Concentrated Risk: Misallocation and Granular Business Cycles*

Tomer Ifergane†

July 15, 2025

Abstract

Does misallocation affect business cycles? This paper argues that when firm-level shocks influence aggregate outcomes, distortions to the firm-size distribution have direct implications for business cycle volatility. In particular, when distortions are positively correlated with productivity, they dampen aggregate fluctuations. Viewed through this lens, size-dependent policies, such as SME subsidies, function as automatic stabilizers when they favor small firms, and as amplifiers when they favor large ones. I develop a portable method to quantify these effects and calibrate it to the U.S. economy. The results suggest that distortions can play a substantial stabilizing role. Moreover, observed dispersion in marginal products may reflect an efficient allocation that internalizes firm-level risk. The analysis highlights a trade-off between production efficiency and macroeconomic stability, with optimal allocations tilting away from the most productive firms to reduce volatility arising from granular shocks.

JEL: E32 D24 O47

Keywords: Business cycles, Misallocation, Granularity, Stabilization policies, Size-dependent policies.

*I am especially grateful to Ben Moll for his extensive feedback and continued support during my work on this paper. I also thank Daniel Albuquerque, Nadav Ben Zeev, Vasco Carvalho, Maarten De Ridder, Wouter den Haan, Xavier Gabaix, Lukas Freund, Jonas Gathen, Basile Grassi, Hugo Hopenhayn, David Lagziel, Ethan Ilzetzki, Guido Menzio, Dmitry Mukhin, Hugo Reichardt, Jane Olmstead Rumsey, Anthony Savagar, Aditya Soenarjo, Yannick Schindler, Roy (JJ) Shmueli, Ricardo Reis, and participants in several conferences and seminars for their helpful feedback, comments and suggestions. I am also grateful to Matan Levintov and Jonah Weiniger for outstanding research assistance. Previous version of this paper had been circulated under the title 'Efficiency and Macroeconomic Stability in Granular Economies'. This research was supported by THE ISRAEL SCIENCE FOUNDATION (grant No.51/22 and No. 922/23).

†Tel Aviv University, The Eitan Berglas School of Economics & London School of Economics and Political Science, Centre for Macroeconomics. iftomer@gmail.com

1 Introduction

How does misallocation affect business cycles? Recent advances in business cycle theory propose that business cycles can originate from micro-level shocks that are translated into the macroeconomy and not only from aggregate shocks, such as monetary shocks or disasters, affecting all micro units similarly. Crucially, this granular view of the business cycle (Gabaix, 2011) hinges on the presence of a skewed firm-size distribution. This paper argues that any policy or element of the economic environment that affects the firm's optimal choice of size alters the firm-size distribution and, therefore, affects business cycle volatility. Thus, misallocation in the form of correlated distortions, affecting more productive firms differently in a systematic way, can modulate business cycle fluctuations. This simple observation has profound implications, as it introduces a welfare-relevant trade-off between maximizing output and stabilizing economic activity and suggests a novel macro-stabilization role for existing size-dependent policy tools.

I begin by proposing a theory of business cycle volatility arising endogenously from firms' optimal choice of size. The theory combines the canonical firm dynamics model à la Hopenhayn (1992), set in a repeated static setting whereby firms are exposed to proportional idiosyncratic shocks that are independent of their size (e.g., Gibrat's law holds). Business cycle fluctuations are governed by the severity of firm-specific shocks and the degree of concentration in the economy, as in the seminal work of Gabaix (2011). In the model, higher concentration implies a higher loading from a shock to a large firm to the economy as a whole. I then allow for the possibility that firm input choices are distorted, which ultimately distorts the size distribution, by introducing implicit taxes or wedges as is commonly done in the misallocation literature. These wedges result in deviations from the equalization of marginal products in equilibrium as conceptualized in Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), allowing me to study the interaction between misallocation and granular business cycles.

The main theoretical result states that the correlation between distortions and firm productivity determines the effect of misallocation on business cycle volatility. Namely, positively correlated distortions reduce the relative size of the most productive firms, thus lowering concentration and the share of aggregate output determined by large firms, i.e., they act as stabilizers. Conversely, negatively correlated distortions increase concentration and raise volatility as a result. Note that the notion of stabilization relevant here is ex-ante stabilization, whereby positively correlated distortions reduce unconditional business cycle volatility.

Using this theory, I formally demonstrate how size-dependent policies act as automatic stabilizers. Such policies include employment protection, small and medium enterprise (SME) subsidies, and other institutions that disproportionately favor small firms or penalize large firms. The proof hinges on a simple intuition: these policies reduce concentration, making the firm-size distribution less skewed. These policies do not constitute a free lunch, of course, since they depress economic activity. Thus, they are only valuable if business cycle risk is relevant for aggregate welfare.

Next, I quantify the business cycle effect of misallocation using estimates from the literature. Quantifying these effects is challenging, as one cannot simply compute expectations and appeal to the law of large numbers when modeling a discrete number of firms. To overcome this, I propose and validate a flexible approximation strategy for computing the effects of correlated distortions on business cycle volatility, allowing for a computationally efficient evaluation even with a realistically large number of firms. Importantly, the method's data requirements are minimal, and it is sufficiently portable, such that it can be easily applied to conventional existing models, even without explicitly modeling a discrete number of firms. I validate the method by replicating existing results from the works of Gabaix (2011), and Carvalho and Grassi (2019), finding that it performs well.

The baseline model suggests a significant stabilizing role for positively correlated distortions. Using my baseline U.S. calibration, firm-specific shocks can explain a total factor productivity (TFP) volatility of 0.38% compared to an empirical counterpart of about 1%, consistent with the literature. Counterfactually removing positively correlated distortions increases TFP volatility from the baseline of 0.38%, reported above, to 0.65%. Alternatively, existing positively correlated distortions dampen business cycle volatility due to firm-level shocks by 42% of its total potential level. This result is robust to a range of parameter values. In appendix B, I also discuss the robustness of the results to various assumptions regarding deviations from Gibrat's law (i.e., introducing a size-volatility relationship) and the potential role of ex-post factor mobility. My analysis finds a consistent stabilizing role for positively correlated distortions across all scenarios explored with economically meaningful effect sizes.

Last, I consider the welfare implications arising from granular shocks in a stochastic environment. I demonstrate how concentration affects welfare in the model economy and show that a risk-averse planner has an incentive not to equalize the marginal products of labor across firms. In an illustrative quantitative example, such a planner optimally induces a positively correlated distortion to reduce consumption risk. In the process, the planner trades off the expected level of output to reduce its volatility and, as a result, lowers consumption risk in the environment.

Ultimately, this work motivates concerns about the presence of large firms in the economy and high concentration for macro stabilization reasons that are independent of traditional ones, such as monopoly power or adverse political incentives. If markets are efficient, this granular risk should be internalized by the firms, leading to an equilibrium in which measured marginal products of labor appear to exhibit positively correlated distortions. However, in the presence of inefficiencies that prevent the markets from fully pricing granular risk, there is room for policy intervention by adopting size-dependent policies.

Related literature. The present work is primarily connected to the literature on granular business cycles, e.g., Gabaix (2011); Carvalho and Gabaix (2013); di Giovanni et al. (2014, 2018); Carvalho and Grassi (2019). The most closely related papers are Gabaix (2011), who was the first to propose the granular hypothesis, and Carvalho and Grassi (2019), who demonstrated that this mechanism is quantitatively meaningful within the context of the canonical heterogeneous firms model. This paper builds on the insights of these two works and contributes by demonstrating how misallocation, and particularly systematic dispersion in marginal products (correlated distortions), affects business cycle volatility. Recent work demonstrated how misallocation is affected cyclically by granular forces such as market power in the case of Burstein et al. (2025) or investment irreversibility as in Senga and Varotto (2024). In contrast, this work examines the opposite direction: how non-cyclical misallocation affects business cycles in a granular environment? Such factors were found to be quantitatively important as indicated by the findings of David and Venkateswaran (2019). Additionally, this paper contributes a tractable framework and portable approximation techniques, allowing for analytical characterizations and a reduced computational burden involved with working on granular business cycles.

This paper also contributes to the misallocation literature. The modern literature on misallocation is vast, starting from the works of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Within it, the most closely related works are those concerned with the interaction between firm or establishment size distribution and systematic dispersion in marginal products or correlated distortions as in the work of Bartelsman et al. (2013); Hsieh and Klenow (2014); Bento and Restuccia (2017); Buera and Fattal Jaef (2018), and Poschke (2018).[1] I contribute to this literature by demonstrating that commonly studied

[1] For a comprehensive review of this literature see the excellent review in Hopenhayn (2014).

counterfactuals have significant business cycle implications. Namely, I show that firm-level shocks are transmitted to the aggregates differently in an economy with and without correlated distortions. When distortions are positively correlated with productivity, as indicated by the empirical literature, these distortions act as stabilizers, reducing business cycle volatility.

The analysis indicates that a risk-averse planner's optimal allocation features positively correlated distortions or markup dispersion as a result of internalizing that firm-level shocks induce aggregate consumption risk. While in many macro models markup dispersion is associated with inefficiency, this is not necessarily the case and factors such as adjustment costs (Asker et al., 2014), consumer search frictions (Menzio, 2024), and non-linear pricing (Bornstein and Peter, 2024) can lead to environments where dispersion in marginal products efficiently arises. The most related work in this strand is the paper by David et al. (2022) demonstrating that exposure to aggregate risk at the firm-level generates systematic deviations in marginal products of capital due to higher risk premiums. This paper argues that since large firms are a source of aggregate risk, idiosyncratic firm-level risk correlates with aggregate consumption risk. In such an environment, an optimal factor allocation is one where expected risk-weighted marginal products are equalized and not expected marginal products due to the planner internalizing the mapping from firm-level shocks to aggregate consumption. In related work Boar et al. (2022); Boar and Midrigan (2024); Di Tella et al. (2025) also explore the aggregate and normative implications of firm-level risk, creating systematic distortions but without discussing its business cycle implications.

Finally, this paper is conceptually and methodologically related to the literature concerning Hulten's theorem (Hulten, 1978) and the transmission of shocks through production networks Acemoglu et al. (2012); Grassi (2018); Baqaee and Farhi (2019) and Baqaee and Farhi (2020). While the present paper abstracts from the presence of production networks, the analysis conducted here can be extended by Hulten's original theorem to be the first-order effect in any arbitrary production structure. My baseline is a model in which Hulten's theorem holds exactly. Doing so allows me to draw sharp predictions from a model involving a discrete number of firms without alluding to the law of large numbers, which would nullify the granular mechanism. However, in appendix B.2 I employ the insights gleaned by Baqaee and Farhi (2019) to demonstrate how and to what extent potential deviations from Hulten's theorem affect the results.

This paper is organized as follows. Section 2 presents the benchmark model. Section 3 presents the main theoretical result of the paper on the effect of misallocation on granular business cycles and shows the stabilization role of size-dependent policies. Section 4 proposes and validates an approximation strategy to assess the role of correlated distortion on business cycle volatility and reports counterfactual effects. Section 5 discusses the welfare implications of commonly conducted misallocation counterfactuals and argues that measured dispersion in the marginal product of labor across firms can be the optimal result of an efficient allocation. The final section concludes. Extensions and additional derivations are relegated to the appendix.

2 Benchmark Environment - Efficient Production

2.1 Statement of the Benchmark Environment

Technology. There are N firms in the economy, each using a decreasing returns to scale production technology to produce a single homogeneous final consumption numeraire good y_i. Each firm i produces its output using labor[2] l_i using the production function

$$y_i = \tilde{z}_i {l_i}^{\gamma}, \quad \log(\tilde{z}_i) = \log(a_i) + \tilde{x}_i, \tag{1}$$

where $0 < \gamma < 1$ denotes the degree of decreasing returns in the economy, and $\tilde{z}_i$ is the productive ability of firm i. $\tilde{z}_i$ is composed of a firm's deterministic ability a_i and a stochastic idiosyncratic component $e^{\tilde{x}_i}$ with $\mathbb{E}\left[e^{\tilde{x}_i}\right] = 1$, such that its log is a random variable with mean $\mathbb{E}[\tilde{x}_i] = \bar{x}$ and variance σ_x^2, common to all N firms. Realizations of $\tilde{x}_i$ are assumed to be iid. The vector $\tilde{x} \in \mathbb{R}^N$ denotes the aggregate state of the economy listing all realized values of $\tilde{x}_i$. Throughout this paper, I adopt the convention that a tilde denotes stochastic variables, boldface letters denote column vectors, and aggregates are denoted in capital letters.

Under these assumptions, the model economy experiences no aggregate shocks in the conventional sense; instead, all business cycle fluctuations arise solely from firm-level shocks manifesting in the aggregate. The model is of a repeated static economy, and thus, volatility arises from differences in the cross-sectional distribution of firm-level shocks. Business cycle volatility will be defined and discussed with respect to the stochastic aggregates.

The form $\log(\tilde{z}_i) = \log(a_i) + \tilde{x}_i$ is particularly convenient since it allows one to consider

[2] Labor represents a mobile factor; replacing labor with a bundle of other factors does not affect the result.

aggregate fluctuations around a stable ergodic firm productivity distribution, or the constant dispersion of a_i. Under most conventional firm heterogeneity models, the distribution of a_i would generate, in equilibrium, the ergodic firm-size distribution and the measures of concentration in the economy. Thus, the shocks $\tilde{x}_i$ can be viewed as fluctuations around that ergodic firm-size distribution. Suppose one contemplates the business cycle as fluctuations around a balanced growth path. In that case, one can still interpret the values of a_i as the de-trended productivity distribution and the shocks $\tilde{x}_i$ as the deviations from trend growth in each firm. Therefore, the model is well-suited to study short-run fluctuations but not long-run dynamics or industry-specific trends.

Decision problem of the firm. The firm chooses how much labor to hire to maximize *expected* profits. The firm's problem is as follows

$$\max_{l_i} \quad \mathbb{E}\left[\tilde{z}_i {l_i}^{\gamma}\right] - w \times l_i. \tag{2}$$

Expectations are taken with respect to $\tilde{x}_i$ since the firm is assumed to hire labor without knowledge of the idiosyncratic state, making the production choice ex-ante efficient. Note that the assumptions on the process governing $\tilde{x}_i$ imply that labor is chosen based on the deterministic component a_i. Each firm takes the wage rate, w, as given. Labor is supplied inelastically in quantity L by the household, and the wage rate is such that the labor market clears according to

$$\sum_{i=1}^{N} l_i = L. \tag{3}$$

Thus stated, production is ex-ante efficient with the expected marginal revenue product of labor equalized ex ante across all production units. The resulting equilibrium would yield the maximum *expected* output in the economy. I will refer to this case as that of efficient production in the output-maximizing sense. This problem is standard, and the resulting production economy exhibits the following aggregate properties.

Lemma 1. *Aggregate properties of the efficient production economy:*

1. *The production economy with N firms and a given labor supply L allocates labor according to the allocation rule*

$$l_j = \frac{{a_j}^{\frac{1}{1-\gamma}}}{\sum_{i=1}^{N} {a_i}^{\frac{1}{1-\gamma}}} L. \tag{4}$$

2. *This labor allocation lends itself to an aggregate production function representation where aggregate output $Y_{\tilde{x}}$ in aggregate state $\tilde{x}$ is given by*

$$Y_{\tilde{x}} = \sum_{i=1}^{N} y_i \;= \frac{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right]^{\gamma}} L^{\gamma} = \underbrace{Z_{\tilde{x}}}_{TFP} \times L^{\gamma}, \quad Z_{\tilde{x}} = \frac{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right]^{\gamma}}. \tag{5}$$

3. *Hulten's theorem: the effect of a one-percent shock to the productivity of the j^{th} firm $\tilde{z}_j$ on $Z_{\tilde{x}}$, denoted by η_j, is given by*

$$\eta_j = \frac{\partial \log Z_{\tilde{x}}}{\partial \log \tilde{z}_j} = \frac{a_j^{\frac{1}{1-\gamma}} e^{\tilde{x}_j}}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}, \tag{6}$$

which incidentally is also the sales share of firm j or $y_j/Y_{\tilde{x}}$.

4. *Business cycle volatility: The standard deviation of log TFP is given by*

$$\sigma_Z \quad \approx \quad \underbrace{\sigma_x}_{\text{micro volatility}} \quad \times \quad \underbrace{\Psi}_{\text{amplification term (HHI)}} \quad , \Psi = \sqrt{\sum_{i=1}^{N} \overline{\eta}_i^2}, \tag{7}$$

where $\overline{\eta}_i = \frac{\partial \log [Z_{\tilde{x}}]}{\partial \tilde{x}_i}\Big|_{\tilde{x}=\overline{x}}$. Given that $\overline{\eta}_i$ is the sales share of i, then Ψ is also the square root of the Hirschman Herfindahl index (HHI) of sales in the economy when all shocks are at their expected level.

For formal proof and step-by-step derivation, see Appendix A.1.

2.2 Discussion of the Benchmark Environment

Lemma 1 establishes several properties of the static stochastic heterogeneous firms model. It is worthwhile to emphasize that the model of this section and the following extensions are set to follow as closely as possible the canonical firm dynamics model à la Hopenhayn (1992), which is commonly employed in the misallocation literature.[3]

Several features of the resulting economy are noteworthy. First, the allocation rule for labor is scale invariant. I.e., scaling up or down the productivity of each firm by a constant factor $(1+g)$, where g is a common growth rate, leaves the allocation of labor identical,

[3]For further discussion see the review in Hopenhayn (2014).

and *relative* productive ability is all that matters. Second, as is standard in the literature, the model generates a non-degenerate firm-size distribution since $l_i > 0, \forall i$. Third, *expected aggregate total factor productivity* (TFP)[4] is given by $\mathbb{E}[Z_{\tilde{x}}] = \left[\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right]^{1-\gamma}$ which is not the same as the realized TFP, given by $Z_{\tilde{x}}$. Finally, the resulting economy aggregates into a single-factor real business cycle (RBC) economy with an inelastic labor supply. Since all the positive results of this paper are related to TFP and not output, focusing on the inelastic case is not a limitation, as the labor allocation rule will remain scale invariant.

Business cycle volatility. Lemma 1 (3) verifies that Hulten's theorem (Hulten, 1978), which states that the effect of a shock to a sector or a firm on aggregate output is the firm's Domar weight or the ratio of its sales to GDP. Note that while conventionally stated in terms of the effect on output, given equation (5), the effect on TFP and output coincide throughout this paper.[5] Unlike Hulten's theorem, the above derivation involves no envelope condition. It is the exact solution to the aggregate representation of the model economy and thus sidesteps the critique of Baqaee and Farhi (2019). Furthermore, results obtained in this environment generalize via Hulten's theorem argument into a more complex setting as the first-order representation of TFP volatility in an economy with any arbitrary production network.

