity in cyclical exposure. Moreover, the presence or absence of differential sensitivities should not be interpreted as direct evidence either supporting or refuting the role of financial frictions or the financial accelerator mechanism. Instead, our goal is to document which characteristics are most strongly associated with heterogeneity in sensitivities, without implying that these variables represent definitive or superior proxies for financial constraints.

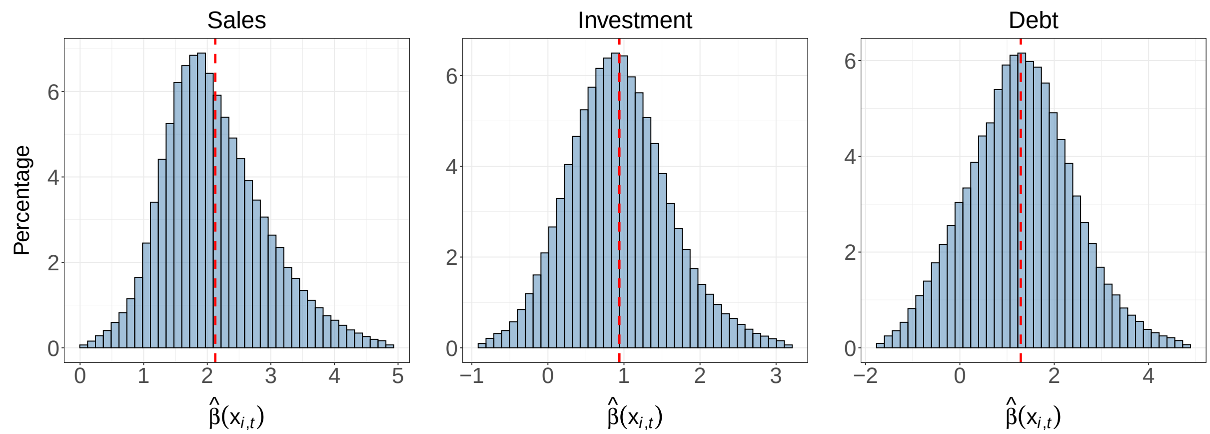
3.2 Estimated Firm-level Sensitivities $\hat{\beta}$

We estimate extremely rich cross-sectional heterogeneity in firms' sensitivity to business cycle fluctuations. Figure 1 shows the distribution of firms' sensitivities to business cycle for each outcome variable estimated using the GRF algorithm. We estimate that a 1 percent increase in real GDP growth is associated, on average, with a 2.1 percent increase in firms' sales, closely matching the 3% reported by Crouzet and Mehrotra (2020) using QFR establishment-level data. Similarly investment and debt issuances increase by 0.9 percent and 1.2 percent, respectively, also in line with economic intuition and previous estimates. However, the average sensitivity masquerades substantial cross-sectional heterogeneity in how firm outcomes co-move with aggregate fluctuations. A large share of firms exhibit sensitivities that are more than twice as large as the median; in the case of investment and debt issuance, some firms even exhibit $\hat{\beta}$ with opposite sign relative to the average, ranging up to -1 and -2, respectively. We use the Chernozhukov et al. (2018) test to assess the statistical significance of the heterogeneity in $\hat{\beta}$ by comparing the prediction from the forest to the predictions that consider only the average effect.[14] We find that we can strongly reject the null hypothesis of no heterogeneity in estimated β for all outcome variable (Figure 14 in Appendix C.1).

Our estimated $\hat{\beta}$ diverge from those obtained using the linear models mainly because of the role that nonlinearities and interactions among firms' characteristics play in driving firms' responses to the business cycle fluctuations. Table 4 in Appendix C compares the individual firm-level sensitivities estimated by GRF to those obtained from the linear panel model. Across all cases, the average sensitivity estimated by GRF and the LPM are statistically identical, suggesting that both methodologies quantify the same average effects of business cycle fluctuations on outcome variables. However, there are substantial differences between GRF and the linear model estimates at the firm-level, particularly in the tails of the distribution. These differences, which can be as large as 100 percent in magnitude and even opposite in sign, suggest that firm characteristics influence the effect

[14] Appendix C.1 provides details on the construction of the test.

Figure 1: Estimated $\hat{\beta}$ to Business Cycle



Notes: The Figure shows the distribution of firm-level sensitivities to business cycle fluctuations estimated using the GRF algorithm. Each subplot represents a specific outcome variable. The vertical dashed line indicates the average sensitivity. Firm-level sensitivities are trimmed at the 0.5% level on both tails.

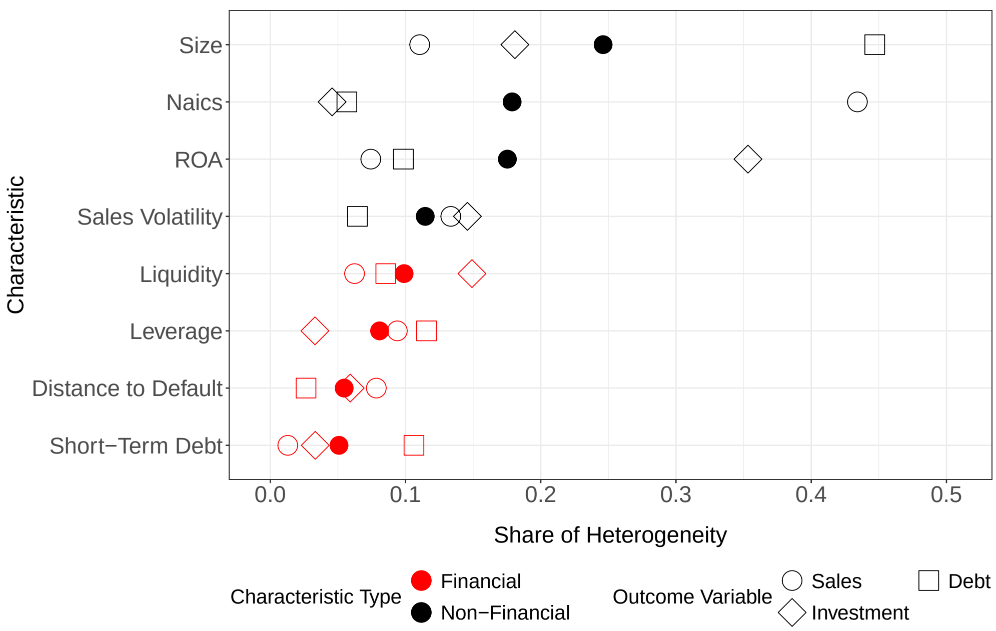
of business cycle fluctuations in complex and nonlinear ways that the LPM fails to capture.[15]

Drivers of Heterogeneity in $\hat{\beta}$ We evaluate the role of firm characteristics in shaping the heterogeneity in firms' sensitivity to business cycle fluctuations using machine learning tools. Unlike traditional parametric models, the GRF algorithm estimates firm-level sensitivities without imposing a predetermined functional form, assigning greater weight to the most relevant covariates. Therefore, we measure the importance of each characteristic for heterogeneity through its role in the moment function, captured by the proportion of splits in which it appears. In a random forest (Breiman, 2001), this corresponds to the depth-weighted frequency of splits where the characteristic is used. This metric measures the contribution of a characteristic X to the heterogeneity in β as it indicates how frequently the algorithm relies on a given characteristic when partitioning the data. Figure 2 reports the importance of each characteristic in explaining the heterogeneity in sensitivities across all outcome variables.

We find that non-financial characteristics—especially industry scope and firm size—are the primary drivers of the heterogeneity in firm sensitivities to the business cycle fluctuations, while financial characteristics contribute little. Industry scope is the dominant driver of the heterogeneity in the sensitivity of sales, accounting for about 45 percent of the algorithm's splits, while firm size accounts for nearly half of the splits used to explain debt issuance and about one-fifth of those for

[15] Figure 13 in Appendix C reports the distribution of errors, defined as the percentage deviation between GRF and linear panel sensitivities, for each outcome variable.

Figure 2: Importance of Individual X_{it} for Heterogeneity in $\hat{\beta}$



Notes: The Figure shows the share of heterogeneity explained by each characteristic across different outcome variables. The share of heterogeneity explained by each characteristics is computed as the depth-weighted frequency of splits in the forest where the characteristic is used. The characteristics on the y-axis are ordered by their average importance share across outcome variables, with filled points representing the average importance share of each characteristic. "Financial" characteristics are depicted in red, while "Non-Financial" characteristics are shown in black. Unfilled shapes represent the importance share for individual outcome variables: circles represent sales, squares represent debt issuance, and diamonds represent investment. The x-axis shows the importance share, where a value of 0.01 corresponds to 1 percent of total heterogeneity.

investment. Taken together, non-financial characteristics guide most of the algorithm's partitioning decisions—roughly 76 percent for sales, 87 percent for investment and 67 percent for debt issuances. Financial variables such as leverage, liquidity, or distance-to-default play instead only a marginal role, exhibiting individual shares lower than 15 percent. These findings are consistent with the conclusions of Crouzet and Mehrotra (2020), who emphasize that firms' sensitivity to business cycle fluctuations is driven by demand-side factors, rather than the cost of external finance.[16]

[16] In Appendix C, we show that quantitatively similar results holds when we use Shapley values to measure the contribution of each characteristics to the heterogeneity in sensitivities. The game-theoretic approach based on Shapley values allows to measure the marginal contribution of a firm characteristic by computing the difference in sensitivity with and without that characteristic. To quantify the average importance of a characteristic to firms' sensitivities, we follow standard practice and compute the mean absolute value of the estimated Shapley values over a grid of 100 points corresponding to the characteristic's percentiles. Figure 19 in Appendix C shows a positive correlation between the share of heterogeneity explained by one

Relationship between $\hat{\beta}$ and Firm Characteristics We use Accumulated Local Effect (ALE) plots to illustrate how the marginal effect of each characteristic on predicted firm sensitivities varies across its distribution. A key advantage of the GRF algorithm over traditional linear panel models (LPM) is that it estimates firm-level sensitivities without imposing a predetermined functional form on how firm characteristics influence them, thereby allowing for nonlinearities and interactions among variables in X_{it}. ALE plots provide a way to visualize these relationships while accounting for correlations among all other characteristics. They display how predicted sensitivities change locally as each feature varies, holding the joint distribution of the remaining variables constant.[17] Steeper ALE slopes indicate stronger marginal effects, whereas flatter segments denote regions where the characteristic has little influence on firm sensitivity. Figure 3 presents the ALE plots for the three most important features by importance share for each outcome.

Firm size, measured by total assets, significantly influences firms' sensitivities in sales, investment, and debt issuance after controlling for all other characteristics. Small and medium-sized firms are less sensitive to business-cycle fluctuations in terms of sales, but they tend to increase investment and issue more debt during booms.[18] [19] Financial characteristics also play an important role: firms with higher leverage are predicted to experience lower increases in sales relative to the average firm, while firms with higher liquidity, by contrast, exhibit stronger increases in investment compared to the average firm during expansions. Similarly, we find that investment responses are larger for firms operating with very negative ROA, suggesting stronger adjustment dynamics among financially weaker firms. Finally, firms with shorter debt maturities and lower leverage tend to expand borrowing more strongly in booms. These patterns point in the opposite direction to the financial accelerator literature (Bernanke et al., 1999; Kiyotaki and Moore, 1997), as firms with weaker balance sheets are not more sensitive to relaxing financial constraints.[20] Figure

characteristic and the Shapley-based characteristic relevance.

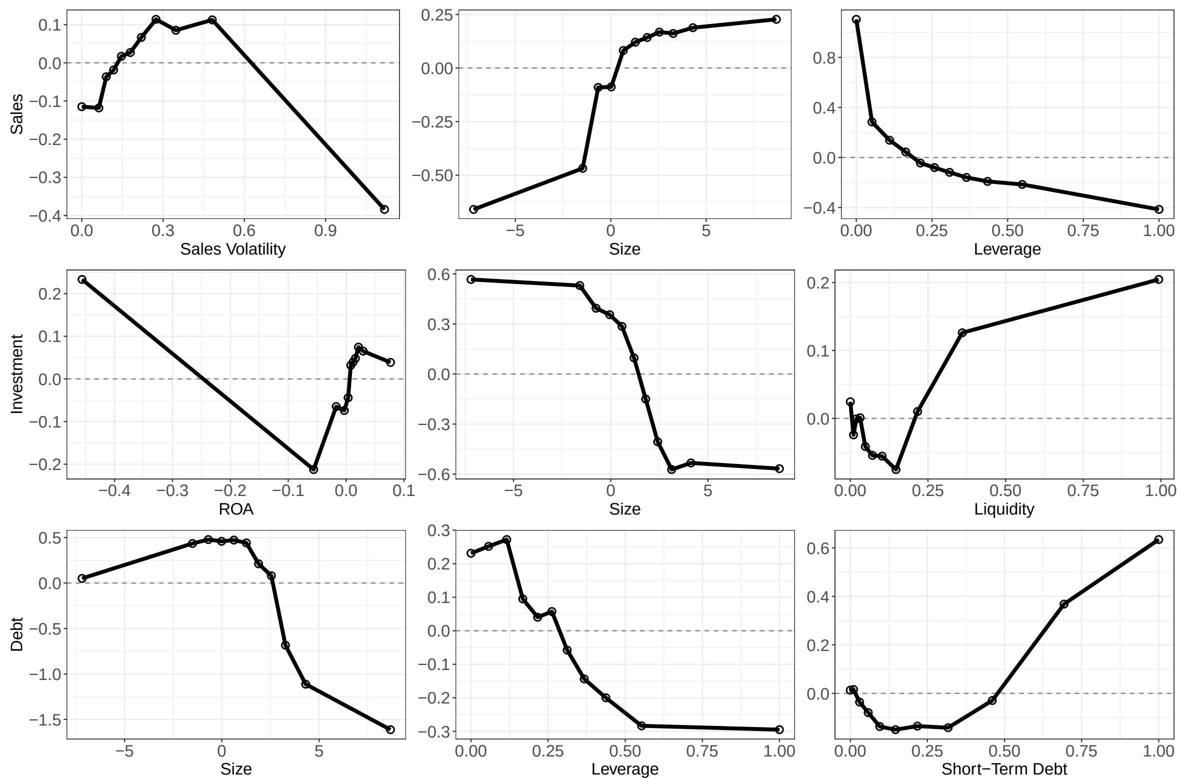
[17] ALE plots offer a more reliable alternative to the commonly used Partial Dependence Plots (PDP). PDPs assume that characteristics are independent, an assumption that does not hold in our empirical setting. In contrast, ALE plots compute local effects within small intervals of a feature's observed range while conditioning on the joint distribution of other variables, thus capturing the true marginal effect of a feature in the presence of correlated covariates.

[18] A potential explanation is that during booms, smaller firms may prefer to expand physical investment in capital stock rather than invest in customer acquisition or intangible assets. This interpretation is consistent with a different role of firm size when comparing the responses of sales and investment (Peters and Taylor, 2017; Favara et al., 2024).

[19] Our results relate to the recent evidence in Crouzet and Mehrotra (2020), who document that smaller firms are more sensitive to business cycle fluctuations. In contrast, we find that sales of larger firms are more cyclical. This difference likely reflects the fact that we control for a richer set of firm characteristics and focus on a restricted sample of publicly listed firms.

[20] In models such as Bernanke et al. (1999) and Kiyotaki and Moore (1997), firms with weaker balance sheets face tighter external financing constraints, which amplify the effects of aggregate shocks on investment and borrowing when financial conditions ease. However, we do not find evidence consistent with these mechanism as firms with higher liquidity ratios and/or firms lower leverage, and shorter debt maturities are

Figure 3: Marginal Effect of Top 3 Individual X_{it} on $\hat{\beta}$



Notes: The Figure presents the Accumulated Local Effects (ALE) plots estimated for each outcome and the three most important characteristics based on their share of importance. Each row corresponds to one outcome variable. The x-axis represents the value of the characteristic, and the y-axis shows the average deviation in the predicted firm sensitivity from the overall mean sensitivity to business-cycle fluctuations. Positive (negative) values indicate firms that are predicted to be more (less) sensitive than the average firm at that point in the distribution of the characteristic.

20 in Appendix C shows the ALE plots for all the other characteristics.

Moreover, firm characteristics influence firms' sensitivities to business-cycle fluctuations in a non-linear way. The ALE plots in Figure 3 display several kinks and changes in slope, indicating that the marginal effect of a characteristic on predicted sensitivity varies sharply across different regions of its distribution.[21] For instance, we find that the marginal effect of size on all outcome variables is steep in the middle of the size distribution and strongly muted at its tails. Similarly, the marginal effect of leverage on sales and debt issuance decreases with firm leverage, while the

those that display the largest increases in investment and borrowing during booms.

[21]The presence of non-linearities in the relationship between firm characteristics and sensitivities is confirmed by several statistical tests, including the Generalized Additive Model (GAM), the Harvey-Collier, and the RESET tests. Additional details are reported in Table 5 in Appendix C.

marginal effects of liquidity and the share of short-term debt exhibit inverted-U patterns. These nonlinearities, which would be ruled out under standard linear specifications, highlight how the distribution of firm characteristics can shape both firm cyclicality and aggregate dynamics.

Finally, Figure 15 in Appendix C shows that a specific form of non-linearities, interactions among characteristics, account for up to 40% of the variance in the outcome variable explained by a given characteristic. We quantify the strength of these interactions relying on Friedman's H-statistic, which measures whether the joint contribution of two or more characteristics to predicted sensitivities exceeds the sum of their individual effects. This approach allows us to assess how much explanatory power arises from interactions rather than from additive effects alone.[22] Firm size emerges as the characteristic with the highest or second-highest H-statistic across the three outcome variables, suggesting that a large part of its relevance stems from its influence on the effect of other characteristics. More broadly, non-financial characteristics display stronger interactions than financial ones, consistent with their key role in driving the heterogeneity of $\hat{\beta}$s.[23] These results support the importance of including high-dimensional characteristic space and flexible machine learning techniques to measures the drivers of $\hat{\beta}$.

3.3 Aggregate Implications of Firm-level Heterogeneity

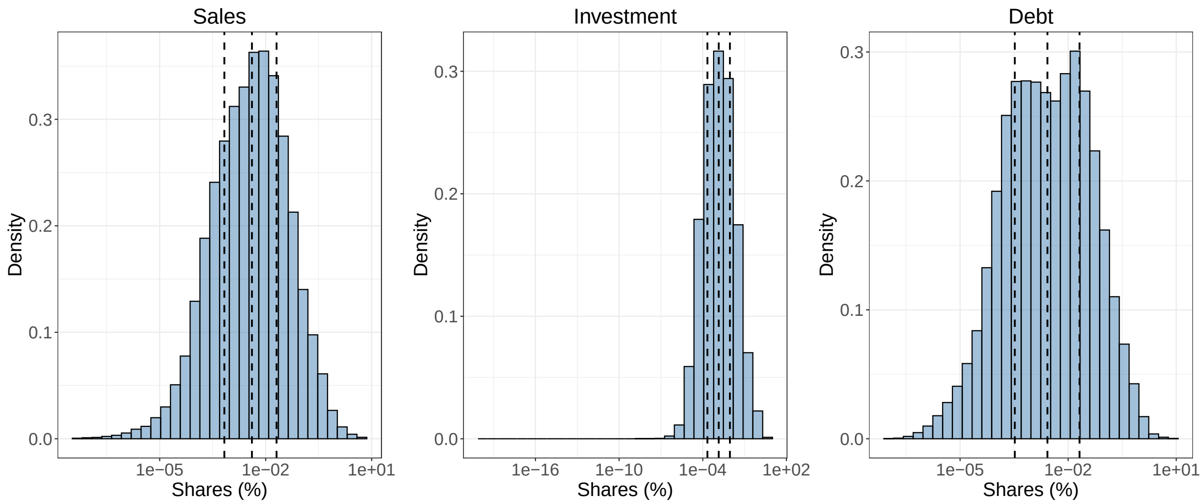
With the set of firm-level sensitivities estimated using machine learning, we can now apply the theory presented in Section 2 to gauge the relevance of firm heterogeneity for the aggregate economy. Figure 4 shows that the distribution of firms' weights, w, is highly unequal, with a small number of firms accounting for a disproportionately large share. Depending on whether larger firms—those accounting for a greater share of sales, investment, or debt issuance—tend to be more or less sensitive than the average firm, the aggregate impact of business cycle fluctuations can be muted or amplified by firm heterogeneity.

