Australian National University | Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Equilibrium Indeterminacy When Only Some Firms Are Scalable

CAMA Working Paper 50/2026
July 2026

Oscar Pavlov
University of Tasmania
Centre for Applied Macroeconomic Analysis, ANU

Mark Weder
Aarhus University
Centre for Applied Macroeconomic Analysis, ANU

Abstract

We study a macroeconomy in which firms differ in the strength of scale economies. A subset of firms operates under increasing returns to scale, which endogenously leads them to become larger, capture greater market shares, and exert more goods-market power than peers. The framework can simultaneously account for observed employment shares and markups of top firms. We characterize how heterogeneity in productivity and scalability affects aggregate outcomes and the propagation of shocks. We show that for empirically relevant parameterizations, the heterogeneity gives rise to equilibrium indeterminacy, allowing for self-fulfilling fluctuations. Using an estimated version of the model, we quantify the extent to which non-fundamental disturbances ("animal spirits") contribute to business cycle fluctuations in key macroeconomic aggregates.

Keywords

scalable technologies, business cycles, indeterminacy, firm heterogeneity

JEL Classification

E32, E71

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

Equilibrium Indeterminacy When Only Some Firms Are Scalable

Oscar Pavlov*
University of Tasmania, CAMA

Mark Weder†
Aarhus University, CAMA

June 26, 2026

Abstract

We study a macroeconomy in which firms differ in the strength of scale economies. A subset of firms operates under increasing returns to scale, which endogenously leads them to become larger, capture greater market shares, and exert more goods-market power than peers. The framework can simultaneously account for observed employment shares and markups of top firms. We characterize how heterogeneity in productivity and scalability affects aggregate outcomes and the propagation of shocks. We show that for empirically relevant parameterizations, the heterogeneity gives rise to equilibrium indeterminacy, allowing for self-fulfilling fluctuations. Using an estimated version of the model, we quantify the extent to which non-fundamental disturbances ("animal spirits") contribute to business cycle fluctuations in key macroeconomic aggregates.

Keywords: Scalable technologies, Business cycles, Indeterminacy, Firm heterogeneity.

JEL codes: E32, E71.

*Corresponding author. Tasmanian School of Business and Economics, University of Tasmania and CAMA, Hobart TAS 7001, Australia, oscar.pavlov@utas.edu.au.

†A substantial part of this research was untertaken while Weder visited the *Universitat Jaume I*. Thanks heaps for the hospitality and for providing an excellent research environment.

1 Introduction

Micro-level data reveal substantial dispersion in firm outcomes. Even within narrowly defined industries, firms differ markedly in output and employment. A well-established empirical regularity is that larger firms tend, on average, to exhibit higher measured productivity. Following the contributions of Lucas (1978) and Hopenhayn (1992), this relationship has commonly been interpreted as reflecting differences in total factor productivity (TFP). While TFP is observed to increase with firm size, this association appears to weaken among the largest firms, yet returns to scale keep rising systematically with size, indicating that larger firms are characterized by greater scalability (Hubmer et al., 2025). What is more, they find that scalability rises sharply at the very top end of firms. This pattern is echoed by Ma et al. (2025) who highlight the adoption of scalable technologies as key for understanding the rise of the largest American companies. Motivated by these observations, this paper proposes a framework in which in returns to scale at the firm level give rise to dispersion in size and in market power. In particular, we build an economy in which only a subset of firms operates technologies with increasing returns to scale. These firms coexist with less scalable counterparts and, in equilibrium, attain greater relative size in terms of output and market share, while exercising higher power in goods markets. We study how such heterogeneity shapes macroeconomic dynamics. In particular, we show that plausible configurations of scalability lead to firm dispersions that give rise to equilibrium indeterminacy, thereby opening the door to endogenous business cycles driven by self-fulfilling beliefs. Within this framework, we assess the quantitative contribution of animal spirits for aggregate fluctuations.

Firm size and employment are highly concentrated. Among the more than five million operating firms in the United States, the vast majority are of very small scale, typically employing between one and four workers. At the same time, the economy contains more than ten thousand firms with employment exceeding one thousand workers, and these firms account for a disproportionate share of aggregate employment: for example, the employment share of the top one percent of firms ranked by employment hovers around 60 percent (Ma et al., 2026). A similar pattern emerges in the distribution of sales. A relatively small set of firms accounts for the bulk of aggregate market activity: the top one percent of firms ranked by sales generate over 70 percent of total sales. Taken together, these patterns suggest that the largest firms derive their scale primarily through productivity rather than through a disproportionately large expansion in employment.