I interpret the shocks described throughout this paper as the usual churn of economic activity "garden-variety fluctuations", if you will, and emphatically not as those induced by catastrophic events like the 2007-2008 financial crisis or the COVID-19 recession, which are arguably "true" aggregate shocks.

Observe that the derivation of the elasticities η_j holds regardless of the distributional assumptions on a_i and $\tilde{x}_i$. However, these assumptions affect the resulting amplification term Ψ, given in equation (7), which is a concentration indicator in the resulting economy, namely the square root of the HHI when $\tilde{x} = \bar{x}$.[6] Ψ governs business cycle volatility in the economy and maps the realizations of micro-level shocks into business cycle volatility, thus providing a possible micro-foundations for the "granular hypothesis". By pinning down the value of Ψ, the distribution of $\overline{\eta_j}$, which is given by the distribution of deterministic ability a_i, in the economy determines its aggregate volatility. This key insight originates from

[4] Lemma 1 implies that expected TFP is increasing in the number of firms reflecting scale effects in the economy. I relate to this expression as TFP rather than use $Z_{\tilde{x}}/N$ as is sometimes done in the literature. My reasons are twofold: first, $Z_{\tilde{x}}$ is the correct theoretical counterpart in my environment to the classical Solow residual, and second I do not discuss entry and exit so the two are proportional, and no insight would be gained from this distinction.

[5] I refer to Hulten's theorem in TFP terms instead of output throughout for the quantification in section 4.

[6] Equation (7) is similar in flavor to equations (4) and (5) in Gabaix (2011).

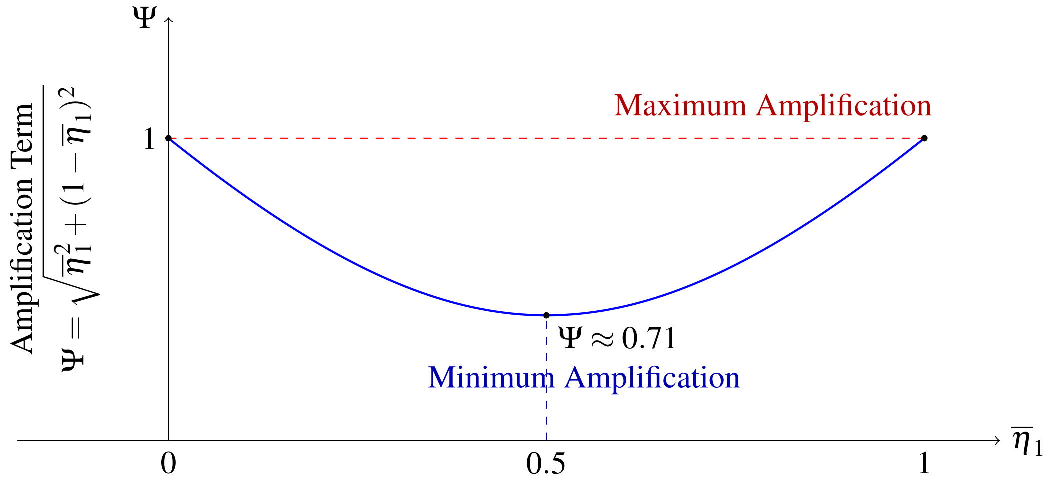


Figure 1: The Amplification Term Ψ as a Function of $\overline{\eta}$ in a Two-Firm Economy

the seminal work of Gabaix (2011), which illustrates that if sales shares are distributed according to a fat-tailed distribution, granular shocks of a reasonable magnitude can generate business cycles of a quantitatively plausible scale, even in the presence of a large number of firms.

Note that all of these elasticities sum up to unity, $\sum_{i=1}^{N} \eta_i = 1$, implying that reducing the productive ability of all firms by 1% is equivalent to a 1% aggregate TFP shock. However, the concentration of these elasticities governs the excess sensitivity of the resulting economy to shocks to one particular firm. To demonstrate this idea, let us briefly consider a two-firm economy.

A Two-Firm Example. Suppose that there were only two firms in the economy $N = 2$. Aggregate volatility in this economy is given by $\Psi = \sqrt{{\overline{\eta}_1}^2 + (1-\overline{\eta}_1)^2}$. Figure 1 reports the value of the amplification term for different levels of the market share of the first firm $\overline{\eta}_1$. Intuitively, aggregate volatility is lowest if both firms are of the same size and highest if one firm controls the entire market. This intuition generalizes to the N firm case. The more concentrated the firm-size distribution is, the more volatile the economy will be, all else being equal.

Observation. The firms' choice of relative size, i.e., the share of labor they can utilize in equilibrium, generates business cycle volatility through the resulting HHI, as demonstrated in Lemma 1. In so doing, *the firm-size distribution maps micro-level shocks to macro-level shock. Any friction or element of the economic environment that distorts the firms' choice of size also changes the volatility of TFP*. This effect is true in addition to the friction's effect on the level of TFP, which is the conventional channel in the misallocation literature. This observation motivates the analysis in the next sections.

To build intuition for the next sections, one can ask how business cycle volatility in the economy with inefficient production would differ from the efficient benchmark. Through the lens of Figure 1, the question amounts to what would be the influence of production inefficiency on the share of the larger firm? An increase in the share of the larger firm results in higher volatility. Conversely, if the largest firm's market share decreases as a result of the distortions, then volatility would decrease. This intuition will be formalized in the next section.

3 Misallocation and Business Cycle Volatility

In this section, I introduce misallocation into the benchmark model and study its implications for business cycles in the resulting economy. I demonstrate that the correlation between distortions and productive ability pinpoints the direction of influence: positively correlated distortions dampen business cycles, while negatively correlated distortions amplify them. The literature favors the former scenario as the more likely case.

3.1 The Distorted Economy: Introducing Misallocation

Compared to the efficient benchmark, the production technology remains unchanged; there are still N firms producing with the same decreasing returns to scale technology. However, the firm now faces an implicit output tax τ_i, which introduces a wedge between marginal revenue and marginal cost. The firm solves the following decision problem:

$$\max_{l_i} \; \mathbb{E}[y_i(1-\tau_i)] - wl_i. \tag{8}$$

Production is still carried out according to equation (1) such that $y_i = \tilde{z}_i l_i^{\gamma}$ and the wedge $(1-\tau_i)$ only affecting input choices. Thus, the wedge alters the firm's input choice and, consequently, its size, compared to the efficient production case, leading to misallocation as in Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and subsequent work. This modeling tool has proven particularly useful as it lends itself to straightforward empirical interpretation whereby observing a higher average revenue product of labor for a firm, sector, or establishment implies a higher τ_i. The aggregate properties of the distorted economy can be summarized as follows.

Lemma 2. *Aggregate properties of the distorted production economy*

1. *The production economy with N firms and a given labor quantity L allocates labor according to the allocation rule*

$$l_j = L \frac{\left(a_j\left(1-\tau_j\right)\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^{N}\left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}}, \tag{9}$$

2. *This labor allocation lends itself to an aggregate production function representation where aggregate output $Y_{\tilde{x}}$ is produced according to*

$$Y_{\tilde{x}}^d = \sum_{i=1}^{N} y_i = Z_{\tilde{x}}^d \times L^\gamma, \quad Z_{\tilde{x}}^d = \frac{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^{N}\left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right]^{\gamma}}. \tag{10}$$

3. *Hulten's theorem: the effect of a one-percent shock to the productivity of the j^{th} firm $\tilde{z}_j$ on $Z_{\tilde{x}}^d$, denoted by δ_j, is given by*

$$\delta_j = \frac{\partial \log\left({Z_{\tilde{x}}}^d\right)}{\partial \log \tilde{z}_j} = \frac{a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_j}}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_i}}, \tag{11}$$

   which incidentally is also the sales share of firm j or $y_j^d / Y_{\tilde{x}}^d$.

4. *Business cycle volatility: The volatility of log TFP is given by*

$$\sigma_Z^d \approx \sigma_x \times \Psi^d, \quad \Psi^d = \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^2}, \tag{12}$$

   where $\overline{\delta}_i = \left.\frac{\partial \log\left[Z_{\tilde{x}}^d\right]}{\partial \tilde{x}_i}\right|_{\tilde{x}=\overline{x}}$ and Ψ^d is the square root of the sales HHI in the distorted economy.

For proof, see Appendix A.2.

To distinguish the two economies, I denote all sizes pertaining to the distorted economy with superscript d and let the distorted counterpart of the elasticity η_i by δ_i. Note that Lemma 1 is a special case of Lemma 2 in which $\tau_i = \tau_j$ for every i and j. The aggregate production function is such that wedges affect TFP by inducing a different allocation of labor. In so doing, they affect the firm-size distribution and market concentration, inducing a different sales HHI, and thus result in different strengths of business cycle volatility σ_Z^d even while holding constant the abilities of the N firms. This environment facilitates analytical characterization since factors are allocated ex ante, thus creating an environment with misallocation in which Hulten's theorem still holds, as is demonstrated in Lemma 2 (3).[7]

3.2 Comparing TFP Volatility in a Distorted Economy with the Efficient Benchmark

To compare business cycle volatility in the efficient production economy to the distorted one really amounts to understanding the difference between the amplification term Ψ or the HHI in both economies. Changing the HHI implies that the wedges distort the firms' size choices unevenly. If misallocation involves drawing more factors into less productive firms, i.e., firms with low $\overline{\eta}_j$, and away from high-productivity firms with high $\overline{\eta}_j$, then misallocation dampens business cycle volatility. The converse is true when misallocation draws factors into more productive firms, thus making them inefficiently large; it also exacerbates business cycles.

To formalize this intuition, I define a *relative distortions vector* $d \in \mathbb{R}^N$. A relative distortion d_j faced by firm j maps the elasticity $\overline{\eta}_j$, which is also its sales share in the efficient production case, into its counterpart in the distorted case $\overline{\delta}_j$, as follows

$$\overline{\delta}_j = \frac{a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}}{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}}} = \frac{a_j^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}}\frac{\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}}{(1-\hat{\tau})^{\frac{\gamma}{1-\gamma}}} \equiv \overline{\eta}_j\sqrt{1-d_j}, \tag{13}$$

where the value of $\hat{\tau}$ is defined as the following weighted average $(1-\hat{\tau})^{\frac{\gamma}{1-\gamma}} = \sum_{i=1}^N \overline{\eta}_i(1-\tau_i)^{\frac{\gamma}{1-\gamma}}$. Given the assumptions on a_j and τ_j, leading $\overline{\eta}_j$ and $\overline{\delta}_j$ to be strictly positive we can conclude that $\sqrt{1-d_j} > 0$.

We can now compare the business cycle volatility in the distorted and efficient produc-

[7] Deviations from this assumption are discussed in appendix B.2.

tion economies by substituting equation (13) into the definition of Ψ^d in equation (12)

$$\Psi^d = \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^{\,2}} = \sqrt{\sum_{i=1}^{N} \overline{\eta}_i^{\,2}(1-d_i)}\,. \tag{14}$$

By examining the squares of the above equation and using the definition of Ψ in equation (7) we can obtain the following:

$$\underbrace{\Psi^2 - \Psi^{d^2}}_{\Delta\text{HHI}} = \sum_{i=1}^{N} \overline{\eta}_i^{\,2} d_i \; = \|d\| \times \|\overline{\eta}^2\| \times \cos(\theta)\,. \tag{15}$$

where $\|\cdot\|$ denotes the Euclidian norm, vector powers denoted elementwise operations, and since $\sum_{i=1}^{N} \overline{\eta}_i^{\,2} d_i$ is the dot product of two vectors we can express the difference in terms of the cosine similarity where θ is the angle between the two vector as follows $\cos(\theta) = \frac{\sum_{i=1}^{N} \overline{\eta}_i^{\,2} d_i}{\|d\| \times \|\overline{\eta}^2\|}$.[8] Cosine similarity measures how aligned two vectors are in direction, irrespective of their magnitude. When both vectors are mean-centered, cosine similarity coincides with the Pearson correlation coefficient. In equation (15), it serves as a measure of how relative distortions are correlated with (squared) relative sizes. In what follows, I will relate to the cosine similarity between $\overline{\eta}_j^2$ to relative distortions d_j as the correlation between relative distortions and productivity.[9] Thus, misallocation affects business cycle volatility as follows:

Proposition 1. *Misallocation dampens (amplifies) business cycle volatility arising from firm-level shocks if relative distortions are positively (negatively) correlated with firms' productive ability. The strength of this dampening (amplification) effect increases with:*

1. *the magnitude of the correlation, as measured by the absolute value of the cosine similarity* $|\cos(\theta)|$*;*

2. *the dispersion in (squared) sales shares, given by* $\|\overline{\eta}^2\|$*;*

3. *the dispersion of relative distortions, measured by* $\|d\|$.

The proof follows immediately from equation (15).

[8]Generally, for any two vectors $v, u \in \mathbb{R}^N$ we can state their dot product in terms of their cosine similarity, or the cosine of the angle α between them, as $\sum_i^N v_i u_i = \|u\| \times \|v\| \times \cos(\alpha)$.

[9]This statement holds since $\overline{\eta}_j$ and its square are both strictly positive and increasing functions of firm-level productive ability a_j.

Note that the change in concentration $\Delta\text{HHI} = \Psi^2 - {\Psi^d}^2$ is proportional to the change in business cycle variance given by $\sigma_Z^2 - (\sigma_Z^d)^2 = \sigma_x^2 \times \left(\Psi^2 - {\Psi^d}^2\right)$. Additionally, observe that when distortions are uncorrelated, or when distortions are all identical, the wedges do not introduce changes to the economy's business cycle behavior.

Positive or negative correlation? In the misallocation literature broadly, the term distortion is usually applied to the wedge $1-\tau_j$, and the concept of positively correlated distortions implies that misallocation disproportionally harms the high-ability firms by more than it does the low-ability firms, i.e., high a_j predicts a high τ_j, and a high measured marginal product of labor. Negatively correlated distortions imply the converse. My newly-defined relative distortion d_j is similar but defined in terms of the elasticities or relative firm size instead of absolute firm size in the economy where $\tilde{x} = \bar{x}$.

Through equation (13), having $d_j > 0$ implies that in the distorted case, firm j is smaller in relative terms than it would have been in the efficient production case where $d_j < 0$ implies the opposite. Observing a positive correlation between relative size and relative distortion requires that a high d_j firm also has a high $\overline{\eta}_j$. Given that $\overline{\eta}_j$ is a strictly increasing function of firm j's productive ability a_j. Therefore, correlated relative distortions in my setting are conceptually equivalent, up to the exact correlation measure used, to the correlated distortions employed by the misallocation literature, which relates to the correlation between τ_j and a_j.[10]

The bulk of the misallocation literature, starting from Restuccia and Rogerson (2008) argues that the positive correlation of distortions and ability is an important feature. Hsieh and Klenow (2014) use plant life-cycle data from India, Mexico, and the United States and find a positive correlation between distortions and ability. They also report model-implied positive elasticities for the United States. Similar positive elasticities are used to match establishment sizes and firm sizes in various countries, e.g., Bartelsman et al. (2013), Buera and Fattal Jaef (2018); Bento and Restuccia (2017); Poschke (2018), and David and Venkateswaran (2019).

An alternative approach to modeling misallocation is to consider its sources directly and micro-found the wedges. Two compelling cases are market power and financial frictions. Consider, for example, the endogenous-markup model of Atkeson and Burstein (2008). In this type of model, more productive firms within a sector can charge higher markups and are underproducing compared to an efficient production benchmark. This would manifest

[10] In section 4.3, I will also demonstrate how the two notions coincide using a parametric form used widely in the misallocation literature.

as a higher τ_i for the high-ability firms or as a positively correlated distortions.[11] For the case of financial frictions á la Kiyotaki and Moore (1997), as applied in Buera and Shin (2013), and Moll (2014), financial frictions limit the amount of inputs a producer can use as a function of their wealth. The higher their production ability, the more capital they wish to employ, and thus, the higher the implicit tax imposed on them by the financial friction. Alternatively stated, high-ability individuals are harmed more by the existence of the same constraint on their input structure, giving rise to positively correlated distortions. To think more clearly about the matter, let us now consider two concrete policy examples that can map cleanly into the benchmark environment of Section 2.

3.3 The Stabilizing or Destabilizing Role of Policies: A Concrete Policy Example

Consider now an economy whereby all firms are uniformly subject to a revenue tax at a rate of t_0 and either (i) the government decided to subsidize small businesses or (ii) large firms get tax credit or subsidies to maintain large local manufacturers and prevent them from moving abroad. The former case will be referred to as 'SME subsidies'[12] and the latter as large-business subsidies. Note that SME subsidies are an instance of positively correlated distortion, and large business subsidies of negatively correlated ones. Unlike the general statement in Proposition 1, these taxes and subsidies are now made explicit. Both of these examples are stylized representations of size-dependent policies more broadly. I now restrict attention to examples with a constant subsidy elasticity as follows

$$s(l_j) = \left(\frac{l_j}{C_0}\right)^{-\nu} - 1, \tag{16}$$

where $s_j(l_j)$ is the size-dependent subsidy rate. Stated in logs $\log(1+s_j) = \alpha_0 - \nu \log(l_j)$, with C_0 denoting a normalizing constant such that $1+s_j(l_j) \geqslant 1$ for every j. Importantly, ν is the elasticity of the subsidy with respect to size. I use size as measured by employment here for tractability. When $\nu > 0$, we are considering SME subsidies, and when $\nu < 0$, we have large-business subsidies. The flat tax rate t_0 is chosen such that it results in the entire tax system generating zero equilibrium revenues, making it revenue-neutral and purely

[11] For more on this logic, see Edmond et al. (2015).

[12] Subsidies geared towards small and medium enterprises (SMEs) are common in the development context and are in place in many countries.

redistributive. This structure streamlines the presentation considerably and implies that the firms are now solving the following modified problem

$$\max_{l_j} \quad E\left[(1-t_0)\left(1+s(l_j)\right)y_j\right] - w \times l_j, \tag{17}$$

where, as previously, $y_i = \tilde{z}_i l_i{}^{\gamma}$. The first-order conditions now require that the firm's behavior follows a modified demand for labor, internalizing the subsidy structure. Note that while the tax rate t_0 affects all firms uniformly and thus will not affect their relative sizes in equilibrium, it is not the case for subsidies that do alter *relative* sizes. The implied allocation rule for labor is now

$$l_j = \frac{a_j{}^{\frac{1}{1-(\gamma-\nu)}}}{\sum_{i=1}^{N} a_j{}^{\frac{1}{1-(\gamma-\nu)}}} L, \tag{18}$$

that is, the subsidy elasticity is isomorphic to a reduction of the span of control parameter γ.[13] These subsidies affect business cycle volatility as follows.

Lemma 3. *SME subsidies dampen business cycle volatility, whereas large-business subsidies amplify it.*

For proof, see Appendix A.3. The result is intuitive; volatility in the granular economy is governed by the degree of concentration as given by the HHI or Ψ^2. SME subsidies reduce concentration and, therefore, reduce Ψ. Conversely, large-business subsidies or favorable tax treatment for big businesses increase concentration, thereby increasing the volatility arising from firm-level shocks. Viewed thus, SME subsidies are ex-ante automatic stabilizers in the sense of McKay and Reis (2021).

