Firm Heterogeneity, Market Power and Macroeconomic Fragility*

Alessandro Ferrari† and Francisco Queirós‡

†University of Zurich

‡University of Naples Federico II and CSEF

October 2021

Latest Version Here

Abstract

We investigate how firm heterogeneity and market power affect macroeconomic fragility, defined as the probability of long-lasting recessions. We propose a theory in which the positive interaction between firm entry, competition and factor supply can give rise to multiple steady-states. We show that when firm heterogeneity is large, even small temporary shocks can trigger firm exit and make the economy spiral in a competition-driven poverty trap. Calibrating our model to incorporate the well-documented trends in increasing firm heterogeneity we find that, relative to 2007, an economy with the 1985 level of firm heterogeneity is 5 to 9 times less likely to experience a very persistent recession. We use our framework to study the 2008-09 recession and show that the model can rationalize the persistent deviation of output and most macroeconomic aggregates from trend, including the behavior of net entry, markups and the labor share. Post-crisis cross-industry data corroborates our proposed mechanism. Firm subsidies can be powerful in preventing quasi-permanent recessions and can lead to a 21% increase in welfare.

JEL Classification: E22, E24, E25, E32, L16

Key words: firm heterogeneity, competition, market power, poverty traps, great recession

*A previous version of this paper circulated under the title "Low Competition Traps". We thank Árpád Ábrahám, Pol Antràs, Fernando Broner, Jesús Bueren, Giacomo Calzolari, Vasco Carvalho, Andrea Colciago, Russell Cooper, David Dorn, Jan Eeckhout, Matteo Escudé, Luca Fornaro, Manuel García-Santana, Matteo Gatti, Basile Grassi, David Hemous, Nir Jaimovich, Tullio Jappelli, Philipp Kircher, Ramon Marimon, Isabelle Mejean, Marco Pagano, Giacomo Ponzetto, Edouard Schaal, Florian Scheuer, Armin Schmutzler, Jaume Ventura, Nic Vincent, Xavier Vives and seminar participants at the EUI, UPF, the EEA, SMYE, CSEF and the University of Konstanz. All errors are our own.

1 Introduction

A growing body of evidence suggests that the US and other advanced economies have been undergoing decades of increasing heterogeneity between firms along several dimensions, such as productivity, size and markups.[1] From a theory point of view, increases in firm heterogeneity can generate opposing forces. On the one hand, making the best firms more productive can result in higher aggregate TFP and output. On the other hand, the higher competitive advantage of larger and more productive firms can generate higher market power and rent extraction. This *static* cost may result in lower output and welfare. In this paper, we highlight a different threat posed by rising firm heterogeneity, which is linked to the dynamics of business cycles. In particular, we show that rising firm heterogeneity can result in a higher probability of deep and long-lasting recessions. We refer to such a probability as *macroeconomic fragility*.

We study an RBC model with oligopolistic competition, endogenous firm entry and factor supply. At the heart of our model is a complementarity between competition and factor supply. First, economies with intense competition in product markets feature low profit shares and high factor shares and factor prices, which induce high factor supply. Second, high factor supply allows more firms to enter the market, which results in greater competition. This complementarity can give rise to multiple competition regimes or (stochastic) steady-states. A key contribution of our theory is to show that rising heterogeneity in idiosyncratic TFP can make transitions from high to low steady-states more likely to occur. When firm heterogeneity increases, large, productive firms expand while small, unproductive firms contract. As a consequence, smaller negative shocks can be enough to trigger firm exit and a transition to a steady-state with lower competition, capital stock and output. We characterize this result formally by showing that the minimum size (negative) shock required to trigger a transition from a high to a low steady-state decreases when firm TFP heterogeneity rises. We also show that such an effect can be obtained by higher fixed production costs.

To quantify these economic forces, we provide two calibrations of our model: one where we infer the degree of TFP heterogeneity and fixed costs from 2007 US data moments, and another one where we match the same moments in 1985. The 1985 economy is characterized by lower TFP differences across firms and lower fixed costs. We subject the two model economies to the same shocks and observe that the 2007 economy exhibits significantly larger amplification

[1] Several studies have documented rising firm differences in terms of i) revenue TFP (Andrews et al., 2015; Kehrig, 2015; Decker et al., 2018), ii) size (Bonfiglioli et al., 2018; Autor et al., 2020), and iii) markups (De Loecker et al., 2020; Calligaris et al., 2018; Díez et al., 2018). See Van Reenen (2018) for a summary of the recent findings.

and propagation. Quantitatively, the 2007 economy is between 5 and 9 times more likely to experience a deep and prolonged recession, relative to the 1985 economy. More precisely, the 2007 economy experiences a recession larger than 10% every 75 years, while the 1985 economy experiences one every 380 years. This suggests that rising firm heterogeneity and fixed costs made the US economy significantly more fragile and more prone to long-lasting recessions.

We also use our model economy as a laboratory to study the Great Recession and its aftermath. The 2008 crisis has been marked by a large and persistent deviation of output and other aggregates from trend, something unusual in the entire postwar period. For example, in 2019, output per capita was 14% below its pre-2007 trend, a deviation far larger and persistent than in previous recessions (Figure A.1). We ask whether our model can replicate such a quasi-permanent drop in aggregate output and other variables. To this end, we feed our 2007 economy with a sequence of shocks calibrated to match the behavior of aggregate TFP in 2008-9. The model generates the observed persistent deviation from trend of GDP as well as of investment, hours and aggregate TFP. It also explains the sharp drop of the labor share after 2008. Importantly, when we subject the 1985 economy with the same shocks, the model does not predict a persistent deviation from trend.

We also provide empirical evidence in favor of our proposed mechanism. Our theory provides sharp predictions in terms of cross-industry responses to the business cycle. In particular, for any two industries with the same number of firms, the one with a higher level of heterogeneity reacts more to a negative shock. We test this prediction in the data using concentration as a proxy for heterogeneity in US 6-digit NAICS industries. Consistent with our model predictions, we show that industries that were more concentrated in 2007 experienced larger cumulative declines in net entry, the labor share and economic activity over the 2008-2017 period.

In terms of policy lessons from our theory, we show that firm subsidies can be effective in preventing deep recessions, leading to a welfare gain of 21% in consumption-equivalent terms. The fundamental intuition behind this result is that in our economy firms entry/exit decisions have externalities as they change market power. A planner may then find beneficial to trade off the efficiency loss attached to less productive firms staying in the market with the reduction in the rents of more productive firms. This policy lesson is specific to our model and to the pivotal role of the extensive margin in shaping the degree of aggregate market power. Its logic, however, is more general. The cyclical response of market power is a source of endogenous amplification of aggregate fluctuations. Policies aiming at keeping the degree of product market competition high can prevent the economy from entering in low output regimes and improve welfare.

Related Literature Our paper speaks to three different strands of the literature. First, it is related to the macroeconomic literature studying models of coordination failures (Cooper and John, 1988; Matsuyama, 1991; Benhabib and Farmer, 1994; Farmer and Guo, 1994; Herrendorf et al., 2000). We are not the first to show how multiple equilibria and/or multiple steady-states can arise in a context of imperfect competition and variable markups (Pagano, 1990; Chatterjee et al., 1993; Galí and Zilibotti, 1995; Jaimovich, 2007).[2] We add to this literature by discussing the role of firm-level heterogeneity in shaping macroeconomic *fragility*; as we discuss, this concept is distinct from *existence* of steady-state multiplicity. We also provide a quantification of our mechanism and link it to the 2008 crisis.[3]

Second, this paper relates to a large and growing literature documenting long-run trends in firm heterogeneity and market power. There are several signs indicating rising market power in the US and other advanced economies. For example, Autor et al. (2020) use data from the US census and document rising sales and employment concentration, while Akcigit and Ates (2019) document a rise in patenting concentration. Other studies have documented a secular rise in price-cost markups. Using data from national accounts, Hall (2018) finds that the average sectoral markup increased from 1.12 in 1988 to 1.38 in 2015. De Loecker et al. (2020) document a steady increase in sales-weighted average markups for US public firms between 1980 and 2016.[4] This was driven by both an increasing share of large firms and by rising dispersion in the markup distribution. Close to our approach, De Loecker et al. (2021) argue that declining dynamism and rising market power can be explained, among other channels, by increasing productivity dispersion and fixed costs. Edmond et al. (2021) estimate that the welfare cost of markups can be large and represent up to a 25% loss in consumption-equivalent terms.[5] We contribute to this literature by investigating the business cycle implications of these trends, and

[2]Without relying on multiple equilibria or multiple steady-states, Cooper and John (2000), Etro and Colciago (2010) and Bilbiie et al. (2012) show that a combination of imperfect competition with endogenous entry can generate endogenous amplification and persistence of aggregate fluctuations.

[3]Our paper is also linked to the literature studying the cyclical properties of markups, which includes Rotemberg and Saloner (1986), Jaimovich and Floetotto (2008), Bils et al. (2018), Nekarda and Ramey (2020) and Burstein et al. (2020).

[4]Edmond et al. (2021) show that a cost-weighted average markup displays a less pronounced upward trend. They show that the cost-weighted average markup is the relevant metric for welfare analysis. See also Gutiérrez and Philippon (2017), Traina (2018), Karabarbounis and Neiman (2019) and Bond et al. (2021) on trends in markups.

[5]Gouin-Bonenfant (2018) builds a search model with heterogeneous firms to study the impact of rising productivity dispersion on the labor share.

in particular their impact on the 2008 crisis and the subsequent *great deviation*. We also highlight that increased firm heterogeneity and market power may have negative welfare consequences through an increase in macroeconomic fragility.

Lastly, this paper relates to the literature focusing on the persistent impact of the 2008 crisis. Schaal and Taschereau-Dumouchel (2015) study a model with endogenous capacity utilization and, like us, they interpret the post-2008 deviation as a transition to a low steady-state. Others have proposed explanations based on the complementarity between firms' innovation decisions and aggregate output (Benigno and Fornaro, 2017; Anzoategui et al., 2019). Finally, Clementi et al. (2017) argue that the persistent decline in firm entry, observed after 2008, is crucial to understand the slow recovery. We add to this literature by arguing that rising firm-level heterogeneity has increased the likelihood of quasi-permanent recessions. We also propose a mechanism based on the interactions between market size and market structure that can jointly match the post crisis behaviour of output, markups, the labor share and the number of active firms.

The rest of the paper is organized as follows. Section 2 sets the general framework and provides the first result relating technology and fragility. Section 3 presents a neoclassical growth model with oligopolistic competition and derives the main theoretical results. Section 4 discusses the calibration and presents the quantitative results. Section 5 focuses on the US great recession and its aftermath. In Section 6, we study the welfare effects of fiscal policy in our model. Section 7 presents the cross-industry empirical evidence. Finally, section 8 concludes.

2 General Framework

In this section we introduce the concept of macroeconomic fragility. We do so in the context of a general economy, featuring a representative household, endogenous factor supply, and a large number of industries (characterized by an endogenous extensive margin of firms, variable markups and love-for-variety). We proceed by first describing the economy. We then define macroeconomic fragility and study how it changes with firm-level heterogeneity.

2.1 Primitives

Demographics The economy contains an infinitely-lived representative household, who has time separable utility $U = \sum_{t=0}^{\infty} \beta^t u(C_t, L_t)$. The household can save by investing in capital, which depreciates at rate δ. The household earns rental and wage rates R_t and w_t, which are determined in competitive factor markets.

Technology There is a final good (used for consumption and investment), which is an aggregate of different varieties. The economy contains $I \in \mathbb{N}^+$ industries and within each industry there is a unit mass of identical submarkets or locations (e.g. different towns where a supermarket can locate).[6] Every industry submarket is characterized by a maximum number of players $N \in \mathbb{N}^+$, each of which can produce a differentiated variety. Operating in a given submarket of industry i requires a per period fixed cost c_i (in units of the final good). Firms purchase factors of production in competitive factor markets, but imperfectly compete in product markets. Firms produce according to $\tilde{F}(\gamma, L, K) = \gamma \cdot F(L, K)$ where γ is idiosyncratic TFP and $F(L, K)$ is a constant returns to scale production function satisfying the Inada conditions. The technology set is further described by a matrix of Hicks-neutral idiosyncratic productivities $\Gamma := [\gamma_{ij}]_{(I \times N)}$ and a vector of fixed costs $\mathcal{C} := [c_i]_{(I \times 1)}$. Given these definitions, the technology set consists of $\Lambda := [\Gamma, \mathcal{C}]$. The aggregate mass of firms can be written as $n = \sum_{i=1}^{I} \sum_{m=1}^{N} \eta_{im} m$, where $\eta_{im} \in [0, 1]$ is the share of locations within industry i with $m = 1, \ldots, N$ firms. We assume that firms enter sequentially and that firms making higher profits enter first. Given these assumptions, the aggregate mass of firms n and the technology set Λ fully characterize the set of active firms. For exposition purposes, we work with a deterministic setting subject to MIT shocks. In particular, we think of the possibility of shocks that destroy a fraction $1 - \chi \in [0, 1]$ of the capital stock. Aggregate uncertainty, by means of stochastic TFP, will be introduced in the model we study in Section 3. We now define the equilibrium in this economy.

Definition 1 (Equilibrium)**.** *A decentralized equilibrium in this economy is a set of policies such that i) all agents optimize; ii) all active firms make no loss; iii) inactive firms would make a loss upon entry and iv) all markets clear.*

We restrict attention economies featuring a unique equilibrium and an arbitrary number of steady-states (to be defined below). We further assume that there exists an aggregate production function.

Definition 2 (Aggregate Production Function)**.** *An aggregate production function is given by*

$$F(\Lambda, K, n(\Lambda, K)) := \Phi(\Lambda, n(\Lambda, K))\, F(K, L),$$

where Λ is the technology set (as defined above), n is the mass of active firms, and $\Phi(\cdot)$ is aggregate TFP. Finally $F(K, L)$ is the CRS production function introduced above.

[6] We adopt this formulation only for tractability reasons, to make the aggregate number of firms a continuously differentiable variable, in spite of the existence of a finite number of industries I.

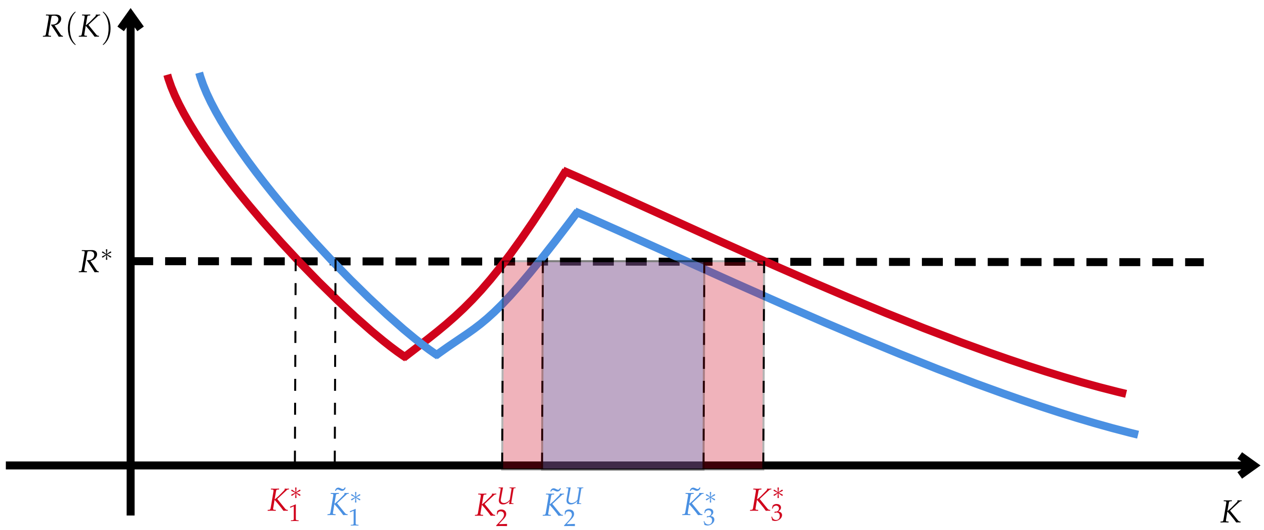


Figure 1: Steady-state multiplicity

Note that L will be also a function of the set of state variables (Λ, K). However, to simply the exposition, we write L. Aggregate TFP $\Phi(\Lambda, n(\Lambda, K))$ will reflect two terms: a weighted average of firm level productivity γ_{ij} and love for variety. From the aggregate production function it is possible to characterize the economy's inverse demand for capital. Let Ω denote the aggregate factor share (i.e. the ratio of total labor and capital payments over gross output Y). Then the equilibrium rental rate is given by

$$R(\Lambda, K) = \Omega(\Lambda, n(\Lambda, K))\, \Phi(\Lambda, n(\Lambda, K))\, F_K(K, L). \tag{1}$$

It will be convenient to define $\Theta(\Lambda, n) := \Omega(\Lambda, n)\, \Phi(\Lambda, n)$, which represents a measure of aggregate factor prices (as it will become clear later).[7] Given the demographic structure of this economy, the long-run supply of capital will be infinitely elastic and given by[8]

$$R^* = \beta^{-1} - (1 - \delta). \tag{2}$$

The economy will feature multiple steady-states if the following equation admits more than one solution

$$R(\Lambda, K) - R^* = 0. \tag{3}$$

[7]Note that this is equal to the Lagrange multiplier of the firms' cost minimization problem.

[8]This can be obtained by evaluating the stationary Euler equation of the representative household.

The red curve in Figure 1 represents the capital demand schedule in equation (1) for a particular economy. Assuming that the Inada conditions are satisfied, we have $R(\Lambda, 0) = \infty$ and $R(\Lambda, \infty) = 0$. Therefore, multiple steady-states will occur only if (1) features at least one increasing part. As explained in Section B.1 of the Supplementary Material, there are different mechanisms that can make $R(\cdot)$ locally increasing in K.[9] For example, suppose that firm entry increases in capital, so that competition and the aggregate factor share $\Omega(\Lambda, n)$ are also increasing in K. If $\Omega(\Lambda, n)$ increases sufficiently fast in K in some region, it can counteract decreasing returns $F_K(\cdot)$ and make $R(\cdot)$ locally increasing. In Figure 1, the economy in red contains two stable steady-states, K_1^* and K_3^*, and an unstable one, K_2^U. The initial capital stock determines the steady-state to which the economy converges.[10] We are particularly interested in the possibility of downward transitions across steady-states. Note that, if the economy starts close to the high steady-state K_3^*, a transition to K_1^* will occur if the economy is hit by a shock that destroys a fraction $1 - K_2^U / K_3^*$ of its capital stock. Thus, the likelihood of a downward transition depends on the distance between K_2^U and K_3^*. For example, under the alternative economy represented in blue, a smaller fraction of the capital stock needs to be destroyed for a downward transition to take place, $1 - \tilde{K}_2^U / \tilde{K}_3^*$. Note that this second economy is characterized by a lower capital demand schedule for sufficiently high values of K. This suggests that a reduction in the aggregate demand for capital can be associated with a higher likelihood of downward transitions across steady-states. We next formalize the concept of *fragility* and explore different mechanisms that, by decreasing the aggregate demand for capital, can result in more likely downward transitions.

2.2 Technological Change and Fragility

Let $\mathcal{K}$ be the ordered set of capital levels in steady-states and $\mathcal{K}^{\mathcal{U}}$ the collection of even elements of $\mathcal{K}$, which represent capital levels in unstable steady-states.[11]

Definition 3 (Steady-States Transitions)**.** *Define an upward transition as the economy moving from the $n-1^{th}$ element of the set $\mathcal{K}$ to the $n+1^{th}$ element, $\forall n$ even. Define a downward transition as the economy moving from the $n+1^{th}$ element of the set $\mathcal{K}$ to the $n-1^{th}$ element, $\forall n$ even.*

[9]In summary, they can be grouped into three broad sets of mechanisms: reallocation, love for variety and market power. We exclude, by assumption, increasing returns in $F(\cdot)$ and backward bending labor supply. See Section B.1 of the Supplementary Material.

[10]When the economy starts on the left of K_2^U it reaches K_1^*, otherwise it achieves K_3^*.

[11]Note that we use even and odd elements as they correspond to unstable and stable steady-states, which are defined based on the local slope of $R(\cdot)$.

We next provide our definition of fragility. Recall that, in our setting, the economy can be hit by MIT shocks that destroy a fraction $1-\chi$ of the capital stock.[12] Our measure of fragility captures the proximity of a *stable* steady-state $\mathcal{K}_{n+1}$ to the preceding *unstable* steady-state $\mathcal{K}_n^U$. Note that the closer these two steady-states are to each other, the lower is the minimum shock needed to trigger a downward transition.

Definition 4 (Fragility)**.** *Let $\chi_{n+1} \in [0,1]$ be defined as $\chi_{n+1} := \mathcal{K}_n^U / \mathcal{K}_{n+1}$. It follows that $1-\chi_{n+1}$ is the minimum size shock needed to trigger a downward transition from $\mathcal{K}_{n+1}$. We call an increase in χ_{n+1} an increase in fragility.*

Two further observations should be made. First, our main focus is on *fragility*, which is different from the *existence* of multiplicity. Although the first requires the second, these are different concepts. Second, the notion of *fragility* is related to but distinct from the idea of *stability* of a steady-state. We think of *fragility* as the possibility of downward transitions only. This is the size of the left partition of the basin of attraction of the steady-state, which is given by $\mathcal{K}_{n+1} - \mathcal{K}_n^U$. The notion of stability, instead, relates to the size of the whole basin of attraction and also accounts for the possibility of upward transitions.

The fundamental question we are interested in is how a change in the technology set, $d\lambda$, affects χ_{n+1}. In Proposition 1 we state a sufficient condition for a technology shift to increase fragility, i.e. $\partial \chi_{n+1}/d\lambda > 0$. This condition says that fragility increases when the rental rate decreases for intermediate values of K (i.e. in the region between steady-states). This depends on the response of $\Theta(\cdot)$ to a technology shift.

Proposition 1 (Comparative Statics)**.** *Let $\mathcal{K}_n^U$ (with n even) be an unstable steady-state and consider a technological shift $d\lambda$. Then $\partial \chi_{n+1}/d\lambda > 0$, i.e. the stable steady-state $\mathcal{K}_{n+1}$ becomes more fragile, if for $K = \{\mathcal{K}_n^U, \mathcal{K}_{n+1}\}$*

$$\frac{\partial}{d\lambda}\Theta\left(\Lambda, n\left(\Lambda, K\right)\right) = \Theta_\lambda\left(\Lambda, n\left(\Lambda, K\right)\right) + \Theta_n\left(\Lambda, n\left(\Lambda, K\right)\right) n_\lambda\left(\Lambda, K\right) < 0.$$

[12]The intuition behind all our results carries through under more general shock structures, with some adjustments. In general the capital level in a steady state depends on the shock realization. This would significantly increase the difficulty in defining a norm between $\mathcal{K}_{n+1}$ and $\mathcal{K}_n$ which will be useful later in defining fragility.

Using the definition of $\Theta(\cdot)$ *we note that this is equivalent to*

$$\underbrace{\Omega(\Lambda,n)\Phi_{\lambda}(\Lambda,n)}_{\Delta\, TFP} + \underbrace{\Omega_{\lambda}(\Lambda,n)\Phi(\Lambda,n)}_{\Delta\, factor\ share} + \underbrace{\left[\Omega_{n}(\Lambda,n)\Phi(\Lambda,n)+\Omega(\Lambda,n)\Phi_{n}(\Lambda,n)\right]n_{\lambda}\left(\Lambda,K\right)}_{\Delta\, number\ of\ firms} < 0. \tag{4}$$

The first two terms characterize the direct effect of $d\lambda$ *on aggregate TFP and the factor share. The last term describes the effect of changes in the number of active firms.*

Proof. See Appendix A.2 ■

Proposition 1 provides a sufficient condition for a (technology-driven) increase in fragility. There are three main channels highlighted in this result: i) changes in aggregate TFP (Φ), ii) changes in the factor share (Ω) and iii) changes in the number of firms (n).[13] To connect this result to the empirical patterns discussed in the introduction, we can think of $d\lambda$ as an increase in firm level heterogeneity. Through the lens of Proposition 1 such an increase can increase fragility by reducing the factor share (due to greater market power), via reallocation of activity towards less productive firms or by reducing the number of active firms. If firms' market shares are increasing in their productivity, as heterogeneity increases three things happen: i) a mechanical efficiency gain due to large firms becoming more productive and through positive reallocation; ii) a compression of factor shares as large firms can exert more market power; iii) entry is harder for firms outside the market, which can generate a loss of varieties and higher market power. Intuitively, if the anti-competitive forces dominate the efficiency gains from higher heterogeneity, the return to capital decreases and it becomes harder to sustain a high steady-state.

The intuition behind the result described in Proposition 1 is similar in spirit to the findings of Baqaee and Farhi (2020) and Edmond et al. (2021) relating dispersion and aggregate TFP. As it will become even clearer in the model presented in Section 3, we bring about the additional effect on the stability of the economy with respect to aggregate fluctuations. Our economy features an endogenous amplification mechanism that can trigger non-linearities in the response to shocks. As firms become more heterogeneous these non-linearities become more salient and the economy can be more likely to experience quasi-permanent recessions.

To dig deeper into the forces underlying the effect of heterogeneity on fragility we specify a RBC model in which firms compete oligopolistically. This additional structure allows us to provide sharper theoretical and quantitative results.

[13] We study these channels in detail in the Supplementary Materials through Remarks B.1, B.2 and B.3.