Estimating Average Aggregate Response In order to quantify the aggregate implications of the heterogeneity in firm-level sensitivities, we first construct the aggregate response, $\hat{G}_t$, by weighting the predicted firm-level responses, $\hat{g}_{i,t}$, based on the estimated sensitivities $\hat{\beta}$, by each firm's share of a total outcome measure, $w_{i,t-1}$, using Equation (1). We then estimate the average

[22]Formally, the H-statistic compares the share of variance in sensitivities explained by interactions to the total variance explained by the characteristics, ranging from zero (purely additive) to one (entirely joint effects). We compute both total interactions—capturing a characteristic's joint effect with all others—and two-way interactions between specific pairs.

[23]Figure 18 in Appendix C reports the ten most significant characteristic pairs for each outcome variable, indicating generally diffuse patterns without a single dominant pair.

Figure 4: Distribution of the shares of outcome variables



Notes: This Figure presents the distribution of firm-level shares across different outcome variables. The x-axis represents the firm-level share on a log scale, while the y-axis denotes the density. The vertical lines indicate the first, second, and third quartiles of the distribution.

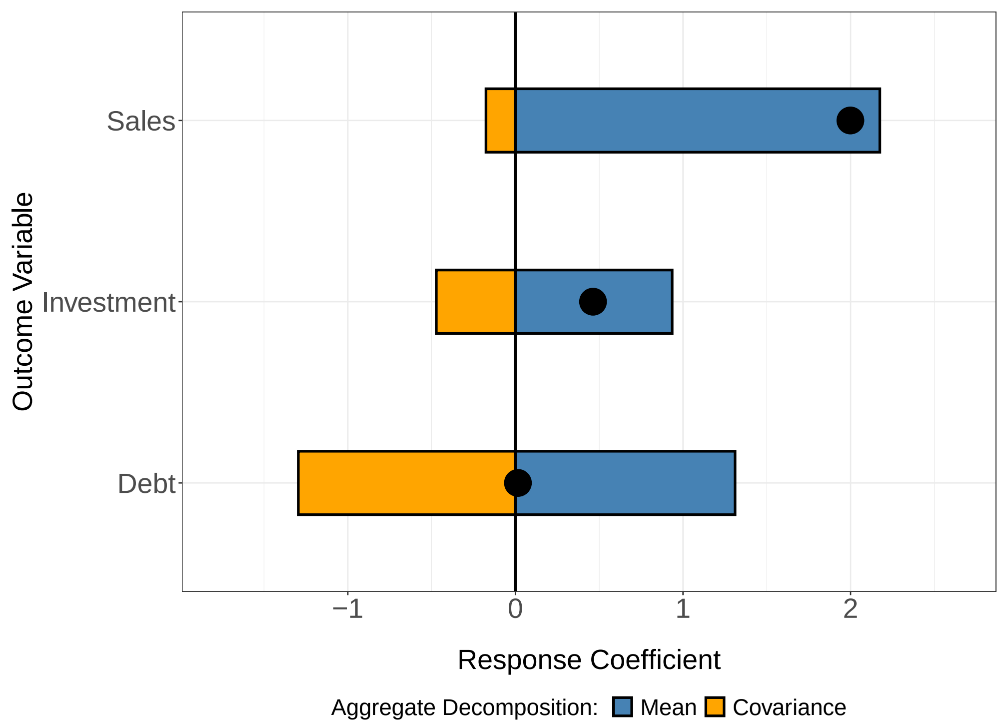
aggregate effect of business cycle fluctuations using the following time-series regression:

$$\hat{G}_t = \alpha + \gamma W_t + \epsilon_t, \tag{8}$$

where the coefficient γ reflects the average aggregate effect on a given outcome variable of a one percent change in the business cycle. This coefficient captures the effect of both the average sensitivity of firms and the contribution of firm-level heterogeneity. In order to separate the contributions of the average and covariance terms, we apply the decomposition from Equation (4) and separately regress the two terms on W_t using the time-series regression in Equation (8).

Decomposition of Aggregate Response We find that firm heterogeneity dampens the aggregate response to the business cycle fluctuations, but its magnitude depends on the outcome variable. Figure 5 shows that the unweighted average firm response to the business cycle aligns with economic intuition, as sales, investment and debt issuance increase over the business cycle. However, firm heterogeneity reduces aggregate responses in sales and investment by about 6% and 53%, respectively, while the response of debt issuance is almost completely offset. The modest dampening in sales reflects that larger firms—which account for most output—are only slightly less cyclical than the average firm, implying a relatively small covariance term. The dampening in investment and debt is more pronounced. For investment, a plausible economic explanation is that firms with larger weights in the aggregate are typically closer to their optimal scale of operation,

Figure 5: Decomposition of Aggregate Response

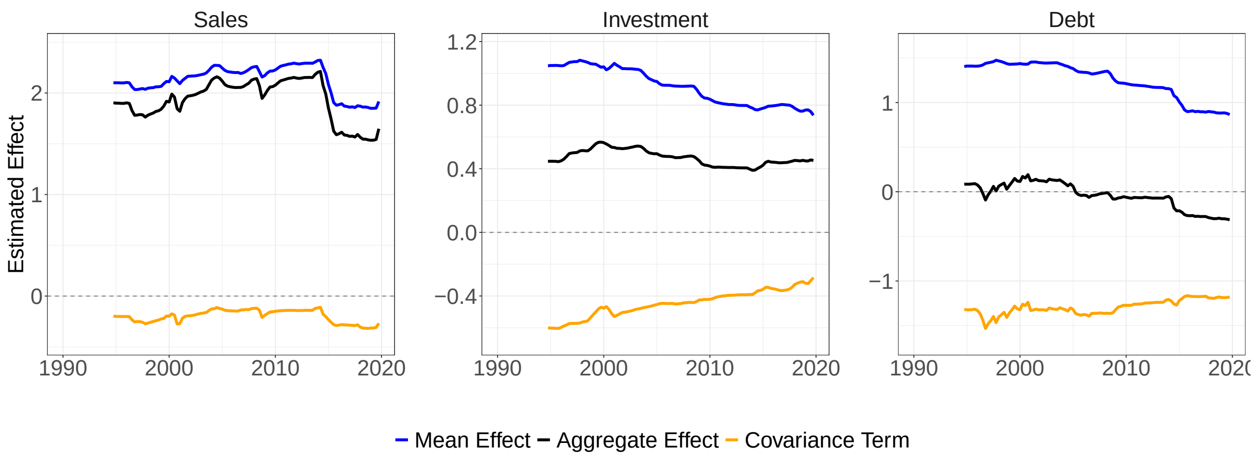


Notes: The Figure illustrates the decomposition of aggregate responses into mean and covariance terms (bars) for each outcome variable and aggregate shock. The black point denotes the total average aggregate response. We estimate Equation (8) using the mean and covariance terms in Equation (4) as dependent variable. The mean and covariance terms are constructed using the set of firm-level sensitivities estimated with the GRF algorithm.

which reduces the marginal benefit of additional investment and limits cyclical fluctuations (Cooper and Haltiwanger, 2006; Winberry, 2021). In contrast, for debt, highly leveraged firms are less likely to issue new liabilities because high debt levels amplify financial frictions, thereby constraining borrowing even under favorable conditions (Hennessy and Whited, 2005). These results have policy implications as firm heterogeneity acts as a natural stabilizer for the aggregate economy.

We study whether the estimated average aggregate responses and its decomposition remain stable over time given the rising concentration (De Loecker et al., 2020) or potential changes in firm sensitivity to business cycle fluctuations. We re-estimate the time-series framework in Equation (8) and decompose the coefficients into mean and covariance terms using a five-year rolling window. Figure 6 shows that, overall, the dynamics of both components are relatively stable only for sales. The covariance term for investment has gradually declined, implying that heterogeneity in sensitivities has become less important for aggregate investment dynamics over the last 30 years. This decline, however, is offset by a change in the mean sensitivity of similar

Figure 6: Decomposition Over Time



Notes: The Figure illustrates the mean and covariance decomposition of the average aggregate response across all outcome variable, utilizing a five-year rolling window version of Equation (8). We estimate the time-serie model with the mean and covariance components, as defined in Equation (4), serving as the dependent variable. Each point in the time series represents the corresponding coefficient estimate, derived from a sample ending at the respective quarter and spanning the preceding five years. The mean and covariance components are calculated based on the set of GRF firm-level sensitivities.

magnitude but opposite sign, leaving the aggregate response largely unchanged. This indicates that, firms' investment has become less sensitive in an unweighted sense, and firms with different weights now respond more homogeneously. This pattern is consistent with evidence that firm-level investment volatility has declined over time, even as aggregate volatility has remained broadly stable (Davis and Kahn, 2008; De Veirman and Levin, 2018).[24] We also observe that, while the covariance term remains broadly stable, the mean sensitivity of debt issuance has declined in the last 15 years. As a result, the corresponding overall aggregate response to business cycle has become less cyclical.

Within and Across Sector Heterogeneity We show that both within-sector and across-sector heterogeneity equally contribute to the dampening of the aggregate response due to firms' heterogeneity. To illustrate this, we consider a counterfactual scenario where the sensitivity of each firm is set to the median sensitivity of all firms within the same sector for a given quarter, where sectors are defined as 5-digit NAICS industries. We then construct a counterfactual aggregate response using these counterfactual sensitivities and the aggregation theory. Re-estimating the time-series framework in Equation (8), we obtain a counterfactual average aggregate response that

[24]See also McConnell and Perez-Quiros (2000), Stock and Watson (2002), and Bachmann and Bayer (2014).

accounts only for across-sector variation in firms' sensitivities. We assess the relative importance of within-sector and across-sector heterogeneity in firms' sensitivities by comparing the counterfactual average aggregate response coefficients and their decomposition into mean and covariance terms with those obtained in the benchmark case.

Figure 23 in Appendix D shows that accounting only for sectoral heterogeneity reduces by half the effect of firms' heterogeneity on the average aggregate response. Not surprisingly, the mean effects estimated when setting firms' sensitivities equal to the median sensitivity with each sector are quantitatively and statistically identical to the benchmark case, as the average effect is usually well approximated by the average sensitivity across firms. However, the covariance term estimated in the counterfactual case is approximately half of the covariance term estimated in the benchmark case across all scenarios. This indicates that the aggregate effect of firms' heterogeneity in Figure 5 is equally due to both within-sector and across-sector variation. In other words, sectors with larger economic shares exhibit lower sensitivities in absolute terms, but firms with larger shares within each sector also exhibit lower sensitivities relative to the sectoral average. The fact that both margins of heterogeneity are equally significant for the aggregate response underscores the importance of accounting for both dimensions of heterogeneity.

Aggregate Role of Non-Linearities in Sensitivity We find that non-linearities in firm-level sensitivities are not only prevalent at the micro level but also significantly influence the aggregate response of outcomes to macroeconomic fluctuations. To evaluate their aggregate impact, we construct the economy-wide response using firm-level sensitivities estimated from both GRF and the LPM. We then compare the average aggregate effect of a shock, γ from Equation (8) across methods. Table 6 in Appendix D shows that the differences between the average aggregate response estimated via GRF and LPM are statistically significant and economically relevant, indicating that non-linearities in firms' sensitivities play a crucial role in shaping the aggregate response to macroeconomic fluctuations. We find that the direction of the bias depends on the outcome variable, with LPM predicting stronger responses of sales and debt issuance to an increase in GDP by 0.3 and 0.1 p.p., respectively, but weaker growth of investments by 0.2 p.p.. Furthermore, Figure 24 in Appendix D shows that the primary driver of the differences in the average aggregate effects obtained from the two methods is the difference in the covariance terms, as both methods estimate similar average sensitivities. These findings suggest that non-linearities in firm sensitivities play a key role in shaping the aggregate role of firm heterogeneity, underscoring the importance of estimating firm-level β allowing for non-linearities in characteristics.

4 Additional Results

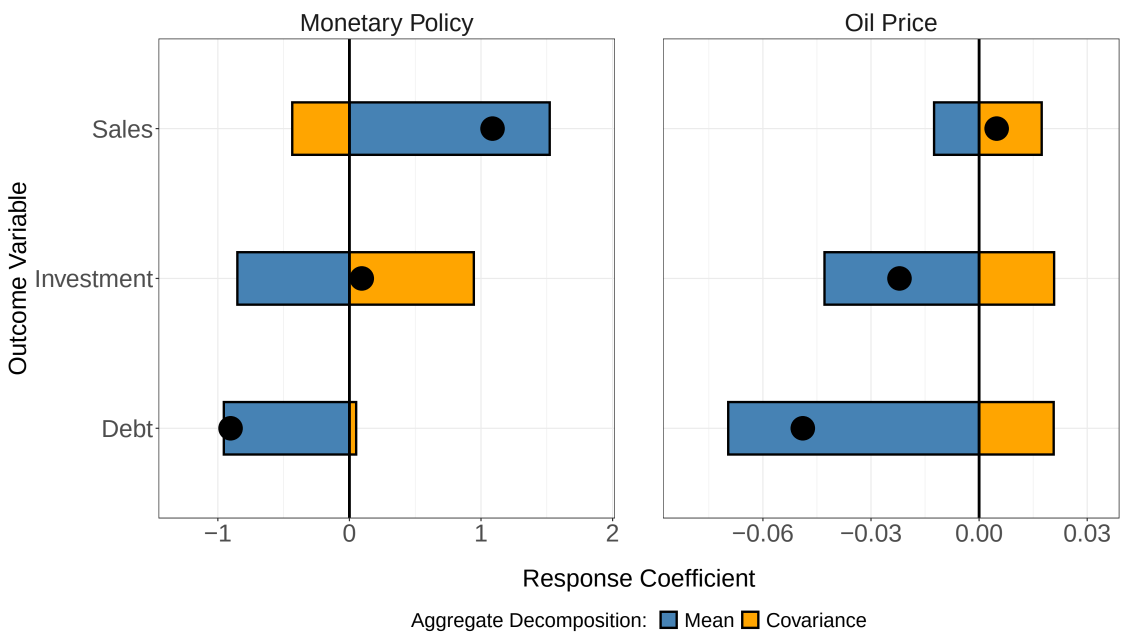
We extend our previous analysis in two directions. First, we study how heterogeneity shapes the response of sales, investment, and debt issuance to two identified exogenous shocks–monetary policy and oil price shocks. Our earlier results were based on business cycle fluctuations, which, by construction, reflect the interaction of multiple unobserved demand and supply disturbances. That approach was informative about the overall importance of heterogeneity, but it did not provide causal estimates. By focusing on monetary policy shocks, widely interpreted as demand-driven, and oil price shocks, commonly regarded as supply-driven, we can isolate more clearly the channels through which heterogeneity matters for aggregate dynamics. Second, we examine the role of heterogeneity for the response of additional outcome variables such as stock market value, extending the analysis to financial variable in addition to real outcomes.

4.1 Response to Exogenous Shocks

We study the implications of firm heterogeneity for the aggregate response to exogenous monetary policy and oil price shocks. Monetary policy shocks are measured using interest rate surprises around Federal Reserve announcements, identified using high-frequency variations in the 3-month federal funds rate futures, and cleaned of past aggregate fluctuations as in Bauer and Swanson (2023). Oil price shocks are proxied with high-frequency changes in oil supply expectations around OPEC announcements from Känzig (2021). To normalize the size of the shocks, we use them as instruments for a set of endogenous variables. Using the exogenous variables as instrument imposes a unit effect normalization of the shocks in terms of a one-unit change in the endogenous variable (Stock and Watson, 2018). Specifically, we use the one-year percentage change in the one-year government bond yield for monetary policy shocks, and the one-year change in the oil price index for oil price shocks. We estimate micro-level sensitivities using the GRF algorithm for each outcome variable, and assess the role of heterogeneity at the aggregate level using the theory of aggregation in Section 2.1.

Monetary Policy Shocks We estimate that, on average, a one-percentage-point increase in interest rates reduces investment and debt issuance by 0.8 percent and 1 percent, respectively, whereas firms' sales increase by 1.4 percent. However, there is substantial heterogeneity in both the magnitude and sign of these effects across firms (Figures 25 in Appendix D). Differently from the results on business cycle fluctuations, we find that financial characteristics are particularly relevant for explaining the heterogeneity in firms' sensitivity to monetary policy shocks: leverage and the share of short-term debt are particularly relevant for the heterogeneity in investment decisions,

Figure 7: Decomposition of Aggregate Response



Notes: The Figure illustrates the decomposition of aggregate responses into mean and covariance terms for each outcome variable-aggregate shock pair. The left-hand side shows the aggregate response and decomposition for monetary policy shocks, while the right-hand side shows the response to oil price shocks. Bars represent the contributions of the mean and covariance terms, while the black point denotes the total average aggregate response. We estimate Equation (8) using the mean and covariance terms in Equation (4) as dependent variable, $\hat{G}_t$. The mean and covariance terms are constructed using benchmark set of firm-level sensitivities estimated with the GRF algorithm.

while default risk and liquidity emerges as the most important characteristics for sales and debt issuance (Figure 16 in Appendix D). These results are consistent with the mechanisms proposed by Ottonello and Winberry (2020) and Jeenas (2018a), which emphasize the role of financial frictions and liquidity in shaping firms' responses to monetary policy shocks. At the same time, industry scope also play an important role in explaining the heterogeneity in sales responses across firms, accounting for 20 percent of the total heterogeneity (Crouzet and Mehrotra, 2020; Pasten et al., 2020; Ozdagli and Weber, 2017).

Figure 21 in Appendix C shows how firm characteristics influence marginal sensitivities to monetary policy shocks using the ALE plots. An unexpected interest rate hike has a more negative marginal effect on investment for firms with lower leverage, suggesting that firms that are less financially constrained display greater sensitivity to monetary policy shocks (Ottonello and Winberry, 2020). Sales of firms with higher distance-to-default are less sensitive than average, indicating that safer firms are more resilient to demand changes due to monetary policy shocks. Lastly, debt issuance of firms with high liquidity drops more relative to the average firms, suggesting that firms leverage existing cash buffers in periods of high interest rates.[25]

The presence of larger firms dampens the overall response of all outcome variables to monetary tightening, reducing the aggregate impact of the shock (left panel of Figure 7). Similar to previous literature Jeenas (2018b), we also observe a positive effect of monetary policy on sales growth in the same year the shock occurs—likely reflecting the slow-moving nature of monetary transmission to output and prices. Similar to investment, firms with higher sales shares exhibit a smaller response, dampening the aggregate decline in sales. On the contrary, firm heterogeneity does not impact the aggregate response of debt issuance. Following a contractionary monetary policy shock, debt issuance decline at impact, but the covariance term is small, indicating that the sensitivity is similar across firms.[26] Over the past fifteen years, firms' sales and debt issuance have, on average, become less sensitive to interest rate changes, but this is partially offset by the fact that larger companies become more sensitive to monetary policy (Figure 27b in Appendix D). We also document a rising role of heterogeneity for investment dynamics, contributing to the positive response of aggregate investment to interest rate hikes.