4 Quantitative Analysis

After formally demonstrating how misallocation in the form of correlated distortions alters business cycle volatility, I now proceed to quantify its effect. This quantification is challenging since one needs to compute Ψ without perfectly observing a and to hold this vector constant to compute the counterfactual economies to guarantee that the baseline and the counterfactual differ only in the severity of distortions and not in firm-level productivity.

[13] To derive this allocation rule, I need to impose that $0 < \gamma - \nu < 1$ or that $\gamma - 1 < \nu < \gamma$ such that the firm problem is well behaved. For a step-by-step derivation of this result, see Appendix A.3.

In this section, I propose a portable method with minimal data requirements. Employing it requires only a statistical or model-generated description of the firm-size distribution. I proceed to validate this method against existing estimates of the volatility attributable to firm-level shocks. Applying the method to study the stabilization role of correlated distortions suggests that positively correlated distortions in the U.S. have a powerful stabilization effect against firm-level shocks.

4.1 Approximation Strategy

I have defined Ψ in Lemma 1 to be $\Psi = \sqrt{\sum_{i=1}^{N} \overline{\eta}_i^2}$. Therefore, Ψ is a function of the firm-level deterministic ability a and not of the stochastic shocks $\tilde{x}$. Suppose that a_i are a random sample of size N drawn in an iid fashion from an ability distribution with CDF $G(a) : \mathbb{R}_{++} \rightarrow [0,1]$.[14] For a given G and N, one might end up with various values for Ψ, depending on the N realizations of a_i. However, we can consider the expectations with respect to a vector a and the resulting amplification term Ψ.

Define the quantile function Q as the inverse of G such that $Q : [0,1] \rightarrow \mathbb{R}_{++}$. For a random sample a we can define the appropriate vector $q \in [0,1]^N$ such that $a_i = Q(q_i)$ or $G(a_i) = q_i$. Without loss of generality, we can order the vectors such that $a_1 \geqslant a_2, \ldots, \geqslant a_N$. Since q_i are quantiles of a random sample of N iid draws, they are uniformly distributed, and their ordering implies that they also correspond to order statistics. The expected order statistics q_i in a random sample are distributed as $q_i \sim \text{Beta}\,(N+1-i, i)$ with expectations given by $\overline{q}_i = \mathbb{E}\,[q_i] = \frac{N+1-i}{N+1}$.[15] With some abuse of notation, letting $\hat{\Psi}(q)$ denote the value of Ψ as a function of q, we can derive the following first-order Taylor series expansion of Ψ around $\overline{q}$ as

$$\hat{\Psi}\,(q) \approx \hat{\Psi}\,(\overline{q}) + \sum_{i=1}^{N} \frac{\partial \hat{\Psi}(\overline{q})}{\partial q_i}(q_i - \overline{q}_i). \tag{19}$$

Thus, taking expectations of the above, and letting $\overline{\Psi} = \mathbb{E}\left[\hat{\Psi}\,(q)\right]$ we obtain an approximation for the amplification term Ψ.

Lemma 4. *Given an ability distribution G, the amplification term Ψ can be approximated*

[14] Note also that firm-size, both in absolute terms and in relative terms is governed by G through the following $l_i, \overline{\eta}_i \propto a_i^{\frac{1}{1-\gamma}}$.

[15] This is a well known result on the distribution of order statistics in finite random samples. For a formal treatment of this, see Gentle (2009). The textbook orders the sample from smallest to largest, which is the opposite of the order presented above.

as

$$\overline{\Psi} \approx \mathbb{E}\left[\hat{\Psi}\left(\overline{q}\right)\right] = \sqrt{\sum_{i=1}^{N}\left[Q\left(\overline{q}_i\right)\right]^{\frac{2}{1-\gamma}}} \times \left[\sum_{i=1}^{N}\left[Q\left(\overline{q}_i\right)\right]^{\frac{1}{1-\gamma}}\right]^{-1}. \tag{20}$$

The proof is immediate and follows from noting that $\mathbb{E}\left[\left(q_j - \overline{q}_j\right)\right] = 0$, along with the fact that as $\tilde{x} = \overline{x}$ we have that $\overline{\eta}_j = a_j^{\frac{1}{1-\gamma}} \times \left(\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right)^{-1}$.[16]

This approximation has some desirable features. Primarily, it accommodates fat-tailed distributions or distributions without finite moments. That is, even if G itself is ill-behaved, statistically speaking, the quantiles vector q still represents order statistics of a uniform distribution, thus allowing us to take expectations around a function of potentially pathological distributions through their inverse. Moreover, the approximation has an unbounded support for G. To demonstrate, generally, $\overline{q} = \left(\frac{1}{N+1}, \ldots, \frac{N}{N+1}\right)$, where the last term monotonically increases with the sample size.

Pareto distributions. Since many works on granular business cycles restrict attention to the empirically-likely case where the firm-size distribution exhibits a Pareto tail, specializing the above approximation to Pareto distributions offers useful insights. Suppose now that the ability distribution in the economy takes Pareto form with a tail parameter $\zeta_a > 0$ and scale parameter that is normalized to unity[17], i.e. the CDF for a_i is given by $G\left(a\right) = \text{Prob}(a_i \leqslant a) = 1 - a^{-\zeta_a}$. The immediate implication of this assumption is as follows

Lemma 5. *The firm-size distribution as given by employment l_i or by relative sizes $\overline{\eta}_i$ is a Pareto distribution with tail parameter $\zeta_{emp} = \zeta_a\left(1-\gamma\right)$.*

This result is well known; for a formal proof, see Appendix A.4. Note that since $1-\gamma < 1$, the tail parameter of the ability distribution is higher than that of the firm-size distribution. Alternatively, the distribution of ability is more equal than the resulting firm-size distribution.

The tail parameter in Lemma 5 is $\zeta_{\text{emp}} = \zeta_a\left(1-\gamma\right)$, which corresponds to the empirically observed tail of the firm-size distribution when size is measured by employment. E.g., for the U.S. Axtell (2001) estimates $\zeta_{\text{emp}} = 1.059$, and more recently Carvalho and Grassi

[16] Higher order approximations are conceptually possible since the second derivatives and covariances of order statistics are also straightforward to obtain, see Gentle (2009) for exact formulas. However, these are computationally infeasible because the variance-covariance matrix required has N^2 entries.

[17] This assumption is without loss of generality because business cycle volatility is governed by the distribution of relative, and not absolute, sizes.

(2019) estimate $\zeta_{\text{emp}} = 1.097$. Given the multiplicative structure, the two parameters ζ_a and γ cannot be separately identified from the firm-size distribution alone. Incidentally, in the model without distortions $\zeta_{\text{emp}} = \zeta_{\text{sales}}$, where ζ_{sales} is the empirically observed tail of the firm-size distribution when size is measured by sales or revenue (or y_i). However, when including distortions $\zeta_{\text{emp}} \neq \zeta_{\text{sales}}$. I use ζ to streamline notations when the two measures coincide.

$\overline{\Psi}$ *for a Pareto distribution.* Combining the result from equation (20) with the fact that the quantile function of a Pareto distribution is given by $G^{-1}(a)$ or $Q(q) = (1-q)^{-\frac{1}{\zeta_a}}$, we can obtain the following approximation for the expected amplification term in an economy populated with N firms:

Lemma 6. *If a_i is Pareto distributed, the amplification term Ψ can be approximated as.*

$$\overline{\Psi} = \sqrt{\sum_{i=1}^{N} i^{-\frac{2}{\zeta}}} \times \left[\sum_{i=1}^{N} i^{-\frac{1}{\zeta}}\right]^{-1}. \tag{21}$$

For proof, see Appendix A.5. Appendix A.5 also proves the following corollary relating the asymptotic properties of this approximation to Proposition 2 in Gabaix (2011), which serves as an important conceptual validation

Corollary 1. *If a_i is Pareto distributed, then as N grows, business cycle volatility decays as follows*

(a) $\sigma_Z \sim$ *constant* *for* $\zeta < 1$.

(b) $\sigma_Z \sim \log(N)$ *for* $\zeta = 1$.

(c) $\sigma_Z \sim N^{1-\frac{1}{\zeta}}$ *for* $1 < \zeta < 2$.

(d) $\sigma_Z \sim \sqrt{N}$ *for* $\zeta \geqslant 2$.

This asymptotic behavior of TFP volatility in cases (2) through (4) replicates the asymptotic decay properties for TFP volatility given in Proposition 2 of Gabaix (2011), whereas case (1) is new to the best of my knowledge. That is, if the tail of the firm-size distribution is $\zeta < 1$, volatility due to micro-level shocks does not vanish even if the number of firms goes to infinity. Now with this approximation and insights at hand, I turn to validating the approximation against estimates in the existing literature on granular business cycles.

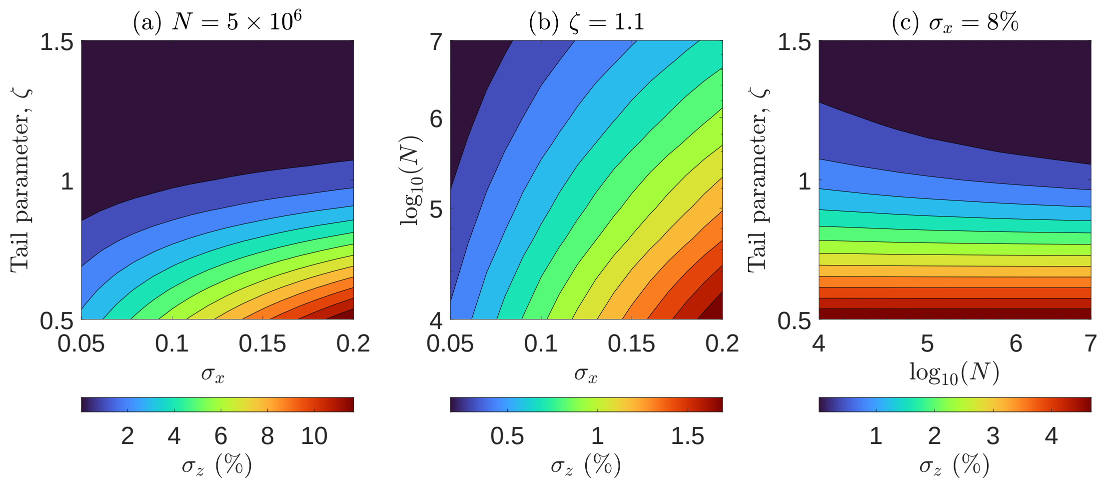


Figure 2: Business Cycle Volatility Due to Granular Shocks

Note: Each panel reports the implied values of business cycle volatility $\sigma_Z = \overline{\Psi} \times \sigma_x$ as a function of the parameters on the axes holding the title parameter fixed.

4.2 Quantification and Validation: Business Cycle Volatility from Granular Shocks?

The above approximation strategy allows one to ask, how much TFP volatility can we expect to arise from micro-level shocks? Equation (7) states that business cycle volatility is a function of: the micro-level volatility σ_x, the number of firms N, and the distribution of relative firm-sizes governed by the tail of the firm-size distribution ζ. Figure 2 demonstrates how business cycle volatility σ_Z changes for given values of σ_x, N, and ζ. Panel (a) shows that business cycle volatility increases when the firm-size distribution is more fat-tailed, i.e., when ζ is lower, and when the micro-level shocks are stronger. Panel (b) reports that holding the tail parameter constant at $\zeta = 1.1$, volatility due to firm-level shocks increases in σ_x and decays as the number of firms grows. Panel (c) fixes σ_x and demonstrates that, as indicated by Corollary 1, volatility does not decay when the tail parameter is small enough.

Validation. To validate the approximation strategy, I utilize estimates from two existing works that employ different approaches to quantify the magnitude of granular business cycles: Gabaix (2011) and Carvalho and Grassi (2019). Results from this exercise are summarized in Table 1.

In his seminal paper, Gabaix cites estimates of the square root of the economy-level HHI (Ψ) in the U.S. as 5.3% and cites firm-level volatility of $\sigma_x = 12\%$. His estimate

Table 1: Validation of Approximation Strategy Against Existing Results

	Paper	Cited estimate σ_Z	Computed σ_Z	σ_x	ζ	Ψ
(1)	Gabaix (2011)	0.630%	0.636%	12% input	1.061 inferred	0.053 input
(2)	Carvalho and Grassi (2019)	0.250%	0.329%	8% input	1.097 input	0.041 inferred

Note: This table recomputes the estimated business cycle volatility from Gabaix (2011) and Carvalho and Grassi (2019) using the approximation strategy in equation (21) with $\gamma = 0.8$ and $N = 4.5 \times 10^6$.

implies that volatility due to such shocks can account for 0.63%. Using Gabaix's HHI as an input, I use equation (21) to infer that $\zeta = 1.06$, which is precisely the estimate of Axtell (2001). Thus, given an HHI, the approximation is consistent with the firm-size distribution.

Next, the work of Carvalho and Grassi (2019) discusses the business cycle properties of a granular economy, somewhat similar to the one described in the present paper, however, relying on different approximation techniques. The authors leverage properties of multinomial distributions and simulate an economy with $N = 4.5 \times 10^6$ firms along a productivity grid consisting of 36 distinct bins. The authors provide a survey of estimates for σ_x, the firm-level Solow residual[18] in the literature citing values between 0.09 and 0.2, using a lower bound estimate of 0.08 as their benchmark. Combined with a tail parameter estimate of $\zeta = 1.097$, they estimate a TFP volatility of 0.25% due to firm-level shocks alone, compared with the total U.S. TFP volatility of 1%.

Using the parameterization of Carvalho and Grassi (2019) and my method, I compute an implied TFP volatility of 0.33%. The two estimates are quite close, and the discrepancy between them is likely due to the bin-based approximation strategy of Carvalho and Grassi (2019), which necessitates truncation of the support, thereby making the large firms slightly too small and suppressing the variance within the very top of the firm-size distribution. My strategy avoids this truncation and yields a slightly higher estimate. To conclude, the two approaches yield very consistent estimates. My approach, however, involves a significantly lower computational burden, requiring only a fraction of a second to compute on a desktop machine, and is more compatible with the case of correlated distortions discussed next. We now turn to quantifying the effects of correlated distortions on business cycle volatility.

[18] In my model, firm-level output is given by $y_j = e^{\tilde{x}_j} a_j l_j^{\gamma}$, thus the firm-level Solow residual or the residual term from regressing output changes on input changes (in logs) in the model is $\tilde{x}_j$.

4.3 How Strong is the Dampening Effect of Positively Correlated Distortions on Volatility from Granular Shocks?

To evaluate quantitatively the effects of correlated distortions on business cycle volatility. I follow Bento and Restuccia (2017), Poschke (2018), and Buera and Fattal Jaef (2018) in specifying that distortions are positively correlated with productivity, taking the following single parameter specification

$$(1-\tau_i) = a_i^{-\phi}, \tag{22}$$

where $\phi \in [0,1]$ corresponds to the elasticity of the distortions with respect to ability. An increase in a_i implies a higher value of τ_i. Observe that using this parametric form, firm-level employment in the model is proportional to $a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{1}{1-\gamma}}$ that is, given equation (22), firm-size is proportional to $a_j^{\frac{1-\phi}{1-\gamma}}$ when the idiosyncratic shocks are at their expected level. Using the logic of Lemma 5, we have that the firm-size distribution has the following observed tail $\zeta_{\text{emp}} = \zeta_a \frac{1-\gamma}{1-\phi}$ which is larger than in the efficient benchmark if ϕ is larger than zero implying a less skewed firm size distribution.

This parametric example provides additional intuition to Proposition 1. Positively correlated distortions reduce concentration and business cycle volatility. Alternatively, reducing ϕ from a positive level to zero or alleviating misallocation increases concentration and volatility due to firm-level shocks. In what follows, I use literature-based estimates to quantify this effect.

Misallocation counterfactuals have sizeable stability implications. To quantify the effect of improving production efficiency or counterfactually removing all positively correlated distortions, I compute the TFP volatility due to firm-level shocks in both cases with and without positively correlated distortions. To do so, I follow Carvalho and Grassi (2019) and calibrate the model to have $N = 4.5 \times 10^6$ firms, with an observed Pareto tail of $\zeta_{\text{emp}} = 1.097$ and set the volatility fo the firm-level Solow residual to $\sigma_x = 8\%$. I also follow Carvalho and Grassi (2019) and set the span of control parameter to $\gamma = 0.8$, which is well within the range accepted in the literature. Additionally, using the estimates from Hsieh and Klenow (2014), I calibrate the distortion elasticity in the U.S. to $\phi = 0.09$, which is an estimate obtained for U.S. establishments. This calibration is summarized in Table 2.

The calibrated model exhibits a TFP volatility of $\sigma_Z = 0.38\%$, more than the economy calibrated without distortions, discussed in the validation exercise and reported in the second row of Table 1. To understand why, note that while both cases have the same tail

Table 2: Calibration of the Baseline Economy with Positively-Correlated Distortions

	Parameter	**Value**	**Source**
N	Number of firms,	4.5×10^6	Carvalho and Grassi (2019)
ζ_{emp}	Observed Pareto tail	1.097	Carvalho and Grassi (2019)
γ	Span of control	0.80	Carvalho and Grassi (2019)
σ_x	Std. of idiosyncratic shock	8%	Carvalho and Grassi (2019)
ϕ	Distortion elasticity	0.09	Hsieh and Klenow (2014)

parameter for employment ζ_{emp}, introducing distortions means that $\zeta_{\text{sales}} \neq \zeta_{\text{emp}}$ and given by

$$\zeta_{\text{sales}} = \zeta_a \frac{1-\gamma}{1-\gamma\phi}, \quad \Longrightarrow \quad \zeta_{\text{sales}} = \zeta_{\text{emp}} \frac{1-\phi}{1-\gamma\phi},$$

the two would only be identical at the limit of $\gamma \to 1$ or without any correlated distortions. With a positive ϕ and $0<\gamma<1$, one obtains that $\zeta_{\text{sales}} < \zeta_{\text{emp}}$ or that sales HHI, which governs volatility due to firm-level shocks, is higher.[19]

Counterfactually removing the distortions increases TFP volatility from the 0.38% reported above to 0.65%, representing a 71% increase in volatility. Alternatively, the presence of positively correlated distortions in the U.S. dampens business cycle volatility due to firm-level shocks by 42% of its potential level.[20] Indicating a large stabilizing role for positively correlated distortions, even with the relatively mild distortion present in the U.S. economy.