3 A Model with Firm Heterogeneity and Variable Markups

The model presented in this section builds upon the neoclassical growth model, with a representative household who supplies labor and capital. The technology side is comprised of large number of industries, where firms compete oligopolistically as in Atkeson and Burstein (2008). We start by describing the demand side and the technology structure. Then we analyze the equilibrium of a particular industry (taking aggregate variables as given). Finally, we characterize the general equilibrium.

3.1 Preferences

Time is discrete and indexed by $t = 0, 1, 2, \ldots$. There is a representative, infinitely-lived household with lifetime utility

$$U_t = \mathbb{E} \sum_{t=0}^{\infty} \beta^t \, U\left(C_t, L_t\right), \tag{5}$$

where $0 < \beta < 1$ is the discount factor, $C_t \geq 0$ is consumption of the final good and $L_t \geq 0$ is labor. We adopt the period utility function as in Greenwood et al. (1988)

$$U\left(C_t, L_t\right) = \frac{1}{1-\psi} \left(C_t - \frac{L_t^{1+\nu}}{1+\nu}\right)^{1-\psi}, \tag{6}$$

where $0 \leq \psi \leq 1$ is the inverse of the intertemporal elasticity of substitution and $\nu > 0$ is the inverse of the Frisch elasticity of labor supply.

The representative household contains many individual members, which will be denoted by j. Each individual member can run a firm in the corporate sector. We assume that if two or more individuals run a firm in the same industry, they will behave in a non-cooperative way – i.e. they will compete against each other and will not collude. Nevertheless, all individuals will pool together the profits they make. Hence there is a single dynamic budget constraint[14]

$$K_{t+1} = \left[R_t + (1-\delta)\right] K_t + W_t L_t + \Pi_t^N - C_t, \tag{7}$$

[14] We assume that the economy contains perfect financial markets. In particular, there is a stock market where individuals can trade firms (whose price equals the NPV of profits). Since stocks and capital must offer the same expected return $E_t\{R_{t+1}\}$, firm transactions among individuals will not affect the aggregate budget constraint.

where K_t is capital, R_t is the rental rate, W_t is the wage rate and $\Pi_t^N = \sum_j \Pi_{jt}^N$ are the profits from all firms j net of fixed costs. Capital depreciates at rate $0 \leq \delta \leq 1$ and factor prices R_t and W_t are taken as given. The representative household maximizes (6) subject to (7). Our choice of GHH preferences implies that the aggregate labor supply is given by $L_t^S = W_t^{1/\nu}$.

3.2 Technology

There is a final good (the *numeraire*), which is a CES aggregate of I different industries $Y_t = \left(\sum_{i=1}^I y_{it}^\rho\right)^{1/\rho}$, where y_{it} is the quantity of industry i, and $\sigma_I = 1/(1-\rho) > 1$ is the elasticity of substitution across industries. I is assumed to be large, so that each individual industry has a negligible size in the economy. The output of each industry $i \in \{1, \ldots, I\}$ is itself a CES composite of differentiated goods or varieties $y_{it} = \left(\sum_{j=1}^{n_{it}} y_{jit}^\eta\right)^{1/\eta}$, where n_{it} is the number of active firms in industry i at time t (to be determined endogenously) and $\sigma_G = 1/(1-\eta) > 1$ is the within-industry elasticity of substitution. Following Atkeson and Burstein (2008), we assume that goods are more easily substitutable within industries than across industries: $0 < \rho < \eta \leq 1$. Given these assumptions, the inverse demand for each variety j in industry i is given by

$$p_{ijt} = \left(\frac{Y_t}{y_{it}}\right)^{1-\rho} \left(\frac{y_{it}}{y_{ijt}}\right)^{1-\eta}. \tag{8}$$

We assume that in every industry i there is a maximum number of entrepreneurs $N \in \mathbb{N}$, so that $n_{it} \leq N$. Entrepreneur j can produce her variety by combining capital k_{ijt} and labor l_{ijt} through a CRS technology

$$y_{it} = \underbrace{A_t \, \gamma_{ij}}_{\tau_{ijt}} \left(k_{ijt}\right)^\alpha \left(l_{ijt}\right)^{1-\alpha}. \tag{9}$$

Note that the productivity of each entrepreneur τ_{ijt} is the product of two terms (i) a time-varying aggregate component A_t (common to all industries and types) and (ii) a time-invariant idiosyncratic term γ_{ij}. We refer to A_t as aggregate productivity and to γ_{ij} as j's idiosyncratic productivity. Aggregate productivity follows an auto-regressive process

$$\log A_t = \phi_A \log A_{t-1} + \varepsilon_t$$

with $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$. Without loss of generality, we order idiosyncratic productivities according to $\gamma_{i1} \geq \gamma_{i2} \geq \cdots$. Labor is hired at the competitive wage W_t and capital at the rental rate R_t.

Entrepreneur j can thus produce her variety with marginal cost Θ_t / τ_{ijt}, where

$$\Theta_t := \left(\frac{R_t}{\alpha}\right)^{\alpha} \left(\frac{W_t}{1-\alpha}\right)^{1-\alpha} \tag{10}$$

is the marginal cost function for a Cobb-Douglas technology with unit productivity. We refer to Θ_t as the *factor price index*. In addition to all variable costs, the production of each variety entails a fixed production cost $c_i \geq 0$ per period (which can be possibly different across industries). Such a cost is in units of the *numeraire*.[15]

3.3 Market Structure

To conclude the description of the model, we specify how firms interact. We assume that all firms that enter (and thus incur the fixed cost c_i) play a static Cournot game: they simultaneously announce quantities, taking the output of the other competitors as given.[16] Therefore, each entrepreneur j solves

$$\max_{y_{ijt}} \quad \left(p_{ijt} - \frac{\Theta_t}{\tau_{ijt}}\right) y_{ijt} \qquad \text{s.t. } p_{ijt} = \left(\frac{Y_t}{y_{it}}\right)^{1-\rho} \left(\frac{y_{it}}{y_{ijt}}\right)^{1-\eta} \quad \text{and} \quad y_{it} = \left(\sum_{k=1}^{n_{it}} y_{kit}^{\eta}\right)^{\frac{1}{\eta}}. \tag{11}$$

The solution to (11) yields a system of n_{it} non-linear equations in $\{p_{ijt}\}_{j=1}^{n_{it}}$ (one for each firm)[17]

$$p_{ijt} = \underbrace{\frac{1}{\eta - (\eta - \rho)\, s_{ijt}}}_{\mu_{ijt}} \frac{\Theta_t}{\tau_{ijt}}, \tag{12}$$

where s_{ijt} is the market share of firm $j = 1, \ldots, n_{it}$ and μ_{ijt} is the markup. Equation (12) establishes a positive relationship between market shares and markups. This follows from firms internalizing the impact of size on the price they charge (p_{ijt}); large firms end up restricting output disproportionately more (relative to productivity), thereby charging a high markup. Rearranging (12), one can also see that market shares are a positive function of revenue TFP $(p_{ijt}\, \tau_{ijt})$. Our model thus features a positive association between revenue productivity, size and

[15]This implies that fixed costs do not change with factor prices. We modify this assumption in Section 5.3.

[16]We follow Atkeson and Burstein (2008) and assume that firms enter sequentially in decreasing order of TFP.

[17]This system of first order conditions admits a close-form solution only in the limit case in which there is no differentiation within an industry ($\eta = 1$), as shown in the Supplementary Material B.2.

markups. Therefore, a shock that increases dispersion in revenue TFP will also be associated with greater dispersion in market shares and markups.

To conclude the description of the industry equilibrium, we need to determine the number of active firms n_{it}. To this end, let $\Pi\left(j, n_{it}, \mathcal{F}_{it}, X_t\right) := \left(p_{ijt} - \Theta_t/\tau_{ijt}\right) y_{ijt}$ denote the gross profits of firm $j \leq n_{it}$ in industry i, when there are n_{it} active firms, given a productivity distribution $\mathcal{F}_{it} := \{\gamma_{i1}, \gamma_{i2}, \ldots\}$ and a vector of aggregate variables $X_t := [A_t, Y_t, \Theta_t]$. The equilibrium number of firms must be such that (i) the profits of each active firm are not lower than the fixed cost c_i and (ii) if an additional firm were to enter, its profits would be lower than the fixed cost. Formally, an interior solution $n_{it}^* < N$ to the equilibrium number of firms must satisfy

$$\left[\Pi\left(n_{it}^*, n_{it}^*, \mathcal{F}_i, X_t\right) - c_i\right]\left[\Pi\left(n_{it}^* + 1, n_{it}^* + 1, \mathcal{F}_i, X_t\right) - c_i\right] \leq 0, \quad \forall i = 1, \ldots, I. \tag{13}$$

In the Supplementary Material Section B.2, Lemma B.1 provides an analytical characterization of the profit function under the special case of $\eta = 1$. In particular, we show that the profits of any firm j i) increase in its own idiosyncratic productivity γ_{ij} and ii) decrease in the idiosyncratic productivity of all the other firms γ_{ik}. This means that, as top firms become more productive, small firms make lower profits and become closer to their exist threshold (*ceteris paribus*). This is key to understand some of the aggregate results that we describe next.

3.4 General Equilibrium

Equilibrium Definition We start by defining an equilibrium for this economy. Denoting the history of aggregate productivity shocks by $A^t = \{A_t, A_{t-1}, ...\}$ we have the following definition.

Definition 5 (Equilibrium)**.** *An equilibrium is a sequence of policies* $\left\{C_t\left(A^t\right), K_{t+1}\left(A^t\right), L_t\left(A^t\right)\right\}_{t=0}^{\infty}$ *for the household, firm policies* $\left\{y_{ijt}\left(A^t\right), k_{ijt}\left(A^t\right), l_{ijt}\left(A^t\right)\right\}_{t=0}^{\infty}$*, and a set of active firms* $\left\{n_{it}\left(A^t\right)\right\}_{t=0}^{\infty}$ *with* $\forall i \in \{1, \ldots, I\}$ *such that i) households optimize; ii) all active firms optimize; iii) the slackness free entry condition in equation (13) holds; iv) capital and labor markets clear.*

3.4.1 Static Equilibrium

We now describe the general equilibrium of this economy. We start by focusing on a static equilibrium, in which we describe production and labor supply decisions, taking the aggregate level of capital K_t as given. Later on, we describe the equilibrium dynamics.

Aggregate Production Function Given a $(I \times N)$ matrix of idiosyncratic productivity draws $\mathbf{\Gamma}$ and a vector of active firms $\mathbf{N}_t := \{n_{it}\}_{i=1}^{I}$, aggregate output can be written as

$$Y_t = A_t \, \Phi\left(\mathbf{\Gamma}, \mathbf{N}_t\right) \; L_t^{1-\alpha} K_t^{\alpha}. \tag{14}$$

The term $\Phi\left(\cdot\right)$ represents the endogenous component of aggregate TFP and is a function of the number of active firms, individual productivities and market shares. An analytic expression for $\Phi\left(\mathbf{\Gamma}, \mathbf{N}_t\right)$ is provided in the Supplementary Material Section B.3.

Aggregate Factor Share Let $\mathbb{C}_t := W_t \, L_t + R_t \, K_t$ represent aggregate variable costs. We can write the aggregate factor share $\Omega\left(\cdot\right) := \mathbb{C}_t / Y_t$ as a function of individual markups and market shares[18]

$$\Omega\left(\mathbf{\Gamma}, \mathbf{N}_t\right) = \sum_{i=1}^{I} \sum_{j=1}^{n_{it}} s_{it} \; s_{ijt} \; \mu_{ijt}^{-1}. \tag{15}$$

Combining the previous equation with the first order condition in (12) we can also write the aggregate factor share as a negative function of all industry-level HHI of concentration

$$\Omega\left(\mathbf{\Gamma}, \mathbf{N}_t\right) = \sum_{i=1}^{I} \; s_{it} \; \left[\eta - \left(\eta - \rho\right) \; HHI_{it}\right]. \tag{16}$$

where $HHI_{it} := \sum_{j=1}^{n_{it}} \; s_{ijt}^2$. This result, which is identical to the findings of Grassi (2017) and Burstein et al. (2020), highlights two important relationships. First, industries with higher concentration have larger markups. When highly concentrated industries have large shares in the economy (large s_{it}) the economy's average markup is also high. As the aggregate factor share is the inverse of the economy-wide markup, there is a negative relationship between average concentration and the aggregate factor share.

Factor Prices and Factor Markets We can write the aggregate demand schedules for labor L_t and capital K_t, the analogs of equation (1), as

$$\begin{aligned} W_t \; &= \left(1 - \alpha\right) \; \Theta\left(\mathbf{\Gamma}, \mathbf{N}_t\right) \; L_t^{-\alpha} \, K_t^{\alpha}, & (17) \\ R_t \; &= \quad \alpha \, \Theta\left(\mathbf{\Gamma}, \mathbf{N}_t\right) \; L_t^{1-\alpha} \, K_t^{\alpha-1}. & (18) \end{aligned}$$

[18]Note that the aggregate factor share is equal to the inverse of the aggregate markup $\mu\left(\cdot\right) := Y_t / \mathbb{C}_t$.

It will be convenient to write the factor price index as the product between the aggregate factor share and aggregate TFP. Combining equations (14) and (18), we have

$$\Theta\left(\mathbf{\Gamma},\mathbf{N}_t\right) \;=\; \underbrace{\Omega\left(\mathbf{\Gamma},\mathbf{N}_t\right)}_{\substack{\text{aggregate}\\ \text{factor share}}}\;\underbrace{A_t\,\Phi\left(\mathbf{\Gamma},\mathbf{N}_t\right)}_{\text{aggregate TFP}}. \tag{19}$$

Aggregating over all firms' first order condition (12), we can express $\Theta\left(\mathbf{\Gamma},\mathbf{N}_t\right)$ as a function of the number of active firms $\left(n_{it}\right)$, markups $\left(\mu_{ijt}\right)$ and individual TFP $\left(\gamma_{ij}\right)$

$$\Theta\left(\mathbf{\Gamma},\mathbf{N}_t\right) = A_t\left\{\sum_{i=1}^{I}\left[\sum_{j=1}^{n_{it}}\left(\frac{\gamma_{ijt}}{\mu_{ijt}}\right)^{\frac{\eta}{1-\eta}}\right]^{\frac{1-\eta}{\eta}\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}}. \tag{20}$$

There are two aspects that are worth highlighting about equation (20). First, factor prices are decreasing in markups (holding the number of firms constant): higher markups drive a reallocation of income from factors of production to rents, resulting in lower $\Theta\left(\cdot\right)$. Second, factor prices are increasing in the number of active firms, which happens through two channels: *directly* through the increase in n_{it} (i.e. holding markups fixed) and *indirectly* through the decrease in markups.

The factor demand schedules in (18) can be combined with the labor and capital supply equations

$$L_t^S = W_t^{1/\nu} \quad \text{and} \quad K_t^S = K_t \tag{21}$$

to determine the factor market equilibrium. Combining equations (18) and (21) with (14), we can write aggregate labor and output as a function of the aggregate capital stock K_t, the productivity distribution $\mathbf{\Gamma}$ and the set of active firms $\mathbf{N}_t$

$$L_t \;=\; \left[\left(1-\alpha\right)\Theta\left(\mathbf{\Gamma},\mathbf{N}_t\right)\right]^{\frac{1}{\nu+\alpha}}\,K_t^{\frac{\alpha}{\nu+\alpha}}, \tag{22}$$

$$Y_t \;=\; A_t\,\Phi\left(\mathbf{\Gamma},\mathbf{N}_t\right)\;\left[\left(1-\alpha\right)\Theta\left(\mathbf{\Gamma},\mathbf{N}_t\right)\right]^{\frac{1-\alpha}{\nu+\alpha}}\;K_t^{\alpha\frac{1+\nu}{\nu+\alpha}}. \tag{23}$$

Both aggregate labor L_t and aggregate output Y_t are increasing in the factor price index. Higher factor prices result into higher wages (through (18)) and hence a larger labor supply (through 21). We conclude the characterization of the static equilibrium by determining the set of active firms $\mathbf{N}_t$.

Equilibrium Set of Firms The number of active firms in each industry i is jointly determined by equations (20), (23) and the of inequalities defined in (13). Such a joint system does not admit a general analytical characterization. Nonetheless, we can characterize the particular case in which $\eta = 1$ and all industries are identical. Proposition 2 states the conditions for a symmetric equilibrium across industries.

Proposition 2 (Existence of a Symmetric Equilibrium)**.** *Suppose that $\eta = 1$ and that all industries have the same distribution of idiosyncratic productivities $\mathcal{F}_i = \mathcal{F}$. Then there exist two positive values $\underline{K}(\mathcal{F}, n), \overline{K}(\mathcal{F}, n)$, with $\underline{K}(\mathcal{F}, n) < \overline{K}(\mathcal{F}, n)$ such that when $K_t \in [\underline{K}(\mathcal{F}, n), \overline{K}(\mathcal{F}, n)]$ the economy can sustain a symmetric equilibrium with n firms in every industry.*

Proof. See Appendix A.2. ∎

Intuitively, K_t must be sufficiently large so that all existing n firms can break even, but cannot be too high, for otherwise an additional firm could profitably enter in at least one industry. Figure 2 illustrates the static equilibrium when all industries are (ex-ante) identical. When capital is within the bounds $[\underline{K}(1), \overline{K}(1)]$, the economy is characterized by a fixed market structure with a monopoly in every industry $(n = 1)$; both labor and capital increase in the capital stock, but in a concave fashion (because of decreasing returns). When capital is above $\overline{K}(1)$, at least one industry can sustain a duopoly $(n = 2)$. The increase in competition translates into a higher factor price index and a higher labor supply. For this reason, output is locally convex on capital when $K_t \in [\overline{K}(1), \underline{K}(2)]$.[19]

3.4.2 Equilibrium Dynamics

We next explore the dynamic properties of our economy. Denoting by s_t is the aggregate savings rate (from the household maximization problem), we can write the capital law of motion as

$$K_{t+1} = (1 - \delta) K_t + s_t \cdot Y_t \tag{24}$$

The left panel of Figure 3 represents the law of motion (24) for a particular parameter combination. For the sake of exposition, we assume again that all industries are identical and that the economy is only subject to MIT shocks (we relax these assumptions in the quantitative model).

[19]In the regions $[\overline{K}(n), \underline{K}(n+1)]$, the economy is characterized by an asymmetric equilibrium in which some industries have n firms, and some others have $n + 1$ firms. The number of industries with $n + 1$ firms is such that the last firm exactly breaks even. See Supplementary Material Section B.3 for details.

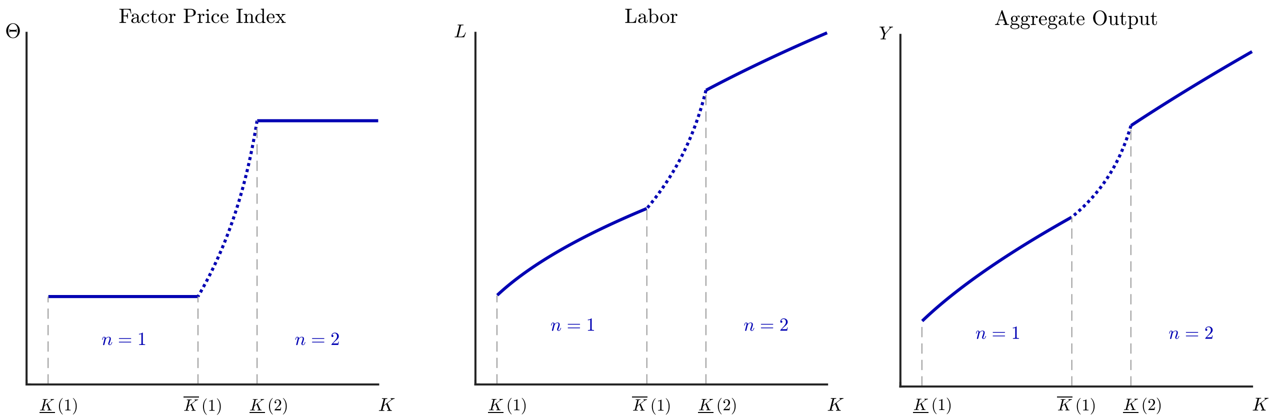


Figure 2: Static equilibrium

Note: the figure shows how the factor price index, aggregate labor and output move with capital. Solid segments represent economies with symmetric market structure (same number of firms), while dotted segments represent non-symmetric equilibria. We use $\alpha = 1/3$, $\rho = 3/4$, $\eta = 1$, $\nu = 2/5$, $\gamma_{ij} = 1$ and $c_i = 0.015$.

First, note that K_{t+1} is not globally concave in K_t. Second, the economy features multiple steady-states: in particular, there are two stable steady-states, K_1^* and K_2^*, and an unstable one, K_U. The shape of the law of motion (and the existence of multiple steady-states) can be explained by the interaction between competition, factor prices and factor supply. Within the colored regions, the economy is characterized by a fixed market structure (constant number of firms n_{t+1}). This ensures that the law of motion is concave within these regions, as in a standard neoclassical growth model. However, in the regions coinciding with changes in the market structure, the law of motion is convex. To understand this, consider again equation (24). The law of motion can be convex in capital if at least one of two conditions holds: i) Y_t is convex in K_t or ii) s_t increases sufficiently fast in K_t. As already highlighted in Figure 2, Y_t can be convex in K_t because of the positive impact of competition on labor supply. On the other hand, more intense competition can also result in a larger savings rate s_t (and hence a larger supply of capital). Even though we cannot provide an analytical characterization of s_t, we can nevertheless establish that, in a steady-state, it will be proportional to the aggregate factor share $\Omega\left(\mathbf{\Gamma}, \mathbf{N}\right)$.

Lemma 1 (Steady-State Savings Rate)**.** *In a steady-state, the following holds:*

$$s^* = \frac{\beta\,\delta}{1-(1-\delta)\,\beta}\;\underbrace{\alpha\,\Omega\left(\mathbf{\Gamma}, \mathbf{N}\right)}_{\textit{capital share}}. \tag{25}$$

Lemma 1 establishes that a more competitive market structure, by resulting in a greater factor

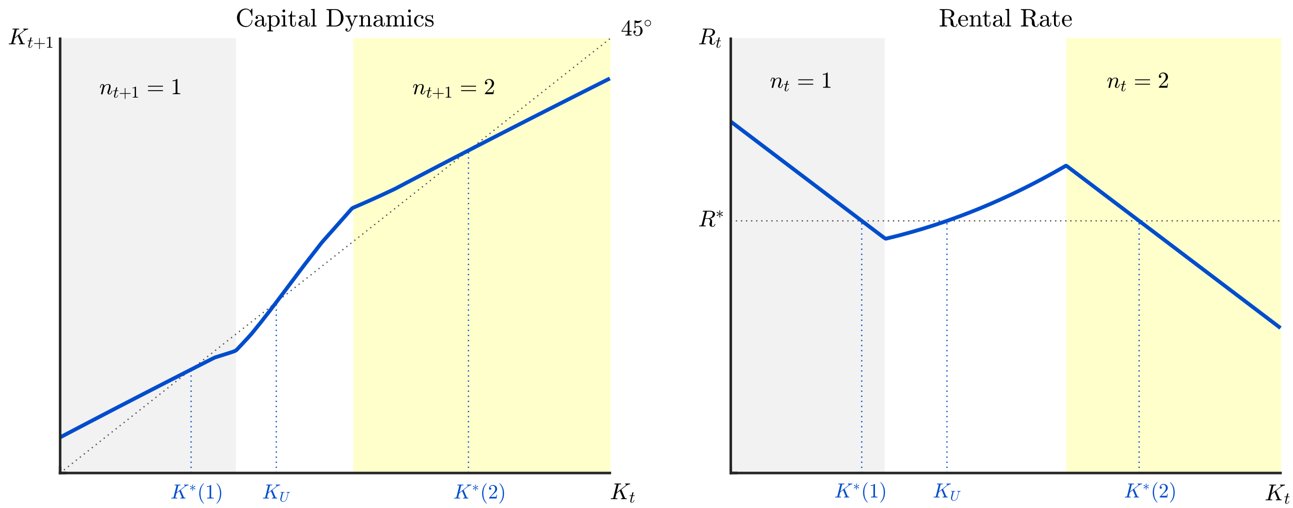


Figure 3: Law of Motion and Rental Rate Map

Note: This example features two stable steady states and an unstable one. We use $\psi = 1$, $\rho = 3/4$, $\eta = 1$, $\alpha = 1/3$, $\delta = 1$, $\nu = 2/5$ and $c_i = 0.015$.

share $\Omega(\mathbf{\Gamma}, \mathbf{N})$, will lead to a higher steady-state savings rate (and hence capital supply). To sum up, relative to K_1^*, steady-state K_2^* is characterized by a more competitive market structure, and hence a larger supply of labor and capital by the representative household.

The right panel of Figure 3 represents the rental rate map of this economy. As discussed in Section 2, multiple steady-states occur whenever this map crosses the steady-state rental rate multiple times. Recall that a steady-state is characterized by a constant rental rate equal to $R^* = \beta^{-1} - (1-\delta)$. Applying the definition of *fragility* from Section 2 to this particular economy, we have $\chi_2 = K_U / K^*(2)$. We next ask how changes in the technology set of the economy can affect *fragility*. We consider two comparative statics exercises: a mean-preserving spread (MPS) to the productivity distribution of active firms and an increase in fixed costs. These technological shifts can explain some of the micro/industry trends that have taken place since the 1980s (e.g. rising concentration or markups). We explore their aggregate consequences, in particular on the likelihood of persistent transitions.