Empirical evidence on price–cost markups has long informed macroeconomic modeling. More recently, De Loecker et al. (2020) document a substantial increase in average markups among U.S. publicly traded firms, rising from approximately 1.2 in 1980 to about 1.6 in 2016, with broadly similar patterns across sectors. The markup specifics regarding magnitudes and trends remain debated (Edmond et al., 2023). Benkard et al. (2026) revisit De Loecker et al. (2020) and highlight the role of sample selection, showing that markup increases are more modest in the full sample and that the median markup may even have declined slightly. Nevertheless, they document a pronounced increase in markup dispersion, driven primarily by firms in the upper tail. Markups for top-percentile firms exceed 1.5, consistent with De Loecker and Eeckhout (2026). This concentration mirrors the "winner-takes-most" pattern emphasized by Autor et al. (2020), whereby a very small subset of firms captures a disproportionate share of output and profits. Taken together, the evidence suggests that aggregate markup trends are largely shaped by large firms with high markups. Understanding the origins and relation of size and markup dispersions appears therefore central.

Hubmer et al. (2025) suggest substantial returns to scale heterogeneity and that such differences in scalability constitute an economically significant dimension of firm size. They find that returns to scale rise systematically with size, in fact, scalability rises sharply for the very top firms. It is this empirical regularity that provides the principal motivation for our modelling strategy. We propose a tractable general equilibrium economy in which heterogeneity in scalability gives rise to differences in firm size and to time-varying market power at the hand of a Kimball (1995) demand system. Concretely, only a subset of firms has access to scalable technologies and, as a consequence, commands large market shares. A realistically calibrated version of the model accounts for the empirically observed heterogeneities in employment and market power. The model further implies that the cyclical behavior of markups is heterogeneous across firms: for large firms, markups are procyclical, reflecting both their increasing market shares during expansions and the curvature of demand. Smaller firms set countercyclical markups. These implications align with firm-level evidence (Burstein et al., 2025, Callebaut and Peersman, 2026).

Our macroeconomic framework admits endogenous fluctuations arising from self-fulfilling expectations. This mechanism is related to the discussion on equilibrium indeterminacy typified by Farmer and Guo (1994) and Galí (1994), in which increasing returns and countercyclical markups generate sunspot equilibria. In contrast to these contributions, the present framework incorporates heterogeneous firms and demonstrates that indeterminacy may arise even when increasing returns are confined to a very small subset of firms—namely, those that are highly scalable.

Furthermore, reallocations of resources between small to large firms reduces the degree of increasing returns required for indeterminacy, suggesting a novel mechanism for multiplicity and also a potentially destabilizing effect of heterogeneity. As a consequence heterogeneous economies may be more susceptible to such endogenous fluctuations than their counterparts with less firm dispersion. Lastly, indeterminacy arises despite the presence of procyclical markups even in the aggregate. In this respect, the model avoids potential conflicts with empirical evidence on aggregate markup cyclicality that remains at best inconclusive (Nekarda and Ramey, 2021).

Finally, we quantify the role of animal spirits using a full-information Bayesian estimation of the model.[1] In line with the findings of medium-scale dynamic stochastic general equilibrium models, such as Justiniano et al. (2011), technology and investment shocks continue to account for a significant portion of U.S. business cycles. However, the disturbances that generate most output fluctuations are demand shocks, reminiscent of Galí and Rabanal (2005). Lastly, we find that non-fundamental shocks—capturing shifts in beliefs—make a non-negligible contribution, particularly with respect to investment.

The remainder of the paper is organized as follows. Section 2 sets out the basic model and its principal mechanisms. Section 3 establishes the conditions under which local indeterminacy arises. Section 4 extends the framework to allow for variable factor utilization. Section 5 presents the quantitative results and Section 6 concludes.

2 Economy

The economy is populated by a unit mass of firms with each producing differentiated intermediate goods using capital and labor. Firms differ in their production technologies, with heterogeneity in scalability—captured by differences in internal returns to scale—serving as a central source of dispersion, in line with the evidence in Hubmer et al. (2025). A perfectly competitive final-goods sector aggregates intermediate inputs into a homogeneous good used for consumption and investment. Households supply labor and capital services on competitive factor markets. Time evolves in discrete steps.[2]

[1] Lubik and Schorfheide (2004), Pintus et al. (2022), Hirose et al. (2023), among others, estimate models with indeterminacy using a similar approach.

[2] The Appendix details the economy's elements and its solution.

2.1 Technology and market structure

Final output Y_t is produced by a representative, perfectly competitive firm that aggregates a continuum of differentiated intermediate inputs $y_{i,t}$ indexed by $i \in [0,1]$. The aggregation technology is given by

$$\int_0^1 G\left(\frac{y_{i,t}}{Y_t}\right) di = 1 \tag{1}$$

in which the functional form of $G(.)$ follows the specification in Kimball (1995), as implemented by Kurozumi and Van Zandweghe (2022):

$$G\left(\frac{y_{i,t}}{Y_t}\right) = \frac{\gamma}{(1+\epsilon)(\gamma-1)}\left((1+\epsilon)\frac{y_{i,t}}{Y_t} - \epsilon\right)^{\frac{\gamma-1}{\gamma}} + 1 - \frac{\gamma}{(1+\epsilon)(\gamma-1)} \tag{2}$$

with $\gamma \equiv \theta(1+\epsilon)$, $\epsilon \leq 0$ and $\theta > 1$ denoting the elasticity of substitution. The constant elasticity of substitution (CES) case obtains in the limit $\epsilon = 0$, in which the aggregator reduces to

$$Y_t = \left(\int_0^1 y_{i,t}^{\frac{\theta-1}{\theta}} di\right)^{\frac{\theta}{\theta-1}}.$$