To assess the robustness of this sizable stabilization effect, Figure 3 also reports how business cycle volatility changes in response to removing the positively correlated distortions for various values of ϕ and γ. The parameter ranges are chosen to be consistent with the existing literature. Estimates for ϕ in the literature are naturally bounded between zero and unity, Hsieh and Klenow (2014) report values of 0.5 for India and 0.66 for Mexico. Bento and Restuccia (2017) report estimates for various countries with the U.S. value of $\phi = 0.09$ as the lower bar, and most other countries ranging between $\phi = 0.3$ and $\phi = 0.7$. I thus consider values ranging between $\phi = 0$ and $\phi = 0.2$ for my U.S.-based calibration. For all parameter values considered in Figure 3, counterfactually reducing $\phi = 0$ (red) increases volatility compared to the baseline (blue), with the effect increasing monotonically

[19]Anecdotally, Axtell (2001) reports that for the U.S. $\zeta_{\text{sales}} = 0.994$ as measured by invoices, and $\zeta_{\text{emp}} = 1.059$, consistent with a positive value of ϕ, however, the estimates are too noisy to be used in any formal comparisons.

[20]For completeness, appendix C.2 demonstrates how to adjust the approximation in (21) for Ψ^d.

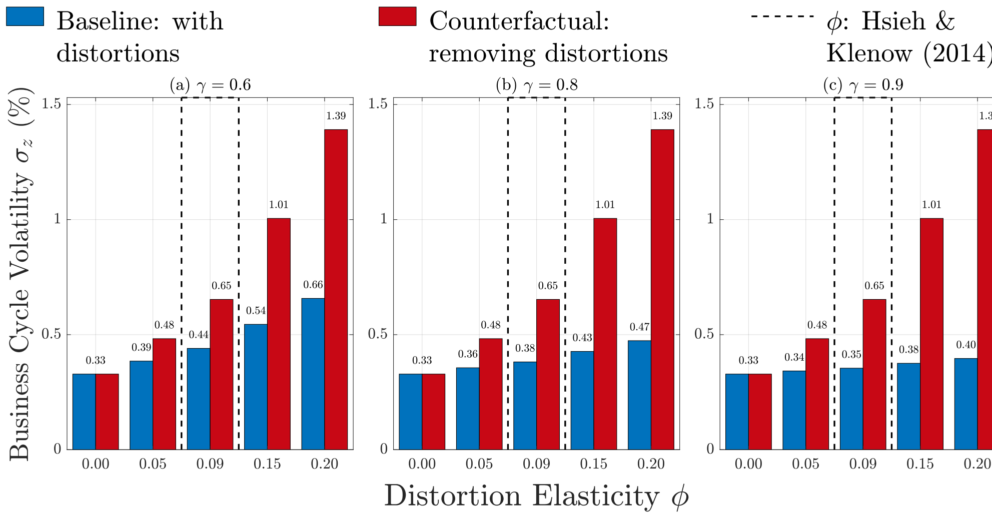


Figure 3: The Effect of Correlated Distortions on Business Cycle Volatility Due to Granular Shocks

Note: Business cycle volatility due to micro-level shocks with positively correlated distortions (benchmark) and when these distortions are counterfactually removed using different assumptions on γ and ϕ.

in ϕ.

The value $\gamma = 0.8$ is a convention in the literature. The admissible range for γ ranges between 0.6 and 0.9. $\gamma = 0.6$ corresponds to the U.S. labor share and serves as a natural lower bound. A natural upper bound for the span of control parameter is one, at the constant returns to scale limit. However, given that corporate pre-tax profits are in the order of magnitude of 10%, $\gamma = 0.9$ serves as a better upper bound since with decreasing returns to scale, the profit share is $1-\gamma$. Figure 3 illustrates that the higher γ is, the stronger the stabilizing effect of correlated distortions. Still, even under the lowest value of $\gamma = 0.6$, a distortion elasticity of $\phi = 0.09$ implies a sizable effect from removing correlated distortions with the volatility due to firm-level shock rising from the benchmark of 0.44% to a counterfactual level of 0.65%.

Appendix B demonstrates that these results are robust to changing two critical assumptions. First, the results are robust when allowing for deviations from Gibrat's law, whereby large firms are less volatile than small ones. Second, the results also hold when allowing firms to re-allocate labor immediately after the shock's realization; however, this case is less amenable to analytical characterization and is handled numerically. I now turn to the normative implications of these results.

5 Welfare Implications

Having established that big businesses are a source of aggregate fluctuations and that empirically observed dispersion in marginal products affects this risk, the final point of this paper is that to mitigate this business cycle risk, a risk-averse social planner or an efficient market mechanism will try to reduce the size of large businesses to some extent. Note that this is so without assuming anything about these large firms' market power or political influence. Instead, the planner is concerned about the presence of large firms in the economy solely because of their size. To formally argue this point, I now derive a formula for the welfare effect of misallocation counterfactuals, including the role of firm-level shocks, highlighting the role of business cycle volatility. I then proceed to solve the optimal factor allocation problem in an economy with a finite number of firms, showing how the planner's choices naturally give rise to an allocation of labor that appears *as if* there are positively correlated distortions in the economy. However, this is not evidence of misallocation but rather of the planner fully internalizing the effect of large firms on the economy and optimizing factor allocations accordingly. Before proceeding, I want to note that my focus in this section is to make a qualitative statement. The model presented exhibits a small welfare cost of business cycles, a property known since the work of Lucas (1987), and uses similar tools. The channel generalizes to more sophisticated frameworks in which concentration maps to business cycle risk.

Consider the welfare of a representative household with flow utility from consumption $u(C_{\tilde{x}})$ that is monotonically increasing and concave, with output evolving as in the previous sections. Due to goods market clearing $Y_{\tilde{x}} = C_{\tilde{x}}$ thus output volatility also implies consumption volatility. Let household preferences exhibit CRRA with $\chi = -C_{\tilde{x}}u''(C_{\tilde{x}}) \big/ u'(C_{\tilde{x}})$, and suppose we normalize $L = 1$. Therefore, $Y_{\tilde{x}} = Z_{\tilde{x}}$ and thus welfare is solely a function of TFP with $u(C_{\tilde{x}}) = u(Z_{\tilde{x}})$. Using standard second-order approximations one can obtain that welfare in the economy with distortions can be approximated by

$$\mathbb{E}\left[u\left(Z^d_{\tilde{x}}\right)\right] = u\left(\overline{Z^d}\right) + \frac{{\sigma_x}^2}{2}\overline{Z^d}\,u'\left(\overline{Z^d}\right) - \chi\frac{1}{2}\left(\Psi^d\sigma_x\right)^2\overline{Z^d}\,u'\left(\overline{Z^d}\right), \tag{23}$$

where $\overline{Z^d} = Z^d_{\tilde{x}}\big|_{\tilde{x}=\bar{x}}$. Note that the efficient production case is analogous but with $Z_{\tilde{x}}$ and Ψ substituted instead of $Z^d_{\tilde{x}}$ and Ψ^d. [21] The log utility case, in which $\chi = 1$ and $\overline{Z^d}\,u'\left(\overline{Z^d}\right) = 1$, is particularly insightful since when we subtract the case with misallocation from the one

[21] For a detailed derivation see appendix C.1.

with efficient production we have that the welfare gain or loss from alleviating misallocation is given by

$$\underbrace{\mathbb{E}\left[u\left(Z_{\tilde{x}}\right)\right]-\mathbb{E}\left[u\left(Z_{\tilde{x}}^{d}\right)\right]}_{\text{Counterfactual welfare gain}}=\underbrace{u\left(\overline{Z}\right)-u\left(\overline{Z^{d}}\right)}_{\substack{\text{Production efficiency}\\ \text{Gain}}}-\underbrace{\frac{\sigma_{x}{}^{2}}{2}\left[\Psi^{2}-\left(\Psi^{d}\right)^{2}\right]}_{\substack{\text{Stability effect}\\ \Delta\text{HHI}}}. \tag{24}$$

The above equations demonstrate that stabilizing business cycle fluctuations due to firm-level shocks is valuable to a risk-averse social planner. Furthermore, it also maps the significant stability effects discussed in section 4 into a welfare-relevant wedge, namely the change in the amplification term or the sales HHI in the economy. Works in the misallocation literature often ignore the second term of equation (24) when evaluating counterfactuals. Proposition 1 implies that this ignored stability effect is signed by the correlation between distortions and productivity. When distortions are positively correlated, we obtain overstated welfare gains. Conversely, the welfare gains are understated when the distortions are negatively correlated.

However, this economy is naïve in the sense that the household does not allocate labor to maximize expected utility but rather expected output, as in the production economies thus described. One can conceptualize this as the result of some agency friction between firm management and the representative household or as an intermediate exercise leading to the result.

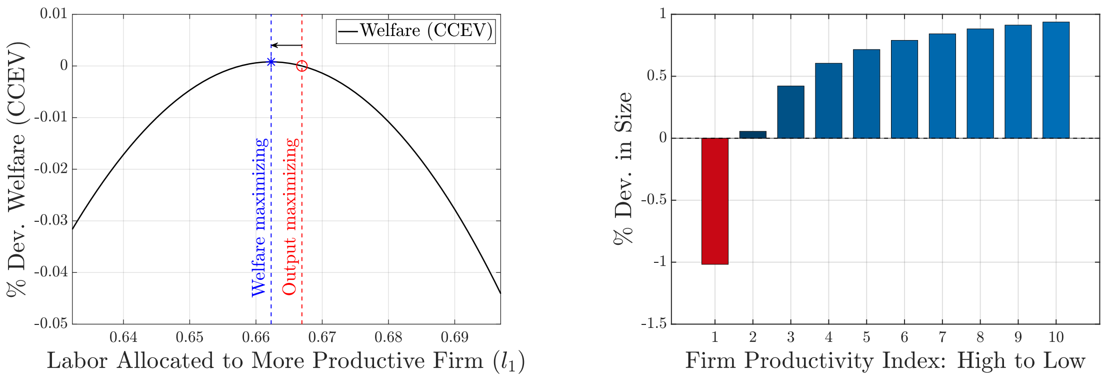
Let us now allow the planner to solve the optimal allocation of labor without knowledge of the shocks $\tilde{x}$. Still, the planner understands how the firm-size distribution governs consumption risk. Formally, let the planner solve

$$\begin{aligned} \max_{l} \quad & \mathbb{E}\left[u\left(Y_{\tilde{x}}\right)\right] \\ \text{s.t.} \quad & Y_{\tilde{x}}=\sum_{i=1}^{N} a_i e^{\tilde{x}_i} l_i^{\gamma}, \quad \sum_{i=1}^{N} l_i = L, \quad l \in \mathbb{R}_{+}^{N} \end{aligned} \tag{25}$$

Note that the first order condition for this problem does not involve equalizing marginal products across firms but rather equalizing risk-weighted expected marginal products as

$$\mathbb{E}[u'\left(Y_{\tilde{x}}\right)\tilde{z}_i\gamma l_i^{\gamma-1}]=\mathbb{E}[u'\left(Y_{\tilde{x}}\right)\tilde{z}_j\gamma l_j^{\gamma-1}], \quad \forall i\neq j. \tag{26}$$

Therefore, allocating more labor to a more productive firm has a larger effect on aggregate



(a) Size of most productive firm ($N = 2$)

(b) Comparing optimal vs output maximizing labor allocation ($N = 10$)

Figure 4: Distorting Production to Stabilize Activity

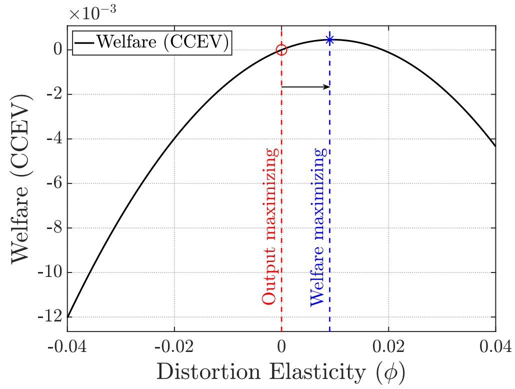
Note: This figure reports the optimal allocation of labor from solving (25) given the calibration in the text.

consumption risk, while low productivity firms can serve as partial insurance. Alternatively, it is the covariance of marginal utility from consumption and firm-level productivity that governs the optimal allocation of labor.

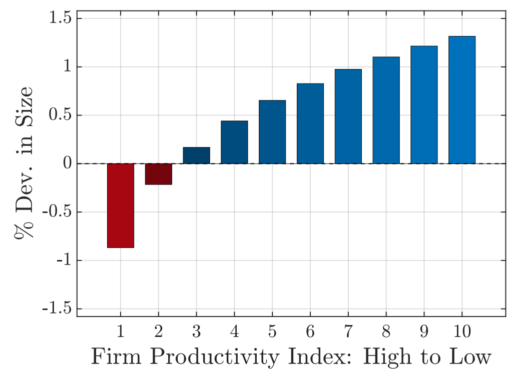
The assumption that $\tilde{x}$ is not known now lends much tractability to the problem, as the planner does not need to allocate labor state-by-state but instead to form a non-state-contingent allocation. While conceptually simple, the problem requires a large state space. To illustrate, if $\tilde{x}_i$ can take k distinct levels, the state space consists of k^N states, which is infeasible for a realistic number of firms.

To numerically study this problem, I calibrate a model economy as follows. Consider an economy with the representative household's utility being CRRA with $\chi = 2$, the span of control parameter is $\gamma = 0.8$, and normalize the total allocation of labor to $L = 1$. Figure 4 demonstrates the optimal allocation of labor resulting from numerically solving (25) with a discrete stochastic process for $\tilde{x}_i$ with $k = 2$ levels such that $\mathbb{E}\left[e^{\tilde{x}_i}\right] = 1$ and $\sigma_x = 8\%$. To be as consistent as possible with previous sections, I choose individual firm productivity levels according to the expected quantiles $\overline{q} = (\frac{1}{N+1}, \ldots, \frac{N}{N+1})$ of a Pareto distribution with a tail parameter such that $\zeta_{\text{emp}} = \zeta_a \frac{1-\phi}{1-\gamma} = 1.097$ yielding a $\zeta_a = 4.99$ when $\phi = 0.09$ as in Hsieh and Klenow (2014).

The simplest case to analyze is the one in which $N = 2$, depicted in Figure 4a, comparing the optimal allocation of labor to the (expected) output maximizing one. The planner understands that the output-maximizing allocation in equation (4) results in suboptimally



(a) Welfare as a function of ϕ



(b) Allocation of labor with optimal ϕ vs output maximizing allocation

Figure 5: Distorting Production to Stabilize Activity: Constant Distortion Elasticity ϕ

Note: This figure reports the optimal allocation of labor obtained from solving (25) with the calibration summarized in the text, but constraining the planner to follow labor allocation rule with a fixed ϕ.

high business cycle volatility and thus allocates less labor to the more productive firm. The planner *trades off production efficiency for economic stability* as the optimal allocation reduces both the expected output level and volatility. Equation (24) is the most straightforward way to illustrate the trade-off between production efficiency and economic stability analytically. The two terms on the right-hand side of the equation demonstrate the trade-off, as raising the first is done at the expense of raising the second. The more risk averse the planner is, the more value they derive from reducing business cycle volatility. This intuition is maintained when we increase the number of firms, as illustrated in Figure 4b, which reports the optimal allocation of labor in the $N = 10$ case compared to the labor allocation in equation (4). The planner again allocates less labor to the high-productivity firms to stabilize economic activity compared with the output-maximizing case.

Note that the above is the result of an efficient welfare-maximizing planner. What if some friction prevents the planner from allocating labor explicitly? Consider a frictional economy where each firm owner maximizes profits due to some un-modeled friction. However, welfare is still given by the same objective. This is a restricted version of (25) in which the planner can only set the institutional environment, or choose ϕ, the distortion elasticity in equation (22) to maximize welfare and not to choose labor at the firm-level. Figure 5 reports the results of this exercise. Importantly, Figure 5a reports that the planner optimally sets a positively correlated distortion $\phi > 0$ to maximize welfare. Note that the resulting allocation of labor given in Figure 5a appears qualitatively similar to the one in

Figure 4b, with the most productive firms under-producing and the least productive firms over-producing compared with the output-maximizing allocation of labor.

To conclude, this exercise challenges the commonly held normative interpretation of positively correlated distortions in the data. Viewed from a static perspective, a dispersion of marginal products is output-decreasing, thus decreasing welfare. However, a stochastic view of the same empirical regularity yields a different interpretation whereby a degree of positively correlated distortion can be optimal as it stabilizes economic activity. I reiterate that these results should be viewed as only qualitative, and a more comprehensive quantitative assessment of 'optimally distorting the firm-size distribution' is left for future work.

6 Concluding remarks

This paper combines two strands of the macroeconomic literature: the literature on granular business cycles and misallocation. The key idea of this paper is that factors affecting the firm-size distribution can affect business cycles through the granularity hypothesis. I develop a framework embedded within the canonical firm-dynamics model à la Hopenhayn (1992) in which correlated distortions dampen or amplify business cycles as a function of the correlation between distortions and productivity, where the empirical literature supports the former. I quantify how much business cycle volatility should respond to this channel in the U.S. economy. I find sizable effects under various model specifications and parameterizations. These effects suggest that positively correlated distortions have a significant stabilizing effect on economic activity.

I demonstrate how size-dependent policies, such as SME subsidies, can reduce business cycle volatility and act as automatic stabilizers in an ex-ante sense. Having SME subsidies in the economy reduces business cycle risk. The final section of this paper challenges the normative interpretation of positively correlated distortions, suggesting that size-dependent policies play a significant role in the policymaker's toolkit when market frictions prevent the allocation of labor from internalizing business cycle risk due to firm-level shocks. The extent to which these policies should be employed depends on the efficiency of the market mechanism and the degree to which business cycle fluctuations resulting from firm-level shocks are detrimental to welfare. Moreover, such policies might introduce new inefficiencies or long-run effects not explored in this paper.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Amaral, L. A. N., S. V. Buldyrev, S. Havlin, P. Maass, M. A. Salinger, H. E. Stanley, and M. H. Stanley (1997). Scaling behavior in economics: the problem of quantifying company growth. *Physica A: Statistical Mechanics and its Applications 244*(1-4), 1–24.

Asker, J., A. Collard-Wexler, and J. D. Loecker (2014). Dynamic inputs and resource (mis)allocation. *Journal of Political Economy 122*(5), 1013–1063.

Atkeson, A. and A. Burstein (2008, December). Pricing-to-market, trade costs, and international relative prices. *American Economic Review 98*(5), 1998–2031.

Axtell, R. L. (2001). Zipf distribution of us firm sizes. *science 293*(5536), 1818–1820.

Baqaee, D. and E. Farhi (2019, 2019). The macroeconomic impact of microeconomic shocks: Beyond hulten's theorem. *Econometrica 87*(4), 1155–1203.

Baqaee, D. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *Quarterly Journal of Economics*.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013, February). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review 103*(1), 305–34.

Bento, P. and D. Restuccia (2017). Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics 9*(3), 267–303.