Oil Price Shocks We estimate that, on average, an unexpected increase in oil prices depresses all firms' outcome variables, in line with its interpretation of supply shock. Differently from monetary policy shocks and business cycle fluctuations, the rich heterogeneity in sensitivity across firms is driven by a more balanced split between financial and non-financial characteristics, and more generally, a broader set of firm characteristics contributes to the heterogeneity (Figure 16 in Appendix D). In fact, leverage and liquidity are the most relevant for explaining the heterogeneity in investment decisions and sales dynamics, respectively, while size for debt issuance. However, we find that, on average, no single characteristic explains more than 20 percent of the heterogeneity in sensitivities. We identify similar nonlinearity patterns as for other shocks, such as the muted response of safer (low distance-to-default) firms and of firms at the tails of the size distribution (Figure 22 in Appendix C), indicating that the marginal effects of certain characteristics exhibit strong similarities across shocks.

At the aggregate level, an unexpected increase in oil prices reduces sales, contracts investment, and leads firms to borrow less in line with economic intuition (right panel of Figure 7). Broadly in line with our results for business cycle fluctuations and monetary policy, firm heterogeneity

[25]Figure 21 in Appendix C shows that, as in business-cycle fluctuations, the sensitivity of outcomes to monetary policy shocks exhibits non-linear marginal effects. Clear thresholds emerge in investment responses to financial characteristics such as liquidity and distance to default, which have been particularly studied in the literature (Ottonello and Winberry, 2020; Jeenas, 2018a; Paranhos, 2024).

[26]Overall, our results suggest that firm heterogeneity increases the economy's resilience to negative shocks by limiting fluctuations in investment and sales. However, this resilience comes at the cost of weaker aggregate monetary policy transmission, as larger firms are less sensitive to interest rate hikes.

dampens the aggregate responses to oil shock as firms that contribute more to the economy are consistently less sensitive. The effects are sizable across all real outcome variables. Firm heterogeneity dampens the aggregate response of investment and debt by roughly 40 percent and 20 percent, respectively. For sales, the dampening is even stronger: heterogeneity nearly offsets the negative average effect.[27] Their positive responses counterbalance the negative responses of other firms, yielding an aggregate sales response that is close to zero or even slightly positive. Over time, the contributions of average sensitivity and heterogeneity are generally very stable—particularly for investment and debt issuance–showing little trend or structural change. Sales is the main exception as the average sensitivity to these shocks has declined noticeably since the early 1990s, consistent with declining role of manufacturing in the aggregate economy.

4.2 Stock Prices and Heterogeneity

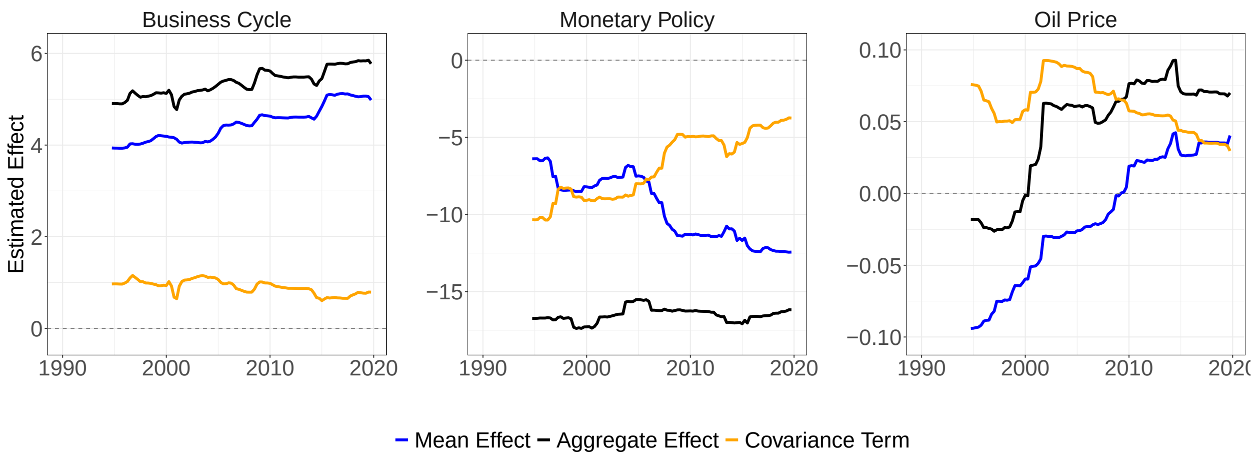
We conclude by studying the role of firm heterogeneity for the aggregate response of stock market value. A large body of literature has documented an increase in stock market volatility over the past few decades, often attributing it to the growing dominance of giant tech firms, which account for a disproportionate share of market capitalization (Nagaram and Phadke, 2025). We estimate the micro-level sensitivities of the one-year change in firms' stock market values, both to business-cycle fluctuations and to the two exogenous shocks discussed above.

We find that market value rises by 4.3 percent when real GDP growth increases by one percent, and declines 9.0 and 0.03 by percent when interest rate or oil price increase by one percent, respectively (Figure 29 in Appendix D). Heterogeneity in the sensitivity of stock market value over the business cycle is mainly driven by differences in profitability and size. Differences in profitability alone account for more than 30% of the heterogeneity we estimate, while firm size explains nearly 20%. When focusing on exogenous shocks, a different picture emerges. Following an unexpected interest rate hike, differences in default risk and size explain approximately 58% and 22% of the heterogeneity in stock market cyclicality, respectively. Financial characteristics also become substantially more relevant in the response of stock market value to oil price shocks, with most of the heterogeneity driven by differences in liquidity and leverage (Figure 30 in Appendix D).

Relative to how firms characteristics influence heterogeneity in sensitivity (Figure 31 in Appendix D), we find that firm size amplifies stock market responses over the business cycle and to contractionary monetary policy shocks: larger firms experience greater increases in stock market value during booms and sharper declines following monetary tightening. In contrast, in response to oil price shocks, larger firms perform better than the average and small- to medium-sized firms. Firms with higher cash holdings have more cyclical stock market value; those with higher liquidity

[27]This likely reflects compositional effects as firms with larger revenue shares are more prevalent in oil-intensive and energy sectors, which tend to benefit from increases in oil prices.

Figure 8: Stock Market Value



Notes: The Figure presents the mean-covariance decomposition of the aggregate market value response to these shocks, estimated with a five-year rolling window of Equation (8). Each point reflects a coefficient from the time-series model using mean and covariance components (Equation (4)) as the dependent variable, calculated from the benchmark GRF firm-level sensitivities.

ratios experience larger increases in market value during booms and smaller declines in response to negative exogenous shocks. Finally, other characteristics matter in specific cases: firms with lower distance to default show smaller average increases in stock value during booms and a significantly weaker reaction to unexpected interest rate hikes.

Finally, Figure 8 reports the mean-variance decomposition over time, estimated using a rolling window.[28] Two results stand out compared to previous results. First, the covariance term is positive throughout most of the sample, indicating that firm heterogeneity amplifies the positive aggregate response of market value to business-cycle and the negative response to monetary policy shocks. This indicates that firms with larger market capitalization exhibit higher sensitivity to these shocks. Second, over the last 30 years, we find that firms' stock market values have become increasingly sensitive to business-cycle fluctuations, mainly reflecting a rise in the average sensitivity across firms. By contrast, the aggregate response of stock prices to monetary policy shocks has remained broadly stable despite an increase in the mean sensitivity. This stability reflects a decline in the covariance term, as larger firms have become less responsive to interest rate changes, thereby offsetting the amplification effect that heterogeneity would otherwise generate.

[28]Figure 28 in Appendix D plots the firm-level shares, while Figure 29 shows the estimated sensitivities across shocks.

5 Conclusions

This paper develops a unified framework to assess how firm-level heterogeneity shapes the transmission of aggregate shocks. By combining a theory of aggregation with machine learning estimates of firm sensitivities, we show that differences in observable characteristics–particularly size and industry scope–generate substantial dispersion in how firms respond to macroeconomic fluctuations. Aggregating these responses reveals that firm heterogeneity dampens the economy's overall sensitivity to business cycle and policy shocks, as larger and more economically relevant firms tend to be less responsive. These findings highlight that macroeconomic dynamics cannot be fully understood from average firm behavior alone; the distribution and interaction of firm characteristics play a central role in shaping aggregate outcomes. More broadly, our results suggest that shifts in the composition of firms–such as changes in the distribution of firm characteristics–can alter the propagation of shocks over time, with important implications for stabilization and macroprudential policies.

References

Alfaro, I., N. Bloom, and X. Lin (2024): "The finance uncertainty multiplier," *Journal of Political Economy*, 132, 000–000.

Athey, S. and G. W. Imbens (2017): "The state of applied econometrics: Causality and policy evaluation," *Journal of Economic perspectives*, 31, 3–32.

Athey, S., J. Tibshirani, and S. Wager (2019): "Generalized Random Forests," *The Annals of Statistics*.

Bachmann, R. and C. Bayer (2014): "Investment dispersion and the business cycle," *American Economic Review*, 104, 1392–1416.

Bauer, M. D. and E. T. Swanson (2023): "A reassessment of monetary policy surprises and high-frequency identification," *NBER Macroeconomics Annual*, 37, 87–155.

Begenau, J. and J. Salomao (2019): "Firm financing over the business cycle," *The Review of Financial Studies*, 32, 1235–1274.

Belloni, A., V. Chernozhukov, I. Fernandez-Val, and C. Hansen (2017): "Program evaluation and causal inference with high-dimensional data," *Econometrica*, 85, 233–298.

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

Breiman, L. (2001): "Random forests," *Machine learning*, 45, 5–32.

Buera, F. J. and B. Moll (2015): "Aggregate implications of a credit crunch: The importance of heterogeneity," *American Economic Journal: Macroeconomics*, 7, 1–42.

Chang, M., X. Chen, and F. Schorfheide (2024a): "Heterogeneity and Aggregate Fluctuations," *Journal of Political Economy*.

——— (2024b): "On the Effects of Monetary Policy Shocks on Earnings and Consumption Heterogeneity," .

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018): "Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India," *National Bureau of Economic Research*.

Cloyne, J., C. Ferreira, M. Froemel, and P. Surico (2018): "Monetary policy, corporate finance and investment," Tech. rep., National Bureau of Economic Research.

Cooley, T. F. and V. Quadrini (2001): "Financial markets and firm dynamics," *American economic review*, 91, 1286–1310.

——— (2006): "Monetary policy and the financial decisions of firms," *Economic Theory*, 27, 243–270.

Cooper, R. W. and J. C. Haltiwanger (2006): "On the nature of capital adjustment costs," *The Review of Economic Studies*, 73, 611–633.

Covas, F. and W. J. D. Haan (2011): "The cyclical behavior of debt and equity finance," *American economic review*, 101, 877–899.

Crouzet, N. (2018): "Aggregate implications of corporate debt choices," *The Review of Economic Studies*, 85, 1635–1682.

Crouzet, N. and N. R. Mehrotra (2020): "Small and large firms over the business cycle," *American Economic Review*, 110, 3549–3601.

Davis, S. J. and J. A. Kahn (2008): "Interpreting the great moderation: Changes in the volatility of economic activity at the macro and micro levels," *Journal of Economic perspectives*, 22, 155–180.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The rise of market power and the macroeconomic implications," *The Quarterly journal of economics*, 135, 561–644.

De Veirman, E. and A. Levin (2018): "Cyclical changes in firm volatility," *Journal of Money, Credit and Banking*, 50, 317–349.

Deng, M. and M. Fang (2022): "Debt maturity heterogeneity and investment responses to monetary policy," *European Economic Review*, 144, 104095.

Durante, E., A. Ferrando, P. Vermeulen, et al. (2020): "How does monetary policy affect investment in the euro area?" *Research Bulletin*.

Farre-Mensa, J. and A. Ljungqvist (2016): "Do measures of financial constraints measure financial constraints?" *The review of financial studies*, 29, 271–308.

Favara, G., C. Minoiu, and A. Perez-Orive (2024): "Zombie lending to US firms," .

Freyberger, J., A. Neuhierl, and M. Weber (2020): "Dissecting characteristics nonparametrically," *The Review of Financial Studies*, 33, 2326–2377.

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

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical asset pricing via machine learning," *The Review of Financial Studies*, 33, 2223–2273.

Gürkaynak, R., H. G. Karasoy-Can, and S. S. Lee (2022): "Stock market's assessment of monetary policy transmission: The cash flow effect," *The Journal of Finance*, 77, 2375–2421.

Hennessy, C. A. and T. M. Whited (2005): "Debt dynamics," *The journal of finance*, 60, 1129–1165.

Jaimovich, N., S. J. Terry, and N. Vincent (2023): "The empirical distribution of firm dynamics and its macro implications," .

Jeenas, P. (2018a): "Firm balance sheet liquidity, monetary policy shocks, and investment dynamics," in *Technical Report*, Working paper.

——— (2018b): "Monetary policy shocks, financial structure, and firm activity: A panel approach," *Financial Structure, and Firm Activity: A Panel Approach (January 5, 2018)*.

Jungherr, J., M. Meier, T. Reinelt, and I. Schott (2024): "Corporate debt maturity matters for monetary policy," *International Finance Discussion Paper*.

Känzig, D. R. (2021): "The macroeconomic effects of oil supply news: Evidence from OPEC announcements," *American Economic Review*, 111, 1092–1125.

Khazra, N. (2021): "Heterogeneities in the House Price Elasticity of Consumption," .

Kiyotaki, N. and J. Moore (1997): "Credit cycles," *Journal of political economy*, 105, 211–248.

Krusell, P., C. Thürwächter, and J. Weiss (2023): "Upwardly Mobile: The Response of Young vs. Old Firms to Monetary Policy," Tech. rep., Working Paper.

Lenza, M. and E. Savoia (2024): "Do we need firm data to understand macroeconomic dynamics?" Tech. rep., Sveriges Riksbank Working Paper Series.

McConnell, M. M. and G. Perez-Quiros (2000): "Output fluctuations in the United States: What has changed since the early 1980's?" *American Economic Review*, 90, 1464–1476.

Merton, R. C. (1974): "On the pricing of corporate debt: The risk structure of interest rates," *The Journal of finance*, 29, 449–470.

Mullainathan, S. and J. Spiess (2017): "Machine learning: an applied econometric approach," *Journal of Economic Perspectives*, 31, 87–106.

Nagaram, R. and A. Phadke (2025): "Volatility-Weighted Concentration and Effective Fragility in US Equity Markets," *Available at SSRN 5395228*.

Nakamura, E. (2005): "Inflation forecasting using a neural network," *Economics Letters*, 86, 373–378.

Narayan, P. K. and S. S. Sharma (2011): "New evidence on oil price and firm returns," *Journal of Banking & Finance*, 35, 3253–3262.

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

Ozdagli, A. K. and M. Weber (2017): "Monetary Policy through Production Networks: Evidence from the Stock Market," Working Paper 17-15, Federal Reserve Bank of Boston, also NBER Working Paper No. 23424.

Paranhos, L. (2024): "How do firms financial conditions influence the transmission of monetary policy? A non-parametric local projection approach," *Journal of Econometrics*, 105886.

——— (2025): "Predicting inflation with recurrent neural networks," *International Journal of Forecasting*.

Pasten, E., R. Schoenle, and M. Weber (2020): "The propagation of monetary policy shocks in a heterogeneous production economy," *Journal of Monetary Economics*, 116, 1–22.

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

Stock, J. H. and M. W. Watson (2002): "Has the business cycle changed and why?" *NBER macroeconomics annual*, 17, 159–218.

——— (2018): "Identification and estimation of dynamic causal effects in macroeconomics using external instruments," *The Economic Journal*, 128, 917–948.

Tsai, C.-L. (2015): "How do US stock returns respond differently to oil price shocks pre-crisis, within the financial crisis, and post-crisis?" *Energy Economics*, 50, 47–62.

Varian, H. R. (2014): "Big data: New tricks for econometrics," *Journal of economic perspectives*, 28, 3–28.

Winberry, T. (2021): "Lumpy investment, business cycles, and stimulus policy," *American Economic Review*, 111, 364–396.

Appendix

A Theoretical Details

A.1 Generalized Random Forest - Algorithm

The GRF relies heavily upon the Random Forests (RF) models, since they both perform random split selection and sub-sampling. To this extent, GRF augments the methodology of RF by allowing the estimated parameters to be a weighted average of predictions, and not a pure simple average as performed in RF.

Formally, the objective of RF models is to estimate the expected value of an outcome $Y_{i,t}$, conditional on covariates $X_{i,t}$ for a given data generating process: $\beta(x) = \mathbb{E}[Y_{i,t}|X_{i,t} = x]$. The GRF aims to estimate the following moment condition:

$$\mathbb{E}[b_{\theta(x),\nu(x)}(O)_{i,t}|X_{i,t} = x] = 0 \quad \forall x \in \mathcal{X}, \text{and } i = 1, \dots, n, \;\; t = 1, \dots, T \tag{9}$$

where $O_{i,t}$ contains the set of observables, both dependent and covariates variables described in the previous section, as well as the set of exogenous shocks (W_t) that we focus on; $X_{i,t}$ represents the set of auxiliary covariates, while $\nu(x)$ is an optional nuisance parameter. Our focus is to estimate the elasticity $\hat{\theta}(x)$ for each dependent variable-shock pair, as function of all covariates.

The GRF model fits the empirical version of condition 9 by minimizing the weighted moment condition:

$$(\hat{\theta}(x), \hat{\nu}(x)) \in \text{argmin}_{\theta,\nu} \left\{ |\sum_{i=1}^{n} \alpha_i(x) b_{\theta,\nu}(O_{i,t})|_2 \right\} \tag{10}$$

The main additional feature of the GRF comes from the weighting function $\alpha_i(x)$: this aims to find firms with similar elasticities - depending on their characteristics $X_{i,t}$ - and associate higher weights to them. The algorithm developed by Athey et al. (2019) grows a set of B trees and defines $L_b(x)$ as the training set falling in the same "leaf" as x.

$$\alpha_{bi}(x) = \frac{\mathbf{1}(\{X_i \in L_b(x)\})}{|L_b(x)|}, \qquad \alpha_i(x) = \frac{1}{B}\sum_{b=1}^{B} \alpha_{bi}(x)$$

By bootstrapping the dataset and growing random forests, the methodology allows estimating the parameters of interest defined on many dimensions, in contrat with linear models (e.g. OLS). The interpretation of the estimated parameters $\hat{\theta}(x)$ is of a conditional local average treatement of the elasticity for a given shock.

We further estimates the average effect in the causal forests via estimates of the average partial effect, i.e. $\mathbb{E}[\text{Cov}(W_t, Y_{i,t})/\text{Var}(W_t|X_{i,t})]$.