3.5 Comparative Statics

Firm Heterogeneity We start by studying the impact of a MPS to the distribution of productivities on *fragility*. The next results will be obtained under two assumptions, which are necessary to obtain an analytical characterization of the model's steady-state(s) i) all industries are identical and ii) $\eta = 1$ (i.e. no within-industry differentiation). We start by characterizing the response of the factor price index $\Theta(\cdot)$ to a productivity spread. Note that, as equation (18) highlights, $\Theta(\cdot)$ will be useful to characterize R_t and then the distance between steady-states.

Lemma 2 (Mean Preserving Spread and Factor Prices)**.** *Let $\eta = 1$ and suppose that all industries are identical and have n firms. Let $\Gamma_n = \{\gamma_1, \ldots, \gamma_n\}$ be the productivity vector of the n active firms and $\tilde{\Gamma}_n$ be a mean-preserving spread on Γ_n. Then*

$$\Theta(\tilde{\Gamma}_n, n) < \Theta(\Gamma_n, n) \tag{26}$$

Proof. See Appendix A.2. ■

This result states that, in a symmetric steady state with n firms, a spread of the productivity distribution decreases the aggregate factor price index. As formalized by the first part of Proposition 3, this is a sufficient condition for the stable steady-state level of capital to decrease.

Proposition 3 (Firm Heterogeneity and Fragility)**.** *Let $\eta = 1$ and suppose that all industries are identical. Let $K^*(n)$ be a stable steady-state with n firms in every industry. Let $\Gamma_n = \{\gamma_1, \ldots, \gamma_n\}$ be the productivity vector of the n active firms and $\tilde{\Gamma}_n$ be a mean-preserving spread on Γ_n. The following holds*

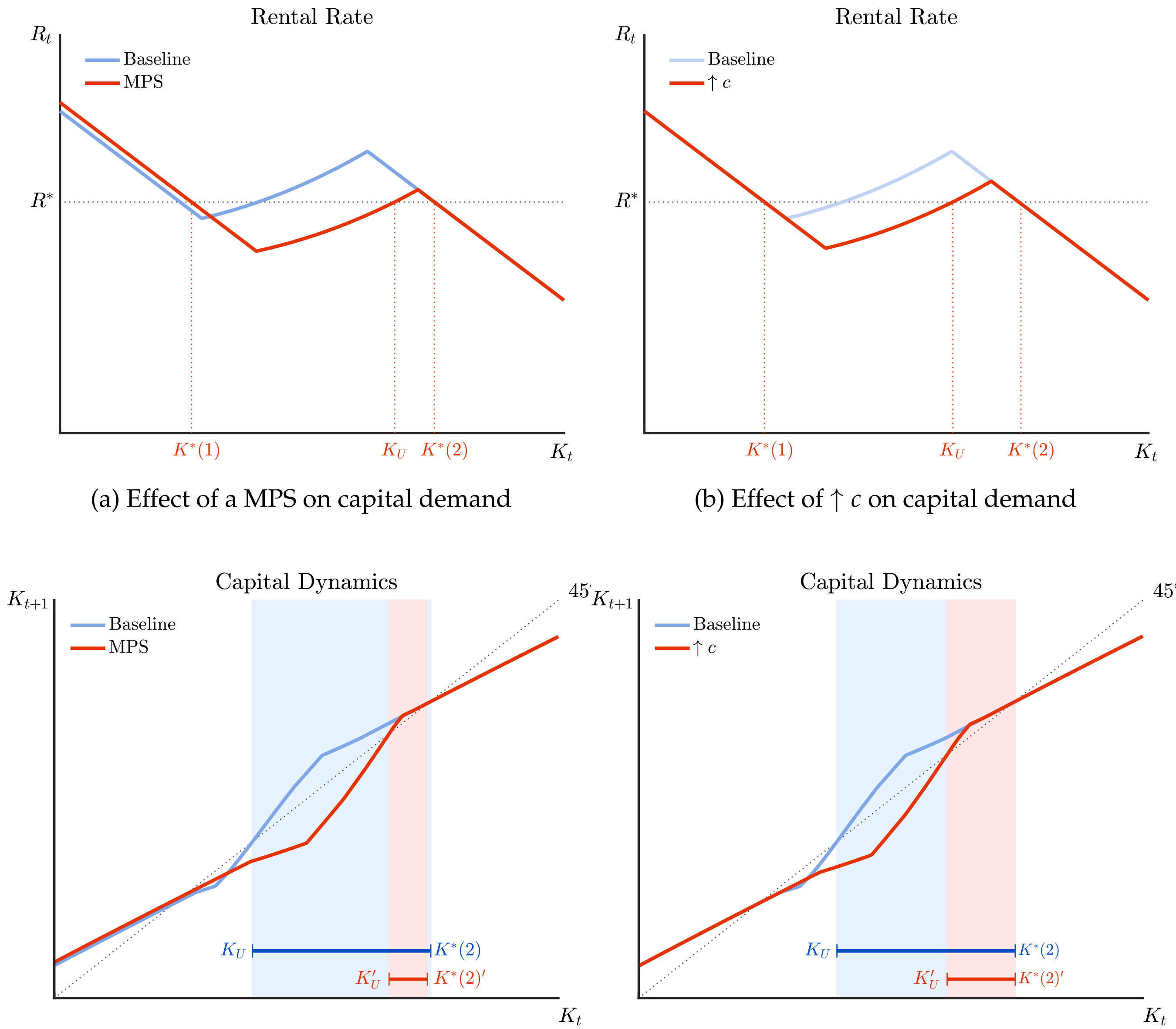
a) $K^*(\tilde{\Gamma}_n, n) < K^*(\Gamma_n, n)$.

b) If $n = 2$ *and* $\rho > 1 - \dfrac{\nu(1-\alpha)}{1+\nu\alpha}$*, then* $K_U(\tilde{\Gamma}_2) > K_U(\Gamma_2)$.

Proof. See Appendix A.2. ■

Corollary 1. *If the condition in Proposition 3b) is satisfied, the highest steady state $K^*(\tilde{\Gamma}_2, 2)$ becomes more fragile after a mean-preserving spread. Formally, $\chi_2(\tilde{\Gamma}_2) > \chi_2(\Gamma_2)$.*

Proposition 3 is the analog of Proposition 1 in Section 2. Before explaining its results, let us start by discussing the distributional consequences of a MPS. First note that, fixing the number of firms, a MPS to idiosyncratic productivities, by increasing industry concentration, will result in a lower factor share (equation (16)). This will push the factor price index down, for a given level of aggregate TFP, as equation (19) highlights. This is a *market power effect* associated with higher firm heterogeneity, which results in lower factor prices. Note, however, that aggregate TFP is likely to increase after a MPS: large, high productivity firms become even more productive and increase their market shares. This will push the factor price index up (equation 19). This is an *allocative efficiency effect*, which results in higher factor prices. The impact of a MPS on factor prices will depend on the relative strength of these two forces. These effects are the more precise



(a) Effect of a MPS on capital demand

(b) Effect of $\uparrow c$ on capital demand

(c) Effect of a MPS on the law of motion

(d) Effect of $\uparrow c$ on the law of motion

Figure 4: Comparative Statics

Panels (a) and (c) represent the effect of a MPS on capital demand and the law of motion of capital (in an economy with two steady-states). Panels (b) and (d) represent the effect of an increase in fixed costs. We use $\rho = 3/4, \eta = 1, \alpha = 1/3$ and $\nu = 2/5$.

characterizations of the elements highlighted in Section 2. Proposition 3 specifies the result from Proposition 1 when the comparative statics is a MPS on productivity.

Let us now analyze Proposition 3. It states two results. The first is that, if there exists a steady-state where all industries are identical and have n firms, the steady-state level of capital necessarily shrinks after a MPS to the productivities of these n firms. This happens because the *market power effect* always dominates, so that $\Theta(\cdot)$ declines. As a result, the steady-state level must shrink, so that equation (18) holds at the same R^*. The second part of Proposition 3 characterizes the behavior of the unstable steady-state K_U. It provides a sufficient condition under which the unstable steady-state rises after a MPS. Note that, given that the rental rate map is upward slopping at K_U, this steady-state will increase whenever the rental rate falls. This is illustrated in the left panels of Figure 4. Panel (a) shows that capital demand decreases for sufficiently high values of K_t. This translates into a higher proximity between the unstable steady-state K_U and the highest stable steady-state $K^*(2)$, as illustrated in panel (c). What Proposition 3b) does is to provide a sufficient condition under which the rental rate falls in the region in which it is increasing. The difficulty in establishing a result for this region has to do with the fact, even when all industries are ex-ante symmetry, there will be ex-post asymmetry: some will have n firms, while others will have $n-1$ firms. For tractability reasons, Proposition 3b) focuses on the case of $n = 2$. Proposition A.1 in Appendix A.2 provides a result for general n. Proposition 3b) states that K_U increases when ρ is sufficiently high. A high value of ρ means that the degree of cross-industry differentiation and of average markups are relatively low. As a result, small productivity differences get magnified and entry becomes more difficult for a second player (which brings down aggregate factor shares and factor prices).

Taken together, as highlighted in Corollary 1, the two parts of Proposition 3 provide a sufficient condition under which fragility increases after a MPS. Note that χ captures the proximity between steady-states. Provided that large shocks are less likely than small ones, then χ is a sufficient statistic for the probability of downward transitions.[20] In summary, if the condition in Proposition 3b) is verified, higher firm heterogeneity makes persistent recessions more likely.

We conclude by making one remark. Proposition 3 focuses on the consequences of a MPS. By construction of a MPS for $n = 2$, we increase the productivity of the top firm and reduce the productivity of the second one, keeping their average unchanged. However, fragility can also increase under productivity spreads that are not mean-preserving.[21]

[20]More formally, this follows from the optimal capital adjustment being a strictly increasing function of the absolute value of the shock. Hence if large shocks are less likely so are large changes in capital and therefore transitions between steady-states.

[21]Figure B.1 in Section B.4 of the Supplementary Material shows one such example. In particular, we increase the productivity of the top firm and keep the productivity of the second one unchanged.

Rising Fixed Costs The decline in product market competition since the 1980s may be also explained, through the lens of our model, by rising fixed costs.[22] The mechanism is simple: higher fixed costs drive out of the market firms that exactly break even. This results in higher average markups and concentration. How can larger fixed costs affect the law of motion of the economy? First, we can show that if a steady-state with n firms exists, and all firms make strictly positive profits, the existence and level of such a steady-state $\left(K_n^{SS}\right)$ will not be affected by a marginal increase in c_f. A larger fixed cost will however result in a higher unstable steady-state. This happens because K_U belongs to a region where some firms are breaking even. These firms can break even with a higher fixed cost only if the capital stock is higher. Proposition 4 and Corollary 2 formalize these results.

Proposition 4 (Fixed Costs and Fragility)**.** *Let $K^*(\mathbf{\Gamma}, \mathbf{N})$ be a stable steady-state with a set of active firms $\mathbf{N}$. Then the following holds*

a) $\dfrac{\partial K^*(\mathbf{\Gamma}, \mathbf{N})}{\partial c} \leq 0$.

b) $\dfrac{\partial K_U(\mathbf{\Gamma}, \mathbf{N})}{\partial c} > 0$.

Proof. See Appendix A.2. ■

Corollary 2. *When fixed costs increase, any stable steady-state becomes more fragile:* $\dfrac{\partial \chi(\mathbf{\Gamma}, \mathbf{N})}{\partial c} > 0$.

Panels (b) and (d) of Figure 4 illustrate the effect of an increase in fixed costs. For simplicity, we represent an economy where all industries are ex-ante identical (same distribution of productivities and identical fixed costs). The two declining segments of the rental rate map represent a situation of full monopoly and full duopoly; since all firms make strictly positive profits in these regions, the equilibrium is unchanged as c increases marginally. In the increasing segment, where some firms are exactly breaking even, some of them will end up being driven out of the market. The rental rate decreases and the unstable steady-state increases. As a result, fragility increases (Corollary 2).

Discussion We conclude by summarizing two keys insights of our theory, which we think are relevant to understand the US growth experience after 2008. The first is that a complementarity between competition and factor supply can generate multiple competition regimes or steady-states. A transition from a high competition to a low competition regime can in many aspects

[22] This channel is also studied in De Loecker et al. (2021).

describe the 2008 recession and the subsequent great deviation. Second, changes in technology that result in larger market power (e.g. larger productivity differences across firms or larger fixed costs) make high competition regimes more difficult to sustain, and transitions to low competition traps more likely. Our model therefore suggests that the US economy, experiencing a long-run increase in markups and concentration since the 1980s, became increasingly vulnerable to transitions like the one observed after 2008.

We next use a calibrated version of our model to ask whether it can quantitatively replicate the behaviour of the US economy in the aftermath of the 2008 crisis. We also perform some counterfactual exercises to quantify the impact of rising market power on the likelihood of persistent recessions.

4 Quantitative Results

The goal of this section is to develop a quantitative version of the model built in Section 3. We will use it to provide a quantification of some of the forces described earlier, and evaluate policy counterfactuals. There are two objects that we need to parametrize: the distributions of idiosyncratic productivities γ_{ij} and of fixed costs c_i. We assume that firms draw their idiosyncratic productivities from a log normal distribution with standard deviation λ, i.e. $\log \gamma_{ij} \sim \mathrm{N}(0, \lambda)$. Recall that each industry i will be characterized by N such draws. Since N is a finite number, industries have different ex-post distributions of idiosyncratic productivities $\{\gamma_{ij}\}_{j=1}^{N}$.

Furthermore, we assume that there are two types of industries: a fraction $f \in (0,1)$ of all industries face a fixed cost $c_i = c > 0$, whereas the remaining fraction $1 - f$ faces a zero fixed production cost $c_i = 0$. We will refer to the first as *concentrated* and the second as *unconcentrated* industries. Note that in industries with a zero fixed cost, the extensive margin will be shut down as all potential N entrants will always be active. However, these industries will not necessarily operate close to perfect competition, as there can be large productivity differences across firms, resulting in high concentration and large markups for productive firms.[23]

4.1 Calibration

The model is calibrated at a quarterly frequency. Our calibration strategy relies on the interpretation that the economy starts in the highest steady-state (as shown below, the economy will

[23] We assume that there is a common fixed cost c among all *concentrated* industries. These industries will however differ in their distributions of idiosyncratic TFP draws $\{\gamma_{ij}\}_{j=1}^{N}$ and may display a different number of active firms.

feature two stochastic steady-states). Some parameters are standard and taken from the literature. For the preference parameters, we set $\beta = 0.99$ and $\psi = 1$ to have log utility. We set $\nu = 0.352$, which implies a Frisch elasticity of 2.84; this corresponds to the average macro elasticity of hours reported by Chetty et al. (2011).

We set the capital elasticity to $\alpha = 0.3$ and the depreciation rate to $\delta = 0.025$. For the two elasticities of substitution, we use $\sigma_I = 1.35$ and $\sigma_G = 13$. These are the values used by Edmond et al. (2021) in an oligopolistic competition setting similar to ours.[24] In general, increasing the elasticity of substitution across industries σ_I depresses markups and weakens the complementarity between capital accumulation and competition, making steady-state multiplicity less likely to arise.[25] We set the number of industries to $I = 10,000$ and the maximum number of firms per industry to $N = 20$ for computational reasons.

There are three parameters that we need to internally calibrate: the fraction of *concentrated* industries (f), the fixed cost for the *concentrated* sector (c) and the standard deviation of the log-normal productivity distribution of the pool of potential entrants (λ). These three parameters are jointly calibrated to target three data moments observed in 2007 (i.e. before the 2008 crisis). To calibrate f, we target the fraction of aggregate employment that is allocated to highly concentrated industries. In the calibrated model, *concentrated* industries have typically less than 4 firms. We define an industry as concentrated if the 4 largest firms represent at least 90% of the output of the 8 largest firms.[26] Using data from the US Census, we find that 6.33% of aggregate employment is allocated to such 6-digit industries in 2007. We calibrate the fixed cost parameter to match the average ratio of fixed to total cost ratio in COMPUSTAT. Following Gorodnichenko and Weber (2016), we define fixed costs as sum of 'Selling, General and Administrative Expenses', 'R&D Expenditures' and 'Advertisement Expenses'. We obtain a target level of 41.4% for 2007 (see Section B.8 of the Supplementary Material). We calibrate the the dispersion parameter by targeting the average sales-weighted markup of public firms in 2007 (as reported by De Loecker

[24] These values are within the bounds typically used in the literature. In identical models with Cournot competition, Atkeson and Burstein (2008) set $\sigma_I \approx 1$ and $\sigma_G = 10$, while De Loecker et al. (2021) estimate $\sigma_I = 1.2$ and $\sigma_G = 5.75$. Using data on French firms, Burstein et al. (2020) estimate $\sigma_I = 1.7$ and $\sigma_G = 7$. We provide an alternative calibration using the parameter values from De Loecker et al. (2021) in Section B.9 of the Supplementary Material.

[25] See Jaimovich (2007) for a similar result.

[26] We would like to think of an industry at the highest possible level of disaggregation (e.g. 10-digit NAICS). However, the US census provides concentration metrics only at the 6-digit NAICS level. This is why we do not look directly at the share of the top 4 firms, but instead scale it by the share of the top 8. We have considered alternative thresholds for the ratio of the top 4 to the top 8 (80% and 95%) and the results were identical.

Description	Parameter	Value	Source/Target
Between-industry ES	σ_I	1.35	Edmond et al. (2021)
Within-industry ES	σ_G	13	Edmond et al. (2021)
Inverse of Frisch elasticity	ν	0.352	Chetty et al. (2011)
Capital elasticity	α	0.3	Standard value
Depreciation rate	δ	0.025	Standard value
Discount factor	β	0.99	Standard value
Coefficient of risk aversion	ψ	1	log utility
Max number of firms per industry	N	20	
Number of industries	I	10,000	
Calibrated Parameters: 2007			
Standard deviation γ	λ	0.305	Sales-weighted average markup
Fixed cost	c	0.0010	Average ratio fixed/total costs
Fraction of industries with $c_i > 0$	f	0.125	Emp share concentrated industries
Persistence of z_t	ρ_z	0.920	Autocorrelation of log Y
Standard deviation of ε_t	σ_ε	0.0035	Standard deviation of log Y
Calibrated Parameters: 1985			
Standard deviation γ	λ	0.170	Sales-weighted average markup
Fixed cost	c	0.0005	Average ratio fixed/total costs
Fraction of industries with $c_i > 0$	f	0.110	Emp share concentrated industries

Table 1: Parameter Values

et al. (2020)). Note that, in our model, the aggregate markup is informative about the degree of TFP dispersion: holding the number of active firms fixed, a higher dispersion parameter λ is associated with a reallocation towards big firms and hence a larger sales-weighted average markup. Finally, we calibrate the two parameters governing the dynamics of aggregate productivity: the autocorrelation parameter ϕ_z and the standard deviation of the innovations σ_ε. We do so by targeting the first order autocorrelation and the standard deviation of output.[27]

To assess the business cycle implications of higher firm level heterogeneity, we also provide an alternative calibration of the model. In particular, we calibrate λ and c to target the same data moments in 1985. Note that both the observed sales-weighted average markup and the fixed to

[27] We compute these moments for the entire postwar period 1947-2019. Consistent with our interpretation that the US economy moved to a different regime after 2008, we remove a linear trend computed for the period 1947-2007.

total cost ratio are lower in 1985 than in 2007 (Table 2). We also recalibrate the fraction of concentrated industries f to target the same employment share in highly concentrated industries.[28] The remaining parameters, including the ones governing aggregate TFP shocks, are kept unchanged.

	1985		2007	
	Data	Model	Data	Model
Sales-weighted average markup	1.27	1.27	1.46	1.43
Average fixed to total cost ratio	0.343	0.371	0.414	0.413
Employment share in *concentrated* industries	-	0.059	0.063	0.055
Autocorrelation log GDP	0.978*	0.976	0.978*	0.969
Standard deviation log GDP	0.061*	0.067	0.061*	0.064

*computed over 1947:Q1-2019:Q4

Table 2: Targeted moments and model counterparts

Table 2 reports our targeted moments, with their model counterparts. The model successfully matches the sales-weighted average markup in both the 1985 and the 2007 economies. We slightly underestimate the employment share of highly concentrated industries in 2007.

Part [A] of Table 3 reports the evolution of some non-targeted moments. When computing aggregate markups with cost-weights, we find an average of 1.22 in the 1985 economy, and an average of 1.32 in 2007. While these values are slightly above the recent estimates by Edmond et al. (2021), the time variation is similar in the model and in the data (10pp increase in the model, against 8pp increase in the data). To characterize some of the aggregate consequences of the change in the three structural parameters (λ, c, f), we also report the evolution of some macroeconomic aggregates. Our model predicts a 18% increase in aggregate TFP (against 23% in the data), and a 23% increase in output per hour (against 21% in the data).

Industries facing positive fixed costs will play an important role in our mechanism. Part [B] of Table 3 provides a characterization of these industries in the two calibrated economies. Industries with positive fixed costs will consist mostly of monopolies and duopolies – the average number of firms is 1.72 in the 1985 economy, and 1.40 in 2007. This implies an average markup of 1.96

[28]The SUBS does not provide employment data by 6 digit industries prior to 1997. For this reason, we decide to keep the same target for the employment share of highly concentrated industries.

	1985		2007	
	Data	Model	Data	Model
[A] Total Economy				
Cost-weighted average markup	1.17	1.22	1.25	1.32
Aggregate TFP (log)	0.00	0.00	0.23	0.18
Output per hour (log)	0.00	0.00	0.21	0.23
[B] Concentrated Industries				
Number of firms per industry		1.72		1.40
Sales-weighted average markup		1.96		2.58

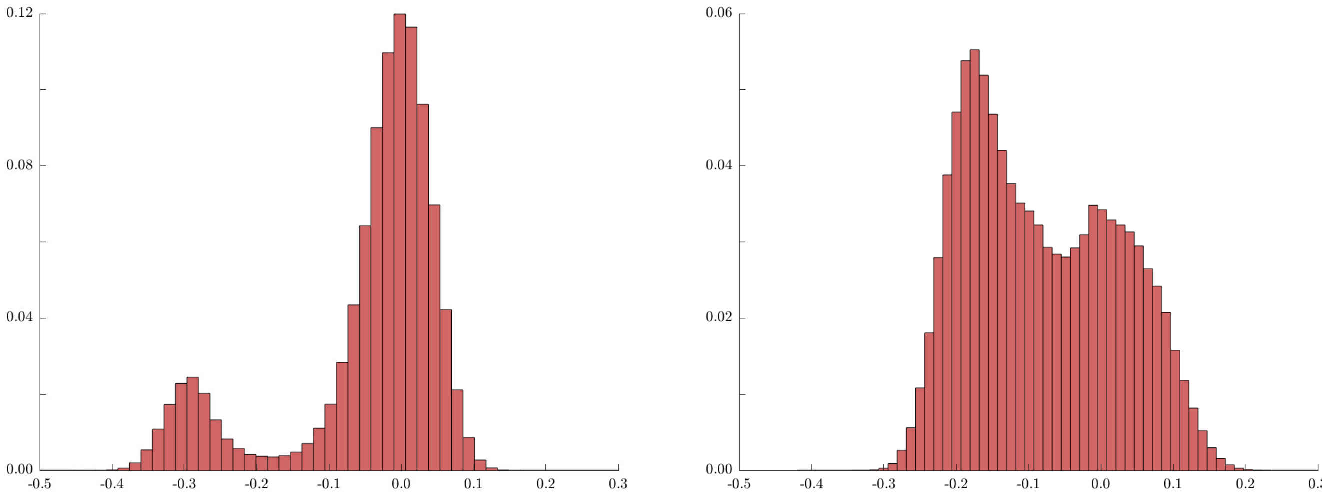
Table 3: Non-targeted moments

Note: The cost-weighted average markup is from Edmond et al. (2021). Aggregate TFP is from Fernald (2012), while output per hour refers to 'Business Sector: Real Output Per Hour of All Persons' from the BLS. The 1985 of the last two series are normalized to zero.

in 1985 and of 2.58 in 2007 in these industries. Note that these values are within the bounds of estimates for the markup distribution of US firms. For example, De Loecker et al. (2020) report that the 90th percentile of the (sales-weighted) markup distribution increased from 1.66 in 1985 to 2.25 in 2007. Note also that concentrated industries represent 5.5.% of aggregate employment in the 2007 model. This means that, in our model, approximately 1/20 of aggregate employment is concentrated in monopolist/duopolist firms.

4.2 Quantitative Results

We start by comparing the dynamic properties of the 1985 and the 2007 economies. Figure 5 shows the ergodic distribution of log output; the distributions are centered around the second mode, so that the horizontal axis represents output in percentage deviation from the high steady-state. Two facts are worth mentioning. First, the two distributions are bimodal, implying that both economies feature (stochastic) steady-state multiplicity. Note that, relative to the 1985 economy, the 2007 model has more mass on the left, meaning that the economy spends on average more time on the low competition regime. Second, in the 2007 distribution, the two steady-states are also closer to each other – a transition from the high to the low regime implies a 17% reduction in output in 2007, as opposed to approximately 30% in 1985. While this fact means that transitions are less pronounced in 2007, it also implies that they will be substantially more likely in 2007 than in 1985 (as discussed below). Table A.1 in Section B.5 of the Supplementary Material provides business cycle moments for the two economies.



(a) Output Distribution: 1985

(b) Output Distribution: 2007

Figure 5: Ergodic distribution of output

Note: This figure shows the distribution of log output for the 1985 and the 2007 economies. We simulate each economy for 10,000,000 periods and plot output in deviation from the high steady state.