Boar, C., D. Gorea, and V. Midrigan (2022). Why are returns to private business wealth so dispersed? Technical report, National Bureau of Economic Research.

Boar, C. and V. Midrigan (2024). Markups and inequality. *Review of Economic Studies*, rdae103.

Bornstein, G. and A. Peter (2024). Nonlinear pricing and misallocation. Technical report, National Bureau of Economic Research.

Buera, F. J. and R. N. Fattal Jaef (2018, June 29). The dynamics of development: Innovation and reallocation. *World Bank Policy Research Working Paper* (8505).

Buera, F. J. and Y. Shin (2013). Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy 121*(2), 221–272.

Burstein, A., V. M. Carvalho, and B. Grassi (2025). Bottom-up markup fluctuations. *The Quarterly Journal of Economics*, qjaf029.

Carvalho, V. and X. Gabaix (2013, August). The great diversification and its undoing. *American Economic Review 103*(5), 1697–1727.

Carvalho, V. M. and B. Grassi (2019, April). Large firm dynamics and the business cycle. *American Economic Review 109*(4), 1375–1425.

Clementi, G. L. and H. A. Hopenhayn (2006). A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics 121*(1), 229–265.

Cooley, T. F. and V. Quadrini (2001). Financial markets and firm dynamics. *American Economic Review 91*(5), 1286–1310.

David, J. M., L. Schmid, and D. Zeke (2022). Risk-adjusted capital allocation and misallocation. *Journal of Financial Economics 145*(3), 684–705.

David, J. M. and V. Venkateswaran (2019, July). The sources of capital misallocation. *American Economic Review 109*(7), 2531–67.

di Giovanni, J., A. A. Levchenko, and I. Mejean (2014). Firms, destinations, and aggregate fluctuations. *Econometrica 82*(4), 1303–1340.

di Giovanni, J., A. A. Levchenko, and I. Mejean (2018, January). The micro origins of international business-cycle comovement. *American Economic Review 108*(1), 82–108.

Di Tella, S., C. Malgieri, and C. Tonetti (2025). Risk markups. Technical report, National Bureau of Economic Research.

Edmond, C., V. Midrigan, and D. Y. Xu (2015, October). Competition, markups, and the gains from international trade. *American Economic Review 105*(10), 3183–3221.

Evans, D. S. (1987). Tests of alternative theories of firm growth. *The Journal of Political Economy 95*(4), 657–674.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Gentle, J. (2009). *Computational Statistics*. Statistics and Computing. Springer New York.

Grassi, B. (2018). IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important. Technical report.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013). Who creates jobs? small versus large versus young. *Review of Economics and Statistics 95*(2), 347–361.

Hopenhayn, H. A. (1992, September). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica 60*(5), 1127–1150.

Hopenhayn, H. A. (2014). Firms, misallocation, and aggregate productivity: A review. *Annual Review of Economics 6*(1), 735–770.

Hsieh, C.-T. and P. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Hsieh, C.-T. and P. Klenow (2014). The life cycle of plants in india and mexico. *The Quarterly Journal of Economics 129*(3), 1035–1084.

Hulten, C. R. (1978, 10). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies 45*(3), 511–518.

Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy 105*(2), 211–248.

Koren, M. and S. Tenreyro (2013, February). Technological diversification. *American Economic Review 103*(1), 378–414.

Lucas, R. E. J. (1987). *Models of Business Cycles*. Yrjö Jahnsson Lectures. London: Blackwell Publishing.

McKay, A. and R. Reis (2021). Optimal automatic stabilizers. *The Review of Economic Studies 88*(5), 2375–2406.

Menzio, G. (2024). Markups: A search-theoretic perspective. Technical report, National Bureau of Economic Research.

Moll, B. (2014, October). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review 104*(10), 3186–3221.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Poschke, M. (2018). The firm size distribution across countries and skill-biased change in entrepreneurial technology. *American Economic Journal: Macroeconomics 10*(3), 1–41.

Restuccia, D. and R. Rogerson (2008, October). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics 11*(4), 707–720.

Senga, T. and I. Varotto (2024). Idiosyncratic shocks and investment irreversibility: Capital misallocation over the business cycle.

Stanley, M., L. Amaral, S. Havlin, H. Leschhorn, P. Maass, M. Salinger, and H. Stanley (1996, 02). Scaling behavior in the growth of companies. *Nature 379*.

Sutton, J. (2002). The variance of firm growth rates: the 'scaling'puzzle. *Physica a: statistical mechanics and its applications 312*(3-4), 577–590.

Yeh, C. (2023). Revisiting the origins of business cycles with the size-variance relationship. *Review of Economics and Statistics*, 1–28.

Appendix A Proofs

A.1 Proof of Lemma 1 and a Step by Step Derivation

Proof. Allocation rule for labor. The first order condition for the production problem 2, is given by

$$\gamma a_j {l_j}^{\gamma-1} = w \quad \Longrightarrow \quad l_j = (a_j)^{\frac{1}{1-\gamma}} \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} \tag{27}$$

Using the market clearing condition for labor we obtain

$$\sum_{i=1}^{N} l_i = \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} \sum_{i=1}^{N} {a_i}^{\frac{1}{1-\gamma}} = L. \tag{28}$$

Substituting the above relationship as $\left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} = \frac{L}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}}$, into 27 yields the labor allocation rule in Lemma 1 (1), and the resulting firm-level output

$$l_j = \frac{a_j^{\frac{1}{1-\gamma}}}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}} L, \quad y_j = \frac{a_j^{\frac{1}{1-\gamma}} e^{\tilde{x}_j}}{\left(\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right)^{\gamma}} L^{\gamma} \tag{29}$$

Aggregate production function representation. To derive the aggregate production function representation we can sum all the realizations of y_j to derive the aggregate production function in Lemma 1 (2)

$$Y_{\tilde{x}} = \sum_{i=1}^{N} y_i = \frac{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right]^{\gamma}} L^{\gamma} = Z_{\tilde{x}} \times L^{\gamma}, \quad Z_{\tilde{x}} = \frac{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right]^{\gamma}}. \tag{30}$$

Hulten's theorem. Given the above structure, the elasticity of the resulting economy's log TFP, to a one-percent shock to the productive ability of the j^{th} firm is straightforward:

$$\eta_j = \frac{\partial \log Z_{\tilde{x}}}{\partial \log \tilde{z}_j} = \frac{\partial \log Z_{\tilde{x}}}{\partial \tilde{x}_j} = \frac{1}{Z_{\tilde{x}}} \frac{a_j^{\frac{1}{1-\gamma}} e^{\tilde{x}_j}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right]^{\gamma}} = \frac{a_j^{\frac{1}{1-\gamma}} e^{\tilde{x}_j}}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}, \tag{31}$$

where the first equality sign comes from the definition of an elasticity, the second from the relationship $\log \tilde{z}_j = \log a_j + \tilde{x}_j$, the third from deriving $Z_{\tilde{x}}$ with respect to $\tilde{x}_j$, and the last from substituting in the definition of $Z_{\tilde{x}}$. Furthermore, observe that the sales share of firm j, denoted by $s_{Y,j} = y_j / Y_{\tilde{x}}$ is given by

$$s_{Y,j} = \frac{y_j}{Y_{\tilde{x}}} = \frac{a_j^{\frac{1}{1-\gamma}} e^{\tilde{x}_j}}{\left(\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right)^{\gamma}} L^{\gamma} \times \left[\frac{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i}}{\left(\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}\right)^{\gamma}} L^{\gamma}\right]^{-1} = \eta_j, \tag{32}$$

thus yielding Lemma 1 (3) and demonstrating that Hulten's theorem holds exactly in the environment. Note that Hulten's theorem holds in TFP and in output terms since $\frac{\partial \log Y_{\tilde{x}}}{\partial \log \tilde{z}_j} = \frac{\partial \log Z_{\tilde{x}}}{\partial \log \tilde{z}_j}$.

Business cycle volatility. To derive the volatility of aggregate TFP, which is the final part of the proof, Lemma 1 (4), we need to examine the stochastic properties of $Z_{\tilde{x}}$. Given the idiosyncratic shock structure introduced, TFP in the economy can be restated with some abuse of notation as

$$Z_{\tilde{x}} = \sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} e^{\tilde{x}_i} \times \left(\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}} \right)^{-\gamma} = Z(\tilde{x}), \tag{33}$$

where $\tilde{x} \in \mathbb{R}^N$ is the vector containing all the N iid realizations of $\tilde{x}_j$. Using a Taylor series expansion around the point $\overline{x} \in \mathbb{R}^N$ in which all the idiosyncratic shocks are at their expected level $\overline{x}$, we can derive the approximation given in Lemma 1 (4) as follows

$$\text{VAR}[\log(Z(\tilde{x}))] \approx \text{VAR}\left[\log(Z(\overline{x})) + \sum_{i=1}^{N} \frac{\partial \log(Z(\overline{x}))}{\partial \tilde{x}_i}(\tilde{x}_i - \overline{x})\right] = {\sigma_x}^2 \sum_{i=1}^{N} \overline{\eta}_i^2\,, \tag{34}$$

where $\overline{\eta}_i = \frac{\partial \log(Z_{\tilde{x}})}{\partial \tilde{x}_i}\Big|_{\tilde{x}=\overline{x}}$, and σ_x^2 is the second moment of $\tilde{x}_j$. The standard deviation of log TFP, σ_Z in equation (7) of Lemma 1 (4), is simply the squared root of the above. This ends the proof. $\square$

A.2 Proof of Lemma 2 and a Step by Step Derivation

Proof. Allocation rule for labor. The first order condition for the production problem (8), is given by

$$\gamma a_j (1-\tau_j) {l_j}^{\gamma-1} = w. \tag{35}$$

we can rearrange this expression and use the market clearing condition for labor to obtain

$$\sum_{i=1}^{N} l_i = \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} \sum_{i=1}^{N} (a_i(1-\tau_i))^{\frac{1}{1-\gamma}} = L. \tag{36}$$

Thus, in equilibrium we would have that

$$l_j = \frac{(a_j(1-\tau_j))^{\frac{1}{1-\gamma}}}{\sum_{i=1}^{N}(a_i(1-\tau_i))^{\frac{1}{1-\gamma}}} L, \quad y_j = \frac{{a_j}^{\frac{1}{1-\gamma}} (1-\tau_j)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_j}}{\left(\sum_{i=1}^{N}(a_i(1-\tau_i))^{\frac{1}{1-\gamma}}\right)^{\gamma}} L^{\gamma}, \tag{37}$$

which is the allocation rule given in equation (9), and firm-level output.

Aggregate production function representation. Using the above firm-level output y_j and summing across all N firms allows one to obtain the aggregate production function in equation (10) as

$$Y_{\tilde{x}}^d = \sum_{i=1}^{N} y_i \ = Z_{\tilde{x}}^d \times L^\gamma, \quad Z_{\tilde{x}}^d = \frac{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}} (1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^N \left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right]^\gamma}. \tag{38}$$

Hulten's theorem. As in the efficient production economy, the elasticity of the resulting economy's log TFP, to a one-percent shock to the productive ability of the j^{th} firm is derived as:

$$\delta_j = \frac{\partial \log\left(Z_{\tilde{x}}{}^d\right)}{\partial \log \tilde{z}_j} = \frac{\partial \log\left(Z_{\tilde{x}}{}^d\right)}{\partial \tilde{x}_j} = \frac{1}{Z_{\tilde{x}}{}^d} \frac{a_j^{\frac{1}{1-\gamma}} (1-\tau_j)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_j}}{\left[\sum_{i=1}^N \left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right]^\gamma} = \frac{a_j^{\frac{1}{1-\gamma}} (1-\tau_j)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_j}}{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}} (1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_i}}, \tag{39}$$

this derivation closely follows the one in the proof of Lemma 1 (3). Furthermore, as was the case in Lemma 1, observe that $s_{Y,j}^d = y_j / Y_{\tilde{x}}^d$ is given by

$$s_{Y,j}^d = \frac{y_j}{Y_{\tilde{x}}^d} = \frac{a_j^{\frac{1}{1-\gamma}} \left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_j}}{\left(\sum_{i=1}^N \left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right)^\gamma} L^\gamma \times \left[\frac{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}} (1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_i}}{\left[\sum_{i=1}^N \left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right]^\gamma} \times L^\gamma\right]^{-1} = \delta_j,$$

thus yielding Lemma 1 (3), and demonstrating that Hulten's theorem holds exactly in the environment.

Business cycle volatility. As was the case in Lemma 1 (4), with some abuse of notation we can express

$$Z_{\tilde{x}}^d = \sum_{i=1}^N a_i^{\frac{1}{1-\gamma}} (1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\tilde{x}_i} \ \times \left[\sum_{i=1}^N \left(a_i(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right]^{-\gamma} = Z^d(\tilde{x}), \tag{40}$$

where $\tilde{x}$ is the vector containing all the N iid. realizations of $\tilde{x}_j$. Using a Taylor series expansion around the mean point $\bar{x} \in \mathbb{R}^N$, we can derive the approximation given in Lemma 2 (4) as follows

$$\text{VAR}\left[\log\left(Z^d(\tilde{x})\right)\right] \approx \text{VAR}\left[\log\left(Z^d(\bar{x})\right) + \sum_{i=1}^N \frac{\partial \log\left(Z^d(\bar{x})\right)}{\partial \tilde{x}_i} (\tilde{x}_i - \bar{x})\right] = \sigma_x{}^2 \sum_{i=1}^N \overline{\delta}_i^2, \tag{41}$$

where $\overline{\delta}_i = \left.\frac{\partial \log(Z_{\tilde{x}}^d)}{\partial \tilde{x}_i}\right|_{\tilde{x}=\overline{x}}$, and σ_x^2 is the second moment of $\tilde{x}_j$. The volatility, σ_Z^d, in equation (7) of Lemma 2 (4) is simply the squared root of the above, thus concluding the proof. □

A.3 Proof of Lemma 3

Proof. To prove that SME subsidies dampen business cycle volatility and that large-business subsidies amplify it, we must first derive the aggregate properties of the economy in question. The firm problem in (17) yields the following first order condition

$$l_j = a_j^{\frac{1}{1-(\gamma-\nu)}} \left(C_0{}^{\nu} \frac{(1-t_0)(\gamma-\nu)}{w} \right)^{\frac{1}{1-(\gamma-\nu)}}. \tag{42}$$

Using labor market clearing as

$$L = \sum_{i=1}^{N} l_i = \left(C_0{}^{\nu} \frac{(1-t_0)(\gamma-\nu)}{w} \right)^{\frac{1}{1-(\gamma-\nu)}} \sum_{i=1}^{N} a_j^{\frac{1}{1-(\gamma-\nu)}}, \tag{43}$$

we obtain the labor allocation rule in equation (18). It is trivial that the resulting allocation of labor is identical to the one in an economy with a different span of control parameter γ. Note that output at the firm level is given by

$$y_j = e^{\tilde{x}_j} a_j^{\frac{1+\nu}{1-(\gamma-\nu)}} \left[\sum_{i=1}^{N} a_i^{\frac{1}{1-(\gamma-\nu)}} \right]^{-\gamma} L^{\gamma}, \tag{44}$$

and therefore, aggregate output and productivity are given by

$$Y_{\tilde{x}}^{\text{SME}} = \sum_{i=1}^{N} y_i = Z_{\tilde{x}}^{\text{SME}} L^{\gamma}, \quad Z_{\tilde{x}}^{\text{SME}} = \sum_{i=1}^{N} \frac{e^{\tilde{x}_i} a_i^{\frac{1+\nu}{1-(\gamma-\nu)}}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-(\gamma-\nu)}} \right]^{\gamma}}. \tag{45}$$

Similarly to the derivations in Appendix A.1, and A.2, we can compute the elasticity of TFP with respect to shocks to firm j as

$$\eta_j^{\text{SME}} = \frac{\partial \log Z_{\tilde{x}}^{\text{SME}}}{\partial \tilde{x}_j} = \frac{1}{Z_{\tilde{x}}^{\text{SME}}} \frac{e^{\tilde{x}_j} a_j^{\frac{1+\nu}{1-(\gamma-\nu)}}}{\left[\sum_{i=1}^{N} a_i^{\frac{1}{1-(\gamma-\nu)}} \right]^{\gamma}} = \frac{e^{\tilde{x}_j} a_j^{\frac{1+\nu}{1-(\gamma-\nu)}}}{\sum_{i=1}^{N} e^{\tilde{x}_i} a_i^{\frac{1+\nu}{1-(\gamma-\nu)}}}. \tag{46}$$

Repeating the approximation in equation (34) allows us to state that business cycle volatility, in this case, is given by

$$\sigma_Z^{\text{SME}} = \sigma_x \times \Psi^{\text{SME}}, \quad \Psi^{\text{SME}} = \sqrt{\sum_{i=1}^{N} \left(\overline{\eta}_i^{\text{SME}} \right)^2} \tag{47}$$

where $\overline{\eta}_j^{\text{SME}}$ again denote the values of η_j^{SME} when all the idiosyncratic shocks $\tilde{x}$ are set to their expected value.

With some abuse of notation, let $p = \frac{1+\nu}{1-(\gamma-\nu)}$ and define the auxiliary function

$$h(p) = \Psi^{\text{SME}}(p) = \sqrt{\sum_{i=1}^{N} {a_i}^{2p}} \times \left[\sum_{i=1}^{N} {a_i}^{p}\right]^{-1}. \tag{48}$$

where the efficient production benchmark in equation (7) is given by $h(p = \frac{1}{1-\gamma}) = \Psi$. To prove that stronger SME subsidies dampen business cycle volatility and large-business subsidies amplify them, it is sufficient to show that Ψ^{SME} is a decreasing function of ν or that h is an increasing function of p since we have that $\frac{dp}{d\nu} = \frac{-\gamma}{(1-(\gamma-\nu))^2} < 0$ for all values of ν obeying the regularity condition $\gamma - 1 < \nu < \gamma$. Given the definitions of γ and ν, p can take values between one at the limit where $\gamma \to 0$ and infinity when $\gamma \to 1+\nu$.[22] We now state technical claim 1.

Claim 1. *Let $a \in \mathbb{R}^N_{++}$. The function: $h(p) = \frac{\|a^p\|_2}{\|a^p\|_1}$ where, a^p is the elementwise power of a by p and $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the L^1 and L^2 norms correspondingly. Then, $h'(p) \geqslant 0, \forall p > 0$ with strict inequality if $\exists j,k \in \{1,\ldots,N\}$ for which $a_j \neq a_k$.*

The proof follows below. This concludes the proof of Lemma 3. □

A.3.1 Proof of Claim 1

Proof. For convenience, we will look at $h(p)^2$ as

$$h(p)^2 = \frac{\|a^p\|_2^2}{\|a^p\|_1^2} = \frac{\sum\limits_i \left(a_i^p\right)^2}{\sum\limits_i \left(a_i^p\right)^2 + \sum\limits_{i\neq j} a_i^p a_j^p} = \frac{\sum\limits_i \left(a_i^p\right)^2 + \sum\limits_{i\neq j} a_i^p a_j^p - \sum\limits_{i\neq j} a_i^p a_j^p}{\sum\limits_i \left(a_i^p\right)^2 + \sum\limits_{i\neq j} a_i^p a_j^p} = 1 - \frac{\sum\limits_{i\neq j} a_i^p a_j^p}{\|a^p\|_1^2}.$$

[22] In the latter case, the bounds also impose that $\nu \to 0$ as $\gamma - 1 < \nu$.