A.2 Monte Carlo simulation

We conduct a Monte Carlo simulation to compare the precision of the GRF and a linear panel regression in estimating heterogeneous responses to aggregate fluctuations. We assume several underlying data-generating processes, incorporating both linear and nonlinear relationships in the conditional effects, with multiple covariates driving the heterogeneity. An econometrician seeking to understand how firms respond to aggregate shocks as a function of their balance-sheet characteristics does not observe the true data-generating process. Instead, they estimate the conditional effects using either a linear panel regression model, as specified in Equation (7), or the GRF algorithm, as described in Equation (5).

Data generating process We generate synthetic data to replicate the econometric setting used in the empirical application studies below. We assume that the simulated economy consists of 6000 firms, indexed by i, over $T = 20$ periods. We denote with $X_{i,t}^j$ denote the j-th characteristic of firm i at time t, where $j = 1, \ldots, 6$. Each covariate follows an independent autoregressive process with a persistence of 0.9, and shocks drawn from a standard normal distribution with mean zero and unit variance. We assume a relatively high value of persistence to be consistent with the balance sheet characteristics in the empirical application. The aggregate shock, W_t, is also drawn from a standard normal distribution. We assume that the outcome variable for firm i at time t, $Y_{i,t}$, depends on the firm's characteristics and the aggregate shock according to the following specification:

$$Y_{i,t} = W_t + \sum_{j=1}^{J} X_{i,t}^j + F\Big(\{X_{i,t}^j\}_{j=1}^{J'}\Big) \cdot W_t + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim N(0,1), \tag{11}$$

where $\varepsilon_{i,t}$ is an independent and identically distributed noise term drawn from a normal distribution with mean zero and variance normalized to one. The aggregate shock, W_t, propagates to $Y_{i,t}$ differently across firms, depending on a subset of firm characteristics, $\{X_{i,t}^j\}_{j=1}^{J'}$. The function $F\Big(\{X_{i,t}^j\}_{j=1}^{J}\Big)$ governs the heterogeneity in firms' responses to aggregate fluctuations. Without loss of generality, we model heterogeneity as a function of the contemporaneous realization of $X_{i,t}^j$, given that W_t is independently drawn by construction and the covariates evolve solely based on their own history.

We consider three scenarios for the function F to evaluate the performance of a linear panel regression and the GRF under different data-generating processes: (i) linear, (ii) non-linear, and

(iii) threshold-based. The corresponding data-generating processes are specified as follows:

i. Linear:

$$F\Big(\{X^j_{i,t}\}^{J'}_{j=1}\Big) \;=\; \sum_{j=1}^{J'} X^j_{i,t}.$$

ii. Non-linear:

$$F\Big(\{X^j_{i,t}\}^{J'}_{j=1}\Big) = \begin{cases} \sum_{j=1}^{J'} X^j_{i,t} + \alpha_1 \sum_{j=1}^{J'} X^{j,2}_{i,t}, & \text{Quadratic} \\ \sum_{j=1}^{J'} X^j_{i,t} + \alpha_1 \sum_{j=1}^{J'} \sum_{k=j+1}^{J'} X^j_{i,t} \cdot X^k_{i,t}. & \text{Interactions} \end{cases}$$

iii. Threshold-based:

$$F\Big(\{X^j_{i,t}\}^{J'}_{j=1}\Big) = \sum_{j=1}^{J'} \Big(\alpha_1 \mathbb{1}_{X^j_{i,t}>0} + \alpha_2 \mathbb{1}_{X^j_{i,t}\leq 0}\Big) \cdot X^j_{i,t}.$$

The heterogeneous effect is estimated using both a linear panel regression and the GRF. To assess the models' ability to recover the true heterogeneity as the dimensionality of the characteristic space increases, we vary the number of covariates relevant for the heterogeneity, J', up to six. The models are evaluated using standard statistical metrics, including Root Mean Squared Error (RMSE), average bias, and explained heterogeneity.[29]

Results. Table 1 shows that GRF consistently outperforms a linear panel regression in capturing heterogeneous sensitivities, particularly when the heterogeneity is nonlinear or exhibits complex patterns. In cases of linear heterogeneity, GRF performs similarly to a correctly specified OLS, provided that the number of characteristics driving the heterogeneity is small. However, as the dimensionality of the heterogeneity increases, GRF's precision declines due to the limitations imposed by the dataset's size. Figure **??** compares the relationship between predicted and true sensitivities for a data-generating process with a single variable driving the heterogeneity. The results show that the loss of precision of the GRF is primarily concentrated in the tails of the distribution, where data are sparse. The advantages of GRF are especially pronounced in nonlinear data-generating

[29] The average bias of an estimator is defined as the expected deviation of the estimator from the true parameter value, averaged over multiple simulation runs. Explained heterogeneity is measured as the ratio of the variance of the predicted treatment effects to the variance of the true treatment effects. A value close to one indicates that the model effectively captures the variability in the true treatment effect, while a value near zero suggests poor performance in identifying heterogeneity. Values greater than one may signal overfitting, where the model captures noise rather than the underlying structure.

Table 1: Results Monte Carlo simulation

	Linear Panel Model			GRF		
DGP of Heterogeneity	Bias	RMSE	Explained	Bias	RMSE	Explained
Panel A: Variables relevant for heterogeneity $\mathrm{J}' = 1$						
Linear	0.01	0.01	1.00	0.01	0.10	1.00
Quadratic	0.18	6.84	0.00	0.01	0.48	0.98
Threshold	0.02	0.66	0.73	0.01	0.08	0.99
Panel B: Variables relevant for heterogeneity $\mathrm{J}' = 3$						
Linear	0.01	0.01	1.00	0.01	0.43	0.91
Quadratic	0.55	12.10	0.00	0.03	2.40	0.88
Interaction	0.05	8.33	0.00	0.02	1.58	0.82
Threshold	0.05	1.57	0.85	0.02	0.87	0.86
Panel C: Variables relevant for heterogeneity $\mathrm{J}' = 6$						
Linear	0.01	0.01	1.00	0.03	1.55	0.64
Quadratic	1.15	17.70	0.00	0.16	6.05	0.70
Interaction	0.13	18.78	0.00	0.08	8.88	0.35
Threshold	0.14	1.74	0.90	0.02	1.41	0.73

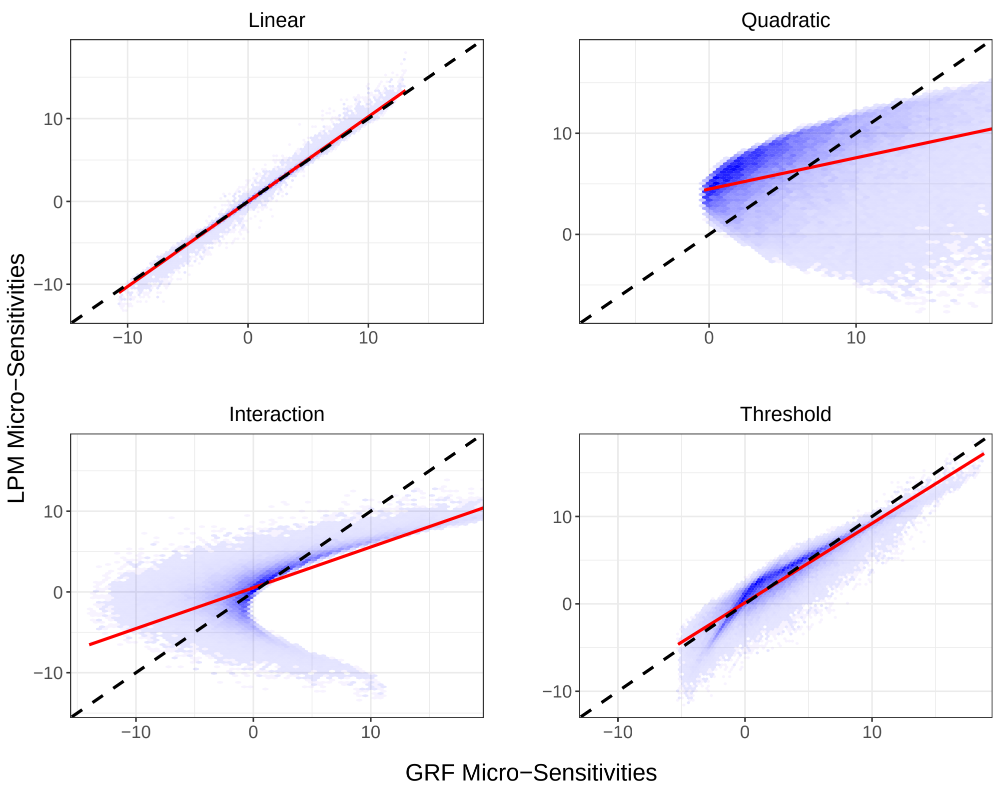
Notes: The table compares the performance of a linear panel regression and the GRF in estimating heterogeneous sensitivities across different data-generating processes and levels of heterogeneity complexity. The evaluation metrics include absolute average bias (Bias), root mean squared error (RMSE), and the proportion of variance in true heterogeneity explained by each model (Explained). Panel A, B, and C report results for a setting where only one, three, and six characteristics drives heterogeneity, respectively. We assume $\alpha_1 = 0.5$ for quadratic and interaction heterogeneity, and set α_1 and α_2 of 0.5 and 1.5 for threshold-based heterogeneity. Results are averaged over 10 simulations of a panel comprising 6,000 firms observed over 20 periods.

processes, where a linear panel regression is misspecified and fails to fully capture heterogeneity. Although both methods experience some loss of precision as the number of covariates, J', increases – reflected in higher RMSE and lower explained variance – GRF remains more robust in high-dimensional settings, effectively capturing more intricate patterns of heterogeneity.

A direct comparison between sensitivities estimated by a linear panel regression and GRF provides a useful diagnostic tool to detect misspecification due to nonlinear heterogeneity in the data.[30] Figure 9 compares the sensitivities estimated by both methods in a Monte Carlo simulation with three sources of heterogeneity ($\mathrm{J}' = 3$). When the true data-generating process is linear, the estimates from both models align closely along the 45-degree line. However, in scenarios with nonlinear or threshold-based heterogeneity, the correlation between the two weakens, and the distribution of predicted sensitivities becomes more dispersed. Depending on the scenario, the

[30]This is not a formal test but rather a graphical check that suggests the presence of unmodeled nonlinear heterogeneity in the estimated model.

Figure 9: Comparison of sensitivities on simulated data



Notes: The Figure presents predicted firm-level sensitivities from a Monte Carlo simulation across four scenarios: linear (top left), nonlinear quadratic (top right), nonlinear interaction (bottom left), and threshold-based (bottom right). The x-axis represents firm-level sensitivity estimates from the GRF, while the y-axis shows estimates from the linear panel regression. Each point corresponds to a simulated firm-time observation. The black dashed line represents the 45-degree reference line, while the red solid line depicts the fitted linear regression. Results are based on a single simulation of a panel with 6,000 firms observed over 20 periods. We assume a $\alpha_1 = 0.5$ for quadratic and interaction heterogeneity, and set α_1 and α_2 of 0.5 and 1.5 for threshold-based heterogeneity. The data-generating process assumes that three characteristics ($J' = 3$) drive heterogeneity.

differences between the sensitivities estimated by the two methods can be as large as 100% or even exhibit opposite signs, underscoring the strong misspecification bias introduced by imposing linearity in firm-level sensitivities.

B Construction of the dataset and cleaning

B.1 Firm-level variables

We construct the firm-level variables in the Compustat database following standard practices. Outcome variables are calculated as a 1-year percentage growth using the Haltiwanger formula. Nominal sales are represented by the variable *saleq* in Compustat. The market value of the firm is the stock price (*prccq*) multiplied by the number of outstanding shares (*cshoq*). The investment rate is the 1-year change in capital stock, with capital stock equal to the book value of capital calculated using the perpetual inventory method. The initial value of a firm's capital stock is measured as the earliest available entry of *ppegtq*, and we then iteratively construct it from *ppentq*. Debt issuances are the percentage change in total debt, calculated as the sum of debt in current liabilities (*dlcq*) and long-term debt (*dlttq*). Inventories are represented by the variable *invtq* in Compustat. Independent variables are always expressed in levels. Leverage is calculated as the ratio of debt in current liabilities (*dlcq*) and long-term debt (*dlttq*) to total assets (*atq*). The liquidity ratio is the ratio of cash and short-term investments (*cheq*) to total assets (*atq*). Sales growth volatility is the standard deviation of firms' real sales growth in a 10-year rolling window. Distance to default is calculated for each firm using the algorithm in Merton (1974). The short-term debt ratio is the ratio of current debt (*dlcq*) to total debt. Size is the log of total assets (*atq*). Return on assets is the ratio of net income (*niq*) to total assets. Finally, industry scope is proxied with industry classification based on the NAICS-5 industry digit. All the independent variables, with the exception of industry classification, are yearly averaged before cleaning.

Additionally, to compute variables in real terms, we deflate capital stock, sales, and total assets using the implied price index of gross value added in the U.S. non-farm business sector.

B.2 Sample selections and cleaning

The sample period is 1990Q1 to 2019Q4. We perform the following cleaning steps:

i) We keep only US-based firms, $fic_{i,t} =$"USA".

ii) To avoid firms with strange production functions, drop regulated utilities and financial companies, we drop all firm-quarters for which the 4-digit sic code is in the range [4900,5000) or [6000,7000).

iii) To get rid of years with extremely large values for acquisitions to avoid the influence of large mergers, we drop all firm-quarters for which the value of acquisitions $acq_{i,t}$ is greater than 5% of total assets $atq_{i,t}$.

iv) We drop all firm-quarters for which the measurement of Total Assets $atq_{i,t}$, Sales $saleq_{i,t}$, Property, Plant and Equipment (Net) $ppentq_{i,t}$, Cash and Short-Term Investments $cheq_{i,t}$, Debt in Current Liabilities $dlcq_{i,t}$, Total Long-Term Debt $dlttq_{i,t}$, Total Inventories $invtq_{i,t}$ are missing or negative.

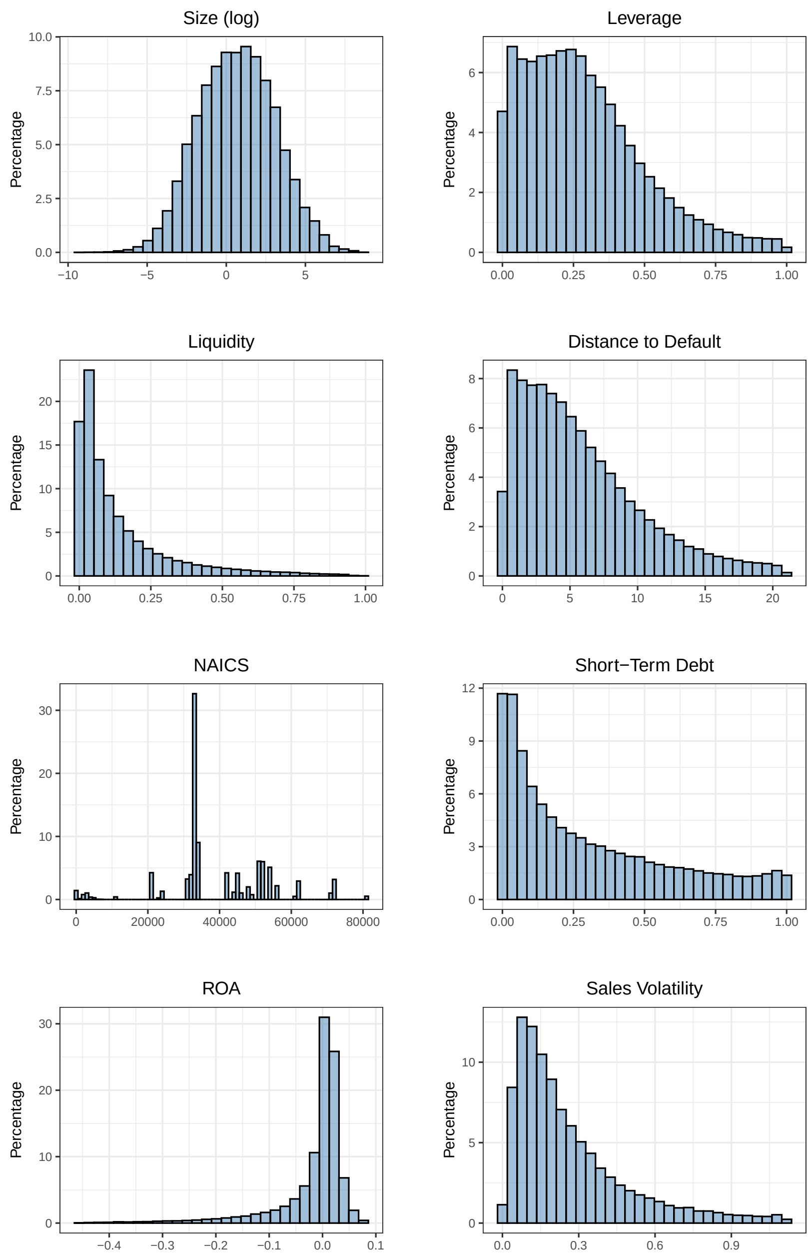
v) We drop all firm-quarters before a firm's first observation of Property, Plant, and Equipment (Gross) $ppegtq_{i,t}$.

Before estimating the models, we trim the variables at the top 1.5% level when the variables are strictly positive, and we trim 1.5% on both sides if the variables can also be negative. To reduce the number of missing values in the GRF, we linearly interpolate each independent variable after completing all cleaning steps.

We further group variables by type, distinguishing between financial and non-financial characteristics. Financial variables include leverage, liquidity, distance to default, and short-term debt. Non-financial variables include size, sales growth volatility, return on assets, and industry classification at the 5-digit NAICS level.