We next study the probabilities of deep recessions in the two economies. To this end, we simulate each economy 100,000 times for 40 and 100 quarters; we then compute the fraction of simulations in which we observe a 10, 15 or 20% deviation from the high steady state level (for at least 4 consecutive quarters). The results are shown in Table 4. When running the 2007 economy for 40 quarters, output drops by 10% in 15.2% of the simulations, whereas the same number for the 1985 economy is 1.6%. This suggests that the former is approximately 9 times more likely to experience a 10% fall in output over a 10-year periods. Over 100 quarters, the 2007 economy appears about 5 times more likely to experience a 10% recession (33.2% probability in 2007, against 6.6% probability in 1985). This implies that, in expectation, the 2007 economy experiences a recession larger than 10% every 75 years, while the 1985 economy does so every 380 years. To illustrate the dynamic properties of our two calibrated economies we next describe the response of several variables to exogenous TFP shocks.

	1985 Model		2007 Model	
	T = 40	T = 100	T = 40	T = 100
Pr[10% recession]	0.016	0.066	0.152	0.332
Pr[15% recession]	0.001	0.012	0.035	0.159
Pr[20% recession]	0.000	0.004	0.001	0.042

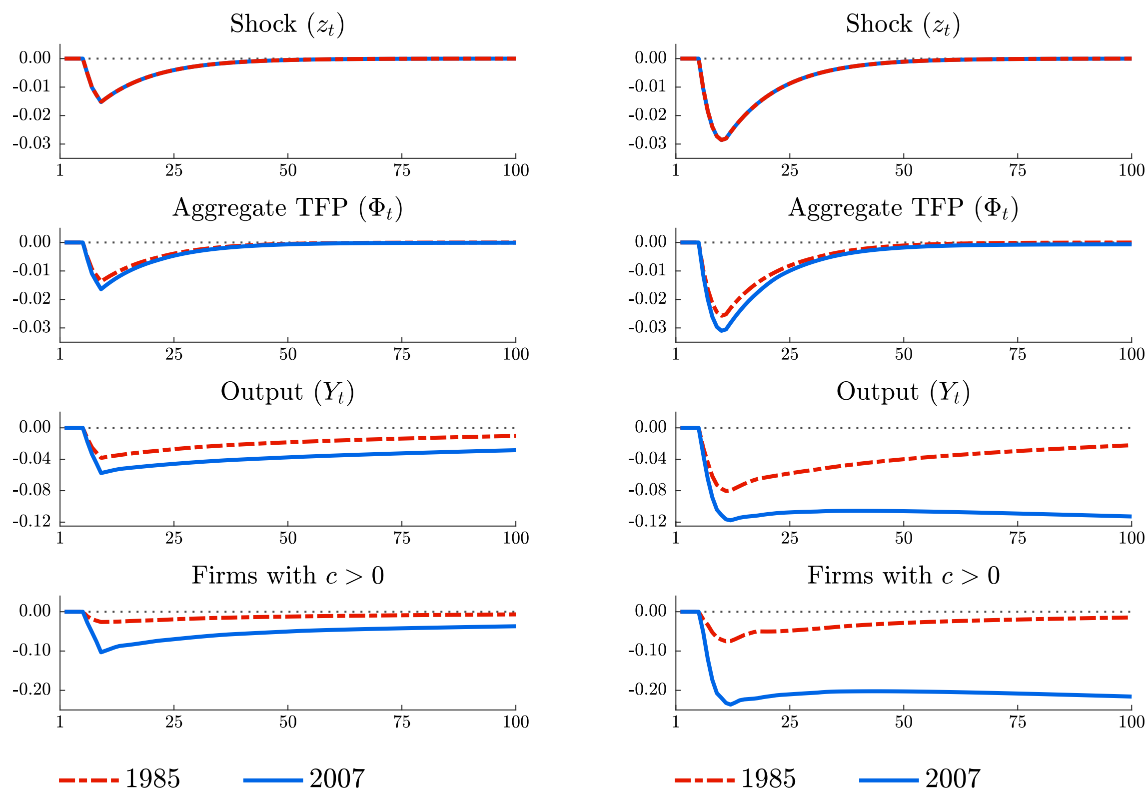
Table 4: Probability of deep recessions

This table shows the probabilities of deep recessions in the 1985 and 2007 economies. Each economy starts in the high steady-state and is simulated for $T = 40$ and $T = 100$ quarters. Each simulation is repeated 100,000 times. The table reports the fraction of simulations in which output is κ% the initial value for at least 4 consecutive quarters.

Impulse Response Functions: Small Negative Shock Next we characterize the reaction of the economy to a small negative shock. We consider a shock to the innovation of the exogenous TFP process that is equal to $\varepsilon_t = -\sigma_\varepsilon$ and lasts for four quarters. Figure 6(a) shows the impulse responses for both the 1985 and the 2007 economy. The simulation of the transition dynamics covers 100 quarters. This shock generates different responses for the two economies, as evidenced by Figure 6(a). The 2007 economy exhibits both greater amplification and greater persistence. First, the 1985 economy experiences a 4% reduction in aggregate output after 5 quarters, against a 6% reduction in the 2007 economy. Second, after 100 quarters, the 1985 economy is 1% below the steady-state, while the 2007 economy has a much more prolonged downturn, being still 3% below pre-crisis output.

The mechanism underlying such increased amplification and persistence can be better understood by looking at the bottom panel, which plots the transition dynamics of the number of firms in *concentrated* industries. In 2007, there is a much more significant reduction in the number of firms, due to the mechanisms outlined above: increased productivity dispersion and larger fixed costs make small, unproductive firms more sensitive to aggregate shocks. This additional action in the extensive margins generates both additional amplification and persistence.

Impulse Response Functions: Large Negative Shock The shock introduced above was small enough to make both economies transition to their initial steady-states. We now study the effect of a larger shock. To this end, we repeat the same exercise for the two economies, but now introduce a negative shock equal to $\varepsilon_t = -2\sigma_\varepsilon$, which lasts for six quarters. The dynamics are shown in Figure 6(b). As before, there is greater amplification and persistence in the 2007 economy. However, this economy now experiences a permanent drop in aggregate output, i.e. it transitions to a lower steady-state. In the example we consider, after 100 quarters, output is 12% below its initial value. Note, however, that the gap is still widening at the end of the sample. This is due to a permanent loss of about 20% of active firms in concentrated industries.



(a) Small Negative Shock

(b) Large Negative Shock

Figure 6: Impulse Response Functions

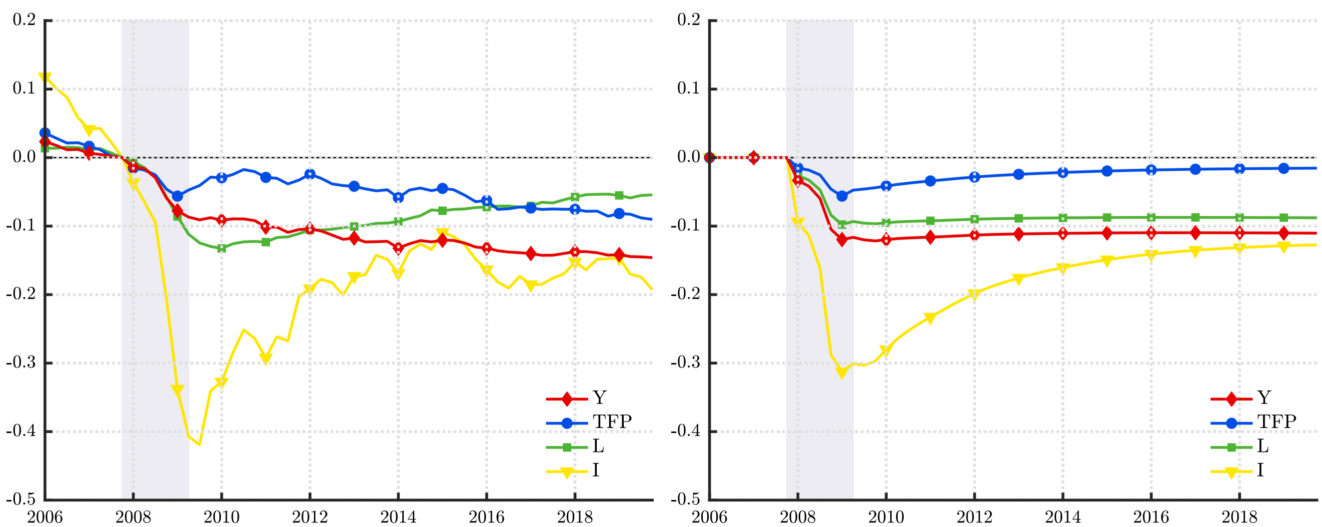
Note: The graphs show the IRFs to an exogenous TFP shock. In panel (a), we feed a shock $\varepsilon_t = -\sigma_\varepsilon$ that lasts for four quarters. In panel (b), we feed $\varepsilon_t = -2\,\sigma_\varepsilon$ during six quarters.

These results suggest rising firm differences and fixed costs are a source non-linearity in the economy's response to aggregate shocks. This may appear to be inconsistent with the idea of a *great moderation* – namely, the fact the the volatility of aggregate output declined between 1985 and 2007. Note, however, that aggregate volatility in our economy is the product of two forces – exogenous volatility (TFP shocks) and endogenous amplification and persistence. If exogenous volatility declined over time, it is possible that aggregate volatility also declined in spite of larger amplification. There are reasons to think that exogenous aggregate volatility may have decreased over time – for example, because of demographic shifts (Jaimovich and Siu, 2009) or a rising share of low-volatility industries (Carvalho and Gabaix, 2013).

We next use our model as a laboratory to study the 2008 recession.

5 The 2008 Recession and Its Aftermath

In this section, we take a deeper look at the 2008 recession and its aftermath. The left panel of Figure 7 shows the behavior of some aggregate variables from 2006 to 2019 – real GDP, real gross private investment and total hours (all in per capita terms), as well as aggregate TFP. All variables are in logs, detrended (with a linear trend computed over 1985-2007) and centered around 2007Q4. The four variables decline on impact and do not seem to rebound to their pre-recession trends. For example, in the first quarter of 2018, real GDP per capita is 13.3% below trend (Table 5). Aggregate TFP experienced a 6.8% negative deviation from trend. Investment declines by more than 40% on impact, and then seems to stabilize at approximately 20% below the pre-crisis trend.



(a) Data (b) 2007 Model

Figure 7: The great recession and its aftermath

Note: The figure shows the evolution of key macro aggregates in the aftermath of the 2008 recession in the data (Panel a) and the model (Panel b). The model is subjected to a sequence of six quarter shocks $\{\varepsilon_t\}$ to match the dynamics of aggregate TFP in the data between 2008Q1:2009Q2. See Appendix A.1 for data definitions.

We then ask whether our model can replicate the behavior of these four variables. To this

	Data			Model		
	2009Q4	2015Q1	2019Q1	2009Q4	2015Q1	2019Q1
Output	-0.088	-0.126	-0.142	-0.122	-0.110	-0.110
TFP	-0.045	-0.026	-0.082	-0.044	-0.020	-0.016
Hours	-0.130	-0.078	-0.055	-0.097	-0.087	-0.088
Investment	-0.340	-0.109	-0.147	-0.297	-0.149	-0.129

Table 5: The great recession and its aftermath

end, we feed the model with a sequence of shocks ε_t (i.e. the innovation of A_t) that lasts for six quarters (2008Q1:2009Q2); these shocks are calibrated so that aggregate TFP in our model (i.e. $A_t\,\Phi_t$) matches the dynamics of its data counterpart over the same period. The economy starts at the high steady-state (with $z_t = 0$). We set the innovations to productivity to zero after 2009Q2 and let the economy recover afterwards. The right panel of Figure 7 shows the implied responses of output, aggregate TFP, employment and investment, generated by our model. As the figure shows, this series of shocks triggers a transition to the low steady-state. Our model provides a reasonable description of the evolution of the four variables. Output experiences a 11.0% decline in the long-run, whereas hours drop by 8.8% (Table 5). Both reactions are of the same order of magnitude as observed in the data. When looking at investment, our model predicts a 29.7% decline on impact, and 12.9% drop by 2019 (14.7% in the data). Finally, our model also predicts a long-run decline of aggregate TFP, representing approximately 1/5 of the observed drop in the data (1.6% in the model, 8.2% in the data). We discuss the mechanisms underlying this result in the next section. In summary, our model can successfully replicate the 2008 crisis and its persistent effects.

Next, as our first counterfactual, we ask whether the sequence of aggregate TFP shocks that we feed in the 2007 economy can also trigger a transition to the low steady-state in the 1985 economy. Figure 8 shows the transition dynamics. Not only does the economy exhibit substantially less amplification, but it also reverts back to the high steady-state.

These results suggest that, in an economy with the 1985 features, a downturn of the magnitude of the 2008-2009 recession would not be large enough to generate a persistent deviation from trend. The economy would have experienced a faster reversal to trend, due to a lower endogenous amplification and persistence. For example, as of 2019, output would be 5.8% below trend (against 11% of the 2007 model and 14.2% as observed in the data). We conclude that the structural differences between the 1985 and the 2007 economies (namely larger productivity

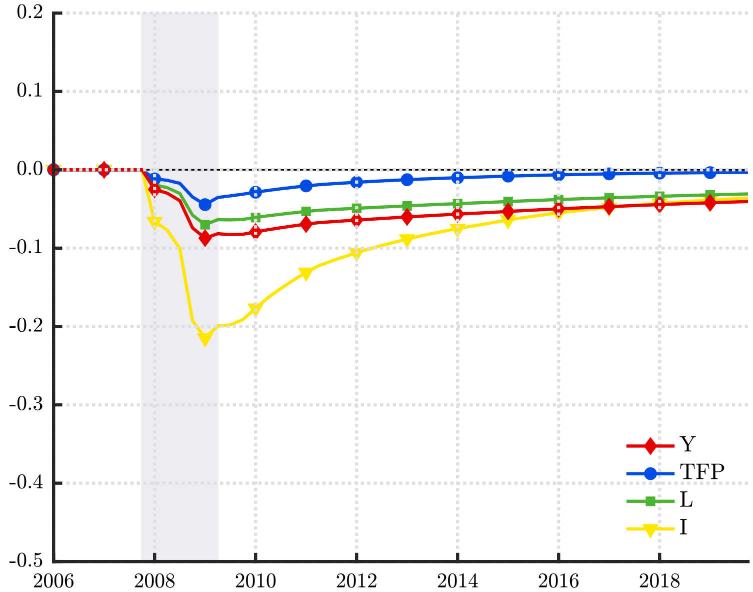


Figure 8: The *great recession* in the 1985 model

Note: This figure shows the response of the 1985 economy to the sequence of shocks used in Figure 7(b)

differences and larger fixed costs) are key to understand the 2008 crisis and the subsequent *great deviation*.

5.1 Other Variables

According to our model, a transition between steady-states is driven by a change in the competitive regime of the economy. We now provide evidence consistent with our model's mechanism. In particular, we show that after 2008 the US economy experienced (i) a persistent decline in the number of active firms, (ii) a persistent decline in the aggregate labor share and (iii) acceleration in the aggregate profit share and markup.

Number of Firms Firms in our model are single-product and operate in only one market. This contrasts with the definition of firms in the data, which are often multi-product and operate in several markets (e.g. geographic segmentation). With this caveat in mind, we ask whether the evolution of the number of firms in the data qualitatively matches the prediction of our model. Figure 9 shows the evolution of the number of active firms (with at least one employee). As the figure shows, the number of active firms experiences a persistent deviation from trend after 2007. Such a persistent decline can also be observed within most sectors of activity (Supplementary Material Section B.7). In Section 7, we show that the decline in the number of firms was more pronounced in industries featuring initially higher concentration.

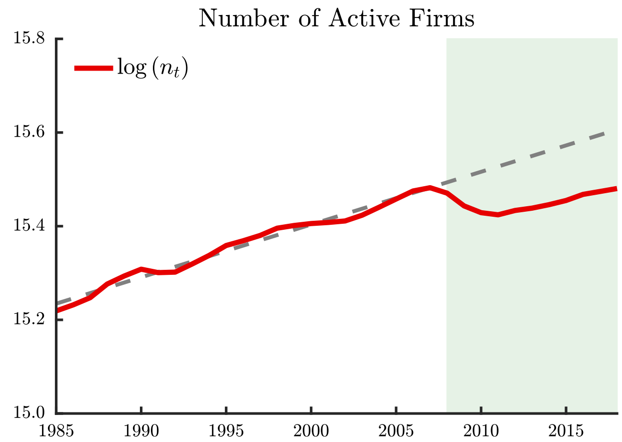


Figure 9: **Number of Firms: 1985-2018**
Note: The red line shows the number of firms with at least one employee (in logs). The dashed grey line shows a linear trend computed for the period 1985-2007. Data is from the US Business Dynamics Statistics.

Aggregate Markups, Labor and Profit Shares We now ask if our model can explain the evolution of labor and profit shares after 2008. Figure 10 shows the evolution of the labor share, the profit share (both computed for the US business sector) and the aggregate markup series for publicly listed firms from De Loecker et al. (2020). The grey dashed lines represent linear trends computed for the period 1985-2007.

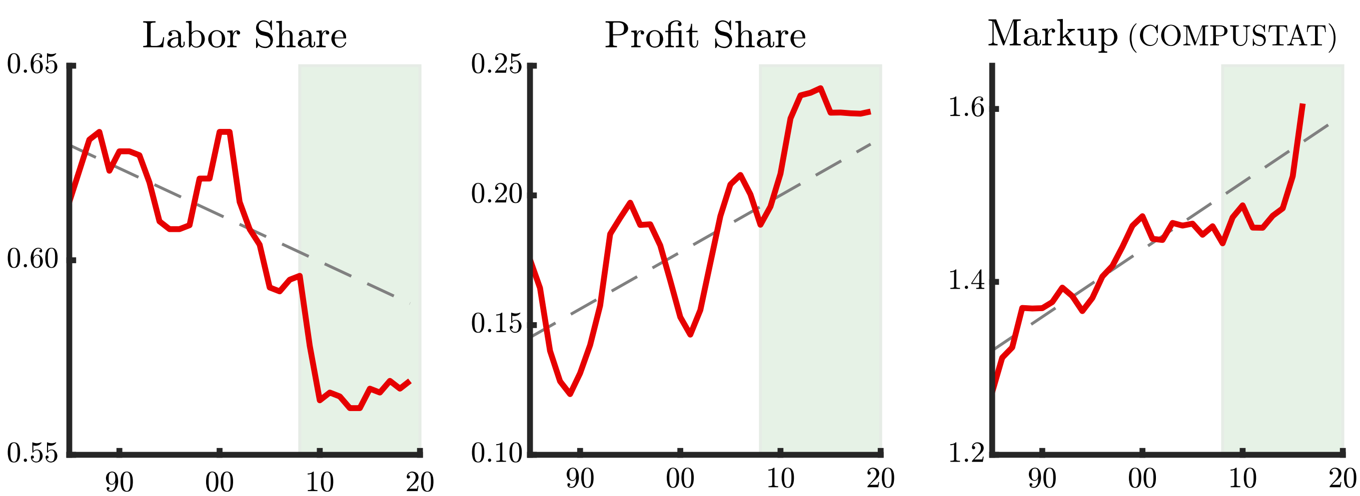


Figure 10: **Aggregate Markup and Labor and Profit Share: 1985-2019**
Note: This figure shows (i) labor share of the US business sector (from the BLS), (ii) the profit share of the US business sector (defined in Appendix A.1) and (iii) the aggregate markup series for COMPUSTAT firms from De Loecker et al. (2020). The dashed grey lines represent linear trends computed for the period 1985-2007.

Table 6 compares, for each series, its evolution between 2007 and 2016 observed in the data and obtained in our model.[29] Overall, our model predicts a 0.6 pp decline in the aggregate labor share, which is approximately 17% of the observed decline between 2007 and 2016. If we account for a pre-crisis trend, we can explain approximately 28% of the deviation in 2016. In the case of the profit share, we can explain 75% of its 0.012 increase from trend. Markups increase by 4.1 points in our model, which represents 29% of the observed increase (14.2 points) and 57% of the deviation from the pre-crisis trend (7.2 points).

	Model	Data	
	$\Delta_{2007-2016}$	$\Delta_{2007-2016}$	$\Delta_{2007-2016} - \Delta_{\text{trend}}$
Labor Share	-0.005	-0.029	-0.018
Profit Share	0.009	0.032	0.012
Aggregate Markup	4.1	14.2	7.2

Table 6: Change in the income shares: model *versus* data

Aggregate Productivity As shown in Figure 7, the transition to the low competition regime is accompanied by a persistent decline in aggregate TFP. This happens in spite of the exit of unproductive firms, which results in higher average firm-level TFP (see Figure B.1 of the Supplementary Material). There are two reasons explaining the decline in aggregate TFP. First, there is a *love-for-variety* effect, associated with the reduction in the number of firms. This can best be seen in the limit case in which all industries have n firms with identical productivity τ. In such a case, aggregate TFP is equal to $\Phi = I^{\frac{1-\rho}{\rho}} n^{\frac{1-\eta}{\eta}} \tau$. Second, in a low competition trap, there is higher cross industry misallocation. This happens because industries with a larger contraction are the ones with positive fixed costs $c > 0$, i.e. whose output is already low.[30]

The macro trends discussed above suggest that, consistent with our model, market power accelerated after 2008. In Section 7 we provide cross-industry evidence that further corroborates our mechanism. We conclude this section by performing some robustness exercises. We ask if our model is capable of replicating another recession (the 1981-82 crisis), and then assess the robustness of different parameters and assumptions.

[29]The aggregate profit share in our model is net of fixed production costs.

[30]Figure B.1 in the Supplementary Material shows that the standard deviation of industry outputs increases.

5.2 The Early 1980s Recession

Through the lens of our model, the 2008 crisis made the US economy transition to a new steady-state. This fact has not been observed after any other postwar recession. This raises a natural question: what was special about the 2008 crisis? Was the shock hitting the economy in 2008 larger than in previous recessions? Or was the economy more fragile in 2008 and therefore more prone to experience a transition even for moderate shocks? To answer these questions, we repeat the experiment of Section 5 using the 1981-1982 crisis. We feed the 1985 economy with a sequence of shocks that replicates the dynamics of aggregate TFP during the 1981-1982 recession (1981Q3:1982Q4). We then take this same sequence of exogenous shocks and apply them to the 2007 economy. The results of this experiment are shown in the Supplementary Material Section B.6. When looking at the response of the 1985 economy, we observe a temporary decline in all variables, but followed by a gradual recovery to the previous steady-state.[31] This contrasts with the response of the 2007 economy, which again experiences a transition to the lower steady-state. These results suggest again that, rather than the consequence of an usually large shock, the post-2008 deviation can be linked to an underlying market structure that made the US economy more fragile to negative shocks.

5.3 Robustness Checks

In the previous section we described how a calibrated version of the model admits multiple stochastic steady states and performs well when tasked with replicating the behaviour of the US economy in the aftermath of the 2008 recession. In this section we want to understand whether the existence of multiplicity, generated by the complementarity between capital accumulation and competition, is robust to alternative modelling assumptions and calibrations.

Alternative values for the elasticity of substitution The first robustness check we carry out is a different parametrization of the key elasticities in our model. In particular, we use the values from De Loecker et al. (2021) and set the cross-industry elasticity of $\sigma_I = 1.2$ and the within-industry elasticity of $\sigma_G = 5.75$. The model retains multiplicity of steady states. As in our benchmark model, when hit with the aggregate shocks to generate the 2007 TFP response, we find that the 1985 converges to the same steady state while the 2007 economy falls in a

[31] The recovery seems to be slower than in the actual data. Note, however, that by design we are shutting down possible positive shocks hitting the economy in 1983 and afterwards.

low competition trap. The calibration, the ergodic distribution and the crisis experiment are all reported in the Supplementary Material Section B.9.

Variable fixed costs In the model presented so far firms pay a fixed cost $c > 0$, which is paid in units of the final good. This assumption implies that the cost of entry is independent of the state of the economy and, hence, of its competition regime and of factor prices. If fixed costs were to change with factor prices, entry could be cheaper (more expensive) in a low (high) competition regime, which could in principle eliminate steady-state multiplicity. To address this concern, we let firms hire labor and capital to pay the fixed cost. In particular we assume production entails a fixed cost $c = k_c^{\alpha}\, l_c^{1-\alpha}$. Under this assumption, firms pay an effective fixed cost $\Theta_t \cdot c$. Recall that the factor cost index Θ_t is increasing in the number of active firms. This implies that entry becomes ever more expensive as firms enter. Appendix B.10 shows the results for this version of the model. The existence of two competition regimes is preserved under this alternative assumption. We also obtain responses for the crisis experiment performed above. The results are qualitatively and quantitatively similar to the benchmark model.

6 Policy Experiment

In this section we evaluate how policy can reduce the distortions introduced by market power. Note that, in this economy, on top of the standard distortionary (static) results associated with markups, market power carries additional negative consequences as it can make the economy trapped in a low competition regime.

We consider a government who grants an entry subsidy equal to a fraction $\tau_f \in [0, 1]$ of the fixed cost, while levying a tax $\tau_\pi \in [0, 1]$ on net profits $\tilde{\Pi}_{ijt} = \Pi_{ijt} - \left(1 - \tau_f\right) c_i$ (i.e. profits in excess of fixed costs) to balance the budget. First, note that by design the entry subsidy only affects *concentrated* industries. Secondly, since the government is only taxing the amount of profits in excess of fixed costs, the tax $\tau_\pi > 0$ does not disincentivize entry. Finally, a positive entry subsidy will reduce the entry productivity thresholds. The planner faces, however, one key trade-off when subsidising entry: while having more firms in the economy depresses markups and increases available varieties, it also implies that more resources are spent in fixed costs and that less productive firms enter the market.