Taking the derivative with respect to p yields

$$\begin{aligned}
\left(h(p)^2\right)' &= -\frac{\left(\sum_{i\neq j} a_i^p \cdot a_j^p \cdot \log(a_i a_j)\right)\cdot \|a^p\|_1^2 - 2\,\|a^p\|_1 \cdot \left(\sum_k a_k^p \cdot \log(a_k)\right)\cdot \sum_{i\neq j} a_i^p \cdot a_j^p}{\|a^p\|_1^4} \\
&= -\frac{\left(\sum_{i\neq j} a_i^p \cdot a_j^p \cdot \log(a_i a_j)\right)\cdot \sum_k a_k^p - \left(\sum_k a_k^p \cdot \log(a_k^2)\right)\cdot \sum_{i\neq j} a_i^p \cdot a_j^p}{\|a^p\|_1^3} \\
&= \frac{1}{\|a^p\|_1^3}\cdot \sum_{i\neq j}\sum_k a_i^p \cdot a_j^p \cdot a_k^p \cdot \log\left(\frac{a_k^2}{a_i a_j}\right),
\end{aligned}$$

Where the first equality sign follows from the quotient rule; the second from canceling out $\|a^p\|_1$ and substituting in the definition of the L^1 norm while noting that the vector a^p is strictly positive, and bringing in the two inside the log sign; and the last from using log rules and taking note of the minus sign outside the fraction. We can perform the following transformation

$$\frac{1}{p\,\|a^p\|_1^3}\cdot p \cdot \sum_{i\neq j}\sum_k a_i^p \cdot a_j^p \cdot a_k^p \cdot \log\left(\frac{a_k^2}{a_i a_j}\right) = \frac{2}{p\,\|a^p\|_1^3}\cdot \sum_{i<j}\sum_k a_i^p \cdot a_j^p \cdot a_k^p \cdot \log\left(\frac{a_k^{2p}}{a_i^p a_j^p}\right),$$

where all we did is multiply and divide by p, used the fact that $\log(a^b) = b\log(a)$, and switched the indexing under the sum from $\neq$ to $<$. The above expression signs the derivative of interest. We now prove that this expression is positive, and strictly so if a contains at least one element that is different than the rest.

For the remainder of the proof, I use the following change of variables, letting $u_i = a_i^p$ and proving that

$$M(N) = \sum_{1\leqslant i<j\leqslant N}\sum_{k=1}^{N} u_i \cdot u_j \cdot u_k \cdot \log\left(\frac{u_k^2}{u_i u_j}\right) \geqslant 0, \quad \forall u \in \mathbb{R}_{++}^N,\;\; u_1 \leqslant \ldots \leqslant u_N,$$

and with strict inequality if at least one of the inequalities $u_1 \leqslant \ldots \leqslant u_N$ is a strict inequality. The proof follows by induction leveraging the two-firm example; if changing p modifies the share of the one firm, $\overline{\eta}_1$, the share of the other is equally affected with an opposite sign. Introducing an SME subsidy ($\nu > 0$) and increasing the share of the smaller firm at the expense of the larger one, thus leading p and Ψ to decrease. Formally, when $N = 2$ the function $h(p)$ is monotonically increasing in p as follows in the base case. I then proceed by induction to demonstrate that this will be true for all integer $N \geqslant 2$.

Base case: $N = 2$:

$$M(N=2) = \sum_{k=1}^{2} u_1 \cdot u_2 \cdot u_k \cdot \log\left(\frac{u_k^2}{u_1 u_2}\right). \tag{49}$$

If $u_1 = u_2$, the above expression is zero. Otherwise, since we have at least one strict inequality, we know that $u_1 < u_2$ and therefore $u_1^2 < u_1 u_2 < u_2^2$. Thus,

$$M(N=2) = -u_1^2 \cdot u_2 \cdot \log\left(\frac{u_2}{u_1}\right) + u_1 \cdot u_2^2 \cdot \log\left(\frac{u_2}{u_1}\right) = u_1 u_2 \log\left(\frac{u_2}{u_1}\right)(u_2 - u_1) > 0,$$

which is true because $u_1 < u_2$.

Inductive step N: We need to prove that if $M(N-1) \geqslant 0$ then $M(N) \geqslant 0 \ \forall u_1 \leqslant \ldots \leqslant u_N$. Note that we can rewrite $M(N)$ recursively as:

$$\begin{aligned} M(N) = & \underbrace{\sum_{1\leqslant i<j\leqslant N-1} \sum_{k=1}^{N-1} u_i u_j u_k \log\left(\frac{u_k^2}{u_i u_j}\right)}_{M(N-1)} \\ & \underbrace{+ \sum_{i=1}^{N-1}\sum_{k=1}^{N-1} u_i u_N u_k \log\left(\frac{u_k^2}{u_i u_N}\right) + \sum_{1\leqslant i<j\leqslant N-1} u_i u_j u_N \log\left(\frac{u_N^2}{u_i u_j}\right) + \sum_{i=1}^{N-1} u_i u_N^2 \log\left(\frac{u_N}{u_i}\right)}_{=R(N)}, \end{aligned}$$

where R includes all the added terms when increasing N by exactly one, or alternatively $R(N) = M(N) - M(N-1)$. To briefly explain the above recursive formulation, $M(N)$ is a recursive expression that is a function of $M(N-1)$ for all $N > 2$, where $M(2)$ is defined via equation (49). Thus, $M(N)$ is a function of four terms: $M(N-1)$; all expressions involving $j = N$ and $k < N$; all expressions involving $j < N$ and $k = N$; and finally, all expressions involving $j = N$ and $k = N$.

To prove that $M(N) > 0$, it is sufficient to show that that $R(N) > 0$ since $M(N-1)$ is positive by the induction hypothesis. Without loss of generality, it is sufficient to show that

$\frac{R(N)}{u_N} > 0$, as $u \in \mathbb{R}^N_{++}$. Let us examine $\frac{R(N)}{u_N} > 0$

$$\frac{R(N)}{u_N} = \underbrace{\sum_{i=1}^{N-1}\sum_{k=1}^{N-1} u_i u_k \cdot \left(\log\left(\frac{u_k}{u_i}\right) - \log\left(\frac{u_N}{u_k}\right)\right)}_{T_1} \tag{50}$$
$$+ \sum_{1\leqslant i<j\leqslant N-1} u_i u_j \cdot \left(\log\left(\frac{u_N}{u_i}\right) + \log\left(\frac{u_N}{u_j}\right)\right) + \sum_{i=1}^{N-1} u_i u_N \log\left(\frac{u_N}{u_i}\right).$$

It is helpful to decompose the first term of the sum T_1 as follows

$$T_1 = \sum_{i=1}^{N-1}\sum_{k=1}^{N-1} u_i \cdot u_k \cdot \left(\log\left(\frac{u_k}{u_i}\right) - \log\left(\frac{u_N}{u_k}\right)\right) = \sum_{1\leqslant k<i\leqslant N-1} u_i u_k \cdot \left(-\log\left(\frac{u_i}{u_k}\right) - \log\left(\frac{u_N}{u_k}\right)\right)$$
$$+ \sum_{1\leqslant i<k\leqslant N-1} u_i \cdot u_k \cdot \left(\log\left(\frac{u_k}{u_i}\right) - \log\left(\frac{u_N}{u_k}\right)\right) + \sum_{i=1}^{N-1} -u_i^2 \cdot \log\left(\frac{u_N}{u_i}\right) - u_i^2 \cdot \log\left(\frac{u_i}{u_i}\right). \tag{51}$$

The first term on the right-hand side accounts for all the pairs for which $k < i$, the second term accounts for all the pairs in which $i < k$, and the last term accounts for the case where $i = k$.

$$\frac{R(N)}{u_N} = \underbrace{\sum_{i=1}^{N-1}\sum_{k=1}^{N-1} u_i u_k \cdot \left(\log\left(\frac{u_k}{u_i}\right) - \log\left(\frac{u_N}{u_k}\right)\right)}_{T_1}$$
$$+ \sum_{1\leqslant i<j\leqslant N-1} u_i u_j \cdot \left(\log\left(\frac{u_N}{u_i}\right) + \log\left(\frac{u_N}{u_j}\right)\right) + \sum_{i=1}^{N-1} u_i u_N \log\left(\frac{u_N}{u_i}\right)$$
$$= \underbrace{\sum_{1\leqslant k<i\leqslant N-1} u_i u_k \cdot \left(-\log\left(\frac{u_i}{u_k}\right) - \log\left(\frac{u_N}{u_k}\right)\right)}_{H1} + \underbrace{\sum_{1\leqslant i<k\leqslant N-1} u_i \cdot u_k \cdot \left(\log\left(\frac{u_k}{u_i}\right) - \log\left(\frac{u_N}{u_k}\right)\right)}_{H2}$$
$$+ \sum_{i=1}^{N-1} -u_i^2 \cdot \log\left(\frac{u_N}{u_i}\right) - \underbrace{u_i^2 \cdot \log\left(\frac{u_i}{u_i}\right)}_{=0} + \underbrace{\sum_{1\leqslant i<j\leqslant N-1} u_i u_j \cdot \left(\log\left(\frac{u_N}{u_i}\right) + \log\left(\frac{u_N}{u_j}\right)\right)}_{H3}$$
$$+ \sum_{i=1}^{N-1} u_i u_N \log\left(\frac{u_N}{u_i}\right) = \sum_{i=1}^{N-1} (u_N - u_i) \cdot u_i \cdot \log\left(\frac{u_N}{u_i}\right) \geqslant 0$$

Where the first equality sign is just restating equation (50), the second comes from substituting in T_1 from equation (51), and the third from noting that $H_1 + H_2 + H_3 = 0$. Finally, we can verify that by our definition of the vector $u \in \mathbb{R}^N_{++}$ as ordered such that $u_1 \leqslant \ldots \leqslant u_N$, we obtain that both the expressions $(u_N - u_i) \geqslant 0$ element by element and that $\log\left(\frac{u_N}{u_i}\right) > 0$ for all i. Thus concluding that $R(N)$ is positive, and strictly so if $(u_N - u_i)\, u_i \log\left(\frac{u_N}{u_i}\right) > 0$ for at least one i, or that $u_1 \leqslant \ldots \leqslant u_N$ holds with at least one strict inequality. This concludes the proof. $\square$

A.4 Proof of Lemma 5

Proof. Given Lemma 1, firm size is proportional to $a_i^{\frac{1}{1-\gamma}}$. Given that a_i is drawn from a Pareto distribution, we can obtain that

$$\text{Prob}\left(a^{\frac{1}{1-\gamma}} > x\right) = \text{Prob}\left(a > x^{1-\gamma}\right) = x^{-\zeta_a(1-\gamma)} = x^{-\zeta}. \tag{52}$$

$\square$

A.5 Proof of Lemma 6

Proof. The amplification term as given by the approximation in equation (20), along with the definition of $\overline{q}_i = \mathbb{E}[q_i] = \frac{N+1-i}{N+1}$, and the relationship $a_i = Q(q_i)$ allows one to derive that given a Pareto ability distribution we have that

$$\overline{\Psi} = \sqrt{\sum_{i=1}^{N} \overline{\eta}_i^{\,2}} = \frac{\sqrt{\sum_{i=1}^{N} a_i^{\frac{2}{1-\gamma}}}}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}} = \frac{\sqrt{\sum_{i=1}^{N} (1-\overline{q}_i)^{-\frac{1}{\zeta_a}\frac{2}{1-\gamma}}}}{\sum_{i=1}^{N} (1-\overline{q}_i)^{-\frac{1}{\zeta_a}\frac{1}{1-\gamma}}} = \frac{\sqrt{\sum_{i=1}^{N} (1-\overline{q}_i)^{-\frac{2}{\zeta}}}}{\sum_{i=1}^{N} (1-\overline{q}_i)^{-\frac{1}{\zeta}}} = \tag{53}$$
$$\frac{\sqrt{\sum_{i=1}^{N} \left(1-\frac{N+1-i}{N+1}\right)^{-\frac{2}{\zeta}}}}{\sum_{i=1}^{N} \left(1-\frac{N+1-i}{N+1}\right)^{-\frac{1}{\zeta}}} = \frac{\sqrt{\left(\frac{1}{N+1}\right)^{-\frac{2}{\zeta}} \sum_{i=1}^{N} i^{-\frac{2}{\zeta}}}}{\left(\frac{1}{N+1}\right)^{-\frac{1}{\zeta}} \sum_{i=1}^{N} i^{-\frac{1}{\zeta}}} = \frac{\sqrt{\sum_{i=1}^{N} i^{-\frac{2}{\zeta}}}}{\sum_{i=1}^{N} i^{-\frac{1}{\zeta}}}.$$

Where the first equality sign is due to equation (7), the second from the definition of as $\overline{\eta}_j = \frac{a_j^{\frac{1}{1-\gamma}}}{\sum_{i=1}^{N} a_i^{\frac{1}{1-\gamma}}}$, the third from substituting in the relationship $a_i = Q(q_i)$, the forth from noting that $\zeta = (1-\gamma)\zeta_a$, the fifth from exploiting that $\overline{q}_i = \mathbb{E}[q_i] = \frac{N+1-i}{N+1}$ and the rest from straightforward algebra. This concludes the proof of Lemma 6.

To obtain Corollary 1, we need to understand the limit behavior of $\overline{\Psi}$ as $N \to \infty$. For that purpose, it is worthwhile introducing the following notations. Let $\zeta_R(p) = \sum_{i=1}^{\infty} i^{-p}$ denote the Riemann zeta function. I restrict attention to real values of $p > 0$. When $p > 1$ this function is bounded, and represents a convergent series. For $p = 1$, this function is divergent and corresponds to the harmonic sum which diverges at the same rate as $\log(N)$ since its limit behavior is given by $\gamma_{EM} + \log(N)$, with $\gamma_{EM} \approx 0.577$ being the Euler Mascheroni constant.[23] For values of $p < 1$ we can observe that the series increases polynomially in N

[23] The constant is in fact defined by the limit of the difference between the harmonic sum up to N and

since $\sum_{i=1}^{N}\frac{1}{i^p}\approx\int_1^N x^{-p}dx=\frac{N^{1-p}-1}{1-p}$. Now, observe that at the limit we have that

$$\lim_{N\to\infty}\overline{\Psi}=\sqrt{\zeta_R\left(\frac{2}{\zeta}\right)}\times\left[\zeta_R\left(\frac{1}{\zeta}\right)\right]^{-1}, \tag{54}$$

thus, the approximate amplification term $\overline{\Psi}$ can be described as the ratio of two Riemann zeta functions. The above expression can be analyzed in cases

1. When $\zeta<1$ both expressions are given by constants since $\frac{1}{\zeta}>1$.
2. When $\zeta=1$ the numerator is a constant since $\frac{2}{\zeta}=2$, but the denominator is given by the harmonic sum such that asymptotically $\overline{\Psi}\sim\frac{\text{Constant}}{\log(N)}$.
3. When $1<\zeta<2$, the numerator is still a constant since $\frac{2}{\zeta}>1$, but the denominator diverges and its limit behavior is given by $\overline{\Psi}\sim\text{Constant}\times N^{-\left(1-\frac{1}{\zeta}\right)}$.
4. When $\zeta\geqslant 2$, the numerator and denominator diverge the former behaves as $\sqrt{N^{1-\frac{2}{\zeta}}}$ and the latter as $N^{1-\frac{1}{\zeta}}$, thus the ratio exhibits decay at a rate of $\frac{1}{\sqrt{N}}$.

$\square$

Supplemental Appendices

Appendix B Extensions

The analysis presented in the main text is centered around a model of a vertical economy, in which shocks to all firms have identical variance, and factors are allocated before the shock is realized. The fact that Hulten's theorem holds in both the efficient production case in Lemma 1 and the distorted case in Lemma 2 guarantees that, to a first order, positively correlated distortions dampen business cycles in economies with arbitrary production networks.[24] However, the other two limitations, namely, equal variance and no reallocation ex-post, merit a more systematic analysis. The tools developed in the main text facilitate an analysis even in cases where clear analytical results are not possible to derive.

$\log(N)$ as $N\to\infty$.

[24]For an analogous argument, see Gabaix (2011), and for a discussion of the limitations of first-order approximations, see Baqaee and Farhi (2019).

B.1 Size-Volatility Relationship

The baseline model assumes that firm-level volatility σ_x is independent of size (Gibrat's law). However, the literature has documented systematic deviations from Gibrat's law, especially for young and small firms, e.g., Evans (1987), Haltiwanger et al. (2013), such that young and small firms are more volatile. Theoretically, one can think of such deviations as resulting from a frictional view of firm dynamics whereby firms find it difficult to attain their optimal size quickly. These can result from adjustment costs to capital and labor or from financial frictions as proposed by Cooley and Quadrini (2001) and Clementi and Hopenhayn (2006). Alternatively, such an outcome might emerge as the result of within-firm diversification as in the model of Koren and Tenreyro (2013).

For the purpose of the present work, it is more important to describe this statistical regularity and see whether it alters the model's prediction on the link between correlated distortions and business cycle volatility. Works such as Stanley et al. (1996), Amaral et al. (1997), Sutton (2002), Koren and Tenreyro (2013), and most recently Yeh (2023) measure the elasticity between volatility and size, or estimate α in the following functional form $\sigma_x(\text{Size}) = \sigma_0 \times \text{Size}^{-\alpha}$. All of these works find positive values of α, i.e., firm size is negatively correlated with firm volatility, with headline estimates ranging between $\alpha = 0.1$ and $\alpha = 0.25$. Yeh (2023) goes even further and compares this constant-elasticity specification to a non-parametric smoother, and finds that the constant-elasticity assumption performs remarkably well. The author also conducts a structural break analysis for the size volatility relationship and finds no evidence of breaks.

To connect this literature to my analysis, I now develop a modified formula for Ψ, adjusted to allow for $\alpha > 0$, and use estimates from the literature to repeat the analysis in Figure 3. I allow for a size-volatility relationship expressed in terms of sales such that $\sigma_x(\overline{y}_j) = \sigma_0 \times \overline{y}_j^{-\alpha}$.[25] I use the parameterization which includes misallocation for generality and to map empirical estimates, based on observed firm size, more easily.