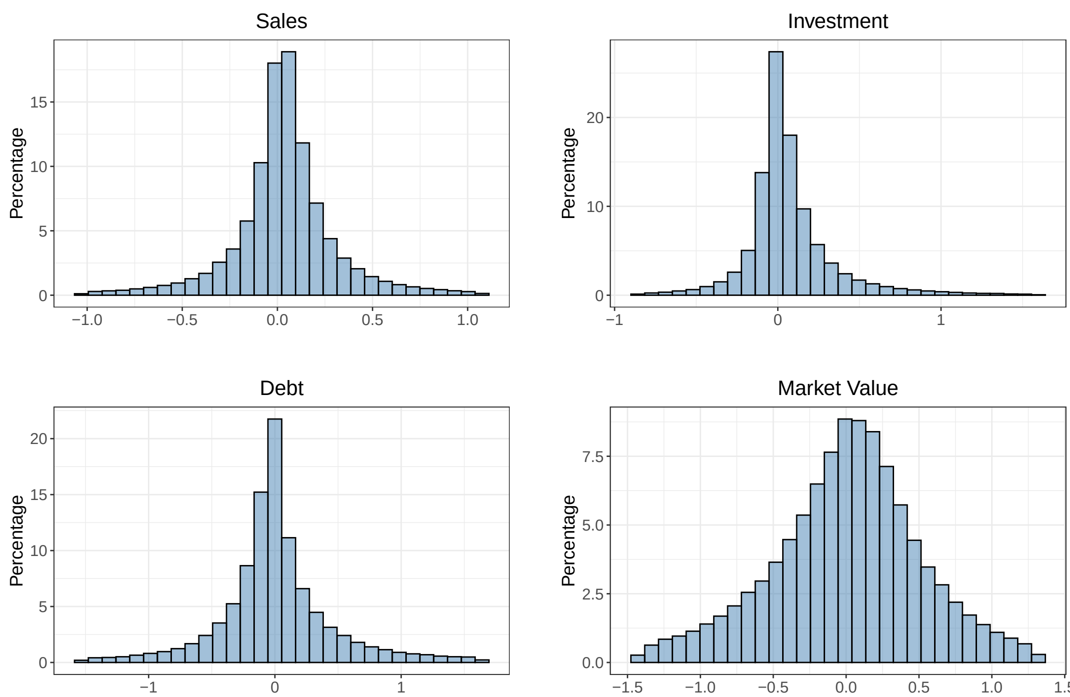
B.3 Distribution of Firms' Variables

Figure 10: Distribution of Independent Variables



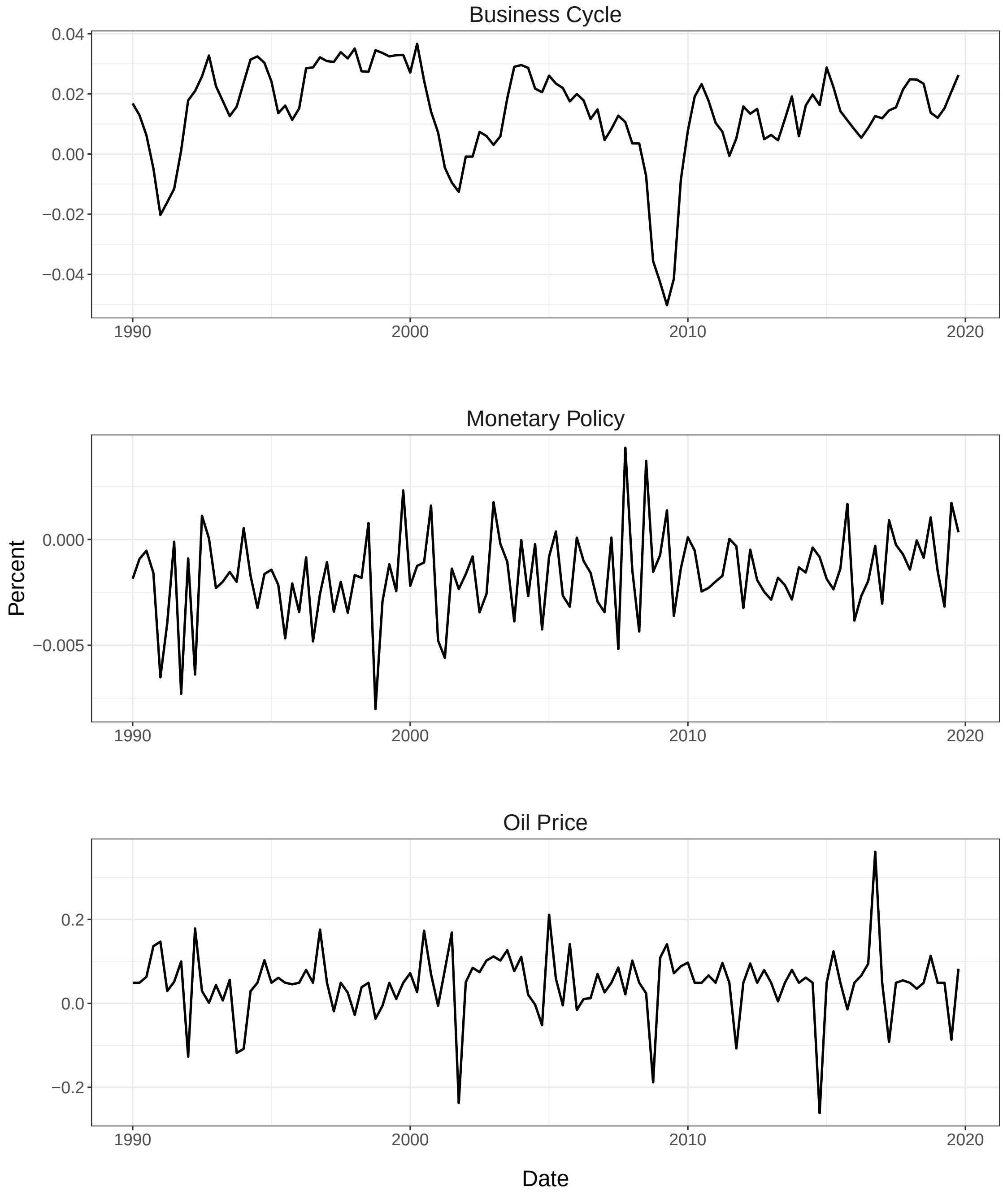
Notes: The Figure shows the distribution of firm-quarter balance-sheet characteristics used as independent variables in the empirical application. Variables are trimmed at the 98.5th percentile and then linearly interpolated before the empirical application. The plot for NAICS shows the percentages at the 4-digits level (i.e. each bin collects 1000 industry codes). Additional details on variable construction and data cleaning are provided in Appendix B.

Figure 11: Distribution of Dependent Variables



Notes: The Figure shows the distribution of firm-quarter balance-sheet characteristics used as dependent variables in the empirical application. The data are from quarterly Compustat, spanning from 1990-Q1 to 2019-Q4. Growth rates are annual and they are calculated using the Haltiwanger formula. Variables are trimmed at the 1.5th and 98.5th percentile before being used in the empirical application. Units of measurement are in percentage points, where 0.01 represents 1%. Additional details on variable construction and data cleaning are provided in Appendix B.

Figure 12: Time Series of Aggregate Shocks



Notes: The Figure shows the time-series of the aggregate shocks used in the empirical application. Units of measurement are in percentage points, where 0.01 represents 1%. Additional information on the variable construction can be found in Appendix B.

B.4 Summary Statistics

Table 2: Summary Statistics

	Statistics							
Variable	Mean	Median	St. Dev.	Min	Max	IQR	Skewness	Obs.
Panel A. Characteristics								
Size	0.69	0.70	2.41	-9.29	8.67	3.40	-0.01	448856
Leverage	0.29	0.26	0.21	0.00	1.00	0.28	0.90	339760
Liquidity	0.14	0.07	0.17	0.00	0.99	0.16	2.04	363361
Distance to Default	5.76	4.74	4.45	0.00	21.03	5.84	1.05	336085
Short-Term Debt	0.30	0.20	0.29	0.00	1.00	0.43	0.89	443857
ROA	-0.02	0.01	0.07	-0.46	0.08	0.04	-2.76	437471
Sales Volatility	0.27	0.20	0.23	0.00	1.12	0.25	1.52	378875
Panel B. Outcome								
Sales Growth	0.04	0.03	0.27	-1.03	1.08	0.22	-0.04	239624
Market Value	0.01	0.03	0.51	-1.43	1.32	0.61	-0.19	214287
Investment Rate	0.07	0.02	0.28	-0.86	1.60	0.20	1.45	418934
Debt Rate	0.01	-0.02	0.45	-1.53	1.64	0.35	0.32	227625

Notes: The first panel contains the summary statistics for quarterly balance-sheet firm characteristics used as independent variables. The second panel contains the summary statistics for the outcome variables. The data are from quarterly Compustat, covering 1990Q1-2019Q4. All dependent variables are trimmed at the 1.5th and 98.5th percentiles, while independent variables are trimmed at the 98.5th percentile when positive. Independent variables are linearly interpolated after cleaning steps. Units of measurement of the outcome variables are in percentage points, where 0.01 represents 1%. Additional information on variable construction can be found in Appendix B.

Table 3: Correlation Matrix of Firm Characteristics

Variable	Size	Leverage	Liquidity	Distance to Default	Short-Term Debt	ROA	Sales Volatility
Size	1.00	0.07	-0.19	0.33	-0.41	0.41	-0.40
Leverage	0.07	1.00	-0.31	-0.37	-0.22	-0.09	0.02
Liquidity	-0.19	-0.31	1.00	0.14	0.13	-0.26	0.31
Distance to Default	0.33	-0.37	0.14	1.00	-0.05	0.28	-0.23
Short-Term Debt	-0.41	-0.22	0.13	-0.05	1.00	-0.21	0.20
ROA	0.41	-0.09	-0.26	0.28	-0.21	1.00	-0.43
Sales Volatility	-0.40	0.02	0.31	-0.23	0.20	-0.43	1.00

Notes: The Table contains the pairwise correlation statistics for quarterly firm balance-sheet characteristics used as independent variables. The data are from quarterly Compustat, covering 1990Q1-2019Q4. All independent variables are trimmed at the 98.5th percentile when positive. Independent variables are linearly interpolated after cleaning steps. Additional information on variable construction can be found in Appendix B.

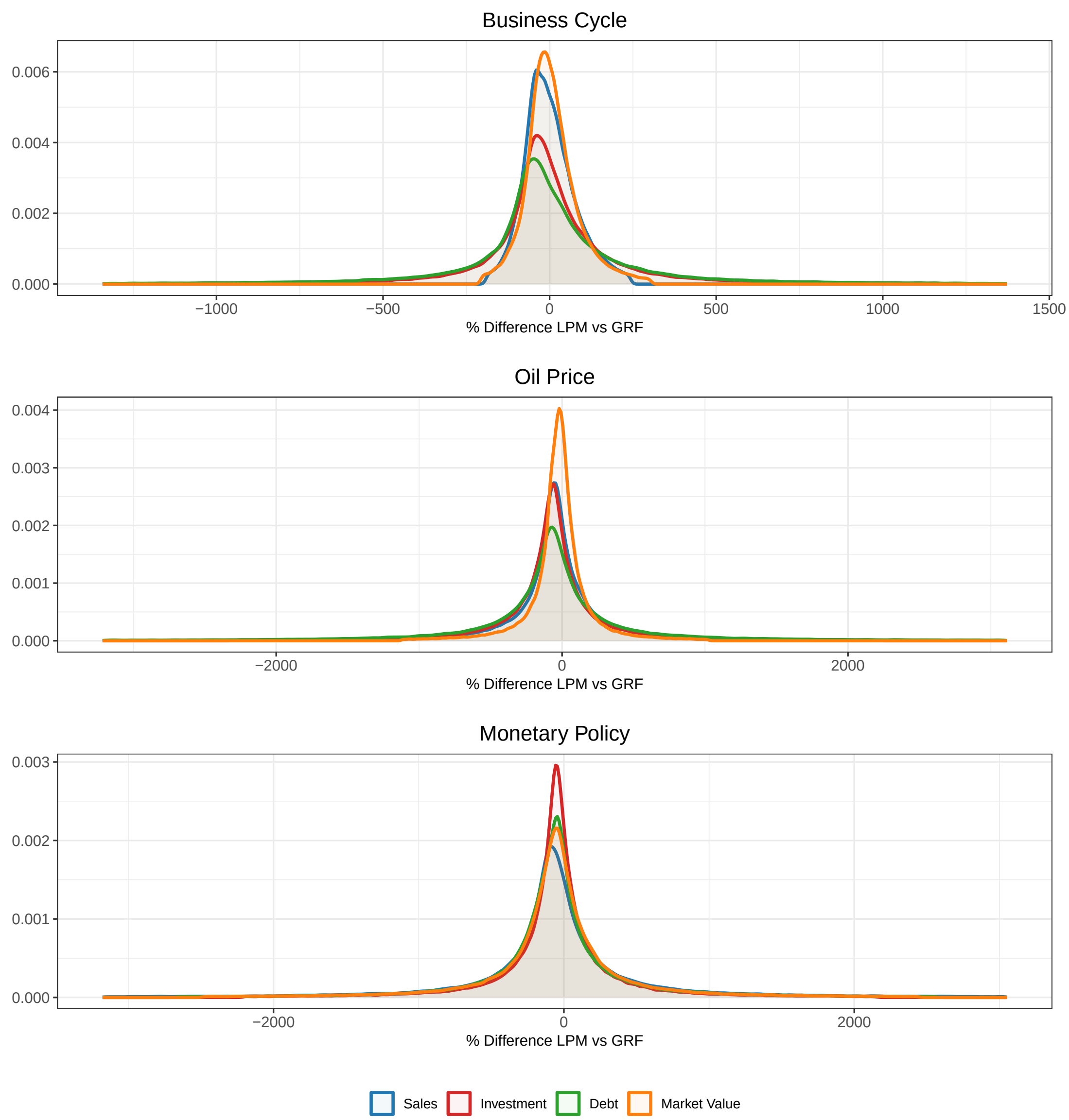
C Additional Figures and Tables - Firm level

Table 4: Summary Statistics of Firm-level Sensitivities

	GRF				Linear Panel Model			
Outcome variable	Mean	St. Dev.	Skewness	Kurtosis	Mean	St. Dev.	Skewness	Kurtosis
Panel A: Business Cycle								
Sales	2.13	0.80	0.54	3.24	2.13	1.85	0.43	3.65
Investment	0.94	0.67	0.30	3.15	0.92	1.35	0.21	5.47
Debt	1.29	1.11	0.11	2.93	1.23	2.41	0.00	5.58
Market Value	4.32	1.57	0.09	2.68	4.33	3.41	0.13	4.26
Panel B: Monetary Policy								
Sales	1.44	3.73	0.31	3.08	1.28	6.38	0.45	4.21
Investment	-0.85	2.58	-0.29	3.30	-1.06	4.43	0.17	6.10
Debt	-1.01	4.42	-0.03	3.76	-0.77	10.06	0.53	5.36
Market Value	-9.05	8.75	0.34	2.57	-10.18	13.32	-0.07	4.26
Panel C: Oil Price								
Sales	-0.02	0.07	0.02	3.49	-0.02	0.18	0.16	6.10
Investment	-0.04	0.06	-0.35	3.32	-0.04	0.12	0.13	7.50
Debt	-0.07	0.11	-0.24	4.03	-0.07	0.27	-0.38	7.00
Market Value	-0.03	0.20	-0.17	3.01	-0.03	0.41	-0.51	5.07

Notes: The Table presents the summary statistics of the estimated firm-level sensitivities obtained from the GRF and the Linear Panel Model (LPM) regression model across different outcome variables and shocks. Metrics include the mean, standard deviation, skewness, and kurtosis for each method. Panels A through C correspond to business cycle fluctuations, monetary policy, and oil price shock, respectively, for all outcome variables analyzed.

Figure 13: Density of differences between GRF and LPM sensitivities



Notes: The Figure presents the kernel density estimates of the percentage difference between LPM and GRF firm-level sensitivities across four dependent variables: Sales, Market Value, Debt, and Investment. The x-axis represents the percentage difference between LPM and GRF estimates, calculated as $(\text{LPM}/\text{GRF} - 1) \times 100$. Each panel corresponds to a specific aggregate shock: business cycle (top), oil price (middle), and monetary policy (bottom). The densities highlight the distribution of deviations for each dependent variable, with colors indicating the specific variable. Differences are trimmed at 2.5% on both sides.

Statistical tests for non-linearity. We complement the machine learning tools by formally testing whether the relationship between the conditional effect estimated using GRF and firms' characteristics is linear. In the LPM, the implied conditional effect of an aggregate shock on firms' outcomes is linear in firms' characteristics, i.e. $b(X_{i,t-1}) = \beta_0 + \sum_{j \in J} \beta_j \cdot X^j_{i,t-1}$, where J is the set of characteristics. We test whether the estimated GRF sensitivities, $\widehat{\beta(X_{i,t-1})}$, are linear in the characteristics, leveraging three different statistical measures commonly used in testing for linearity: the estimated degrees of a Generalized Additive Model (GAM henceforth), and the Harvey-Collier and Regression Specification Error Test (RESET henceforth) tests.

We estimate a GAM of the firms' sensitivities on firms' characteristics. In a GAM, a univariate dependent variable depends linearly on unknown smooth functions of some predictor variables. Formally, this translates into estimating the following GAM: $\widehat{\beta(X_{i,t-1})} = \sum_{j \in J} f_j(X^j_{i,t-1})$, where J is the set of characteristics and f_j is a smooth function of characteristic j.[31] The effective degrees of freedom estimated by the GAM for each smooth function f_j can be interpreted as a proxy for the degree of non-linearity in the relationship between dependent and predictor variables: an EDF around one indicates a linear relationship, while an EDF larger than one indicates a non-linear relationship. The last column of Table 5 reports the minimum estimated degree of freedom across characteristics for each outcome variable - aggregate shock pair. In all cases, the minimum EDF is around six, well above the threshold value of one, indicating the presence of strong non linearities in firms' characteristics, in line with partial dependence analysis.

As an alternative, we run the RESET to check for misspecification in a linear OLS regression of $\widehat{\beta(X_{i,t-1})}$ onto the complete set of firms' characteristics as explanatory variables. The test adds higher-order terms or interaction terms of the independent variables to the regression. If these added terms are statistically significant, it suggests that the model may be misspecified. Columns (3) and (4) of Table 5 report the test statistics and the corresponding p-values for each outcome variable - aggregate shock pair, respectively. Also, in this case, linearity is rejected as no outcome variable - aggregate shock pair accepts the null hypothesis of correct model specification.

Lastly, the Harvey-Collier test for linearity involves a t-test on the mean of the recursive residuals between dependent and independent variables, which should be equal to zero under the null hypothesis that their relationship is linear. We perform the test for each aggregate shock-outcome variable pair by testing the linearity between the firm-level sensitivities estimated using GRF and firms' characteristics. Formally, we consider a linear OLS regression of $\widehat{\beta(X_{i,t-1})}$ onto the complete set of firms' characteristics as explanatory variables The first two columns of Table 5 report the test statistics and the corresponding p-values, respectively. As expected, linearity is strongly rejected, in line with previous statistical measures.

[31] We exclude industry scope from the set of characteristics because it is unreasonable to assess whether the conditional effect is non linear in 5-digit NAICS. We include 5-digit NAICS fixed effects to control for heterogeneity in industry scope.

Table 5: Statistical Test for Non-linearity

	Harvey-Collier Test		RESET Test		GAM
Outcome variable	Statistic	P-Value	Statistic	P-Value	Min EDF
Panel A: Business Cycle					
Sales	4.08	0.00	776.81	0.00	7.36
Investment	13.99	0.00	2299.81	0.00	7.78
Debt	8.67	0.00	3174.57	0.00	6.59
Market Value	14.50	0.00	2682.65	0.00	7.90
Panel B: Monetary Policy					
Sales	5.98	0.00	11157.75	0.00	7.15
Investment	3.40	0.00	286.47	0.00	7.11
Debt	5.74	0.00	272.00	0.00	7.73
Market Value	17.03	0.00	6294.30	0.00	5.86
Panel C: Oil Price					
Sales	2.92	0.00	293.06	0.00	6.63
Investment	2.13	0.03	223.91	0.00	7.63
Debt	3.64	0.00	40.49	0.00	7.40
Market Value	12.45	0.00	3587.47	0.00	7.55

Notes: The Table reports the results of three different linear specification tests between covariates and the conditional average sensitivities produced by GRF for each outcome variable across shocks. We assess the linearity of the conditional effect of an aggregate shock on firms' outcome in firms' characteristics, i.e. $b(X_{i,t-1}) = \beta_0 + \sum_{j \in J} \beta_j \cdot X^j_{i,t-1}$, where J is the set of characteristics. The null hypothesis of Harvey-Collier Test and the RESET Test is that the model is linear. For both tests, we report the test statistics and the p-value of the test. We estimate a GAM model that includes all characteristics. For each characteristic we estimate the effective degrees of freedom (EDF). We report the minimum effective degrees of freedom among characteristics in each outcome variable - aggregate shock. Results are presented for debt, investment, market value, and sales under each aggregate shock (business cycle, monetary policy, and oil price).