We report welfare calculations for different levels of the entry subsidy in Figure A.2 (Appendix A.3). The analysis suggests that the government would find a subsidy of around 70% of the fixed cost optimal, implying a welfare gain of approximately 21% (in consumption equivalent terms).

This value is within the range of estimates reported by the literature for the cost of markups. For example, Bilbiie et al. (2019) find a cost as high as 25%, while Edmond et al. (2021) find a welfare loss of 23.6% for an average cost-weighted markup of 1.25.

It should be noted that the policy experiment we consider does not implement the first best allocation. We evaluate a simple firm subsidy and do not consider size-dependent taxes/subsidies that might be necessary to eliminate markup distortions. Furthermore, as mentioned earlier, we highlight another cost of market power – the fact that it can generate quasi-permanent recessions. Interestingly, Figure A.2 suggests that welfare is very steep around $\tau_f = 0$, implying that relatively small entry subsidies can have a sizeable impact on welfare. For example, a 10% entry subsidy would be enough to generate a 8% gain in consumption equivalent terms. The reason is that even a relatively small subsidy can be enough to significantly shift the probability mass from the low to the high steady state.

Lastly, one can think about the optimal state-dependent subsidy. As the economy is hit by a large negative shock that might trigger a steady state transition, the welfare benefit of such a subsidy can be very large. On the other hand, during a recession, profits are reduced, thereby making the budget balanced constraint tighter. If the government could borrow intertemporally it would have large incentives to do so and finance entry during downturns and pay back debt during booms. This suggests that, through the lens of our model, countercyclical firm subsidies can alleviate downturns by preventing the economy from falling in quasi-permanent recessions.

7 Empirical Evidence

Our model offers cross-industry predictions that can be tested in the data. In particular, according to our theory, industries featuring larger concentration in 2007 should have experienced a larger contraction in 2008. This prediction follows from equation (12). Recall that this equation establishes a positive link between productivity, market shares and markups (for a given number of active firms). Therefore, if we take two industries with the same number of firms, the one featuring a more uneven distribution of productivities will have larger dispersion in market shares and hence larger concentration. In these industries, firms at the bottom of the distribution will be smaller and charge lower markups, and will hence be more likely to exit upon a negative shock. Note that this prediction holds for a given number of firms $n_{i,t}$; when measuring the correlation between concentration in 2007 and the size of the contraction in 2008, we must therefore control for the number of firms in the industry.

We build a dataset combining the 2002 and 2007 US Census data on industry concentration to the Statistics of US Businesses (SUSB) and the Bureau of Labor Statistics (BLS) to obtain outcomes as employment, total wage bill and the number of firms at the industry level (6-digits NAICS). The final dataset includes 791 6-digit industries. In 2016, the median industry had 1,316 firms, 36,910 workers and a total payroll of \$1,880 million.

To assess whether industries with a larger concentration before the crisis experienced a larger post-crisis decline, we estimate the following model

$$\frac{\Delta y_{i,07-16}}{y_{i,07}} = \beta_0 + \beta_1 \text{ concent}_{i,07} + \beta_2 \text{ log firms}_{i,07} + a_s \mathbb{1}\{i \in s\} + u_i.$$

y_i is an outcome for industry i (for example total employment, total wage bill or total number of firms) and $concent_i$ is the share of the 4 largest firms (scaled by the share of the largest 50); we also control for the number of firms before the crisis (firms$_{i,07}$). The outcomes always take the form of the annualized growth rate between 2007 and 2016 in a specific industry. In all regressions, we will also include sector fixed effects as a control (a_s). The unit of observation is a 6-digit industry.

We start by studying the correlation between the change in employment between 2008 and 2016 and concentration in 2007. The results, showed in Table A.2 in Appendix A.4, suggest that more concentrated industries experienced lower employment growth in the aftermath of the great recession. Quantitatively, a 1pp higher pre-crisis concentration is associated with a 2pp lower employment growth rate between 2007 and 2016. This pattern holds irrespective of the inclusion of the number of firms in 2007. To address the concern that industries with larger concentration in 2007 could exhibit lower growth already before the crisis, we include cumulative employment growth between 2003 and 2007 as a control (column 3); the results do not change. Finally, the results are also robust to the inclusion of sector fixed effects (column 4). While these results concern the evolution of employment growth, a similar pattern is found if we use total wage bill instead (Table A.3). We also study the correlation between the measure of concentration and net entry after the crisis (Table A.4). Our finding suggests that a 1 percentage point increase in the concentration measure is associated with a 2 to 3 percentage points decrease in the post crisis net entry. These results suggest that industries with larger concentration in 2007 experienced a larger contraction in activity after the crisis.

We conclude this section by providing evidence on the evolution of the labor share across industries. While the US census of firms provides data on total employment and the total number of firms for all 6-digit industries, it does not contain data on the labor share. We rely on data from the BLS 'Labor Productivity and Cost' programme (see Appendix A.1 for details). This database, however, only provides data on the labor share for a restricted group of industries. The results are shown in Table A.5. Overall, there is a negative relationship between the post-crisis change in the labor share and the pre-crisis level of concentration. Industries with larger concentration in 2007 experienced a larger drop in labor share between 2008 and 2016.

In summary, these results suggest that the structure of US product markets in 2007 is important to understand the consequences of the 2008 crisis. The results presented are, strictly speaking, cross-sectional – industries with larger concentration in 2007, displayed a larger post-crisis contraction. We think, however, that they can also be used to support one of the main insights of the model – namely, that rising concentration can have made the US economy more vulnerable to aggregate shocks.

8 Conclusion

The US economy appears to have experienced a fundamental change over the past decades, with several studies and data sources indicating a reallocation of activity towards large, high markup firms. This observation has raised concerns in academic and policy circles about increasing market power, and it has been proposed as an explanation for recent macroeconomic *puzzles* – such as low aggregate investment, low wage growth or declining labor shares. Besides their impact on factor shares and factor prices, our model suggests that rising firm differences and greater market power can also have an impact on business cycles and provide an amplification and persistence mechanism to aggregate fluctuations. In particular, larger firm heterogeneity may have rendered the US economy more vulnerable to aggregate shocks and more likely to experience quasi-permanent recessions. Through the lens of our theory, such increased fragility may have been difficult to identify, as it manifests itself only in reaction to large shocks.

Our work also provides further motive for policies curbing market power. As we have shown, the endogenous response of the market structure to aggregate shocks act as an accelerant. On top of the standard static effects, any policy that reduces market power can have dynamic benefits in terms of the persistence and amplitude of aggregate fluctuations. These effects are particularly large if the economy is at risk of quasi-permanent recessions.

To keep the analysis simple, we have abstracted from a number of important features. One such example is that, in our model, firms solve a static problem and their productivity is permanent. The interaction between endogenous growth through innovation, market power and fragility is a natural next step for this area of research. Further, our model features one-sided market power. Recent models of oligopoly (see Azar and Vives, 2021) lend themselves to the study of the interaction between two-sided market power and the likelihood of quasi-permanent recessions. Lastly, recent work studies the interplay between competition and monetary policy (see Mongey, 2019; Wang and Werning, 2020; Fabiani et al., 2021). The question on how monetary policy, by changing the market structure, shapes the dynamic properties of an economy is an avenue for future research.

Supplementary Material

The Online Appendix and the Supplementary Material can be found here and here.

References

Akcigit, Ufuk and Sina T. Ates, "What Happened to U.S. Business Dynamism?," NBER Working Paper 25756 2019.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal, "Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries," OECD Productivity Working Papers 2015.

Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez, "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, July 2019, *11* (3), 67–110.

Atkeson, Andrew and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, December 2008, *98* (5), 1998–2031.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," *The Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Azar, José and Xavier Vives, "General Equilibrium Oligopoly and Ownership Structure," *Econometrica*, 2021, *89* (3), 999–1048.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

Benhabib, Jess and Roger E. A. Farmer, "Indeterminacy and Increasing Returns," *Journal of Economic Theory*, June 1994, *63* (1), 19–41.

Benigno, Gianluca and Luca Fornaro, "Stagnation traps," *The Review of Economic Studies*, 2017, *85* (3), 1425–1470.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz, "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy*, 2012, *120* (2), 304–345.

_ , _ , and _ , "Monopoly Power and Endogenous Product Variety: Distortions and Remedies," *American Economic Journal: Macroeconomics*, October 2019, *11* (4), 140–174.

Bils, Mark, Peter J. Klenow, and Benjamin A. Malin, "Resurrecting the Role of the Product Market Wedge in Recessions," *American Economic Review*, April 2018, *108* (4-5), 1118–1146.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch, "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data," *Journal of Monetary Economics*, 2021.

Bonfiglioli, Alessandra, Rosario Crinò, and Gino Gancia, "Betting on Exports: Trade and Endogenous Heterogeneity," *Economic Journal*, March 2018, *128* (609), 612–651.

Burstein, Ariel, Basile Grassi, and Vasco Carvalho, "Bottom-Up Markup Fluctuations," Working Paper 2020.

Calligaris, Sara, Chiara Criscuolo, and Luca Marcolin, "Mark-ups in the digital era," OECD Science, Technology and Industry Working Papers 2018.

Carvalho, Vasco and Xavier Gabaix, "The Great Diversification and Its Undoing," *American Economic Review*, August 2013, *103* (5), 1697–1727.

Chatterjee, Satyajit, Russell Cooper, and B Ravikumar, "Strategic complementarity in business formation: Aggregate fluctuations and sunspot equilibria," *The Review of Economic Studies*, 1993, *60* (4), 795–811.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 2011, *101* (3), 471–475.

Clementi, Gian Luca, Berardino Palazzo, and Peifan Wu, "Firm Demographics and the Great Recession," Technical Report, Society for Economic Dynamics 2017.

Cooper, Russell and Andrew John, "Coordinating Coordination Failures in Keynesian Models," *The Quarterly Journal of Economics*, 1988, *103* (3), 441–463.

— **and** — , "Imperfect competition and macroeconomics : Theory and quantitative implications," *Cahiers d'Économie Politique*, 2000, *37* (1), 289–328.

Decker, Ryan A, John C Haltiwanger, Ron S Jarmin, and Javier Miranda, "Changing business dynamism and productivity: Shocks vs. responsiveness," Technical Report, National Bureau of Economic Research 2018.

Díez, Federico J, Daniel Leigh, and Suchanan Tambunlertchai, "Global Market Power and its Macroeconomic Implications," IMF Working Papers, International Monetary Fund June 2018.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How costly are markups?," Technical Report 2021.

Eggertsson, Gauti B, Jacob A Robbins, and Ella Getz Wold, "Kaldor and Piketty's facts: The rise of monopoly power in the united states," NBER Working Paper 24287 2018.

Etro, Federico and Andrea Colciago, "Endogenous Market Structures and the Business Cycle," *Economic Journal*, 2010, *120* (549), 1201–1233.

Fabiani, Andrea, Luigi Falasconi, and Janko Heineken, "Monetary Policy and Corporate Debt Maturity," Technical Report 2021.

Farmer, Roger and Jang-Ting Guo, "Real Business Cycles and the Animal Spirits Hypothesis," *Journal of Economic Theory*, June 1994, *63* (1), 42–72.

Fernald, John G., "A quarterly, utilization-adjusted series on total factor productivity," Working Paper Series 2012-19, Federal Reserve Bank of San Francisco 2012.

Galí, Jordi and Fabrizio Zilibotti, "Endogenous Growth and Poverty Traps in a Cournotian Model," *Annals of Economics and Statistics*, 1995, *37-38*, 197–213.

Gorodnichenko, Yuriy and Michael Weber, "Are Sticky Prices Costly? Evidence from the Stock Market," *American Economic Review*, January 2016, *106* (1), 165–199.

Gouin-Bonenfant, Emilien, "Productivity Dispersion, Between-firm Competition and the Labor Share," Technical Report 2018.

Grassi, Basile, “Io in io: Size, industrial organization, and the input-output network make a firm structurally important,” *Work. Pap., Bocconi Univ., Milan, Italy*, 2017.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W Huffman, “Investment, Capacity Utilization, and the Real Business Cycle,” *American Economic Review*, June 1988, *78* (3), 402–417.

Gutiérrez, Germán and Thomas Philippon, “Declining Competition and Investment in the US,” NBER Working Paper 23583 2017.

Hall, Robert E, “Using empirical marginal cost to measure market power in the us economy,” Technical Report, National Bureau of Economic Research 2018.

Herrendorf, Berthold, Akos Valentinyi, and Robert Waldmann, “Ruling out multiplicity and indeterminacy: the role of heterogeneity,” *The Review of Economic Studies*, 2000, *67* (2), 295–307.

Jaimovich, Nir, “Firm dynamics and markup variations: Implications for sunspot equilibria and endogenous economic fluctuations,” *Journal of Economic Theory*, November 2007, *137* (1), 300–325.

— **and Henry E. Siu**, “The Young, the Old, and the Restless: Demographics and Business Cycle Volatility,” *American Economic Review*, June 2009, *99* (3), 804–826.

— **and Max Floetotto**, “Firm dynamics, markup variations, and the business cycle,” *Journal of Monetary Economics*, 2008, *55* (7), 1238–1252.

Karabarbounis, Loukas and Brent Neiman, “Accounting for factorless income,” *NBER Macroeconomics Annual*, 2019, *33* (1), 167–228.

Kehrig, Matthias, “The cyclical nature of the productivity distribution,” *Earlier version: US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15*, 2015.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, “The rise of market power and the macroeconomic implications,” *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

— **,** — **, and Simon Mongey**, “Quantifying market power and business dynamism,” *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2021, (2021-54).

Matsuyama, Kiminori, “Increasing Returns, Industrialization, and Indeterminacy of Equilibrium,” *The Quarterly Journal of Economics*, 1991, *106* (2), 617–650.

Mongey, Simon, "Market structure and monetary non-neutrality," *mimeo*, 2019.

Nekarda, Christopher J and Valerie A Ramey, "The Cyclical Behavior of the Price-Cost Markup," *Journal of Money, Credit and Banking*, 2020, *52* (S2), 319–353.

Pagano, Marco, "Imperfect Competition, Underemployment Equilibria and Fiscal Policy," *Economic Journal*, June 1990, *100* (401), 440–463.

Rotemberg, Julio and Garth Saloner, "A supergame-theoretic model of price wars during booms," *New Keynesian Economics*, 1986, *2*, 387–415.

Schaal, Edouard and Mathieu Taschereau-Dumouchel, "Coordinating business cycles," *Working Paper*, 2015.

_ **and** _ , "Aggregate Demand and the Dynamics of Unemployment," *Working Paper*, 2019.

Traina, James, "Is aggregate market power increasing? production trends using financial statements," *Production Trends Using Financial Statements (February 8, 2018)*, 2018.

Van Reenen, John, "Increasing Differences Between Firms: Market Power and the Macro-Economy," *Working Paper*, September 2018.

Wang, Olivier and Iván Werning, "Dynamic oligopoly and price stickiness," Technical Report, National Bureau of Economic Research 2020.

Appendix A - For Online Publication Only

A.1 Data Appendix

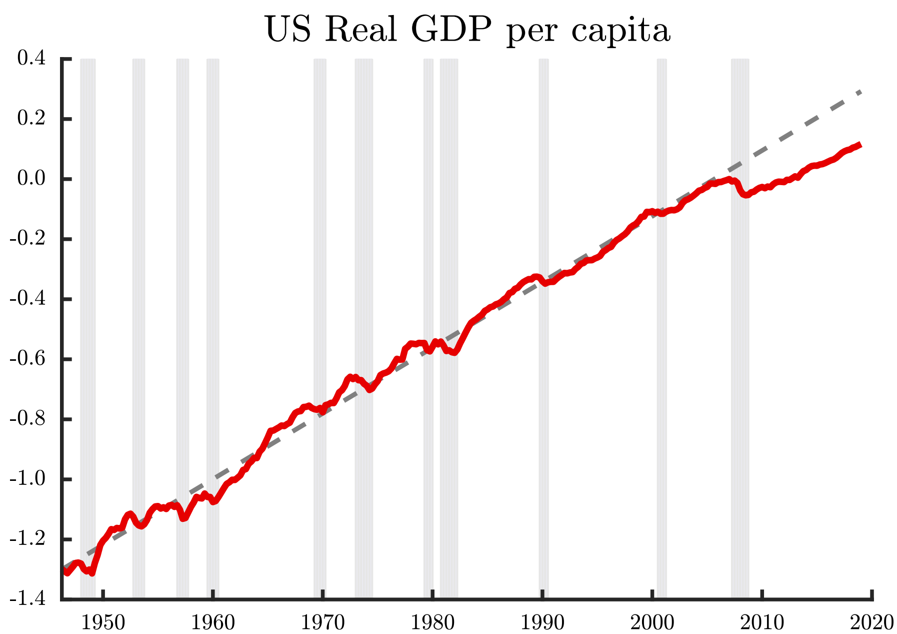


Figure A.1: The *Great Deviation*

Note: This figure shows real GDP per capita (from BEA). The series is in logs, undetrended and centered around 2007. The linear trend is computed for the 1947-2007 period.

Data Definition Table A.1 provides information on all the data sources used in Section 5.

Variable	Source
Real GDP	BEA – NIPA Table 1.1.3 (line 1)
Real Personal Consumption Expenditures	BEA – NIPA Table 1.1.3 (line 2)
Real Gross Private Domestic Investment	BEA – NIPA Table 1.1.3 (line 7)
Total Hours	BLS – Nonfarm Business sector: Hours of all persons
Aggregate TFP	Fernald (2012): Raw Business Sector TFP
Population	BEA – NIPA Table 2.1 (line 40)

Table A.1: Data sources

Aggregate Profit Share The aggregate profit share is computed as

$$\text{profit share}_t \;=\; 1 \;-\; \text{labor share}_t \;-\; \underbrace{\frac{R_t \cdot \mathrm{K}_t \;-\; \mathrm{DEP}_t}{\mathrm{VA}_t}}_{\text{capital share}}$$

where labor share$_t$ is the labor share of the US business sector (from BLS), VA$_t$ is the total value added of the US business sector (NIPA Table 1.3.5, line 2). K$_t$ is the value of private fixed assets (including intangibles) of the US business sector (NIPA Table 6.1, line 1 - line 9 - line 10) and DEP$_t$ is depreciation (NIPA Table 6.4, line 1 - line 9 - line 10). Finally, R$_t$ is the required rate of return. We follow Eggertsson et al. (2018) and compute it as the difference between Moody's Seasoned BAA Corporate Bond Yield and a 5-year moving average of past CPI inflation (from BLS, used as a proxy for expected inflation).

Industry-level Labor Share We obtain data on the labor share at the industry level from the BLS 'Labor Productivity and Costs' (LPC) database. We calculate the labor share as the ratio of 'Labor compensation' to 'Value of Production'. Note that this ratio gives the share of labor compensation in total revenues, and not in value added.[32]

A.2 Proofs and Additional Results

Proof of Proposition 1

Proof. First, note that, by Definition 4 $\frac{\partial \chi_{n+1}}{\partial \lambda} = -\frac{\partial \mathcal{K}_n^U/\mathcal{K}_{n+1}}{\partial \lambda}$. Next, by Definition 3, a sufficient condition for the ratio $\mathcal{K}_n^U/\mathcal{K}_{n+1}$ to decrease in λ is that $R_\lambda < 0$ at $\mathcal{K}_n^U$ and $\mathcal{K}_{n+1}$. It therefore suffices to characterize a sufficient condition for the $R(\cdot)$ map to be locally decreasing in λ. To this end, we begin by using the analog of equation 3 for the wage: $W = \Theta(\Lambda, n)F_L$. By substituting in the labor supply L^S and inverting we have $L^S F_L^{-1} = \Theta$. As both functions on the LHS are non-decreasing in L it follows that $\text{sgn}\{L_\lambda\} = \text{sgn}\{\Theta_\lambda\}$. We are interested in $\partial R/\partial\lambda$ which is given by

$$R_\lambda = \Theta_\lambda F_K + \Theta F_{K\lambda}.$$

As $\text{sgn}\{F_{K\lambda}\} = \text{sgn}\{L_\lambda\} = \text{sgn}\{\Theta_\lambda\}$, $\Theta > 0$, $\text{sgn}\{R_\lambda\} = \text{sgn}\{\Theta_\lambda\}$, the statement follows. ∎

[32]This ratio coincides with the 'Labor cost share' provided by the BLS. This variable is, however, available just for a restricted number of industries.

Proof of Proposition 2

Proof. Suppose $\eta = 1$.When there are n active firms in a given industry, the profits of a firm with productivity γ_j are equal to

$$\Pi\left(\gamma_j, n, \mathcal{F}, \Theta, Y\right) = \Lambda\left(\gamma_j, n, \mathcal{F}\right)\Theta^{-\frac{\rho}{1-\rho}}Y \tag{27}$$

where $\Lambda\left(\gamma_j, n, \mathcal{F}\right)$ has been defined in Appendix B.2. A symmetric equilibrium with n firms per industry is possible provided that

$$\Lambda\left(\gamma_n, n, \mathcal{F}\right)\Theta^{-\frac{\rho}{1-\rho}}Y \geq c$$
$$\Lambda\left(\gamma_{n+1}, n+1, \mathcal{F}\right)\Theta^{-\frac{\rho}{1-\rho}}Y \leq c$$

Using equation (23), we can write the above inequalities as

$$\underline{K}\left(\mathcal{F}, n\right) \leq K_t \leq \overline{K}\left(\mathcal{F}, n\right), \tag{28}$$

where

$$\underline{K}\left(\mathcal{F}, n\right) = \left\{\frac{c}{\Lambda\left(\gamma_n, n, \mathcal{F}\right)}\left(1-\alpha\right)^{-\frac{1-\alpha}{\nu+\alpha}}\left[\Phi\left(\mathcal{F}, n\right)\right]^{-1}\left[\Theta\left(\mathcal{F}, n\right)\right]^{\frac{\rho}{1-\rho}-\frac{1-\alpha}{\nu+\alpha}}\right\}^{\frac{\nu+\alpha}{\alpha(1+\nu)}} \tag{29}$$

$$\overline{K}\left(\mathcal{F}, n\right) = \left\{\frac{c}{\Lambda\left(\gamma_{n+1}, n+1, \mathcal{F}\right)}\left(1-\alpha\right)^{-\frac{1-\alpha}{\nu+\alpha}}\left[\Phi\left(\mathcal{F}, n\right)\right]^{-1}\left[\Theta\left(\mathcal{F}, n\right)\right]^{\frac{\rho}{1-\rho}-\frac{1-\alpha}{\nu+\alpha}}\right\}^{\frac{\nu+\alpha}{\alpha(1+\nu)}}. \tag{30}$$

■

Proof of Lemma 1

Proof. In a steady-state we have a constant rental rate

$$R^* = \beta^{-1} - (1-\delta) \tag{31}$$

and

$$\delta K = sY \tag{32}$$

Combining these two equations with equation (18) we obtain

$$
\begin{aligned}
&\beta^{-1} - (1-\delta) = \alpha\, \Omega\left(\mathbf{\Gamma}, \mathbf{N}\right)\, \frac{Y^*}{K^*} \\
\Leftrightarrow &\beta^{-1} - (1-\delta) = \alpha\, \Omega\left(\mathbf{\Gamma}, \mathbf{N}\right)\, \frac{\delta}{s^*} \\
\Leftrightarrow &s^* = \frac{\delta\, \alpha}{\beta^{-1} - (1-\delta)} \Omega\left(\mathbf{\Gamma}, \mathbf{N}\right)
\end{aligned} \tag{33}
$$

■

Proof of Lemma 2

Proof. When all industries are identical and have n firms, the factor price index is equal to

$$
\Theta\left(\Gamma_n, n\right) = \frac{n - (1-\rho)}{\sum_{k=1}^{n} \frac{1}{\gamma_k}} \tag{34}
$$

As Θ is a concave function of γ_k, by the definition of MPS, we have that if $\tilde{\Gamma}$ is a MPS of Γ, then $\Theta(\tilde{\Gamma}, n) < \Theta(\Gamma, n)$. ■

Proof of Proposition 3

Proof. a) Let K_n^* be a steady-state where all industries are have n firms and common productivity distribution Γ_n. Using equation (18), we can define K_n^* as

$$
R^* \; = \; \alpha \; (1-\alpha)^{(1-\alpha)/(\nu+\alpha)} \; \Theta\left(\Gamma_n, n\right)^{(\nu+1)/(\nu+\alpha)} \; \left(K_n^*\right)^{-\nu(1-\alpha)/(\nu+\alpha)} \tag{35}
$$

Recall from Lemma 2 that $\Theta\left(\Gamma_n, n\right)$ declines after a MPS on Γ_n. Then K_n^* must also decline.
b) We provide a sufficient condition under which the unstable steady-state increases after an MPS. Note that the unstable steady-state increases whenever the increasing segment of the rental rate map lies strictly underneath the original one.
We know from the proof of Proposition 3.a that the new rental rate at $\underline{K}\left(2\right)$ is strictly lower than before. The proof involves two steps
[A] We derive a sufficient condition under which the new rental rate at $\overline{K}\left(1\right)$ is lower than before
[B] We show that the increasing segment of the rental rate map after an MPS cannot cross the previous one more than once. Thus, if the new segment starts and ends below the previous one, it can never go above it.