Lemma 7. *Business cycle volatility in an economy with size-volatility relationship such that $\sigma_x(\overline{\delta}_j) = \sigma_0 \times \overline{\delta}_j^{-\alpha}$ is approximately given by*

$$\sigma_Z^d \approx \frac{\sigma_0}{\overline{Y^d}^{\alpha}} \times \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^{2(1-\alpha)}} = \frac{\sigma_0}{\overline{Y^d}^{\alpha}} \times \Psi^d(\alpha), \tag{55}$$

[25] Given the repeated static nature of my model, I simplify the definition of size to the expected size when all idiosyncratic shocks are switched off, so as to make the model tractable.

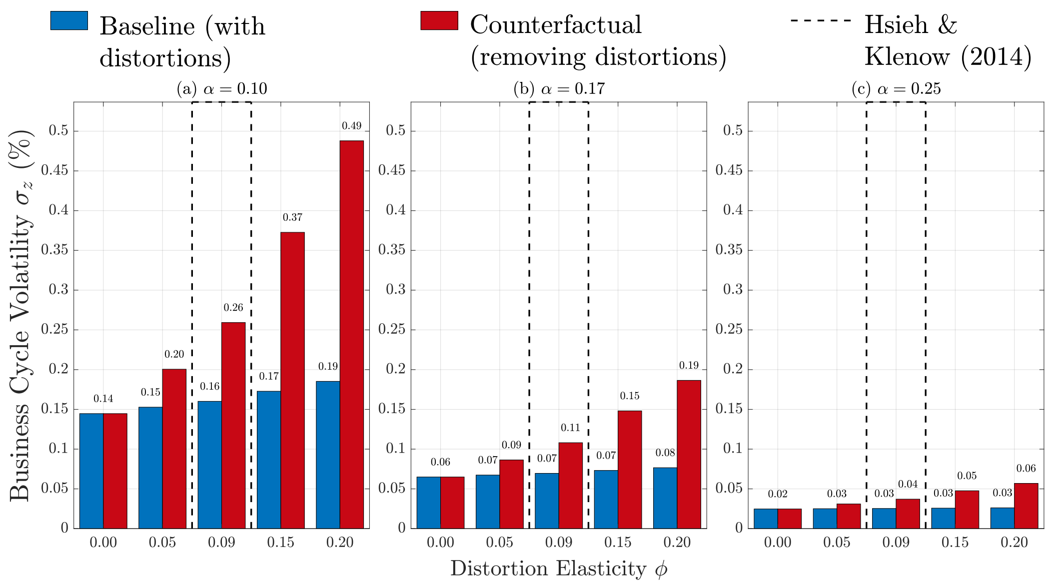


Figure 6: The Effect of Correlated Distortions on Business Cycle Volatility Due to Granular Shocks Allowing for a Size-Volatility Relationship

Note: This figure compares the volatility of business cycle fluctuations due to micro-level shocks, while allowing for deviations from Gibrat's law, with positively correlated distortions (benchmark) vs when these distortions are counterfactually removed using different assumptions on α and ϕ. All computations are done using $\gamma = 0.8$ and $N = 4.5 \times 10^6$ while targeting $\zeta_{\text{emp}} = 1.097$.

where $\overline{Y^d} = Y^d_{\tilde{x}}\big|_{\tilde{x}=\overline{x}}$.

Proof in Appendix C.3. When $\alpha = 0$ this lemma collapses into the structure given by Lemma 2, equation (12), with $\Psi^d = \Psi^d(\alpha = 0)$. Note that $\Psi^d(\alpha)$ is no longer the HHI, but an adjusted measure whereby larger firms also have lower volatility and are thus weighted down. Additionally, the volatility is decaying in total output in the case where all shocks are at their expected level $\overline{Y^d}$, which affects the asymptotic properties for large N in this case, when compared with the baseline.

I follow the calibration strategy of Yeh (2023), and target an *average volatility* of firm-level Solow residual of 12% for different values of α, within the empirically likely range. The technical procedure to conduct this calibration exercise is specified in detail in Appendix C.3. Results are reported in Figure 6 for values of α ranging between 0.1 and 0.25. All else being equal, the level of volatility one can attribute to firm-level shocks is lower, the higher the value of α. However, in all cases considered, the existence of positively correlated distortions in the benchmark stabilizes volatility compared with the efficient production counterfactual. The increase in volatility is monotonically increasing in the strength of the distortions. The basic intuition behind this result is that as long as the elasticity α is strictly below unity, the empirically likely case, as a large firm grows, in relative terms, by one percent, the economy is more exposed to its idiosyncratic shock by $1-\alpha$ percent, and is thus less diversified. It is possible that removing the distortions raises output so much that volatility decreases, since output, given by $\overline{Y^d}$ in equation (55), increases. However, for this effect to be significant, one must have that both α is large and the distortions themselves are severe. This effect is present but not dominant in all examined scenarios reported in Figure 6.

B.2 Ex-post Reallocation

An assumption that underlies the entire analysis thus far is that the allocation of labor is done without knowledge of the realized values of $\tilde{x}_i$. In the main text, I assume that factors are allocated ex ante, and it is impossible to reallocate them after learning the realizations of the shocks. From the perspective of a short-run analysis, I believe this is a likely scenario in practice.[26] However, the stylized nature of my model also allows me to gauge the importance of relaxing this assumption.

To illustrate the key differences that arise compared to my analysis in the main text, suppose now that we were to solve a problem similar to (8) but with full knowledge of the values of $\tilde{x}$. For ease of comparison between the two cases, I use the same notations when possible; for completeness, the full derivation is included in Appendix C.4. In both the efficient production case characterized in Lemma 1 and the inefficient case described in Lemma 2, Hulten's theorem holds exactly. Thus, letting the output share of firm j in the efficient case by $s_{Y,j}$, for the inefficient case by $s^d_{Y,j}$, and their corresponding values when $\tilde{x} = \bar{x}$ by $\bar{s}^d_{Y,j}$ and $\bar{s}_{Y,j}$. I compared the volatility in both cases by exploiting the following relationship

$$\overline{\delta}_j = \bar{s}^d_{Y,j} = \bar{s}_{Y,j} \times \sqrt{1-d_j} = \overline{\eta_j} \times \sqrt{1-d_j}.$$

However, allowing the choice of labor to be made with full knowledge of the firm-level

[26]For example, Google's CEO in announcing major layoffs in January 2023 stated that: "Over the past two years we've seen periods of dramatic growth. To match and fuel that growth, we hired for a different economic reality than the one we face today." The full statement is available at https://blog.google/inside-google/message-ceo/january-update/. Thus, hinting that hiring decisions had been made in the absence of information about the present state of the economy.

shocks, we obtain that

$$\overline{\delta}_j = \overline{s}^d_{Y,j} + \frac{\gamma}{1-\gamma}\left(\overline{s}^d_{Y,j} - \overline{s}^d_{L,j}\right), \tag{56}$$

where $\overline{s}^d_{L,j}$ denotes the input share of firm j in the inefficent production case where $\tilde{x} = \overline{x}$. For the explicit derivation of the above, see Appendix C.4. In the efficient production case, input and output shares are identical, so it would always be the case that $\overline{s}_{Y,j} - \overline{s}_{L,j} = 0$ and Hulten's theorem would hold exactly. However, in the presence of misallocation and ex-post reallocation, input shares and output shares are not necessarily aligned and we have the extra higher-order term $\frac{\gamma}{1-\gamma}\left(\overline{s}^d_{Y,j} - \overline{s}^d_{L,j}\right)$.[27]

The expression in equation (56) is not necessarily positive and depends on the exact values for the implicit taxes or the severity of misallocation. When the difference between input and output shares is sufficiently pronounced, the elasticity $\overline{\delta}_j$ might be negative. Such a negative elasticity has profound implications for understanding the effects of granular shocks on aggregate volatility as follows. Suppose that a firm has a negative elasticity, it implies that we are so far removed from the efficient production case that the input share of this firm is sufficiently high compared to its output share, or that $\overline{s}^d_{Y,j} < \gamma \overline{s}^d_{L,j}$ to be exact. Suppose further that this firm experiences a positive shock. Such a shock draws more resources into this firm at the expense of other, more productive firms, thus reducing TFP as a result. The converse also holds since a negative shock to such a firm frees up inputs to be employed elsewhere.

These negative elasticities render the previously introduced transformation impractical. This is because we cannot map between the ability of the firm and the size of the squared elasticities. These might be high for large negative or positive values. Thus, breaking the previously established link in this case. This rationale implies that understanding the effects of firm-level shocks on business cycle volatility in an environment with ex-post reallocation hinges on the covariance between input and output shares. Specifically, recall that earlier we had

$$\Psi^d = \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^{\,2}} = \sqrt{\sum_{i=1}^{N} \left(\overline{s}^d_{Y,i}\right)^2}, \tag{57}$$

which is also the HHI of the economy computed using output or sales shares. However,

[27] For formal discussion of Hulten's theorem and higher-order effects in that context see Baqaee and Farhi (2019). The particular expression derived here and the effects it generates concerning the propagation of shocks described in the next paragraphs are similar to the effects detailed in Baqaee and Farhi (2020) concerning the propagation of shocks in a horizontal economy in relation to the inverse harmonic mark-up. Examining the formulas derived in Appendix C.4 will show similar ratios between weighted averages of the implicit taxes.

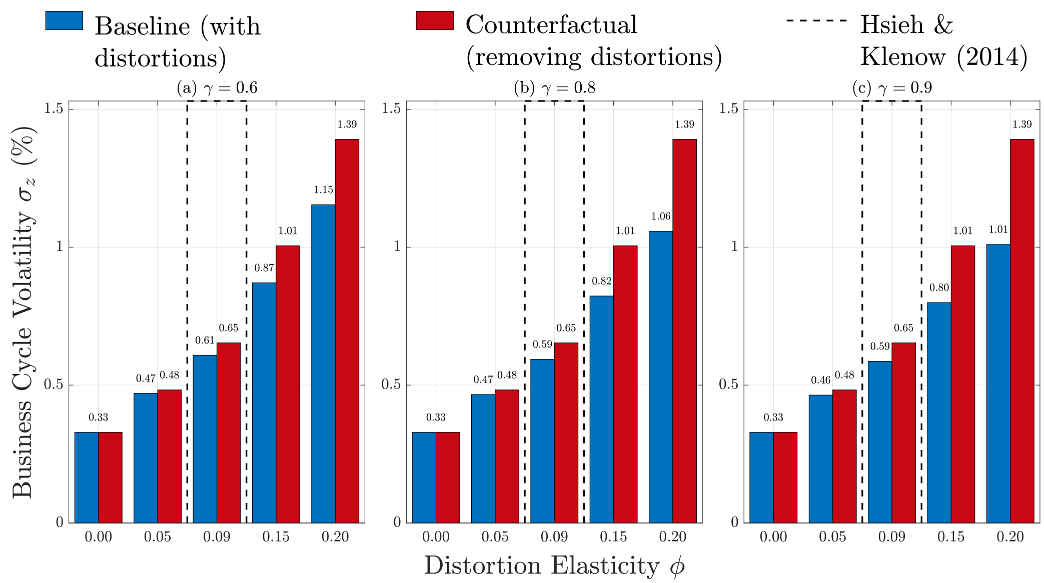


Figure 7: The Effect of Correlated Distortions on Business Cycle Volatility Due to Granular Shocks Allowing for Ex-Post Reallocation

Note: This figure presents the volatility of business cycle fluctuations due to micro-level shocks, while allowing for ex-post reallocation. All computations are done targeting $\zeta_{\text{emp}} = 1.097$ and $N = 4.5 \times 10^6$.

with ex-post reallocation, it follows from equation (56) that

$$\left(\Psi^d\right)^2 = \left(\frac{1}{1-\gamma}\right)^2 \underbrace{\sum_{i=1}^{N} \left(\bar{s}_{Y,i}{}^d\right)^2}_{\text{Sales HHI}} + \left(\frac{\gamma}{1-\gamma}\right)^2 \underbrace{\sum_{i=1}^{N} \left(\bar{s}_{L,i}{}^d\right)^2}_{\text{Input HHI}} - \frac{2\gamma}{(1-\gamma)^2} \underbrace{\sum_{i=1}^{N} \bar{s}_{Y,i}{}^d \bar{s}_{L,i}{}^d}_{\text{Olley-Pakes term}} . \quad (58)$$

Thus, the total volatility in this economy is a function of input and output concentration and the relationship between the two, which is related to the covariance term in the Olley-Pakes decomposition (Olley and Pakes, 1996).[28] The total effect of misallocation on volatility in this case depends crucially on the third term.

Figure 7 uses $\Psi^d = \sqrt{\sum_{i=1}^{N} \bar{\delta}_i{}^2}$ and the formula in equation (56) to evaluate how positively correlated distortions affect business cycle volatility when allowing firms to reallocate factors flexibly in response to the shock. Compared to the results reported in Figure 3,

[28]The above term is stated with respect to the input and output share and using a non-centered measure, whereas the Olley-Pakes formula uses the covariance between firm-level productivity and output shares.

allowing for ex-post reallocation implies a higher degree of volatility at the baseline level. Using the same calibration as in Table 2 implies that volatility due to firm-level shocks in this case is 0.59% compared to 0.38% without reallocation ex post. Still, removing correlated distortions raises this volatility even further to 0.65%, implying a strong dampening effect. The result holds in the same direction for all parameter values explored in this section.

Appendix C Proofs and Additional Derivations for Supplemental Appendix

C.1 Step by step derivation of the approximation in Equation (23)

Consider the following second-order Taylor series approximation for welfare, where the approximation is taken around $\overline{Z^d} = Z^d_{\tilde{x}}\big|_{\tilde{x}=\bar{x}}$

$$u\left(Z^d_{\tilde{x}}\right) \approx u\left(\overline{Z^d}\right) + \left[u'\left(\overline{Z^d}\right)\sum_{i=1}^{N}\frac{\partial Z^d_{\tilde{x}}}{\partial \tilde{x}_i}\right](\tilde{x}_i - \bar{x}) \\ + \frac{1}{2}\left[\sum_{j=1}^{N}\sum_{i=1}^{N}\left(u''\left(\overline{Z^d}\right)\frac{\partial Z^d_{\tilde{x}}}{\partial \tilde{x}_i}\frac{\partial Z^d_{\tilde{x}}}{\partial \tilde{x}_j} + u'\left(\overline{Z^d}\right)\frac{\partial^2 Z^d_{\tilde{x}}}{\partial \tilde{x}_i \partial \tilde{x}_j}\right)(\tilde{x}_i - \bar{x})\left(\tilde{x}_j - \bar{x}\right)\right]. \tag{59}$$

Taking expectations around the above and exploiting the facts that $\mathbb{E}[(\tilde{x}_i - \bar{x})] = 0$, $\mathbb{E}\left[(\tilde{x}_i - \bar{x})^2\right] = {\sigma_x}^2$ and all shocks are i.i.d. thus all the covariance terms are zero yields

$$\mathbb{E}\left[u\left(Z^d_{\tilde{x}}\right)\right] \approx u\left(\overline{Z^d}\right) + \frac{{\sigma_x}^2}{2}\left[\sum_{i=1}^{N}\left(u''\left(\overline{Z^d}\right)\left[\frac{\partial Z^d_{\tilde{x}}}{\partial \tilde{x}_i}\right]^2 + u'\left(\overline{Z^d}\right)\frac{\partial^2 Z^d_{\tilde{x}}}{(\partial \tilde{x}_i)^2}\right)\right]. \tag{60}$$

We can use the elasticities δ_i to express the derivatives as $\frac{\partial Z^d_{\tilde{x}}}{\partial \tilde{x}_i} = \delta_i Z^d_{\tilde{x}}$. This expression allows one to derive that $\frac{\partial^2 Z^d_{\tilde{x}}}{(\partial \tilde{x}_i)^2} = \frac{\partial \delta_i}{\partial \tilde{x}_i} Z^d_{\tilde{x}} + \delta_i \frac{\partial Z^d_{\tilde{x}}}{\partial \tilde{x}_i}$. Additionally, one can compute $\frac{\partial \delta_i}{\partial \tilde{x}_i}$ and obtain that $\frac{\partial^2 Z^d_{\tilde{x}}}{(\partial \tilde{x}_i)^2} = \delta_i Z^d_{\tilde{x}}$.[29] We can combine those derivatives and the definition of Ψ^d to obtain that

$$\mathbb{E}\left[u\left(Z^d_{\tilde{x}}\right)\right] = u\left(\overline{Z^d}\right) + \frac{1}{2}\left(\Psi^d \sigma_x\right)^2 \overline{Z^d}^2 u''\left(\overline{Z^d}\right) + \frac{{\sigma_x}^2}{2}\overline{Z^d}\, u'\left(\overline{Z^d}\right). \tag{61}$$

Finally, by exploiting the CRRA utility specification, i.e., $\chi = -\frac{\overline{Z^d}\, u''\left(\overline{Z^d}\right)}{u'\left(\overline{Z^d}\right)}$, we have that

$$\mathbb{E}\left[u\left(Z^d_{\tilde{x}}\right)\right] = u\left(\overline{Z^d}\right) + \frac{{\sigma_x}^2}{2}\overline{Z^d}\, u'\left(\overline{Z^d}\right) - \chi\frac{1}{2}\left(\Psi^d \sigma_x\right)^2 \overline{Z^d}\, u'\left(\overline{Z^d}\right). \tag{62}$$

Log utility case. When utility takes log form we have that $\chi = 1$ and $\overline{Z^d}\, u'\left(\overline{Z^d}\right) = 1$ thus,

$$\mathbb{E}\left[u\left(Z^d_{\tilde{x}}\right)\right] = u\left(\overline{Z^d}\right) + \frac{{\sigma_x}^2}{2} - \frac{1}{2}\left(\Psi^d \sigma_x\right)^2, \tag{63}$$

whereas in the case without misallocation, one can analogously derive that

$$\mathbb{E}[u(Z_{\tilde{x}})] = u\left(\overline{Z}\right) + \frac{{\sigma_x}^2}{2} - \frac{1}{2}(\Psi \sigma_x)^2. \tag{64}$$

Equation (24) is the result of subtracting the two.

[29]This can be shown directly from

$$\frac{\partial \delta_j}{\partial \bar{x}_j} = \frac{a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_j}\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_i}\right] - \left[a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_j}\right]^2}{\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_i}\right]^2} =$$
$$\frac{\left(a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_j}\right)}{\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_i}\right]} \frac{\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_i}\right] - \left[a_j^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_j}\right]}{\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}} e^{\bar{x}_i}\right]} = \delta_j(1-\delta_j),$$

which implies $\frac{\partial^2 Z^d_{\tilde{x}}}{(\partial \bar{x}_i)^2} = \frac{\partial \delta_i}{\partial \bar{x}_i} \times Z^d_{\tilde{x}} + \delta_i \frac{\partial Z^d_{\tilde{x}}}{\partial \bar{x}_i} = \frac{\partial \delta_i}{\partial \bar{x}_i} \times Z^d_{\tilde{x}} + {\delta_i}^2 \times Z^d_{\tilde{x}} = Z^d_{\tilde{x}} \times \left[\delta_i(1-\delta_i) + {\delta_i}^2\right] = \delta_i Z^d_{\tilde{x}}$.