C.1 Chernozhukov et al. (2018) Test for Heterogeneity

The test creates two synthetic variables, C_i and D_i:

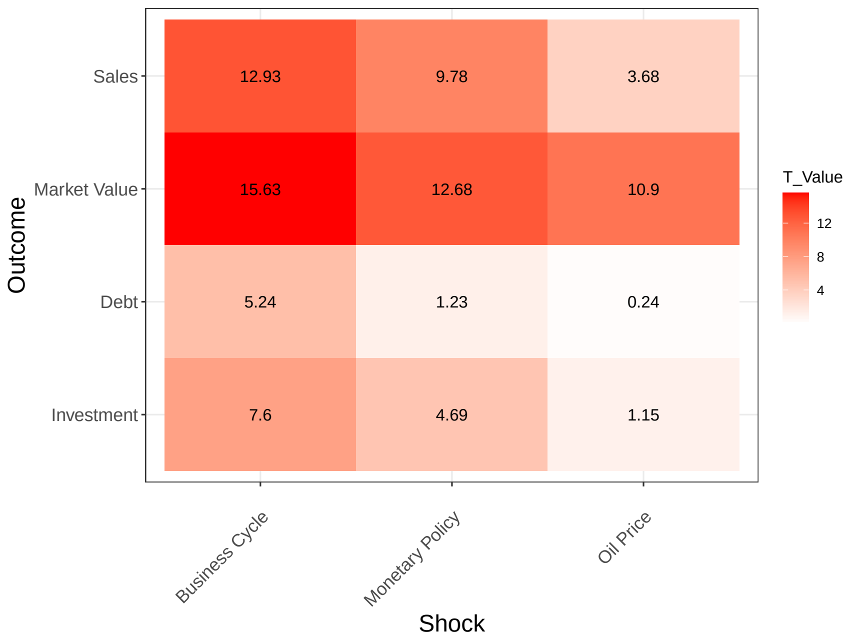
$$C_i = \bar{\beta}(W_i - \hat{W}_i),$$
$$D_i = (\hat{\beta}^{cf} - \bar{\beta})(W_i - \hat{W}_i),$$

where the former uses only the average treatment effect while the latter is the prediction that takes into account the heterogeneity as predicted by the casual forest. The test consists in running the following regression of residuals in treatment on C_i and D_i:

$$Y_i - \hat{Y}_i = \gamma C_i + \delta D_i \tag{12}$$

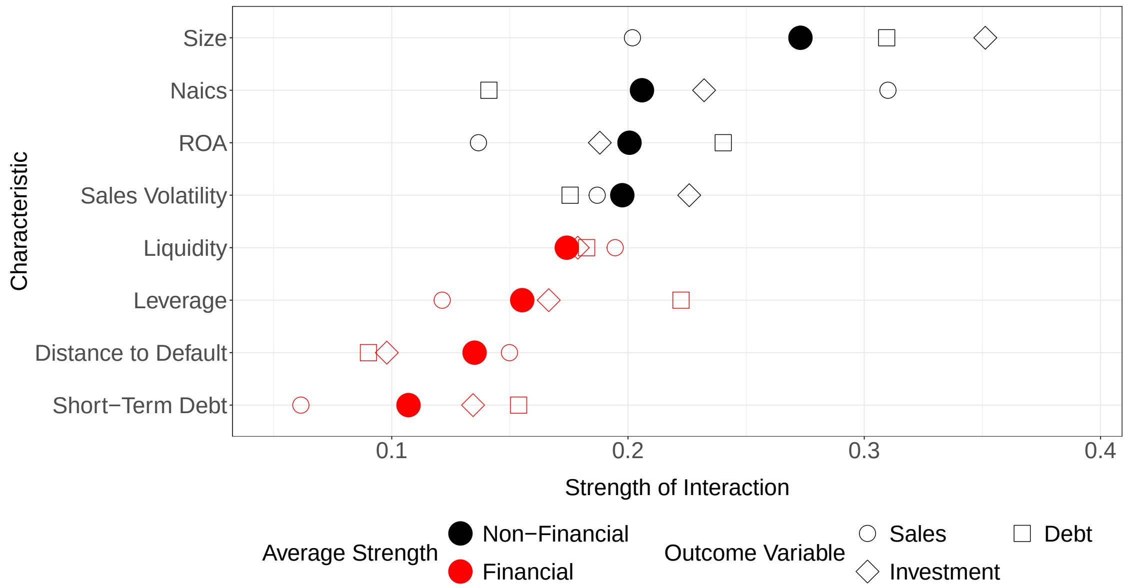
The null hypothesis of the test is $\delta = 0$, which indicates that the casual forest does not capture any heterogeneity. We find that we can reject the null hypothesis of no heterogeneity in treatment effects for almost all aggregate shock-outcome variable pairs.

Figure 14: Test for Heterogeneity in Sensitivity



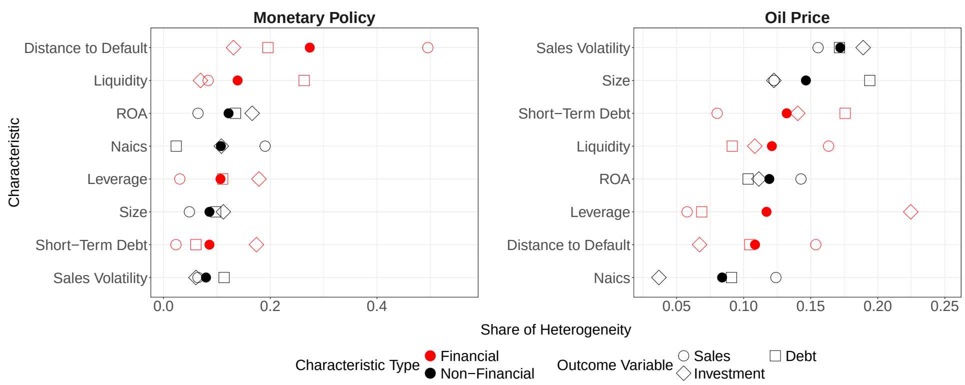
Notes: The Table reports the t-statistic of the Chernozhukov et al. (2018) test for each aggregate shock - outcome variable pair. An absolute t-statistic value below 1.648 indicates no particular degree of heterogeneity, while a value above the threshold of 1.648 suggests a statistical high level of heterogeneity in firm sensitivity at a 90% confidence interval.

Figure 15: Strength of Interactions - Business Cycle



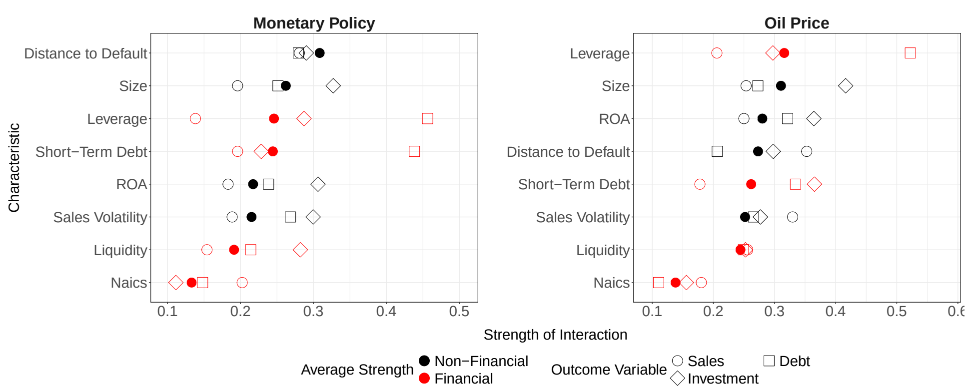
Notes: The Figure presents the strength of interaction between firm characteristics for business cycle fluctuations. The characteristics on the y-axis are ordered by their average strength of interaction within each aggregate shock, with filled points representing the average strength of interaction for each characteristic. "Financial" characteristics are depicted in red, while "Non-Financial" characteristics are shown in black. Unfilled shapes overlay the interaction strength for individual outcome variables: circles represent sales, squares represent debt, and diamonds represent investment. The x-axis reports the interaction strength, where a value of 0.01 corresponds to 1%.

Figure 16: Importance of Individual X_{it} for Heterogeneity in $\hat{\beta}$



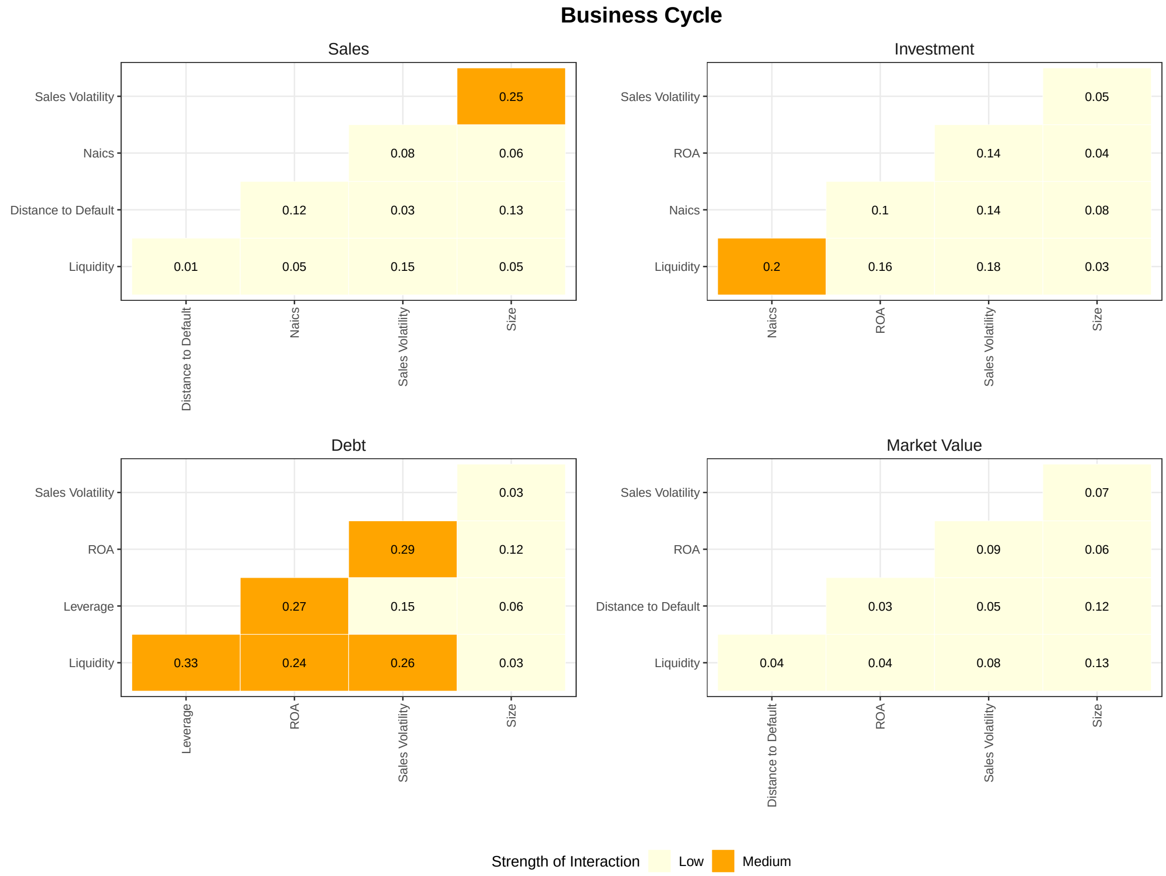
Notes: The Figure visualizes the share of heterogeneity explained by each characteristic across different outcome variables. The characteristics on the y-axis are ordered by their average importance share within each shock, with filled points representing the average importance share of each characteristic. "Financial" characteristics are depicted in red, while "Non-Financial" characteristics are shown in black. Unfilled shapes represent the importance share for individual outcome variables: circles represent sales, squares represent debt, and diamonds represent investment. The x-axis shows the importance share, where a value of 0.01 corresponds to 1% of total heterogeneity.

Figure 17: Strength of Interactions - Exogenous Shocks



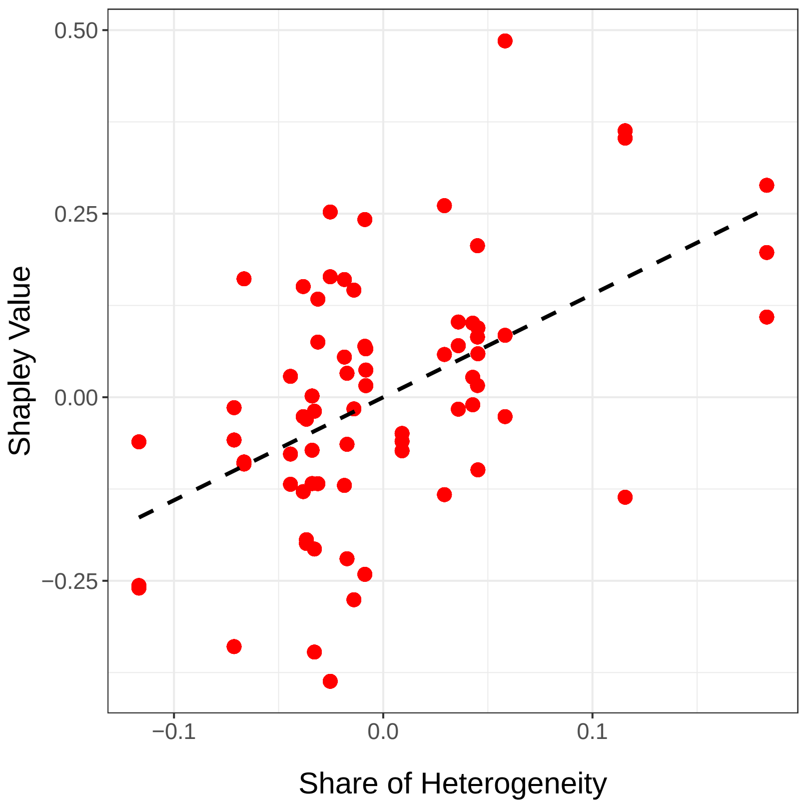
Notes: The Figure presents the strength of interaction between firm characteristics for monetary policy and oil price shock. The characteristics on the y-axis are ordered by their average strength of interaction within each aggregate shock, with filled points representing the average strength of interaction for each characteristic. "Financial" characteristics are depicted in red, while "Non-Financial" characteristics are shown in black. Unfilled shapes represent the interaction strength for individual outcome variables: circles represent sales, squares represent debt, and diamonds represent investment. The x-axis reports the interaction strength, where a value of 0.01 corresponds to 1%.

Figure 18: Pairwise Strength of Interactions



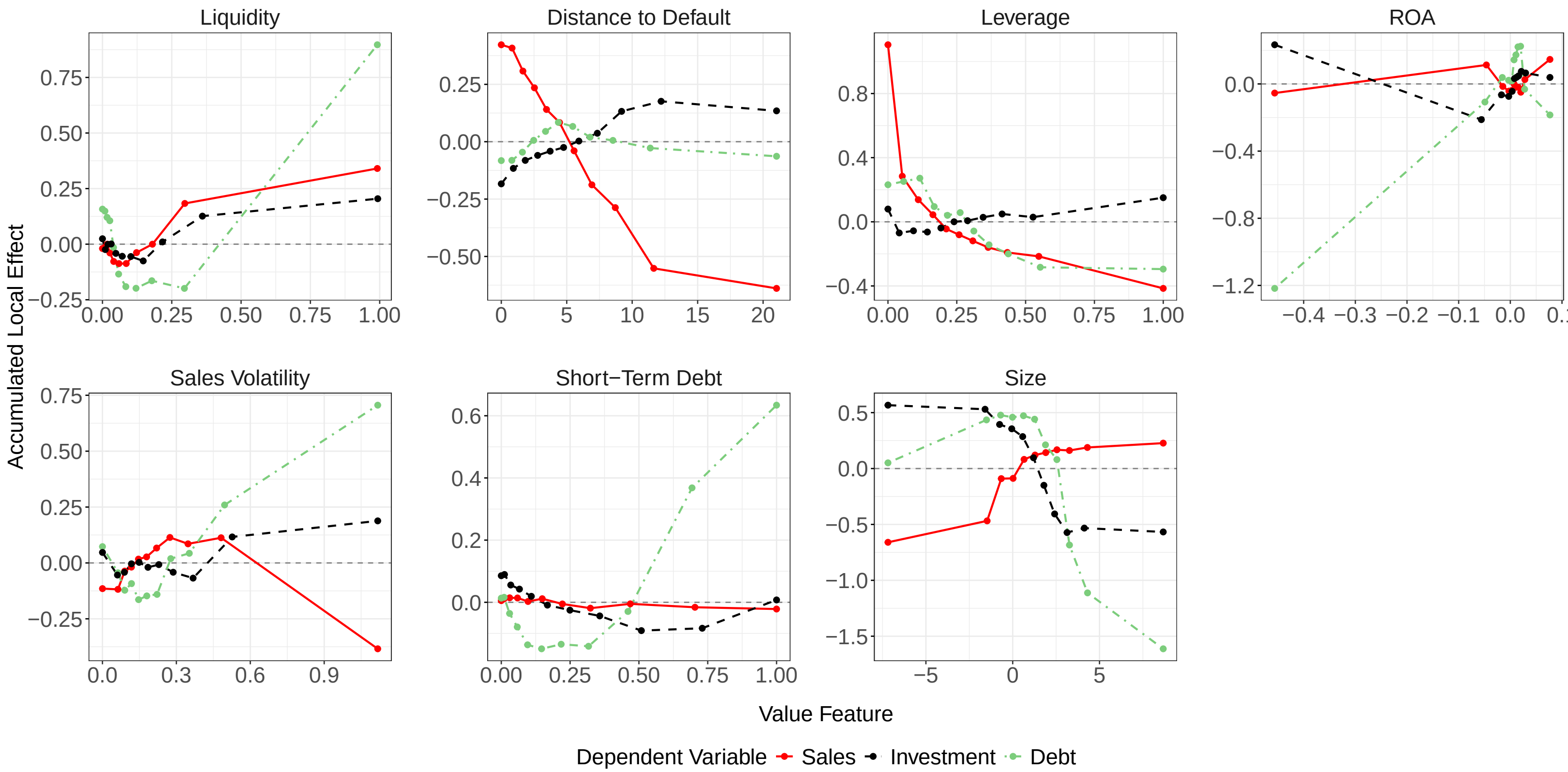
Notes: This heatmap visualizes the average pairwise strength of interaction between firm characteristics for each outcome variable. We measure the strength of interaction of each pair of characteristic using the pairwise Friedman's H-statistic. Each panel corresponds to a specific outcome variable. For each pair of characteristics, interaction values are averaged across outcome variables. For each outcome variable, we consider the ten strongest pairwise interactions. Interaction strength is categorized into three ranges: low (0–0.2), medium (0.2–0.5), and high (0.5+). The ranges are determined based on commonly observed thresholds in machine learning literature and are tailored to highlight meaningful variation in the dataset. The x-and y-axes denote the interacting characteristics, and the color scale indicates the strength of the interaction.