Proof of Part A

The free entry condition is

$$\Lambda\left(n\right)\Theta\left(n-1\right)^{-\frac{\rho}{1-\rho}}\ \Phi\left(n-1\right)K^{\alpha}\,L^{1-\alpha}\ =\ c_{f} \tag{36}$$

Using

$$L=\left[\left(1-\alpha\right)\Theta\left(n-1\right)\right]^{\frac{1}{\nu+\alpha}}\,K^{\frac{\alpha}{\nu+\alpha}} \tag{37}$$

we can rewrite the free-entry condition as

$$\begin{aligned}
&\Lambda\left(n\right)\Theta\left(n-1\right)^{-\frac{\rho}{1-\rho}}\ \Phi\left(n-1\right)\ K^{\alpha}\ \left[\left(1-\alpha\right)\Theta\left(n-1\right)\right]^{\frac{1-\alpha}{\nu+\alpha}}\,K^{\alpha\frac{1-\alpha}{\nu+\alpha}}\ =\ c_{f}\\
\Leftrightarrow&\Lambda\left(n\right)\Theta\left(n-1\right)^{-\frac{\rho}{1-\rho}}\ \Phi\left(n-1\right)\left[\left(1-\alpha\right)\Theta\left(n-1\right)\right]^{\frac{1-\alpha}{\nu+\alpha}}\,K^{\alpha\frac{\nu+1}{\nu+\alpha}}\ =\ c_{f}\\
\Leftrightarrow&K\ =\ \left[\frac{c_{f}\left(1-\alpha\right)^{-\frac{1-\alpha}{\nu+\alpha}}}{\Lambda\left(n\right)\Theta\left(n-1\right)^{\frac{1-\alpha}{\nu+\alpha}-\frac{\rho}{1-\rho}}\ \Phi\left(n-1\right)}\right]^{\frac{1}{\alpha}\frac{\nu+\alpha}{\nu+1}}
\end{aligned} \tag{38}$$

The interest rate is

$$\begin{aligned}
R=&\alpha\,\Theta\left(n-1\right)\ K^{\alpha-1}\,L^{1-\alpha}\\
=&\alpha\,\Theta\left(n-1\right)\ K^{\alpha-1}\ \left[\left(1-\alpha\right)\Theta\left(n-1\right)\right]^{\frac{1-\alpha}{\nu+\alpha}}\,K^{\alpha\frac{1-\alpha}{\nu+\alpha}}\\
=&\alpha\ \left(1-\alpha\right)^{\frac{1-\alpha}{\nu+\alpha}}\ \Theta\left(\mathbf{\Gamma},\mathbf{N}_{t}\right)^{\frac{\nu+1}{\nu+\alpha}}\ K^{\nu\frac{\alpha-1}{\nu+\alpha}}
\end{aligned} \tag{39}$$

Putting the two together

$$R=\alpha\ \left(1-\alpha\right)^{\frac{1-\alpha}{\nu+\alpha}}\ \Theta\left(n-1\right)^{\frac{\nu+1}{\nu+\alpha}}\ \left(\frac{\Lambda\left(n\right)\Theta\left(n-1\right)^{\frac{1-\alpha}{\nu+\alpha}-\frac{\rho}{1-\rho}}\ \Phi\left(n-1\right)}{c_{f}\left(1-\alpha\right)^{-\frac{1-\alpha}{\nu+\alpha}}}\right)^{\frac{\nu}{\nu+1}\frac{1-\alpha}{\alpha}} \tag{40}$$

implying

$$R^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}}=\propto\ \Theta\left(n-1\right)^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}}\ \Lambda\left(n\right)\Theta\left(n-1\right)^{\frac{1-\alpha}{\nu+\alpha}-\frac{\rho}{1-\rho}}\ \Phi\left(n-1\right) \tag{41}$$

where

$$\Theta(n) = g(n) \tag{42}$$

$$\Phi(n) = \frac{1}{\sum_{j=1}^{n} \frac{s_j}{\pi_j}} \tag{43}$$

$$\Lambda(n) = s_n^2 \left[g(n)\right]^{\frac{\rho}{1-\rho}} \tag{44}$$

We can thus rewrite (40) as

$$R^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}} = g(n-1)^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}+\frac{1-\alpha}{\nu+\alpha}} \frac{s_n^2\left[\frac{g(n)}{g(n-1)}\right]^{\frac{\rho}{1-\rho}}}{\sum_{j=1}^{n-1}\frac{\hat{s}_j}{\pi_j}} \tag{45}$$

where $\hat{s}_j$ is the market share of firm j in an industry with $n-1$ firms and s_j is the market share of that firm when there are n player in the industry. We want to show that the expression in (45) goes down when we do an MPS on n firms. The challenge is in the fact that the expression involves terms that refer to $n-1$ industries.
Under $n = 2$, we have

$$g(1) = \rho\, \pi_1 \tag{46}$$

$$g(2) = \frac{1+\rho}{\frac{1}{\pi_1}+\frac{1}{\pi_2}} \tag{47}$$

so that

$$\begin{aligned} R^{\frac{\alpha}{1-\alpha}} &= (\rho\,\pi_1)^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}+\frac{1-\alpha}{\nu+\alpha}-\frac{\rho}{1-\rho}} \left(\frac{1+\rho}{\frac{1}{\pi_1}+\frac{1}{\pi_2}}\right)^{\frac{\rho}{1-\rho}} \frac{\left[1-\frac{1+\rho}{\frac{1}{\pi_1}+\frac{1}{\pi_2}}\frac{1}{\pi_2}\right]^2}{\frac{1}{\pi_1}} \\ &= \propto \pi_1^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}+\frac{1-\alpha}{\nu+\alpha}-\frac{\rho}{1-\rho}-1} \left[\pi_1-\pi_1\frac{1+\rho}{\frac{2x}{\pi_1}}\right]^2 \left[\frac{1+\rho}{\frac{1}{\pi_1}+\frac{1}{2x-\pi_1}}\right]^{\frac{\rho}{1-\rho}} \end{aligned} \tag{48}$$

The last term is decreasing on an MPS, since it is simply $g(2)$. The first term is decreasing on an

MPS provided that

$$1+\frac{\rho}{1-\rho}>\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}+\frac{1-\alpha}{\nu+\alpha}\Leftrightarrow\alpha<\frac{\nu+\rho-1}{\nu\left(2-\rho\right)} \tag{49}$$

since this MPS must result in higher π_1. We just need to evaluate the term in the middle. Note that we can rewrite it as

$$\pi_1-\pi_1^2\frac{1+\rho}{2x} \tag{50}$$

where $2x\equiv\pi_1+\pi_2$ is fixed by construction. The derivative of the expression above is

$$\frac{\partial}{\partial\pi_1}=1-2\pi_1\frac{1+\rho}{2x} \tag{51}$$

$$=1-\underbrace{\frac{\pi_1}{x}}_{>1}\left(1+\rho\right)<0 \tag{52}$$

Therefore, for $n=2$, the interest rate is always declining on an MPS provided that

$$\rho>1-\nu\frac{1-\alpha}{1+\nu\alpha} \tag{53}$$

Proof of Part B

Recall that the free entry condition is

$$\Lambda_j\Theta^{-\frac{\rho}{1-\rho}}\,\Phi\,K^{\alpha}\,L^{1-\alpha}=c_f \tag{54}$$

Aggregate TFP can be written as

$$\begin{aligned}\Phi=&\frac{\left\{(1-m)\left[g\left(n-1\right)\right]^{\frac{\rho}{1-\rho}}+m\left[g\left(n\right)\right]^{\frac{\rho}{1-\rho}}\right\}^{\frac{1}{\rho}}}{(1-m)\left[g\left(n-1\right)\right]^{\frac{1}{1-\rho}}h\left(n-1\right)+m\left[g\left(n\right)\right]^{\frac{1}{1-\rho}}h\left(n\right)}\\=&\frac{\Theta^{\frac{1}{1-\rho}}}{(1-m)\left[g\left(n-1\right)\right]^{\frac{1}{1-\rho}}h\left(n-1\right)+m\left[g\left(n\right)\right]^{\frac{1}{1-\rho}}h\left(n\right)}\end{aligned}$$

where

$$g(n) = \frac{n - (1-\rho)}{\sum_{j=1}^{n} \frac{1}{\pi_j}} \tag{55}$$

$$h(n) = \sum_{j=1}^{n} \frac{s_j}{\pi_j} \tag{56}$$

Now suppose that we do the MPS and have $\tilde{\Theta} = \Theta$ at the same K.[33] From the free entry condition and the expression for Φ, this is possible if

$$\begin{aligned} \frac{\Lambda_j}{\tilde{\Lambda}_j} &= \frac{(1-m)\left[g(n-1)\right]^{\frac{1}{1-\rho}} h(n-1) + m\left[g(n)\right]^{\frac{1}{1-\rho}} h(n)}{(1-\tilde{m})\left[\tilde{g}(n-1)\right]^{\frac{1}{1-\rho}} \tilde{h}(n-1) + \tilde{m}\left[\tilde{g}(n)\right]^{\frac{1}{1-\rho}} \tilde{h}(n)} \\ \Leftrightarrow \frac{\Lambda_j}{\tilde{\Lambda}_j} &= \frac{\left[g(n-1)\right]^{\frac{1}{1-\rho}} h(n-1) + m\left\{\left[g(n)\right]^{\frac{1}{1-\rho}} h(n) - \left[g(n-1)\right]^{\frac{1}{1-\rho}} h(n-1)\right\}}{\left[\tilde{g}(n-1)\right]^{\frac{1}{1-\rho}} \tilde{h}(n-1) + \tilde{m}\left\{\left[\tilde{g}(n)\right]^{\frac{1}{1-\rho}} \tilde{h}(n) - \left[\tilde{g}(n-1)\right]^{\frac{1}{1-\rho}} \tilde{h}(n-1)\right\}} \end{aligned} \tag{57}$$

Rearranging this equation, we can write

$$\tilde{m} = a_1 + b_1 \cdot m \tag{58}$$

where a_1 and b_1 are some numbers (independent of K).
Furthermore, from $\tilde{\Theta} = \Theta$ we have

$$\begin{aligned} &(1-m)\left[g(n-1)\right]^{\frac{\rho}{1-\rho}} + m\left[g(n)\right]^{\frac{\rho}{1-\rho}} = (1-\tilde{m})\left[\tilde{g}(n-1)\right]^{\frac{\rho}{1-\rho}} + \tilde{m}\left[\tilde{g}(n)\right]^{\frac{\rho}{1-\rho}} \\ \Leftrightarrow &\left[g(n-1)\right]^{\frac{\rho}{1-\rho}} + m\left\{\left[g(n)\right]^{\frac{\rho}{1-\rho}} - \left[g(n-1)\right]^{\frac{\rho}{1-\rho}}\right\} = \left[\tilde{g}(n-1)\right]^{\frac{\rho}{1-\rho}} + \tilde{m}\left\{\left[\tilde{g}(n)\right]^{\frac{\rho}{1-\rho}} - \left[\tilde{g}(n-1)\right]^{\frac{\rho}{1-\rho}}\right\} \end{aligned} \tag{59}$$

Rearranging this equation, we can write

$$\tilde{m} = a_2 + b_2 \cdot m \tag{60}$$

Combining (58) and (60), there is at most one pair $(m, \tilde{m})$ such that $\tilde{\Theta} = \Theta$. This establishes that $\tilde{\Theta}$ cannot cross Θ twice. ■

Proposition A.1 (Mean Preserving Spread and Fragility, general n). *Let $\eta = 1$ and suppose that all*

[33] We also have $\tilde{L} = L$, since L is a function of Θ and K.

industries are identical to start ($\gamma_{ij} = \gamma_j \; \forall i$). Let $K^*(n)$ be a steady-state with n firms. Let λ be a mean-preserving spread on the distribution $\{\gamma_1, \ldots, \gamma_n\}$ of active firms, such that for any $j = 1, \ldots, n-1$, γ_1/γ_j is unchanged. Then, if

$$\rho > \frac{1+\nu\alpha}{1+\nu} \tag{61}$$

we have that

$$\frac{\partial \underline{B}(n)}{\partial \lambda} > 0. \tag{62}$$

Proof of Proposition A.1. Let us now provide a sufficient condition for general n. First note that we can write (45) as

$$R^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}} = \underbrace{g(n-1)^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}+\frac{1-\alpha}{\nu+\alpha}-\frac{\rho}{1-\rho}} g(n)^{\frac{\rho}{1-\rho}}}_{v} \underbrace{\frac{1}{\sum_{j=1}^{n-1}\frac{\hat{s}_j}{\pi_j}}}_{z} s_n^2 \tag{63}$$

$$R^{\frac{\alpha}{1-\alpha}} = \underbrace{g(n-1)^{\frac{\alpha}{1-\alpha}-\frac{\rho}{1-\rho}} g(n)^{\frac{\rho}{1-\rho}}}_{v} \underbrace{\frac{1}{\sum_{j=1}^{n-1}\frac{\hat{s}_j}{\pi_j}}}_{z} s_n^2 \tag{64}$$

The first term v is always decreasing on an MPS provided that

$$\frac{\rho}{1-\rho} > \frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha} + \frac{1-\alpha}{\nu+\alpha} \Leftrightarrow \alpha < \frac{\rho(1+\nu)-1}{\nu} \tag{65}$$

To see it note that $g(n)$ is decreasing on an MPS. If $g(n-1)$ is increasing on an MPS, it immediately follows that v decreases on an MPS when the above condition is satisfied. If $g(n-1)$ is instead decreasing on an MPS, just rewrite v as

$$v = g(n-1)^{\frac{\nu+1}{\nu}\frac{\alpha}{1-\alpha}\frac{\nu+1}{\nu+\alpha}+\frac{1-\alpha}{\nu+\alpha}} \left[\frac{g(n)}{g(n-1)}\right]^{\frac{\rho}{1-\rho}} \tag{66}$$

and note that $\dfrac{g(n)}{g(n-1)}$ decreases on an MPS.

Thus, all we need to show is that z is also decreasing on an MPS. Note that we can write z as

$$z = \frac{s_n^2}{\sum_{j=1}^{n-1} \frac{\hat{s}_j}{\pi_j}} = \frac{1}{\sum_{j=1}^{n-1} \frac{\hat{s}_j}{\pi_j} \frac{1}{s_n^2}} \tag{67}$$

We know that s_n is decreasing on an MPS. A sufficient condition for z to be decreasing on an MPS is that

$$\frac{\hat{s}_j}{\pi_j} \frac{1}{s_n^2} \tag{68}$$

is increasing on an MPS for every $j = 1, 2, \ldots, n-1$.
Consider an MPS such that

$$\widetilde{\pi}_j = \gamma\, \pi_j \quad \forall j = 1, 2, \ldots, n-1 \tag{69}$$

and

$$\gamma \sum_{j=1}^{n-1} \pi_j + \widetilde{\pi}_n = \sum_{j=1}^{n-1} \pi_j + \pi_n \tag{70}$$

$$\Leftrightarrow \widetilde{\pi}_n = \pi_n - (\gamma - 1) \sum_{j=1}^{n-1} \pi_j \tag{71}$$

In this case we have

$$\frac{1}{\gamma} \underbrace{\frac{\hat{s}_j}{\pi_j}}_{\text{const}} \frac{1}{\widetilde{s}_n^2} \tag{72}$$

We need to show that

$$\gamma \underbrace{\left[1 - \frac{n - (1-\rho)}{\sum_{j=1}^{n-1} \frac{1}{\gamma \pi_j} + 1} \frac{1}{\pi_n - (\gamma - 1) \sum_{j=1}^{n-1} \pi_j} \right]^2}_{\widetilde{s}_n} \tag{73}$$

decreases in γ. Note that we can rewrite it as

$$\gamma\underbrace{\left[1-\frac{n-(1-\rho)}{\sum\limits_{j=1}^{n-1}\frac{\pi_n-(\gamma-1)\sum\limits_{j=1}^{n-1}\pi_j}{\gamma\,\pi_j}+1}\right]^2}_{\tilde{s}_n}=\gamma\underbrace{\left[1-\frac{n-(1-\rho)}{\left[\frac{1}{\gamma}\left(\pi_n+\sum\limits_{j=1}^{n-1}\pi_j\right)-\sum\limits_{j=1}^{n-1}\pi_j\right]\sum\limits_{j=1}^{n-1}\frac{1}{\pi_j}+1}\right]^2}_{\tilde{s}_n} \tag{74}$$

The derivative with respect to γ is

$$\tilde{s}_n^2+\gamma 2\tilde{s}_n\left\{-(-1)\left[n-(1-\rho)\right]\frac{\left(\pi_n+\sum\limits_{j=1}^{n-1}\pi_j\right)\sum\limits_{j=1}^{n-1}\frac{1}{\pi_j}\left(-\frac{1}{\gamma^2}\right)}{\left(\left[\frac{1}{\gamma}\left(\pi_n+\sum\limits_{j=1}^{n-1}\pi_j\right)-\sum\limits_{j=1}^{n-1}\pi_j\right]\sum\limits_{j=1}^{n-1}\frac{1}{\pi_j}+1\right)^2}\right\} \tag{75}$$

which is lower than zero if

$$\tilde{s}_n-\frac{2}{\gamma}\left[\frac{n-(1-\rho)}{\sum\limits_{j=1}^{n}\frac{\tilde{\pi}_n}{\gamma\pi_j}}\frac{1}{\sum\limits_{j=1}^{n}\frac{\tilde{\pi}_n}{\gamma\pi_j}}\left(\pi_n+\sum_{j=1}^{n-1}\pi_j\right)\sum_{j=1}^{n-1}\frac{1}{\pi_j}\right]<0\Leftrightarrow 1-\underbrace{\frac{n-(1-\rho)}{\sum\limits_{j=1}^{n}\frac{\tilde{\pi}_n}{\pi_j}}}_{(a)}\left[1+2\gamma\underbrace{\frac{\sum\limits_{j=1}^{n}\pi_j}{\tilde{\pi}_n}}_{>n}\underbrace{\frac{\sum\limits_{j=1}^{n-1}\frac{1}{\pi_j}}{\sum\limits_{j=1}^{n}\frac{1}{\pi_j}}}_{(b)}\right]<0 \tag{76}$$

If suffices to prove that $(a)>1/3$ and that $(b)>1/n$.

To prove the first, note that

$$\tilde{s}_n=\frac{1}{1-\rho}\left[1-\frac{n-(1-\rho)}{\sum\limits_{j=1}^{n}\frac{\tilde{\pi}_n}{\pi_j}}\right]<\frac{1}{3}\Leftrightarrow\underbrace{\frac{n-(1-\rho)}{\sum\limits_{j=1}^{n}\frac{\tilde{\pi}_n}{\pi_j}}}_{(a)}>\frac{2+\rho}{3}>\frac{1}{3} \tag{77}$$

To prove the second, note that

$$\frac{\sum_{j=1}^{n-1} \frac{1}{\pi_j}}{\sum_{j=1}^{n} \frac{1}{\pi_j}} > \frac{1}{n} \Leftrightarrow \sum_{j=1}^{n} \frac{\pi_n}{\pi_j} > \frac{n}{n-1} \tag{78}$$

The last equation is implied by the fact that

$$1 - \frac{n-(1-\rho)}{\sum_{j=1}^{n} \frac{\pi_n}{\pi_j}} > 0 \Leftrightarrow \sum_{j=1}^{n} \frac{\pi_n}{\pi_j} > \underbrace{n-(1-\rho)}_{>2 \text{ under } n \geq 3} \tag{79}$$

which is needed for $s_n > 0$. This completes the proof. ∎

Proof of Proposition 4

Proof. From equation (18) we can write

$$R_t \;=\; \alpha \;(1-\alpha)^{(1-\alpha)/(\nu+\alpha)} \;\Theta\left(\mathbf{\Gamma}, \mathbf{N}_t\right)^{(\nu+1)/(\nu+\alpha)} \;K_t^{-\nu(1-\alpha)/(\nu+\alpha)} \tag{80}$$

where $\Theta\left(\mathbf{\Gamma}, \mathbf{N}_t\right)$ is increasing in the number of active firms (as explained above). For a given steady-state K^*, the slackness free entry condition may or may not hold exactly. If it does hold exactly then, in response to a marginal increase in c, the number of firms will necessarily decrease and so will the level of capital at the steady-state. If it does not hold exactly then the level of capital will be unchanged as no firm will leave the market. The statement of part a) follows.

Second, at an unstable steady-state K_U, the rental rate is increasing in the capital stock. For this to happen, $\Theta\left(\mathbf{\Gamma}, \mathbf{N}_t\right)$ must be increasing in K at that point. Assuming I large, this only happens if some firm is exactly breaking even. Therefore, the rental rate at an unstable steady-state K_U necessarily declines after an increase in c, as stated in part b).

∎

A.3 Policy Evaluation

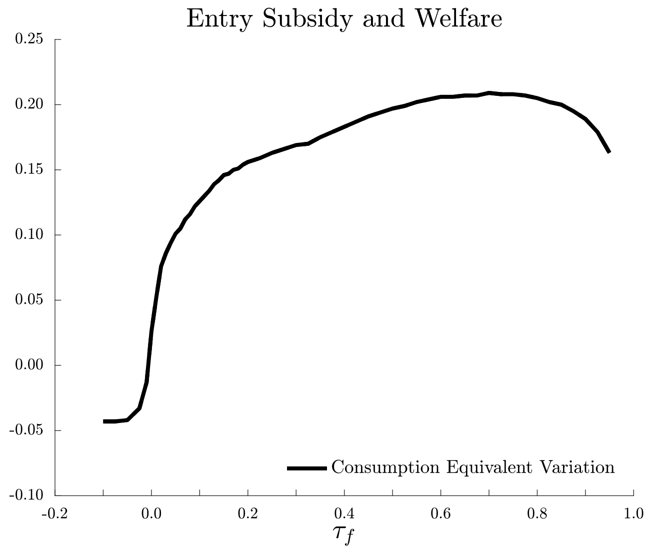


Figure A.2: Welfare: consumption equivalent gain

Note: the figure shows the welfare impact (in consumption equivalent gains) of an entry subsidy equal to a fraction τ_f of fixed costs. For each level of τ_f, we simulate the economy 100,000 times and calculate average welfare.

A.4 Regression Tables

VARIABLES	(1) Δ log emp$_{07-16}$	(2) Δ log emp$_{07-16}$	(3) Δ log emp$_{07-16}$	(4) Δ log emp$_{07-16}$
concentration$_{07}$	-0.0223***	-0.0160**	-0.0177***	-0.0178**
	(0.00667)	(0.00688)	(0.00682)	(0.00732)
log firms$_{07}$		0.00239***	0.00193***	0.00151
		(0.000705)	(0.000706)	(0.000983)
Δ log emp$_{03-07}$			0.0984***	0.0901***
			(0.0241)	(0.0247)
Observations	770	770	769	761
R-squared	0.014	0.029	0.050	0.064
Sector FE	NO	NO	NO	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.2: Change in Employment: 2007-2016

Note: the table shows the results of regressing the growth rate of sectoral employment between 2007 and 2016 on the measure of concentration in 2007. The table presents the results of progressively adding controls and, in the last column, sector fixed effects.

VARIABLES	(1) Δ log payroll$_{07-16}$	(2) Δ log payroll$_{07-16}$	(3) Δ log payroll$_{07-16}$	(4) Δ log payroll$_{07-16}$
concentration$_{07}$	-0.0231***	-0.0177**	-0.0189***	-0.0194***
	(0.00679)	(0.00702)	(0.00697)	(0.00749)
log firms$_{07}$		0.00203***	0.00164**	0.000991
		(0.000724)	(0.000725)	(0.00101)
Δ log payroll$_{03-07}$			0.0823***	0.0697***
			(0.0219)	(0.0225)
Observations	774	774	773	765
R-squared	0.015	0.025	0.043	0.054
Sector FE	NO	NO	NO	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.3: Change in Total Payroll: 2007-2016

Note: the table shows the results of regressing the growth rate of sectoral total payroll between 2007 and 2016 on the measure of concentration in 2007. The table presents the results of progressively adding controls and, in the last column, sector fixed effects.

VARIABLES	(1) Δ log firms$_{07-16}$	(2) Δ log firms$_{07-16}$	(3) Δ log firms$_{07-16}$	(4) Δ log firms$_{07-16}$
concentration$_{07}$	-0.0432***	-0.0391***	-0.0406***	-0.0231***
	(0.00608)	(0.00637)	(0.00635)	(0.00666)
log firms$_{07}$		0.00137**	0.00119*	0.00449***
		(0.000663)	(0.000661)	(0.000897)
Δ log firms$_{03-07}$			0.0881***	0.0808***
			(0.0270)	(0.0273)
Observations	791	791	791	782
R-squared	0.060	0.065	0.078	0.151
Sector FE	NO	NO	NO	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.4: Change in Number of Firms: 2007-2016

Note: the table shows the results of regressing the growth rate of the industry number of firms between 2007 and 2016 on the measure of concentration in 2007. The table presents the results of progressively adding controls and, in the last column, sector fixed effects.

VARIABLES	(1) Δlabor share$_{07-16}$	(2) Δlabor share$_{07-16}$	(3) Δlabor share$_{07-16}$	(4) Δlabor share$_{07-16}$
concentration$_{07}$	-0.0314*	-0.0319*	-0.0314*	-0.0301
	(0.0167)	(0.0168)	(0.0167)	(0.0196)
log firms$_{07}$		-0.00111	-0.00120	-0.00255
		(0.00240)	(0.00240)	(0.00335)
Δlabor share$_{03-07}$			0.169*	0.146*
			(0.0867)	(0.0871)
Observations	99	99	98	97
R-squared	0.035	0.037	0.075	0.111
Sector FE	NO	NO	NO	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.5: Change in Labor Share: 2007-2016

Note: the table shows the results of regressing the growth rate of sectoral labor share between 2007 and 2016 on the measure of concentration in 2007. The table presents the results of progressively adding controls and, in the last column, sector fixed effects.