C.2 Approximating Ψ^d Using the Quantile Function

To approximate the value of Ψ^d in equation (12), with positively correlated distortions as in equation (22), we begin by substituting in the expression δ_j from equation (11) such that

$$\overline{\Psi}^d = \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^2} = \frac{\sqrt{\sum_{i=1}^N \left(a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}}\right)^2}}{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}}} = \frac{\sqrt{\sum_{i=1}^N \left(a_i^{\frac{1}{1-\gamma}}(a_i^{-\phi})^{\frac{\gamma}{1-\gamma}}\right)^2}}{\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(a_i^{-\phi})^{\frac{\gamma}{1-\gamma}}} \tag{65}$$

$$= \frac{\sqrt{\sum_{i=1}^N (Q(q_i))^{2\frac{1-\phi\gamma}{1-\gamma}}}}{\sum_{i=1}^N (Q(q_i))^{\frac{1-\phi\gamma}{1-\gamma}}} = \frac{\sqrt{\sum_{i=1}^N \left(1-\frac{N+1-i}{N+1}\right)^{-\frac{2}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}}}{\sum_{i=1}^N \left(1-\frac{N+1-i}{N+1}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}} = \frac{\sqrt{\sum_{i=1}^N i^{-\frac{2}{\zeta_{\text{sales}}}}}}{\sum_{i=1}^N i^{-\frac{1}{\zeta_{\text{sales}}}}}.$$

This equation follows closely the transformations in equation (53) in the proof of Lemma 6. However, note that, as discussed in the main text, in the presence of correlated distortions as in equation (22), the observed tail of the firm-size distribution maps to the fundamentals as $\zeta_{\text{sales}} = \zeta_a \frac{1-\gamma}{1-\phi\gamma}$.

C.3 Proof of Lemma 7 and Related Required Approximations

Proof. The economy with misallocation and size-dependent volatility behaves identically to the one described in Lemma 2, up to the derivation of business cycle volatility. Starting from equation (41), but this time with $\sigma_x(\overline{y}_j) = \sigma_0 \times \overline{y}_j^{-\alpha}$

$$\text{VAR}\left[\log\left(Z^d(\tilde{x})\right)\right] \tag{66}$$

$$\approx \text{VAR}\left[\log\left(Z^d(\tilde{x}=\overline{x})\right) + \sum_{i=1}^{N} \frac{\partial \log\left(Z^d(\tilde{x}=\overline{x})\right)}{\partial \tilde{x}_i}(\tilde{x}_i - \overline{x})\right] = \sum_{i=1}^{N} \overline{\delta}_i^2 \times \sigma_0^2 \overline{y}_i^{-2\alpha}$$

$$= \left(\frac{\sigma_0}{\overline{Y^d}^{\alpha}}\right)^2 \sum_{i=1}^{N} \overline{\delta}_i^{\,2} \times \left(\frac{\overline{y}_i}{\overline{Y^d}}\right)^{-2\alpha} = \left(\frac{\sigma_0}{\overline{Y^d}^{\alpha}}\right)^2 \sum_{i=1}^{N} \overline{\delta}_i^{\,2(1-\alpha)} = \left(\sigma_0 \overline{Y^d}^{-\alpha} \times \Psi^d(\alpha)\right)^2.$$

The approximation above follows from a first-order Taylor series expansion around the mean point; the first equality sign is the result of substituting in the size-adjusted volatility of each firm, with $\overline{y}_i = y_i|_{\tilde{x}=\overline{x}}$ and $\overline{Y^d} = Y^d_{\tilde{x}}\big|_{\tilde{x}=\overline{x}}$; the second equality sign from multiplying and dividing by $\overline{Y^d}^{-2\alpha}$; the third from noting that $\overline{\delta}_i = \frac{\overline{y}_i}{\overline{Y^d}}$; and the last from letting $\sqrt{\sum_{i=1}^N \overline{\delta}_i^{\,2(1-\alpha)}} = \Psi^d(\alpha)$. Taking the square root of the above concludes the proof. $\square$

Approximating $\Psi^d(\alpha)$ Using the Quantile Function. To approximate the value of $\Psi^d(\alpha)$ in equation (55), with positively correlated distortions as in equation (22) and Pareto distributed ability, we begin by substituting in the expression δ_j from equation (11) such that

$$\overline{\Psi}^d(\alpha) = \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^{2(1-\alpha)}} = \frac{\sqrt{\sum_{i=1}^N \left(a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}}\right)^{2(1-\alpha)}}}{\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(1-\tau_i)^{\frac{\gamma}{1-\gamma}}\right]^{1-\alpha}}$$

$$= \frac{\sqrt{\sum_{i=1}^N \left(a_i^{\frac{1}{1-\gamma}}(a_i^{-\phi})^{\frac{\gamma}{1-\gamma}}\right)^{2(1-\alpha)}}}{\left[\sum_{i=1}^N a_i^{\frac{1}{1-\gamma}}(a_i^{-\phi})^{\frac{\gamma}{1-\gamma}}\right]^{1-\alpha}} = \frac{\sqrt{\sum_{i=1}^N \left(a_i^{\frac{1-\phi\gamma}{1-\gamma}}\right)^{2(1-\alpha)}}}{\left[\sum_{i=1}^N a_i^{\frac{1-\phi\gamma}{1-\gamma}}\right]^{1-\alpha}} = \frac{\sqrt{\sum_{i=1}^N (Q(q_i))^{2(1-\alpha)\frac{1-\phi\gamma}{1-\gamma}}}}{\left[\sum_{i=1}^N (Q(q_i))^{\frac{1-\phi\gamma}{1-\gamma}}\right]^{1-\alpha}}$$

$$= \frac{\sqrt{\sum_{i=1}^N \left(1-\frac{N+1-i}{N+1}\right)^{-\frac{2(1-\alpha)}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}}}{\left[\sum_{i=1}^N \left(1-\frac{N+1-i}{N+1}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}\right]^{1-\alpha}} = \frac{\sqrt{\sum_{i=1}^N i^{-\frac{2(1-\alpha)}{\zeta_{\text{sales}}}}}}{\left[\sum_{i=1}^N i^{-\frac{1}{\zeta_{\text{sales}}}}\right]^{1-\alpha}}.$$

This equation follows closely the transformations in equation (53) in the proof of Lemma 6, with $\zeta_{\text{sales}} = \zeta_a \frac{1-\gamma}{1-\phi\gamma}$.

Calibrating σ_0. To obtain the correct counterpart of σ_0, one needs to use the first-order quantile-based approximation introduced in equation (20), to calibrate to the correct average volatility $\overline{\sigma_x}$ in a sample of N firms. To do so amounts to computing:

$$\overline{\sigma_x} = \frac{\sum_{i=1}^N \sigma_x(\overline{y}_i)}{N} = \frac{\sum_{i=1}^N \sigma_0 \overline{y}_i^{-\alpha}}{N}. \tag{67}$$

Observe that

$$\overline{y_j} = a_j^{\frac{1}{1-\gamma}}(1-\tau_j)^{\frac{\gamma}{1-\gamma}} \times \underbrace{\left[\frac{L}{\sum_{i=1}^N (a_i(1-\tau_i))^{\frac{1}{1-\gamma}}}\right]^{\gamma}}_{C_y} = a_j^{\frac{1-\phi\gamma}{1-\gamma}} C_y, \tag{68}$$

where the last equality sign follows from substituting in $(1-\tau_j) = a_j^{-\phi}$. Next, using the quantile-based approximation method, and $\zeta_{\text{emp}} = \zeta_a \frac{1-\gamma}{1-\phi}$, we have that

$$\sum_{i=1}^{N} (a_i(1-\tau_i))^{\frac{1}{1-\gamma}} = \sum_{i=1}^{N} (a_i)^{\frac{1-\phi}{1-\gamma}} = \sum_{i=1}^{N} (Q(q_i))^{\frac{1-\phi}{1-\gamma}} = \sum_{i=1}^{N} (1-q_i)^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}}$$
$$= \sum_{i=1}^{N} \left(1 - \frac{N+1-i}{N+1}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}} = \sum_{i=1}^{N} \left(\frac{i}{N+1}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}} = (1+N)^{\frac{1}{\zeta_{\text{emp}}}} \sum_{i=1}^{N} i^{-\frac{1}{\zeta_{\text{emp}}}},$$

and thus we can approximate C_y with

$$C_y = \left[\frac{L}{(N+1)^{\frac{1}{\zeta_{\text{emp}}}} \sum_{i=1}^{N} i^{-\frac{1}{\zeta_{\text{emp}}}}}\right]^{\gamma}. \tag{69}$$

Finally, we can target a particular $\overline{\sigma_x}$ and obtain σ_0 via

$$\overline{\sigma_x} = \frac{\sum \sigma_0 \overline{y_j}^{-\alpha}}{N} = C_y{}^{-\alpha}\sigma_0 \left[\sum_{i=1}^{N} \frac{1}{N} a_i{}^{-\alpha\frac{1-\phi\gamma}{1-\gamma}}\right] = C_y{}^{-\alpha}\sigma_0 \left[\sum_{i=1}^{N} \frac{1}{N} Q^{-\alpha\frac{1-\phi\gamma}{1-\gamma}}(q_i)\right]$$
$$= C_y{}^{-\alpha}\sigma_0 \sum_{i=1}^{N} \frac{1}{N}\left(\frac{i}{N+1}\right)^{\alpha\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}} = \frac{C_y{}^{-\alpha}\sigma_0}{N(N+1)^{\frac{\alpha}{\zeta_{\text{sales}}}}} \sum_{i=1}^{N} i^{\frac{\alpha}{\zeta_{\text{sales}}}},$$

implying that

$$\sigma_0 = \overline{\sigma_x} \frac{N(N+1)^{\frac{\alpha}{\zeta_{\text{sales}}}} C_y{}^{\alpha}}{\sum_{i=1}^{N} i^{\frac{\alpha}{\zeta_{\text{sales}}}}}. \tag{70}$$

C.4 The Economy With Misallocation When Labor is Allocated Ex post

Allocation rule for labor. The first order condition for the production problem (8) when the value of $\tilde{x}_i$ is known, is given by

$$\gamma a_i e^{\tilde{x}_i}(1-\tau_i) l_i{}^{\gamma-1} = w. \tag{71}$$

we can again rearrange this expression and use the labor market clearing condition to obtain

$$\sum_{i=1}^{N} l_i = \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} \sum_{i=1}^{N} \left(a_i e^{\tilde{x}_i}(1-\tau_i)\right)^{\frac{1}{1-\gamma}} = L. \tag{72}$$

Thus, the allocation rule for labor is as follows

$$l_j = \frac{\left(a_j e^{\tilde{x}_j}\left(1-\tau_j\right)\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^{N}\left(a_i e^{\tilde{x}_i}\left(1-\tau_i\right)\right)^{\frac{1}{1-\gamma}}} L. \tag{73}$$

Aggregate production function representation. Firm level output is given by

$$y_j = \frac{\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}}{\left(\sum_{i=1}^{N}\left(a_i e^{\tilde{x}_i}\left(1-\tau_i\right)\right)^{\frac{1}{1-\gamma}}\right)^{\gamma}} L^{\gamma}. \tag{74}$$

Aggregating this by summing across all production units allows us to obtain

$$Y_{\tilde{x}} = \sum_{i=1}^{N} y_i = \frac{\sum_{i=1}^{N}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}}{\left(\sum_{i=1}^{N}\left(a_i e^{\tilde{x}_i}\left(1-\tau_i\right)\right)^{\frac{1}{1-\gamma}}\right)^{\gamma}} L^{\gamma}. \tag{75}$$

Transmission of firm-level shocks. To derive the value of the TFP elasticities, observe that TFP is now given by

$$Z_{\tilde{x}}^{d} = \frac{\sum_{i=1}^{N}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}}{\left[\sum_{i=1}^{N}\left(a_i e^{\tilde{x}_i}\left(1-\tau_i\right)\right)^{\frac{1}{1-\gamma}}\right]^{\gamma}}. \tag{76}$$

Therefore we can derive the elasticity by

$$\frac{\partial \log Z_{\tilde{x}}^{d}}{\partial \tilde{x}_j} =$$

$$\frac{1}{Z_{\tilde{x}}^{d}} \frac{\frac{1}{1-\gamma}\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}} - \gamma\left[\sum_{i=1}^{N}\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\right]^{-1}\left((1-\tau_j)a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}\frac{1}{1-\gamma}\sum_{i=1}^{N}\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}{\left[\sum_{i=1}^{N}\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\right]^{\gamma}},$$

which can be condensed into

$$\frac{1}{Z^d_{\tilde{x}}}\frac{\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{1-\gamma}\frac{\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}-\gamma\left[\sum_{i=1}^N\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\right]^{-1}\left(1-\tau_j\right)^{\frac{1}{1-\gamma}}\sum_{i=1}^N\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}{\left[\sum_{i=1}^N\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\right]^{\gamma}}=$$

$$\frac{\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{1-\gamma}\frac{\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}-\gamma\left(\sum_{i=1}^N\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\right)^{-1}\left(1-\tau_j\right)^{\frac{1}{1-\gamma}}\sum_{i=1}^N\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}=$$

$$\frac{\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{1-\gamma}\left[\frac{\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}}{\sum_{i=1}^N\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}-\gamma\frac{\left(1-\tau_j\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}\right]=$$

$$\frac{1}{1-\gamma}\left[\frac{\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}-\gamma\frac{\left(1-\tau_j\right)^{\frac{1}{1-\gamma}}\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}\right]=\delta_{j,\tilde{x}}.$$

Now, observe that the output share or sales share of firm j is given by

$$s_{Y,j}{}^d=\frac{y_j}{Y_{\tilde{x}}}=\frac{\left(1-\tau_j\right)^{\frac{\gamma}{1-\gamma}}\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}, \tag{77}$$

and its input share is given by

$$s_{L,j}{}^d=\frac{l_j}{L}=\frac{\left(1-\tau_j\right)^{\frac{1}{1-\gamma}}\left(a_j e^{\tilde{x}_j}\right)^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N\left((1-\tau_i)a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}}, \tag{78}$$

where the dependence of $s_{L,j}{}^d$ and $s_{Y,j}{}^d$ upon the aggregate state $\tilde{x}$ is suppressed to economize on notation. Thus, we obtain equation (56)

$$\delta_{j,\tilde{x}}=s_{Y,j}{}^d+\frac{\gamma}{1-\gamma}\left(s_{Y,j}{}^d-s_{L,j}{}^d\right). \tag{79}$$

Business cycle volatility. To derive the aggregate effect of firm-level shocks, we can proceed as in the proofs of Lemmas 1 and 2, and use the first-order Taylor series approximation of the variance of log TFP. Letting TFP as a function of $\tilde{x}$ by

$$Z^d_{\tilde{x}}=\frac{\sum_{i=1}^N\left(a_i e^{\tilde{x}_i}\right)^{\frac{1}{1-\gamma}}\left(1-\tau_i\right)^{\frac{\gamma}{1-\gamma}}}{\left[\sum_{i=1}^N\left(a_i e^{\tilde{x}_i}(1-\tau_i)\right)^{\frac{1}{1-\gamma}}\right]^{\gamma}}=Z^d(\tilde{x}), \tag{80}$$

one can use a Taylor series expansion around the point $\tilde{x} = \overline{x}$ to derive that

$$\text{VAR}\left[\log\left(Z^d(\tilde{x})\right)\right] \approx \tag{81}$$
$$\text{VAR}\left[\log\left(Z^d(\tilde{x}=\overline{x})\right) + \sum_{i=1}^{N} \frac{\partial \log\left(Z^d(\tilde{x}=\overline{x})\right)}{\partial \tilde{x}_i}(\tilde{x}_i - \overline{x})\right] = \sigma_x{}^2 \sum_{i=1}^{N} \overline{\delta}_i^2 \,,$$

where $\overline{\delta}_i = \left.\frac{\partial \log\left(Z^d_{\tilde{x}}\right)}{\partial \tilde{x}_i}\right|_{\tilde{x}=\overline{x}}$, and σ_x^2 is the second moment of $\tilde{x}_j$.

Approximating Ψ^d *with Ex-Post Reallocation.* To apply the approximation method of equation (20), we can proceed as follows. The object of interest is

$$\Psi^d = \sqrt{\sum_{i=1}^{N} \overline{\delta}_i^2} \,, \tag{82}$$

it is sufficient to apply the approximation for each value of $\overline{\delta}_j$ and sum them. The approximation is given by

$$\begin{aligned}
\overline{\delta}_j &= \frac{1}{1-\gamma}\left[\frac{(1-\tau_j)^{\frac{\gamma}{1-\gamma}} a_j{}^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N (1-\tau_i)^{\frac{\gamma}{1-\gamma}} a_i{}^{\frac{1}{1-\gamma}}} - \gamma \frac{(1-\tau_j)^{\frac{1}{1-\gamma}} a_j{}^{\frac{1}{1-\gamma}}}{\sum_{i=1}^N (a_i(1-\tau_i))^{\frac{1}{1-\gamma}}}\right] \\
&= \frac{1}{1-\gamma}\left[\frac{a_j{}^{\frac{1-\phi\gamma}{1-\gamma}}}{\sum_{i=1}^N a_i{}^{\frac{1-\phi\gamma}{1-\gamma}}} - \gamma \frac{a_j{}^{\frac{1-\phi}{1-\gamma}}}{\sum_{i=1}^N a_i{}^{\frac{1-\phi}{1-\gamma}}}\right] = \frac{1}{1-\gamma}\left[\frac{\left(Q(\overline{q}_j)\right)^{\frac{1-\phi\gamma}{1-\gamma}}}{\sum_{i=1}^N \left(Q(\overline{q}_i)\right)^{\frac{1-\phi\gamma}{1-\gamma}}} - \gamma \frac{\left(Q(\overline{q}_j)\right)^{\frac{1-\phi}{1-\gamma}}}{\sum_{i=1}^N \left(Q(\overline{q}_i)\right)^{\frac{1-\phi}{1-\gamma}}}\right] \\
&= \frac{1}{1-\gamma}\left[\frac{\left(\frac{j}{1+N}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}}{\sum_{i=1}^N \left(\frac{i}{1+N}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}} - \gamma \frac{\left(\frac{j}{1+N}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}}}{\sum_{i=1}^N \left(\frac{i}{1+N}\right)^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}}}\right] \\
&= \frac{1}{1-\gamma}\left[\frac{j^{-\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}}{\sum_{i=1}^N i^{-\frac{1}{\zeta_a}\frac{1-\phi\gamma}{1-\gamma}}} - \gamma \frac{j^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}}}{\sum_{i=1}^N i^{-\frac{1}{\zeta_a}\frac{1-\phi}{1-\gamma}}}\right].
\end{aligned}$$