Figure 19: Correlation between Importance Measures



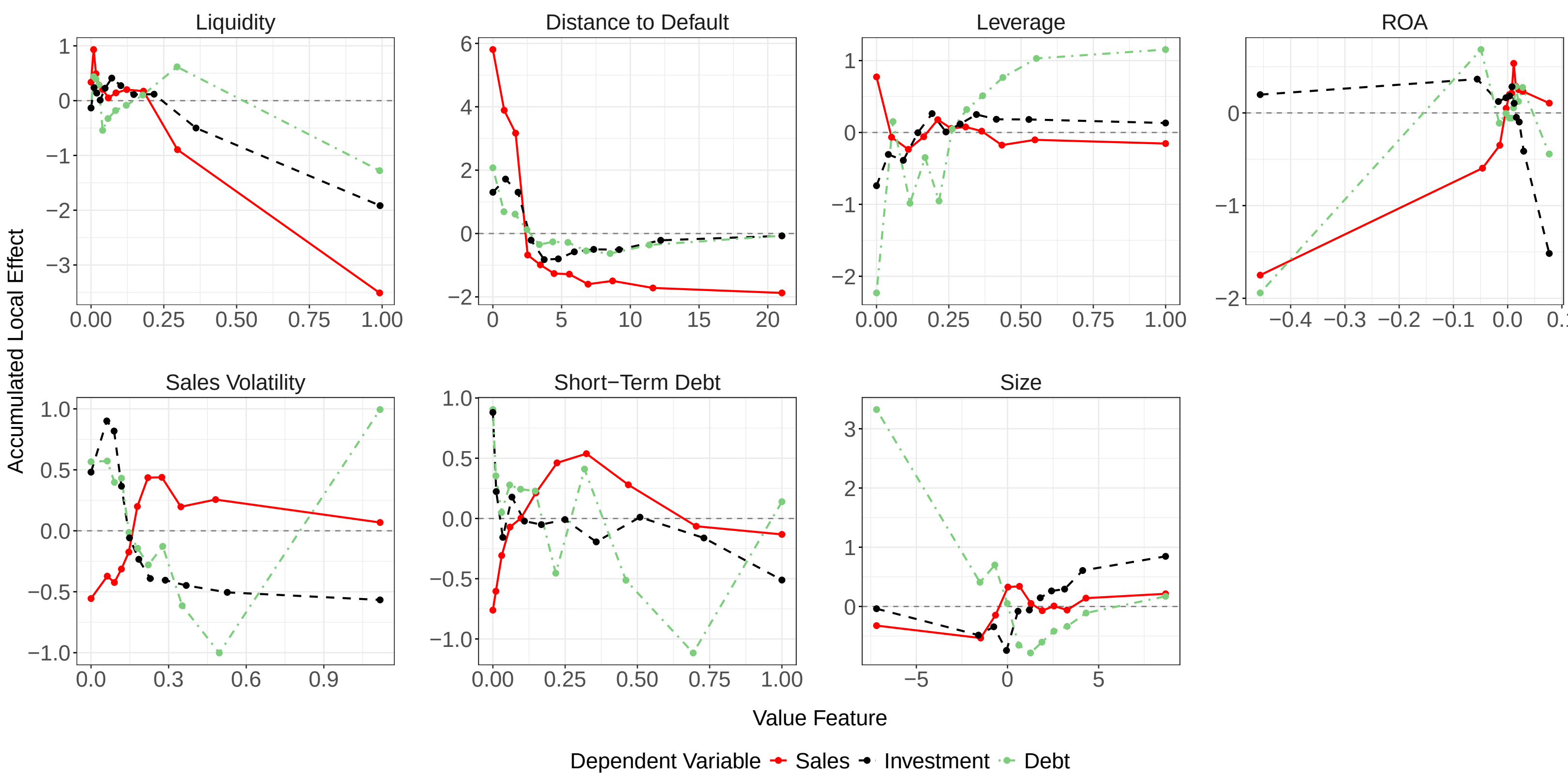
Notes: The Figure reports the scatter plots between the share of heterogeneity explained by each characteristics and the Shapley-based measure of relevance. We absorb aggregate shocks, outcome variables, and characteristic fixed effects. Black dashed lines represent a linear fit.

Figure 20: Accumulated Local Effects - Business Cycle



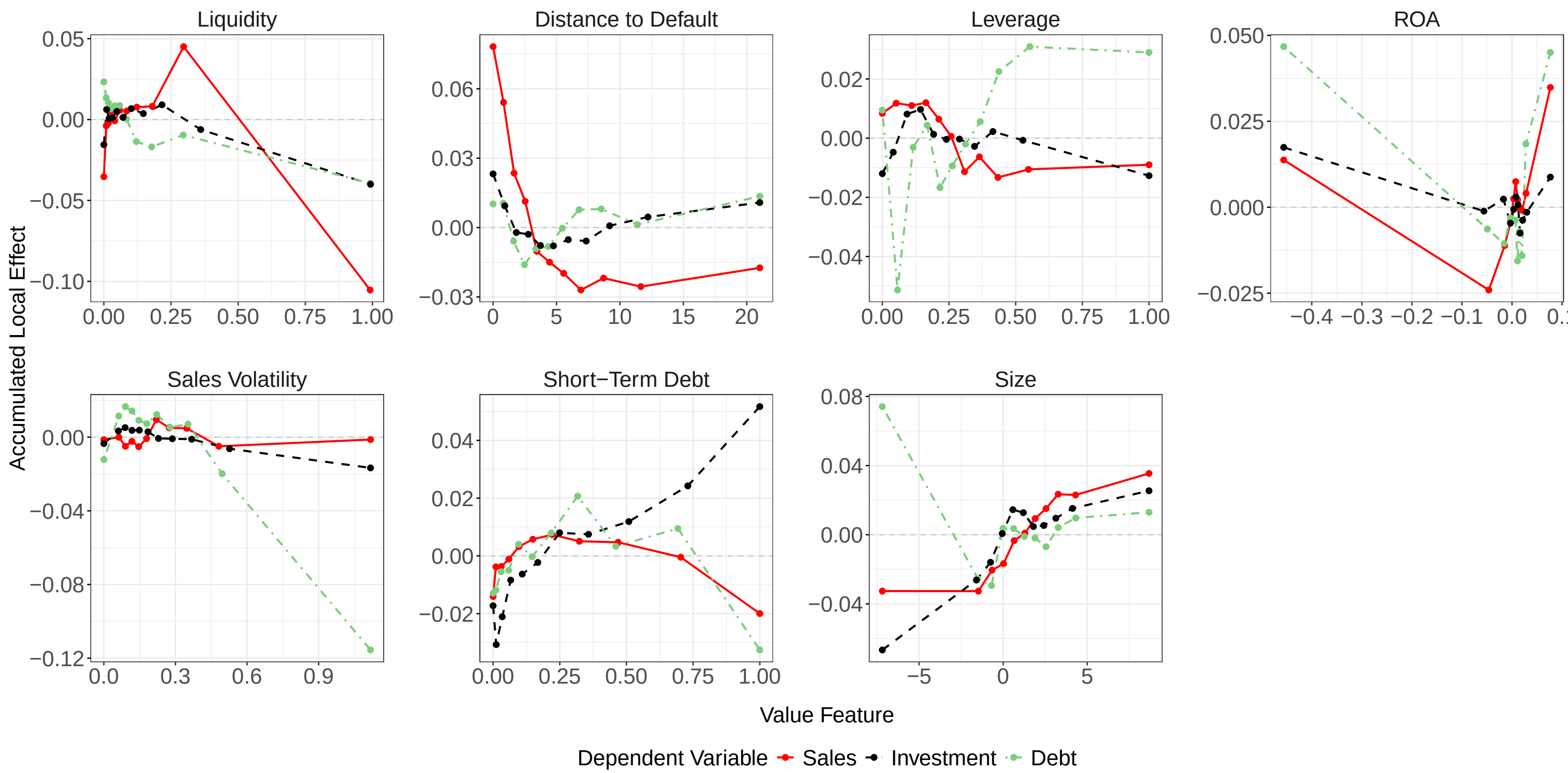
Notes: The Figure presents Accumulated Local Effects (ALE) plots estimated for each firm characteristic across different outcomes variable over business cycle fluctuations. The solid red lines represent sales, the black dash lines represent investment, and the light green dash-dot lines represent debt. The y-axis shows the estimated difference in firm sensitivity relative to the average firm sensitivity.

Figure 21: Accumulated Local Effects - Monetary Policy Shock



Notes: The Figure presents Accumulated Local Effects (ALE) plots estimated for each firm characteristic across different outcomes variable over monetary policy shock. The solid red lines represent sales, the black dash lines represent investment, and the light green dash-dot lines represent debt. The y-axis shows the estimated difference in firm sensitivity relative to the average firm sensitivity.

Figure 22: Accumulated Local Effects - Oil Price Shock

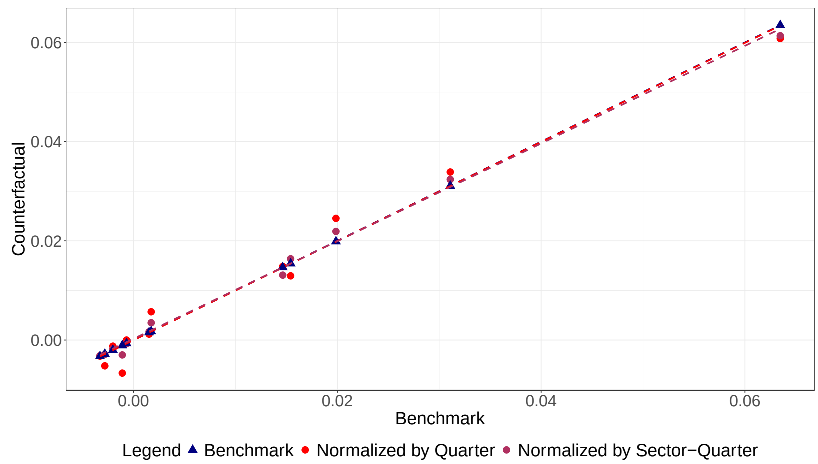


Notes: The Figure presents Accumulated Local Effects (ALE) plots estimated for each firm characteristic across different outcomes variable over oil price shock. The solid red lines represent sales, the black dash lines represent investment, and the light green dash-dot lines represent debt. The y-axis shows the estimated difference in firm sensitivity relative to the average firm sensitivity.

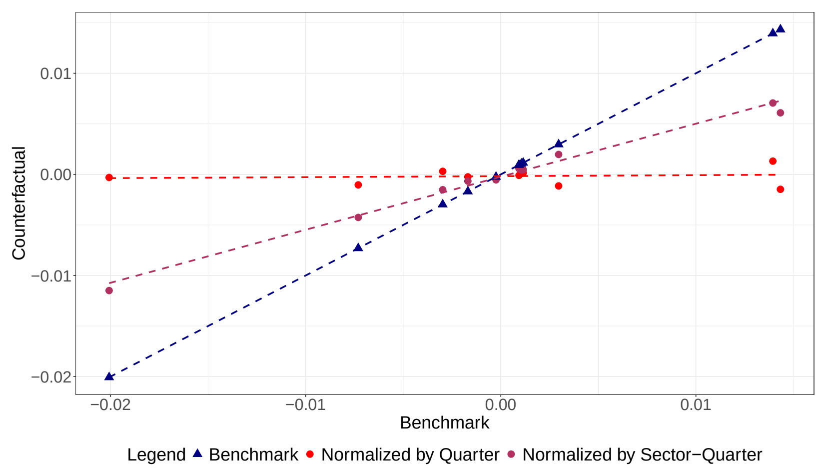
D Additional Figures and Tables - Aggregate

Figure 23: Within and across sector heterogeneity

(a) Mean

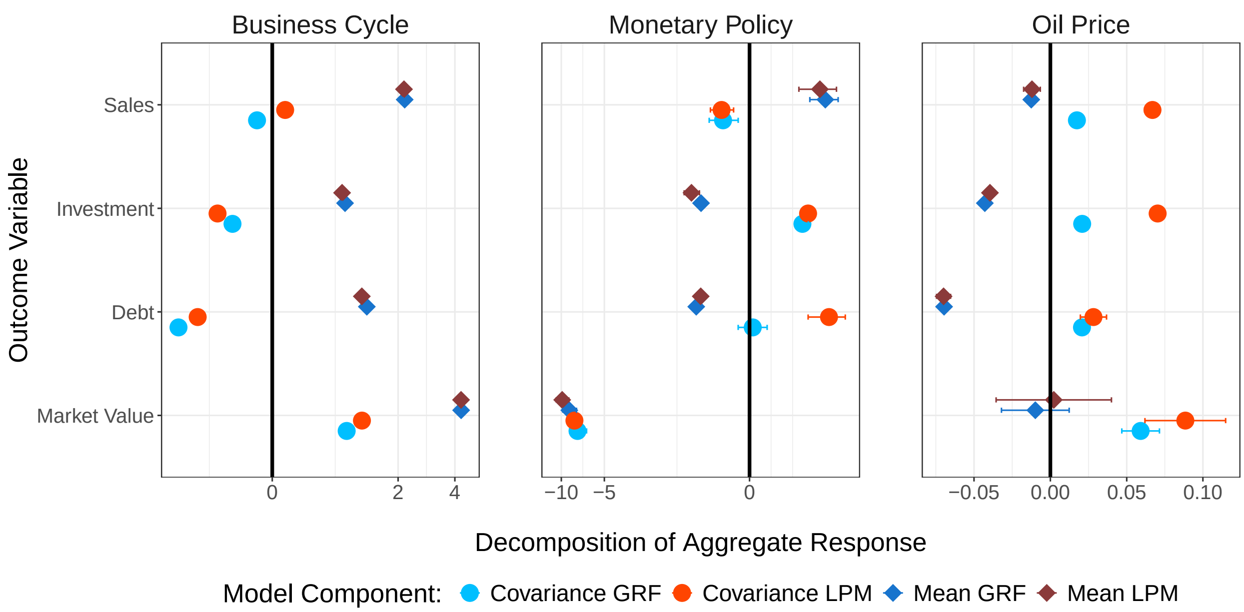


(b) Covariance



Notes: This Figure presents a comparison of firm-level sensitivity estimates under different counterfactual scenarios. The benchmark sensitivity is computed using the benchmark firm-level estimates, while the counterfactual sensitivities are obtained by normalizing firm responses across different dimensions. We compare the covariance components across the benchmark, normalized by quarter, and normalized by sector-quarter specifications. The fitted lines represent linear approximations of the relationship between the benchmark and counterfactual estimates. A lower covariance in the counterfactual scenarios indicates that firm-level heterogeneity plays a smaller role in shaping aggregate responses.

Figure 24: Comparison Mean - Covariance Decomposition



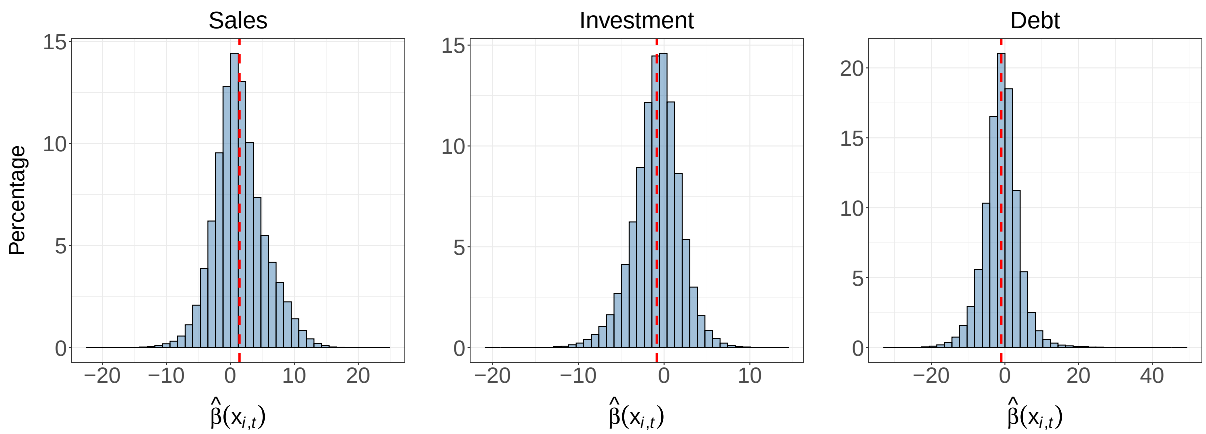
Notes: The Figure decomposes the average aggregate response into mean and covariance components for the benchmark estimates and the estimates from the linear panel model. Each point represents the estimated coefficient from Equation (8), with the mean term capturing the average firm-level sensitivity (diamond markers) and the covariance term reflecting the interaction between firm shares and sensitivities (circle markers). Blue markers denote estimates from the benchmark estimates, while red markers correspond to estimates from the LPM. Error bars indicate 95% confidence intervals based on robust standard errors.

Table 6: Comparing Aggregate Response - GRF vs LPM

	GRF		Linear Panel Model		
Outcome variable	Coefficient	StD. Error	Coefficient	StD. Error	Difference
Panel A: Business Cycle					
Sales	2.00	0.03	2.30	0.04	-0.305***
Investment	0.46	0.01	0.22	0.04	0.244***
Debt	0.02	0.03	0.25	0.06	-0.239***
Market Value	5.27	0.06	5.52	0.08	-0.255***
Panel B: Monetary Policy					
Sales	1.09	0.11	0.91	0.26	0.177
Investment	0.09	0.04	0.01	0.15	0.083
Debt	-0.90	0.07	0.77	0.27	-1.678***
Market Value	-16.51	0.17	-17.96	0.40	1.451***
Panel C: Oil Price					
Sales	0.00	0.00	0.05	0.00	-0.05***
Investment	-0.02	0.00	0.03	0.00	-0.053***
Debt	-0.05	0.00	-0.04	0.01	-0.007
Market Value	0.05	0.01	0.09	0.01	-0.041**

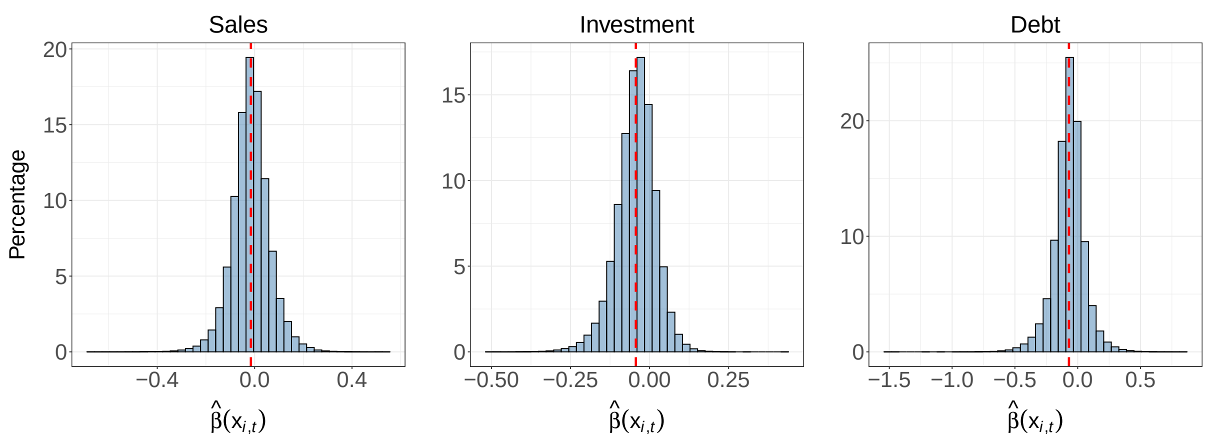
Notes: The Table presents, for each outcome variable - aggregate shock pair, the estimated average aggregate response from Equation (8) using GRF and LPM, along with their respective standard errors. Coefficients are estimated using the time-series regression in Equation (8), using the aggregate response series from Equation (3). Panels A through C correspond to business cycle fluctuations, monetary policy, and oil price shocks, respectively, for all analyzed outcome variables. We also report the statistical significance of the differences at the following levels: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure 25: Estimated $\hat{\beta}$ to Monetary Policy Shock



Notes: The Figure shows the distribution of firm-level sensitivities to monetary policy shock estimated using the GRF algorithm. Each subplot represents a specific outcome variable. The vertical dashed line indicates the average sensitivity. Firm-level sensitivities are trimmed at the 0.5% level on both tails.