Appendix B

Supplementary Material for

Firm Heterogeneity, Market Power and Macroeconomic Fragility

Not for Publication

B.1 General Framework

In this section we provide a taxonomy of the possible mechanisms behind multiplicity in our economy. We then provide to prove an intermediate lemma used in Proposition 1 and conclude by specifying three Remarks to exemplify the effect of heterogeneity on the fragility of the economy.

We start by noting that a necessary condition for multiplicity of steady states is that $\exists K^* : R_K(K^*) > 0$. It follows that the necessary condition can be rewritten as $\exists K^* : (\partial/\partial K)\, \Omega(\Lambda, n(\Lambda, K^*))\Phi(\Lambda, n(\Lambda, K^*))F_K(K^*, L(\Lambda, n, K^*)) > 0$. There are three main mechanisms underlying the possible locally increasing returns to capital.

Average Firm TFP In an economy with heterogeneous technologies and no love for variety, aggregate TFP can be written as a weighted average of firm-level productivities. The weights will depend on market shares. Average firm-level TFP can be increasing in K if a larger capital favors a reallocation towards more productive types.

Love for Variety In models with product differentiation, aggregate TFP typically increases in the number of available varieties. This reflects the fact that utility/welfare are themselves increasing in the number of available goods (i.e. there is *love for variety*). Take for simplicity an economy where firms operate with the same level of productivity $\gamma_{ij} = \gamma$, but with possibly different fixed costs c_{ij}. Each firm produces a differentiated good. A larger capital stock K can increase the incentives for the entry of new firms/goods, thereby making Φ (weakly) increasing in K. Examples of papers highlighting this channel as a source of multiple equilibria/steady-states include Schaal and Taschereau-Dumouchel (2019).[34]

[34] Without relying on multiple equilibria or multiple steady-states, Cooper and John (2000) and Bilbiie et al. (2012) show that a combination of imperfect competition with endogenous entry can generate endogenous amplification and persistence of aggregate fluctuations.

Market Power In models featuring imperfect competition and variable markups, changes in the number of active players can have an impact on the distribution of income across factors of production and oligopoly rents. Take for example an economy where firms have identical fixed costs $c_{ij} = c$ but possibly different productivities π_{ij}. Assume further that firms enter sequentially in reverse order of productivity. If profit levels are increasing in the aggregate capital stock k (for a given set of players), a larger capital stock will result in a larger number of firms and lower markups. Lower markups in turn translate in a higher factor share Ω. This can establish a positive relationship between Ω and k. Examples of papers highlighting this channel as a source of multiple equilibria/steady-states include Pagano (1990), Chatterjee et al. (1993), Galí and Zilibotti (1995) and Jaimovich (2007).

Let us consider some special cases. The first remark discusses a mean-preserving spread to the distribution of idiosyncratic productivities in an economy with a fixed set of producers (i.e. no adjustment along the extensive margin). More precisely, we consider special case of mean-preserving spread in the transformation is monotonically increasing away from the median in the set of active producers.[35] We then consider a marginal change in γ.

Remark B.1 (**Allocative efficiency**). *In an economy with a fixed set of active producers, a technological shift $d\lambda$ increases fragility if*

$$\Omega_\lambda(\Lambda, n)\Phi(\Lambda, n) + \Omega(\Lambda, n)\Phi_\lambda(\Lambda, n) < 0 \quad \textit{for } K = \{\mathcal{K}_n, \mathcal{K}_{n+1}\}, n \textit{ even}. \tag{81}$$

Suppose for example that we consider a mean-preserving spread to the distribution of idiosyncratic productivities. Assume further that markups and market shares are both positive functions of productivities. If market shares are an non-decreasing function of productivities, then the second term is non-negative. To see this note that a mean-preserving spread, fixing the market share distribution, implies that the average productivity increases. Additionally the positive relationship between productivities and market share implies a reallocation from low to high productivity firms, reinforcing the increase in aggregate productivity. Secondly, if markups are themselves non-decreasing in market shares, then the first term is always non-positive. This result comes from the reallocation effect. As large firms become larger, they compress output to extract higher rents. In doing so they compress the factor share. Therefore fragility increases if the anti-competitive effect (first term) dominates the efficiency gains (second term) from increasing

[35] For example, we consider going from Π to $\tilde{\Pi}_{\gamma,S_\Pi} = (1 + \gamma \cdot S_\Pi) \circ \Pi$, with $\gamma > 0$, S_Π monotonically decreasing within a row and row-wise zero-sum, and $\circ$ denoting the Hadamard product. For a similar approach see Herrendorf et al. (2000).

the dispersion of firm-level productivity. Through the lens of the taxonomy the first term in Remark B.1 represents the market power channel, while the second term is the average firm TFP channel.

We are also interested in exploring the consequences of an increase in fixed costs. To simplify the exposition, suppose that the economy only contains one industry type $(I = 1)$, that all producers have identical productivity and there is no love for variety. In that case, aggregate productivity Φ is fixed and changes in the equilibrium rental rate will only happen through the aggregate factor share Ω.

Remark B.2 (Market power). *Consider an economy with one industry type $(I = 1)$, identical producers and no love for variety. This economy will feature constant aggregate TFP, which we normalize to $\Phi(\Lambda, n) = \Phi$. Furthermore, because firms are identical, larger fixed costs only affect the aggregate factor share through changes in the mass of active firms, i.e. $\Omega_\lambda(\Lambda, n) = 0$. Therefore, in this economy a larger fixed cost generates greater fragility if*

$$\Omega_n(\Lambda, n) n_\lambda(\Lambda, k) < 0 \quad \textit{for } k = \{\mathcal{K}_n, \mathcal{K}_{n+1}\}, n \textit{ even}. \tag{82}$$

First, note that the aggregate factor share should be non-decreasing in the aggregate mass of firms, i.e. $\Omega_n(\Lambda, n) \geq 0$. As firms exit due to the higher fixed cost the surviving firms increase their market shares. In doing so they are able to increase their markups and compress the factor share. Second, the aggregate mass of firms should must be non-increasing in fixed costs $n_\lambda(\Lambda, k) \leq 0$. This effect comes from firms being unable to cover the increased fixed costs and exiting. Therefore, larger fixed costs should generate increase fragility.

We also consider an economy with a constant markups and factor shares, to highlight how changes in the mass of firms can affect aggregate TFP.

Remark B.3 (Love for variety). *Consider an economy with constant markups and aggregate factor share $\Omega(\Lambda, n) = \Omega$. A technological shift $d\lambda$ increases fragility if*

$$\Phi_\lambda(\Lambda, n) + \Phi_n(\Lambda, n) n_\lambda(\Lambda, k) < 0 \quad \textit{for } k = \{\mathcal{K}_n, \mathcal{K}_{n+1}\}, n \textit{ even}. \tag{83}$$

Suppose for example that we consider a mean-preserving spread to idiosyncratic productivities. In that case, if there is a reallocation towards more productive firms, we have $\Phi_\lambda(\Lambda, n) \geq 0$, as well as $\Phi_n(\Lambda, n) \geq 0$ (love for variety) and $n_\lambda(\Lambda, n) \leq 0$ (if less productive firms are driven out of the market). If the second effect dominates (i.e. loss in number of varieties is stronger than the increase in average technical efficiency), fragility increases.

B.2 Industry Equilibrium with $\eta = 1$

Equilibrium Price and Output Suppose that $\eta = 1$. When n firms produce, we have a system of n first order conditions

$$p\left[1-(1-\rho)\,s_j\right] = \frac{\Theta}{\gamma_j} \tag{84}$$

Dividing the first order condition of firm j by that of firm 1 we obtain

$$s_j = \frac{1}{(1-\rho)}\left\{1 - \frac{\gamma_1}{\gamma_j}\left[1-(1-\rho)\,s_1\right]\right\} \tag{85}$$

Note that

$$\sum_{k=1}^{n} s_k = 1 \Rightarrow \frac{n-(1-\rho)}{\sum\limits_{k=1}^{n}\frac{1}{\gamma_k}} = \gamma_1\left[1-(1-\rho)\,s_1\right] \tag{86}$$

Plugging the last equation into the first order condition of firm 1 we obtain

$$p = \frac{\sum\limits_{k=1}^{n}\frac{1}{\gamma_k}}{n-(1-\rho)}\Theta \tag{87}$$

Total output is hence equal to

$$y = p^{-\frac{1}{1-\rho}}Y = \left[\frac{\sum\limits_{k=1}^{n}\frac{1}{\gamma_k}}{n-(1-\rho)}\Theta\right]^{-\frac{1}{1-\rho}} Y \tag{88}$$

Market Shares Plugging the previous equation into the first order condition of firm j we have

$$s_j = \frac{1}{1-\rho}\left[1 - \frac{n-(1-\rho)}{\sum\limits_{k=1}^{n}\frac{1}{\gamma_k}}\frac{1}{\gamma_j}\right] \tag{89}$$

It is easy to verify that each firm's market share decreases in the total number of active firms. To see this, suppose that the number of firms increases from n to $n+1$. The new entrant will have a market share

$$s_{n+1} = \frac{1}{1-\rho}\left[1 - \frac{n+1-(1-\rho)}{\sum_{k=1}^{n+1}\frac{1}{\gamma_k}}\frac{1}{\gamma_{n+1}}\right] \tag{90}$$

which is non-negative provided that

$$\gamma_{n+1}\sum_{k=1}^{n+1}\frac{1}{\gamma_k} > n+1-(1-\rho) \tag{91}$$

and below one given that

$$\gamma_{n+1}\sum_{k=1}^{n+1}\frac{1}{\gamma_k} < \frac{1}{\rho}\left[n+1-(1-\rho)\right] \tag{92}$$

If we compare the market share of firm j when there n and $n+1$ firms in the market, we have

$$s_j\,|_{n+1} < s_j\,|_n \Leftrightarrow \gamma_{n+1}\sum_{k=1}^{n+1}\frac{1}{\gamma_k} > n-(1-\rho) \tag{93}$$

Note that the last condition is implied by (91).

Profits When there are n active firms, type γ_j makes production profits

$$\Pi\left(\gamma_j, n, \mathcal{F}, \Theta, Y\right) = \underbrace{\frac{1}{1-\rho}\left[1 - \frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\frac{1}{\gamma_j}\right]^2\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\right]^{\frac{\rho}{1-\rho}}}_{\equiv\Lambda\left(\gamma_j,n,\mathcal{F}\right)}\Theta^{-\frac{\rho}{1-\rho}}Y \tag{94}$$

Lemma B.1. *When $\eta = 1$, the profit function $\Pi\left(j, n_{it}, \mathcal{F}_i, X_t\right)$ satisfies*

$$1)\frac{\partial\Pi\left(j, n_{it}, \mathcal{F}_i, X_t\right)}{\partial Y_t} > 0 \qquad 2) \qquad \frac{\partial\Pi\left(j, n_{it}, \mathcal{F}_i, X_t\right)}{\partial n_{it}} < 0 \quad, \; n_{it} > j \tag{95}$$

$$3)\frac{\partial\Pi\left(j, n_{it}, \mathcal{F}_i, X_t\right)}{\partial \gamma_{ij}} > 0 \qquad 4) \qquad \frac{\partial\Pi\left(j, n_{it}, \mathcal{F}_i, X_t\right)}{\partial \gamma_{ik}} < 0 \quad, \; \forall k \neq j. \tag{96}$$

Proof of Lemma B.1. We start by showing that $\Pi(\cdot)$ increases in γ_j

$$2\left[1-\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\frac{1}{\gamma_j}\right]^{-1}\left\{-\frac{-[n-(1-\rho)]\left[-\left(\frac{1}{\gamma_j}\right)^2\right]}{\left(\sum_{k=1}^{n}\frac{1}{\gamma_k}\right)^2}\frac{1}{\gamma_j}+\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\left(\frac{1}{\gamma_j}\right)^2\right\}+ \tag{97}$$

$$\frac{\rho}{1-\rho}\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\right]^{-1}\frac{-[n-(1-\rho)]\left[-\left(\frac{1}{\gamma_j}\right)^2\right]}{\left(\sum_{k=1}^{n}\frac{1}{\gamma_k}\right)^2}>0 \tag{98}$$

$$\Leftrightarrow 2\left[1-\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\frac{1}{\gamma_j}\right]^{-1}\left(\sum_{k\neq j}\frac{1}{\gamma_k}\right)+\frac{\rho}{1-\rho}\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\right]^{-1}>0 \tag{99}$$

To prove points (ii) and (iii) it suffices to show that $\Lambda(\cdot)$ is decreasing in $[n-(1-\rho)]/\left[\sum_{k=1}^{n}\frac{1}{\gamma_k}\right]$

$$2\left[1-\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\frac{1}{\gamma_j}\right]^{-1}\left(-\frac{1}{\gamma_j}\right)+\frac{\rho}{1-\rho}\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_k}}\right]^{-1}<0 \tag{100}$$

$$\tag{101}$$

$$\Leftrightarrow \gamma_j\sum_{k=1}^{n}\frac{1}{\gamma_k}<\frac{2-\rho}{\rho}[n-(1-\rho)] \tag{102}$$

The last condition is implied by (92). ∎

B.3 Derivations: General Equilibrium

B.3.1 Aggregate TFP

Aggregate TFP is given by

$$\Phi\left(\mathbf{\Gamma},\mathbf{N}_t\right)=\left[\sum_{i=1}^{I}\left(\sum_{j=1}^{n_{it}}\omega_{ijt}^{\eta}\right)^{\frac{\rho}{\eta}}\right]^{\frac{1}{\rho}}\left(\sum_{i=1}^{I}\sum_{j=1}^{n_{it}}\frac{\omega_{ijt}}{\tau_{ijt}}\right)^{-1}, \tag{103}$$

where

$$\omega_{ijt}:=\left[\sum_{k=1}^{n_{it}}\left(\frac{\mu_{ikt}}{\tau_{ikt}}\right)^{\frac{\eta}{1-\eta}}\right]^{\frac{\eta-\rho}{\eta}\frac{1}{1-\rho}}\left(\frac{\tau_{ijt}}{\mu_{ijt}}\right)^{\frac{1}{1-\eta}}. \tag{104}$$

B.3.2 Factor Prices and Factor Shares

We can aggregate firms' best responses, given by equation (12), to find an expression for the aggregate factor cost index. Given a $(I\times N)$ matrix of productivity draws $\mathbf{A}_t$ and a vector of active firms $\mathbf{N}_t\equiv\{n_{it}\}_{i=1}^{I}$, the equilibrium factor cost index is equal to

$$\Theta\left(\mathbf{A}_t,\mathbf{N}_t\right)=\left\{\sum_{i=1}^{I}\left[\sum_{j=1}^{n_{it}}\left(\frac{\tau_{ijt}}{\mu_{ijt}}\right)^{\frac{\eta}{1-\eta}}\right]^{\frac{1-\eta}{\eta}\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}}. \tag{105}$$

The aggregate factor share $\Omega\left(\cdot\right)=\left(W_t\,L_t+R_t\,K_t\right)/Y_t$ is equal to

$$\Omega\left(\mathbf{A}_t,\mathbf{N}_t\right)=\frac{\Theta\left(\mathbf{A}_t,\mathbf{N}_t\right)}{\Phi\left(\mathbf{A}_t,\mathbf{N}_t\right)}. \tag{106}$$

B.3.3 Asymmetric Equilibrium

When

$$\overline{K}\left(\mathcal{F},n\right)<K<\underline{K}\left(\mathcal{F},n+1\right) \tag{107}$$

there will be an asymmetric equilibrium at time $t+1$: some industries will contain n firms, whereas some industries will contain $n+1$ firms. The fraction of industries with $n+1$ will be pinned down by a zero profit condition for the marginal entrant in an industry with $n+1$ firms

$$\Lambda\left(\mathcal{F},\gamma_{n+1},n+1\right)\Theta^{-\frac{\rho}{1-\rho}}Y=c_i \tag{108}$$

The equilibrium is characterized by 4 variables: the fraction of the industries with $n+1$ firms (η), aggregate output (Y), aggregate productivity (Φ) and the aggregate cost index (Θ). These 4 variables are pinned down by the following 4 equations

$$Y=\Phi\left[\left(1-\alpha\right)\Theta\right]^{\frac{1-\alpha}{\nu+\alpha}}K^{\alpha\frac{1+\nu}{\nu+\alpha}} \tag{109}$$

$$\Phi=\frac{\left\{(1-\eta)\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_{1k}}}\right]^{\frac{\rho}{1-\rho}}+\eta\left[\frac{n+1-(1-\rho)}{\sum_{k=1}^{n+1}\frac{1}{\gamma_{2k}}}\right]^{\frac{\rho}{1-\rho}}\right\}^{\frac{1}{\rho}}}{(1-\eta)\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_{1k}}}\right]^{\frac{1}{1-\rho}}\left(\sum_{k=1}^{n}\frac{s_{1k}}{\gamma_{1k}}\right)+\eta\left[\frac{n+1-(1-\rho)}{\sum_{k=1}^{n+1}\frac{1}{\gamma_{2k}}}\right]^{\frac{1}{1-\rho}}\left(\sum_{k=1}^{n+1}\frac{s_{2k}}{\gamma_{2k}}\right)} \tag{110}$$

$$\Theta=\left\{(1-\eta)\left[\frac{n-(1-\rho)}{\sum_{k=1}^{n}\frac{1}{\gamma_{k}}}\right]^{\frac{\rho}{1-\rho}}+\eta\left[\frac{n+1-(1-\rho)}{\sum_{k=1}^{n+1}\frac{1}{\gamma_{k}}}\right]^{\frac{\rho}{1-\rho}}\right\}^{\frac{1-\rho}{\rho}} \tag{111}$$

$$\Lambda\left(\mathcal{F},\gamma_{n+1},n+1\right)\Theta^{-\frac{\rho}{1-\rho}}Y=c_i \tag{112}$$

s_{1k} is the market share of firm k in an industry with n firms, whereas s_{2k} is the market share of firm k in an industry with $n+1$ firms. They are defined in Appendix B.2.

B.4 The Baseline Model

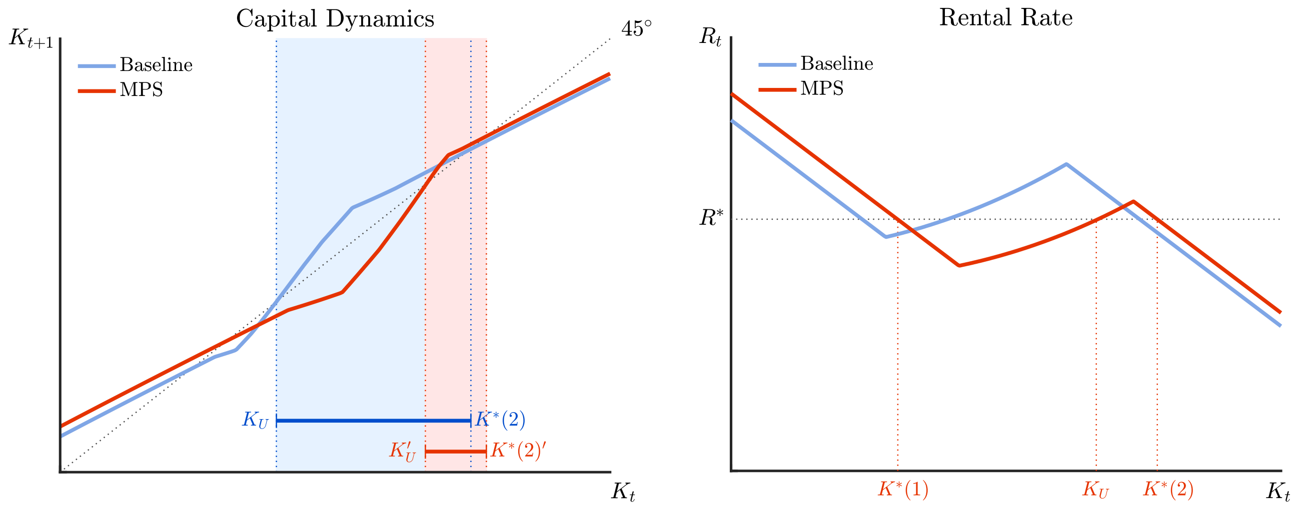


Figure B.1: Law of Motion and Rental Rate Map
This example features two stable steady states and an unstable one. We use $\psi = 1, \rho = 3/4, \eta = 1, \alpha = 1/3, \delta = 1, \nu = 2/5$ and $c_i = 0.015$.

B.5 The Quantitative Model

B.5.1 Calibration

Steady-State We perform two different calibrations of our model – to match the average level of markups and its dispersion in 1985 and in 2007. We need to calibrate three technology parameters: the Pareto tail λ, the fixed production cost c and the fraction of industries with zero fixed cost f_{comp}.

We specify a grid of possible candidates for λ, c and f_{comp}. We also specify a grid with values for the aggregate capital stock K. We then compute the aggregate equilibrium for each parameter combination (λ, c, f_{comp}) and for each value K.[36] We start by assuming that all firms are active, so that there are N firms in each of the I industries. We compute the aggregate equilibrium using equations (103) and (20). We then compute the profits net of the fixed cost that each firm makes

$$\left(p_{ijt} - \frac{\Theta_t}{\tau_{ijt}}\right) y_{ijt} - c_i$$

and identify the firm with the largest negative value. We exclude this firm and recompute the

[36] Aggregate TFP e^{z_t} is assumed to be constant and equal to one.

aggregate equilibrium. We repeat this iterative procedure until all firms have non-negative profits (net of the fixed production cost). For most parameter combinations, our model admits a unique equilibrium. However, if equilibrium multiplicity arises, this algorithm allows us to consistently select the equilibrium that features the largest number of firms.

For each triplet (λ, c, f_{comp}), we then have the general equilibrium computed for all possible capital values. The steady-state(s) of our economy correspond to the value(s) of K for with the rental rate R_t is equal to $\frac{1}{\beta} - (1 - \delta)$.

Given our interpretation that the US economy was in a competition regime in both 1985 and 2007, we compute model moments in the highest steady-state.

Data Definitions For the sales weighted-average markup, we use the series computed by De Loecker et al. (2020). The authors calculate price-cost markups for the universe of public firms, using data from COMPUSTAT. The markup of a firm j in a 2-digit NAICS sector s at time t is calculated as

$$\mu_{sjt} = \xi_{st} \cdot \frac{\text{sale}_{sjt}}{\text{cogs}_{sjt}}$$

where ξ_{st} is the elasticity of sales to the total variable input bundle, sale_{sjt} is sales and cogs_{sjt} is the cost of the goods sold, which measures total variable costs.

B.5.2 Solution Algorithm for the Dynamic Optimization Problem

We now explain the algorithm we use for the dynamic optimization problem of the representative household. We take the calibrated parameters (λ, c) and form a grid for aggregate capital with $n_K = 70$ points. This grid is centered around the highest steady-state K_H^{ss}, with a lower-bound $0.5 \times K_H^{ss}$ and upper bound $1.5 \times K_H^{ss}$. We also form a grid for (log) aggregate TFP, z. We use Tauchen's algorithm with $n_z = 11$ points, autocorrelation parameter ϕ_Z and standard deviation for the innovations σ_ε (the last two parameters are calibrated, as explained in the main text). We compute the aggregate equilibrium for each value of K and z.

We next compute a numerical approximation for the household policy function, by iterating on the Euler equation.

B.5.3 Business Cycle Moments

	Output	Consumption	Investment	Hours	TFP
			Correlation with Output		
Data: 1947-2019	1.00	0.95	0.76	0.67	0.71
Model: 1985 calibration	1.00	1.00	0.96	1.00	0.34
Model: 2007 calibration	1.00	1.00	0.94	1.00	0.53
			Standard Deviation Relative to Output		
Data: 1947-2019	1.00	0.90	2.04	0.98	0.95
Model: 1985 calibration	1.00	0.95	1.19	0.78	0.09
Model: 2007 calibration	1.00	0.95	1.27	0.78	0.10

Table A.1: Business Cycle Moments. All variables are in logs. Data variables are in per capita terms (except TFP) and in deviation from a linear trend computed over 1947-2007.

Table A.1 shows some business cycle moments for our two calibrated economies, as well as their data counterparts. To be consistent with our interpretation that the US economy transitioned to a lower steady-state after 2008, all data variables are in deviation from a linear trend computed over 1947-2007. This fact explains the large empirical correlation between consumption and output. Comparing our two calibrated economies, we see that both economies display the same correlations of consumption and hours with output. The 2007 economy displays, however, a significantly lower correlation of investment with output.