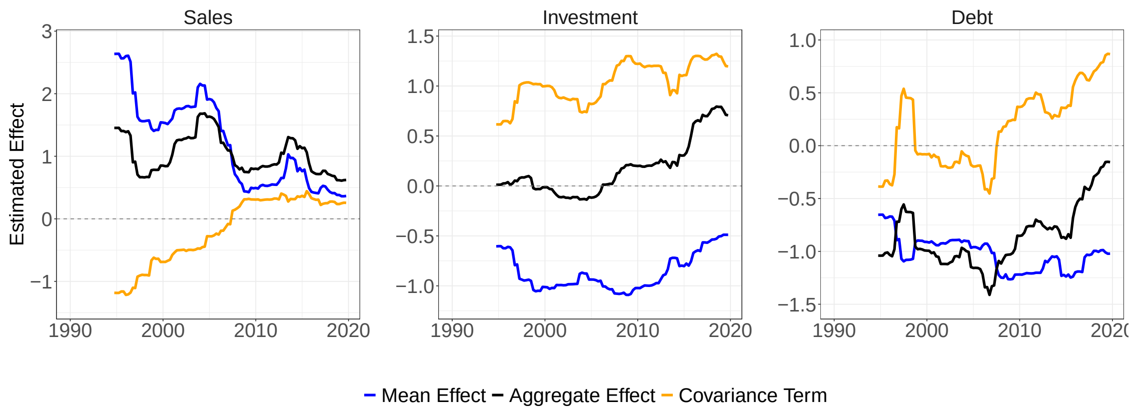
Figure 26: Estimated $\hat{\beta}$ to Oil Price Shock



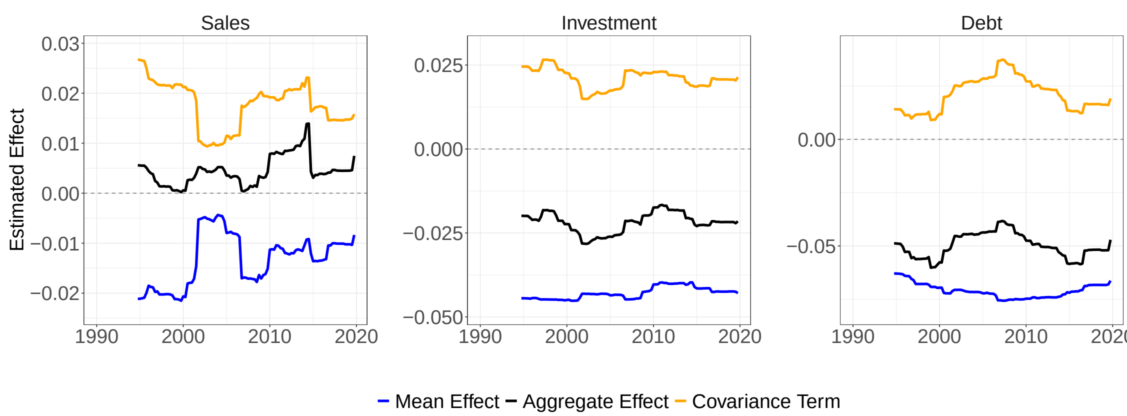
Notes: The Figure shows the distribution of firm-level sensitivities to oil price shock estimated using the GRF algorithm. Each subplot represents a specific outcome variable. The vertical dashed line indicates the average sensitivity. Firm-level sensitivities are trimmed at the 0.5% level on both tails.

Figure 27: Aggregate Response Decomposition Over Time - Exogenous Shocks

(a) Monetary Policy Shock

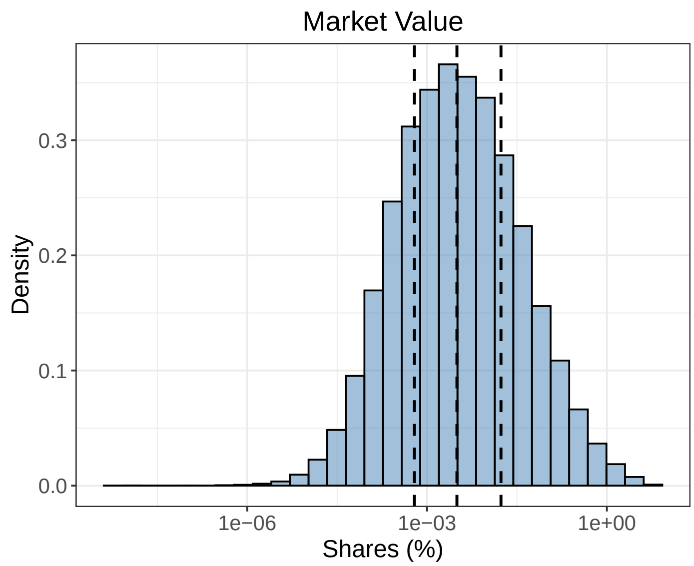


(b) Oil Price Shock



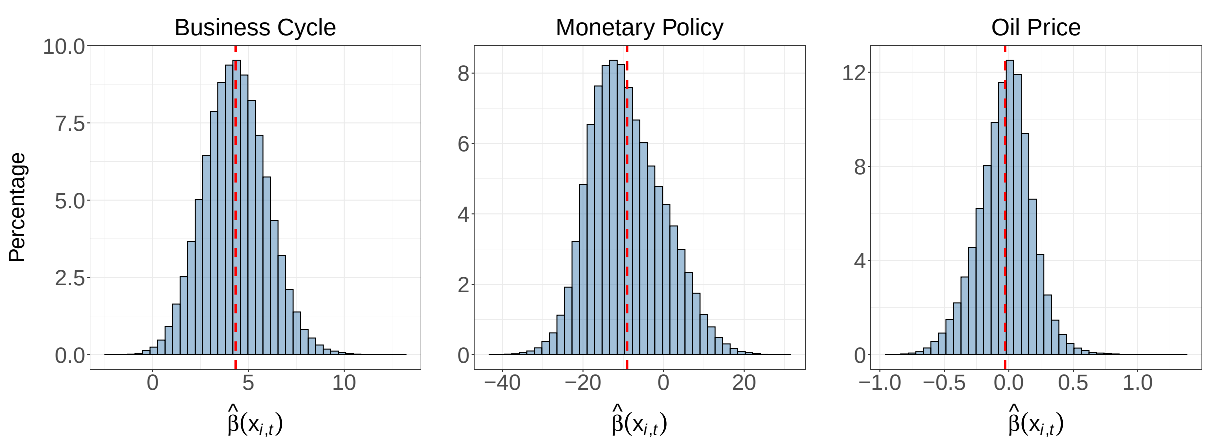
Notes: The Figure illustrates the mean and covariance decomposition of the average aggregate response to a monetary policy and oil price shock across all outcome variables, utilizing a five-year rolling window version of Equation (8). We estimate the time-serie model with the mean and covariance components, as defined in Equation (4), serving as the dependent variable. Each point in the time series represents the corresponding coefficient estimate, derived from a sample ending at the respective quarter and spanning the preceding five years. The mean and covariance components are calculated based on the benchmark GRF set of firm-level sensitivities.

Figure 28: Distribution of the Shares of Stock Market Value



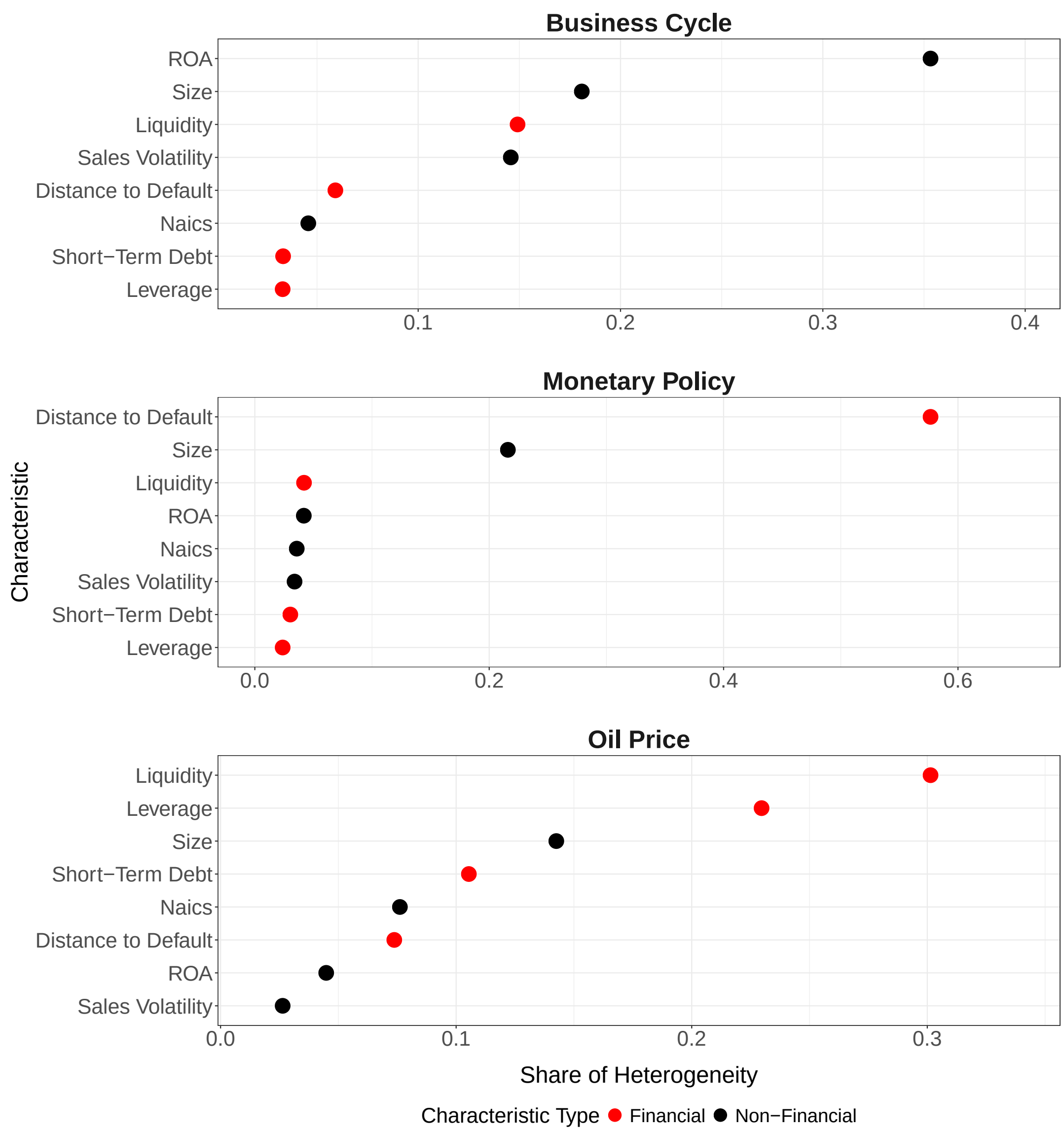
Notes: This Figure presents the distribution of firm-level shares of market value. The x-axis represents the firm-level share on a log scale, while the y-axis denotes the density. The vertical lines indicate the first, second, and third quartiles of the distribution.

Figure 29: Estimated $\hat{\beta}$ of Stock Market Value



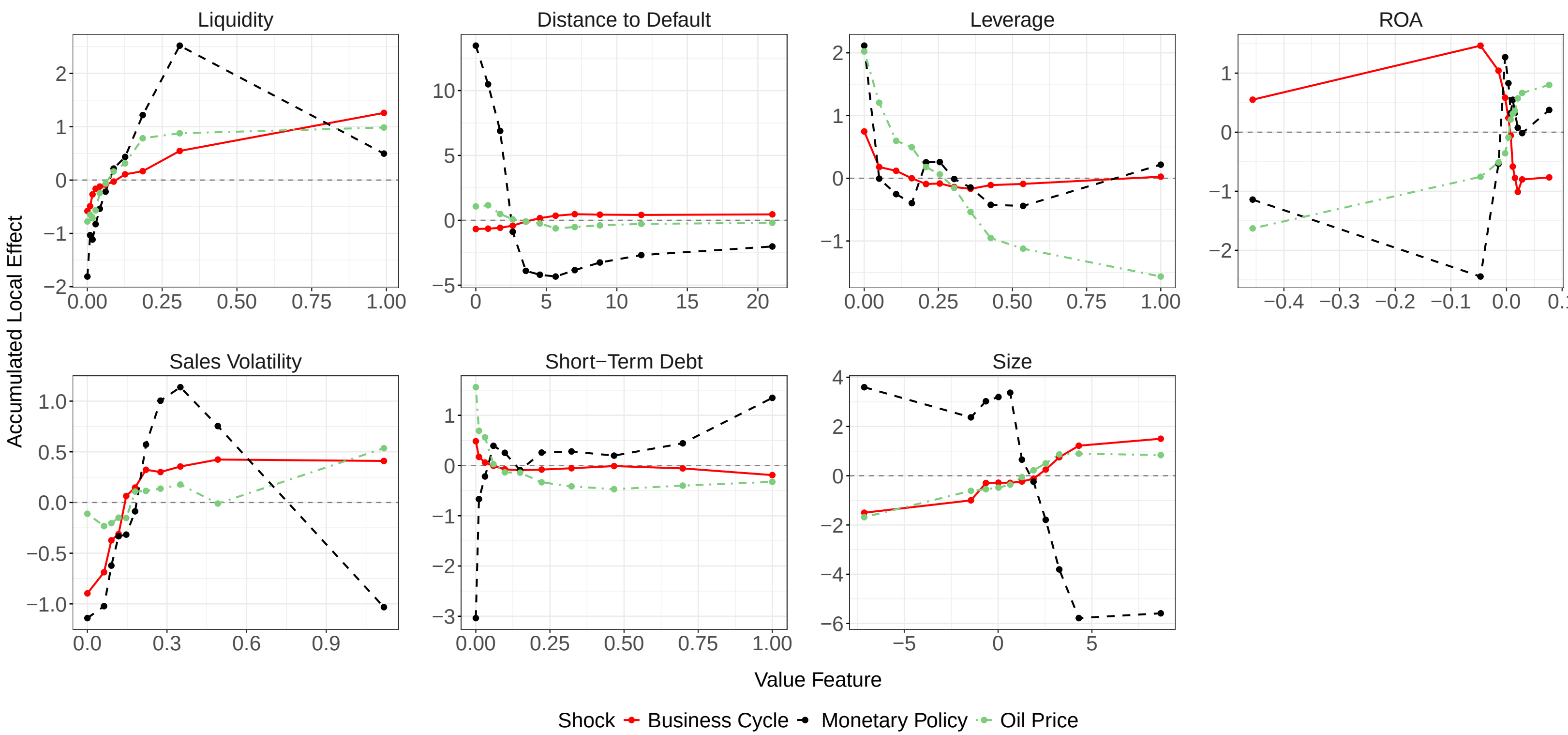
Notes: The Figure shows the distribution of firm-level sensitivities of market value estimated using the GRF algorithm. Each subplot represents a specific shock. The vertical dashed line indicates the average sensitivity. Firm-level sensitivities are trimmed at the 0.5% level on both tails.

Figure 30: Importance of Individual X_{it} for Heterogeneity in $\hat{\beta}$ - Stock Market Value



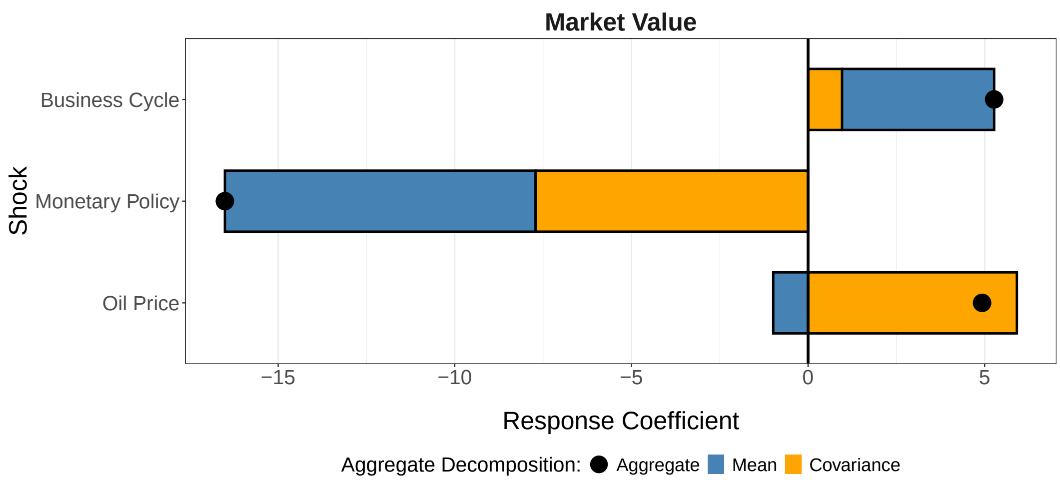
Notes: The Figure visualizes the share of heterogeneity explained by each characteristic across business cycle fluctuations, monetary policy and oil price shock for market value. The characteristics on the y-axis are ordered by their average importance share within each shock. "Financial" characteristics are depicted in red, while "Non-Financial" characteristics are shown in black. The x-axis shows the importance share, where a value of 0.01 corresponds to 1% of total heterogeneity.

Figure 31: Accumulated Local Effects - Stock Market Value



Notes: The Figure presents Accumulated Local Effects (ALE) plots estimated for each firm characteristic across different shocks over stock market value. The solid red lines represent business cycle, the black dash lines represent monetary policy, and the light green dash-dot lines represent oil price shock (rescaled by 10). The y-axis shows the estimated difference in firm sensitivity relative to the average firm sensitivity.

Figure 32: Decomposition of Average Aggregate Responses - Stock Market Value



Notes: The Figure illustrates the decomposition of aggregate responses into mean and covariance terms for market value over business cycle fluctuations, monetary policy and oil price shock. Bars represent the contributions of the mean and covariance terms, while the black point denotes the total average aggregate response. We estimate Equation (8) using the mean and covariance terms in Equation (4) as dependent variable. The mean and covariance terms are constructed using benchmark set of firm-level sensitivities estimated with the GRF algorithm. The aggregate response for oil price shock has been rescaled by 100.