Aggregate Productivity

Average Firm Level TFP

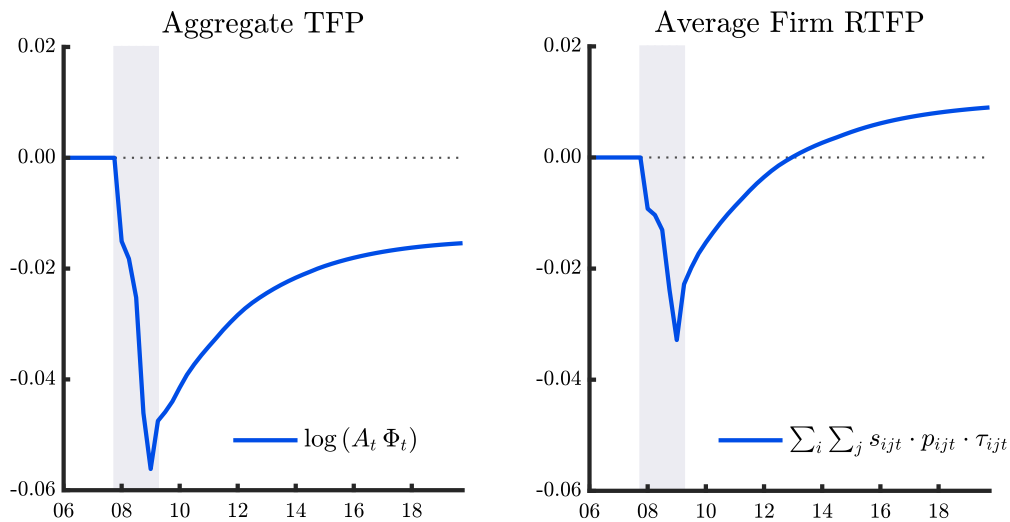


Figure B.1: Aggregate TFP versus Average Firm Level TFP

Note: The left panel shows aggregate TFP. The right panel shows a sales-weighted average of firm level revenue TFP $p_{ijt} \cdot \tau_{ijt}$.

Figure B.1 reports a sales-weighted average of firm level revenue TFP. A similar pattern emerges if one uses physical TFP instead.

Dispersion in Industry Output

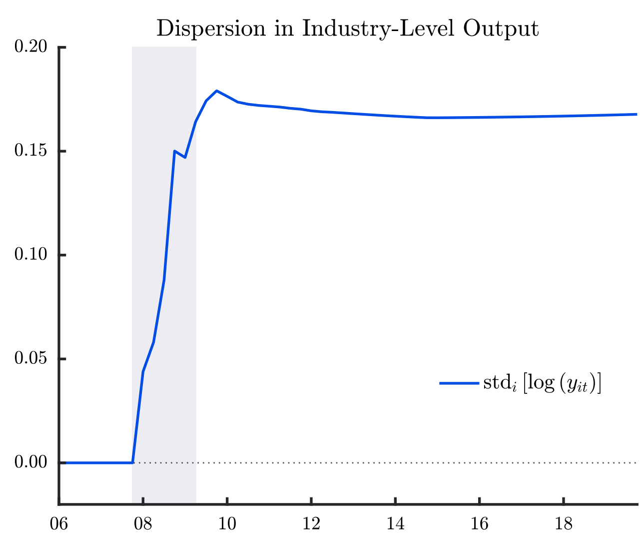
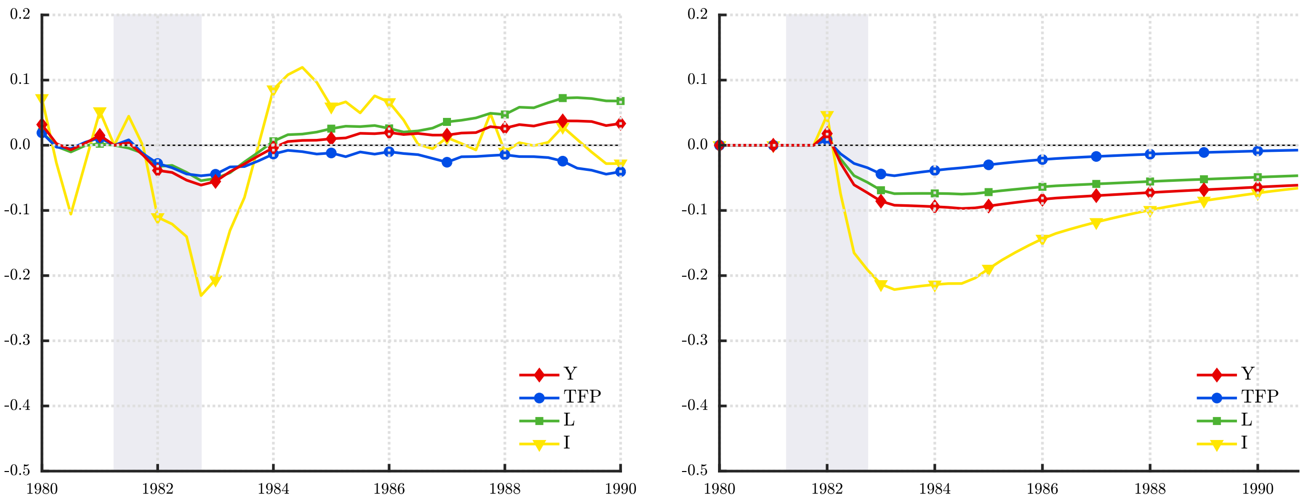


Figure B.1: Dispersion in $\log(y_{it})$

B.6 The 1981-1982 Recession

The response in the 1985 economy



(a) 1981-1982 recession (data) (b) The 1981-1982 shock in the 1985 model

Figure B.1: The 1981-1982 recession

The response in the 2007 economy

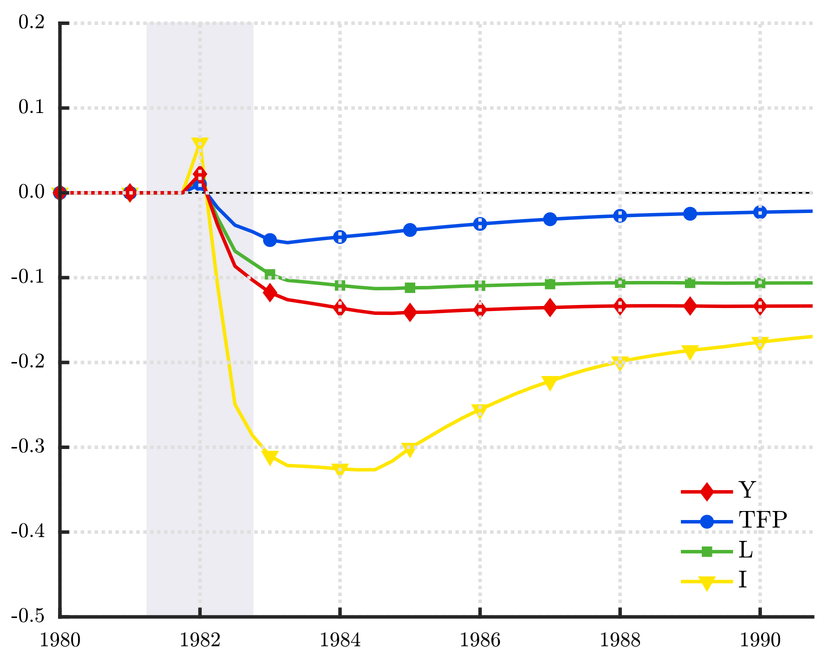


Figure B.1: The 1981-1982 shock in the 2007 model

B.7 Number of Firms per Sector

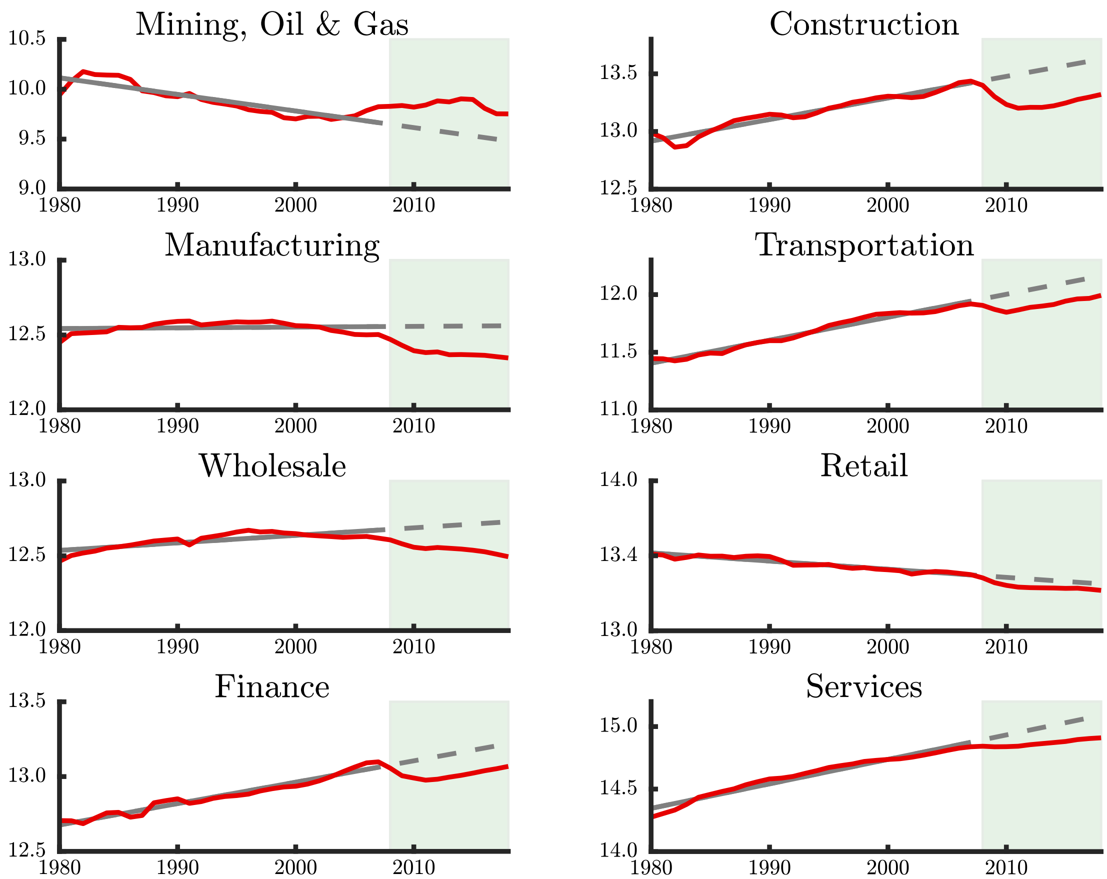


Figure B.1: **Number of Firms per Sector: 1980-2018**

Each panel shows the number of firms with at least one employee in each sector (in logs). For each series, the dashed grey line shows a linear trend computed over the 1980-2007 period. Data is from the US Business Dynamics Statistics

B.8 Fixed Costs

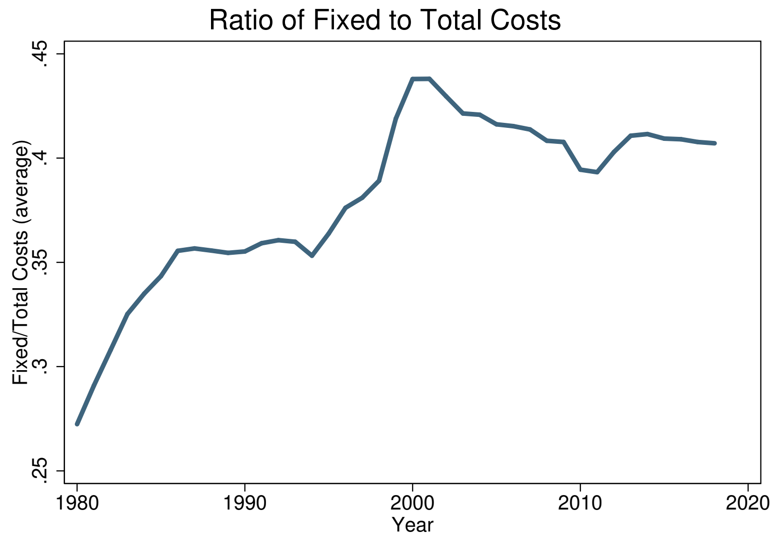


Figure B.1: Ratio of fixed to total costs

This figure shows the average ratio of fixed to total costs for COMPUSTAT firms. Following Gorodnichenko and Weber (2016), we define fixed costs as the sum of 'Selling, General and Administrative Expenses' (COMPUSTAT item XSGA), 'Advertising Expenses' (Compustat item XAD) and 'R&D Expenditures' (Compustat item XRD). Total costs are the sum of fixed costs and variable costs, where the latter correspond to the 'Cost of Goods Sold' (Compustat item COGS).

B.9 Robustness: Different Elasticities of Substitution

Calibration

Parameters not reported are as in the baseline calibration (Table 1).

Description	Parameter	Value	Source/Target
Between-industry ES	σ_I	1.2	De Loecker et al. (2021)
Within-industry ES	σ_G	5.75	De Loecker et al. (2021)
Calibrated Parameters: 2007			
Standard deviation γ	λ	0.280	Sales-weighted average markup
Fixed cost	c	0.0034	Average ratio fixed/total costs
Fraction of industries with $c_i > 0$	f	0.080	Emp share concentrated industries
Persistence of z_t	ρ_z	0.900	Autocorrelation of log Y
Standard deviation of ε_t	σ_ε	0.0035	Standard deviation of log Y
Calibrated Parameters: 1985			
Standard deviation γ	λ	0.140	Sales-weighted average markup
Fixed cost	c	0.0004	Average ratio fixed/total costs
Fraction of industries with $c_i > 0$	f	0.110	Emp share concentrated industries

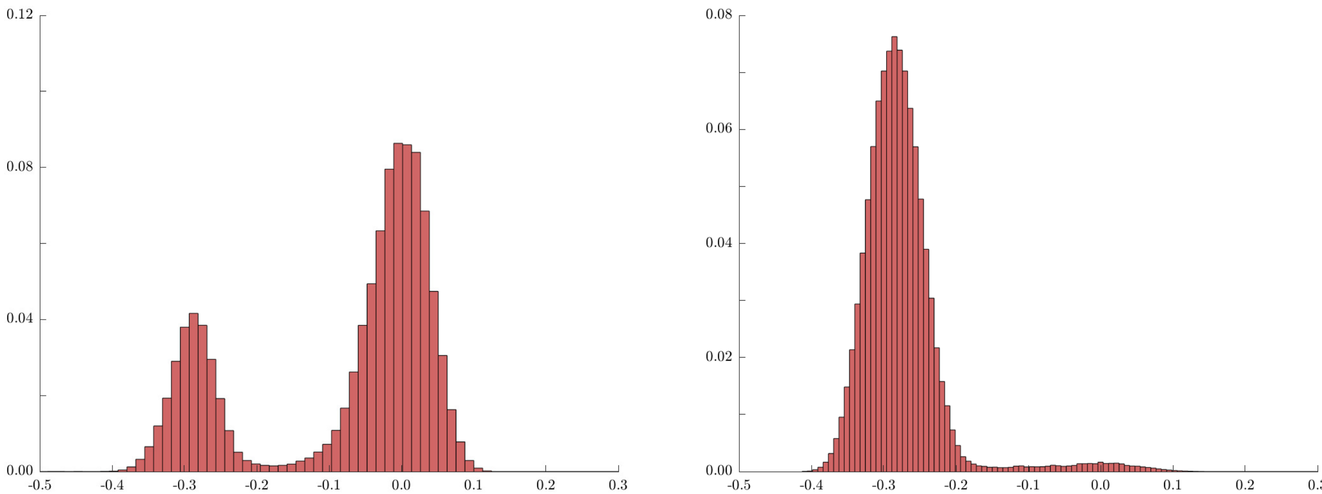
Table A.2: Parameter Values

	1985		2007	
	Data	Model	Data	Model
Sales-weighted average markup	1.27	1.25	1.46	1.43
Average fixed to total cost ratio	0.343	0.370	0.414	0.479
Employment share in *concentrated* industries	-	0.058	0.063	0.037
Autocorrelation log GDP	0.978*	0.953	0.978*	0.956
Standard deviation log GDP	0.061*	0.040	0.061*	0.059

*computed over 1947:Q1-2019:Q4

Table A.3: Targeted moments and model counterparts

Ergodic Distributions

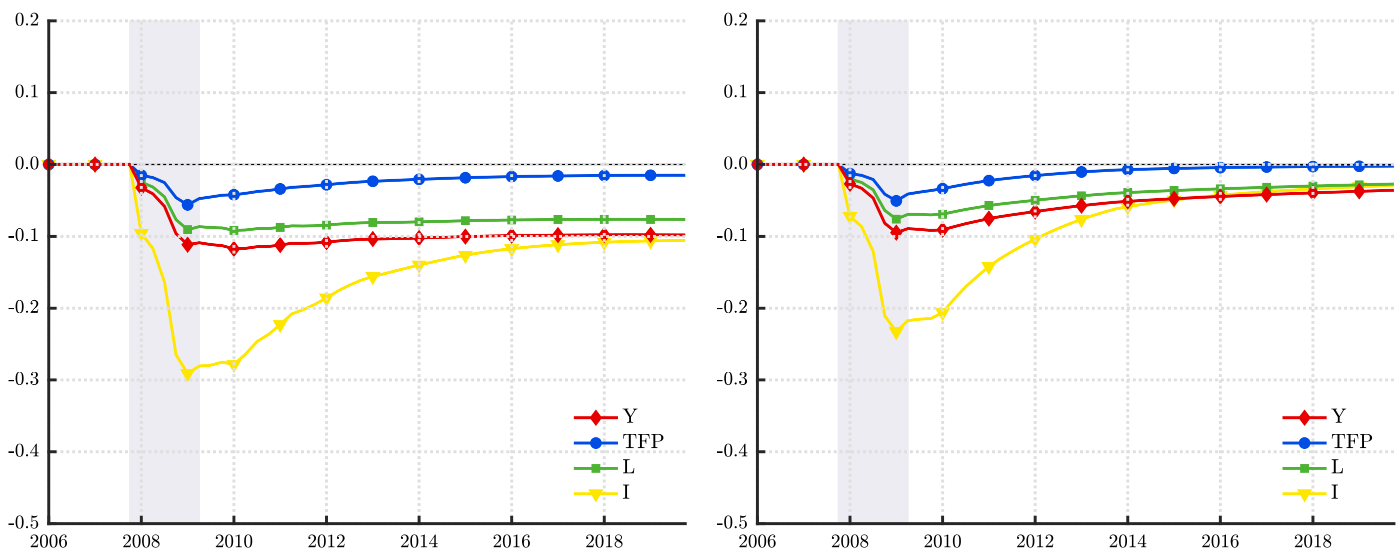


(a) Output Distribution: 1985 (b) Output Distribution: 2007

Figure B.1: Ergodic distribution of output

This figure shows the distribution of log output for the 1985 and the 2007 economies. We simulate each economy for 10,000,000 periods and plot output in deviation from the high steady state.

The 2008 Crisis



(a) 2007 Model
This figures replicates Figure 7(b)

(b) 1985 Model
This figures replicates Figure 8

Figure B.1: The *great recession* and its aftermath

B.10 Robustness: Variable Fixed Costs

We assume that, each period, a fixed amount c_f of firms' output is lost

$$c_f = k_c^{\alpha} l_c^{1-\alpha}$$

Given these assumptions, firms need to pay a per per period fixed cost

$$\Theta_t \cdot c_f$$

where Θ_t is the factor price index.

Denoting by L_{yt} and K_{yt} the aggregate stocks of labor and capital used in the production, we have the following market clearing conditions for labor and capital

$$L_t = L_{yt} + N_t^c \cdot l_c$$
$$K_t = K_{yt} + N_t^c \cdot k_c$$

where N^c denotes the number of firms incurring c_f. Note that the optimal mix of l_c and k_c chosen by each individual firm satisfies

$$\frac{k_c}{l_c} = \frac{K_{yt}}{L_{yt}}$$

Calibration

Parameters not reported are as in the baseline calibration (Table 1).

Description	Parameter	Value	Source/Target
Calibrated Parameters: 2007			
Standard deviation γ	λ	0.300	Sales-weighted average markup
Fixed cost	c	0.0007	Average ratio fixed/total costs
Fraction of industries with $c_i > 0$	f	0.150	Emp share concentrated industries
Persistence of z_t	ρ_z	0.920	Autocorrelation of log Y
Standard deviation of ε_t	σ_ε	0.0035	Standard deviation of log Y
Calibrated Parameters: 1985			
Standard deviation γ	λ	0.160	Sales-weighted average markup
Fixed cost	c	0.0004	Average ratio fixed/total costs
Fraction of industries with $c_i > 0$	f	0.130	Emp share concentrated industries

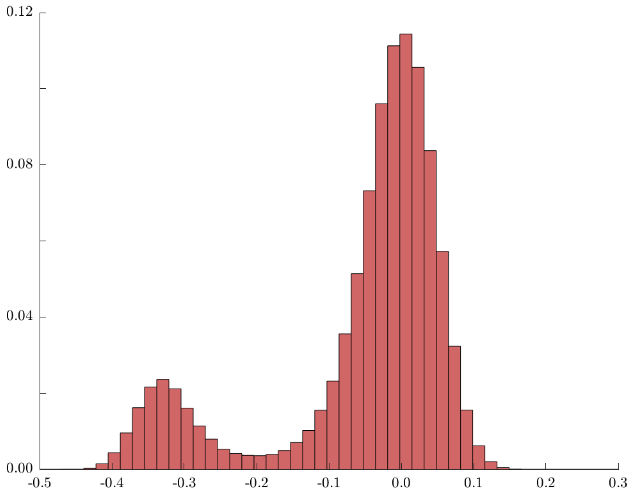
Table A.4: Parameter Values

	1985		2007	
	Data	Model	Data	Model
Sales-weighted average markup	1.27	1.28	1.46	1.47
Average fixed to total cost ratio	0.343	0.377	0.414	0.431
Employment share in *concentrated* industries	-	0.068	0.063	0.066
Autocorrelation log GDP	0.978*	0.980	0.978*	0.972
Standard deviation log GDP	0.061*	0.087	0.061*	0.059

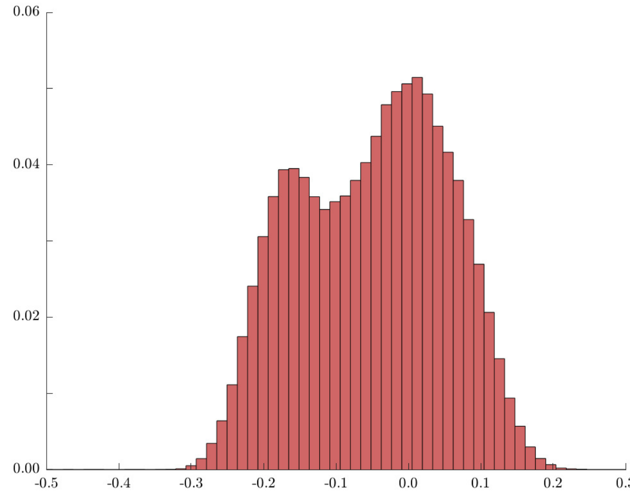
*computed over 1947:Q1-2019:Q4

Table A.5: Targeted moments and model counterparts

Ergodic Distributions



(a) Output Distribution: 1985

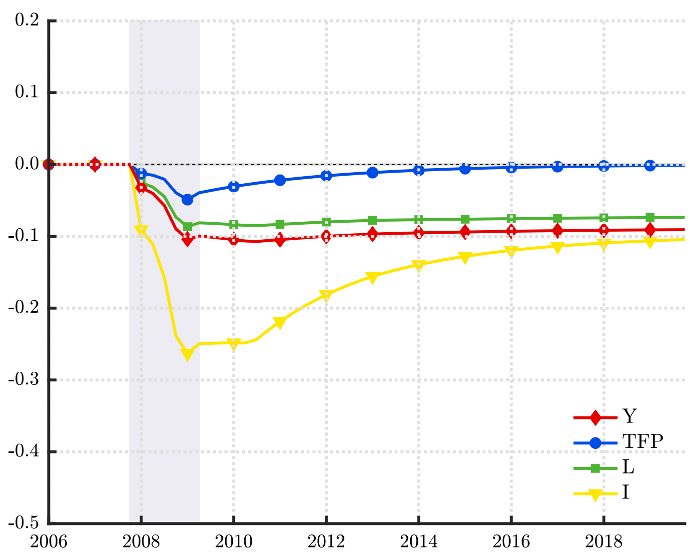


(b) Output Distribution: 2007

Figure B.1: Ergodic distribution of output

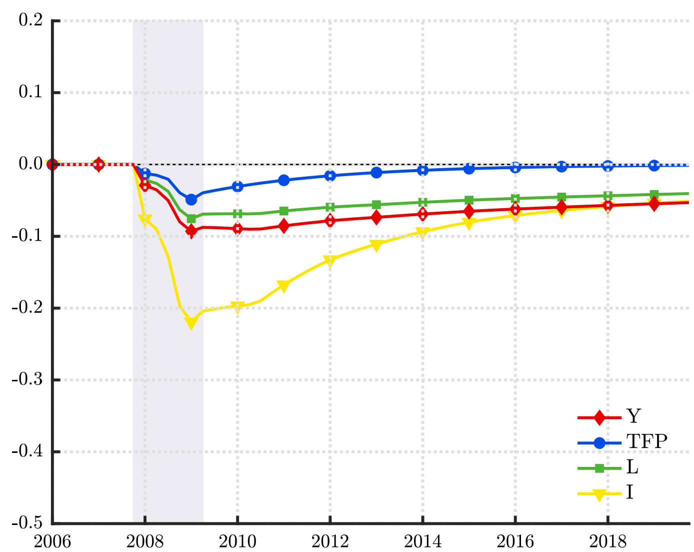
This figure shows the distribution of log output for the 1985 and the 2007 economies. We simulate each economy for 10,000,000 periods and plot output in deviation from the high steady state.

The 2008 Crisis



(a) 2007 Model

This figures replicates Figure 7(b)



(b) 1985 Model

This figures replicates Figure 8

Figure B.1: The *great recession* and its aftermath