The final-good firm solves a static cost-minimization problem subject to the aggregation constraint. The associated first-order conditions imply the demand

$$\frac{y_{i,t}}{Y_t} = \frac{\left(\frac{p_{i,t}}{P_t}\right)^{-\gamma} v_t^{\gamma} + \epsilon}{1+\epsilon} \tag{3}$$

in which $p_{i,t}$ is firm i's price, P_t stands for the aggregate price index

$$P_t v_t = \left(\int_0^1 p_{i,t}^{1-\gamma} di\right)^{\frac{1}{1-\gamma}} \tag{4}$$

and v_t represents the multiplier on the final good firm's constraint to (1)

$$v_t = 1 + \epsilon - \epsilon \int_0^1 \frac{p_{i,t}}{P_t} di. \tag{5}$$

The intermediate goods are provided by a measure one of firms that employ both capital $k_{i,t}$ and labor $h_{i,t}$ as inputs. Factor services are bought on competitive markets at prices r_t and w_t. A fraction m of firms ("large" firms or m-firms, indexed by $i \in (n,1]$), operates a production function

$$y_{i,t} = z\left(k_{i,t}^{\alpha} h_{i,t}^{1-\alpha}\right)^{\omega} - \phi_m \qquad 0 < \alpha < 1,\ \omega > 0 \tag{6}$$

in which $z > 0$ is relative TFP and $\phi_m > 0$ stands for firm-level fixed costs that help to pin down steady state profits. The remaining firms ("ordinary" firms or n-firms, indexed by $i \in [0,n]$) produce according to

$$y_{i,t} = \left(k_{i,t}^{a} h_{i,t}^{1-a}\right)^{\eta} - \phi_n \qquad 0 < a < 1,\ \eta > 0. \tag{7}$$

Firm-level fixed cost $\phi_n > 0$ are calibrated such that these firms emerge with zero profit in the steady state.

The asymmetry in technologies aims to captures a key empirical regularity. Namely, as suggested by Hubmer et al. (2025), while big firms are often more productive simply because they operate with higher TFP, they are also more scalable especially at the top of the distribution. We would calibrate these two features at the hands of $z > 1$ and $\omega > \eta$. In most of our analysis we will set $\eta = 1$ and, as such, $\omega > 1$ can also describe relative scalability. The number of firms remains fixed to concentrate on effects that are unrelated to possible endogenous productivity changes arising from firm entry and exit. Intermediate good producers maximize profits subject to their respective demands and production technologies (6) and (7). The profit-maximizing condition for each firm is the familiar equality between marginal revenue and marginal costs $\mathbf{MC}_{i,t}$, and the markup $\mu_{i,t}$ takes on the form

$$\mu_{i,t} = \frac{p_{i,t}}{\mathbf{MC}_{i,t}} = \frac{\gamma\left(\frac{y_{i,t}}{Y_t} - \frac{\epsilon}{1+\epsilon}\right)}{(\gamma-1)\frac{y_{i,t}}{Y_t} - \gamma\frac{\epsilon}{1+\epsilon}}. \tag{8}$$

When $\epsilon < 0$, markups move with with relative demands, while $\epsilon = 0$ brings about constant markups $\mu_{i,t} = \mu = \theta/(\theta-1)$.

We focus on symmetric equilibria within each firm type, thus, $(y_{m,t}, p_{m,t})$ and $(y_{n,t}, p_{n,t})$ denote allocations for large and ordinary firms. For large firms, marginal cost is given by

$$\mathbf{MC}_{m,t} = \frac{1}{\omega}\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}w_t^{1-\alpha}r_t^{\alpha}\left(y_{m,t}+\phi_m\right)^{\frac{1}{\omega}-1}z^{\frac{-1}{\omega}}$$

which is decreasing in output whenever $\omega > 1$. Differences in returns to scale, ω and η, capital intensities, α and a, and relative total factor productivity z cause asymmetries in prices and quantities. In addition, the Kimball demand system brings to pass asymmetries in markup levels as well as in markup elasticities that would not occur had we assumed the more standard CES aggregator. As m-firms' outputs expand, two opposing effects on relative prices come into play. First, marginal costs fall that allow m-firms to charge lower prices, which creates even more demand for their commodities. On the other hand, as large firms' market shares rise at the expense of small firms, their markups $\mu_{m,t}$ increase, putting upward pressure on their prices again. In all cases that we consider, the markup-effect is dominated. You can also see this from $m-$firms' demands, which can be written as

$$\frac{y_{m,t}}{Y_t} = \frac{\left(m + n\left(\frac{p_{n,t}}{p_{m,t}}\right)^{1-\gamma}\right)^{\frac{\gamma}{1-\gamma}} + \epsilon}{1+\epsilon}$$

where big firms' relative size $y_{m,t}/Y_t$ follows the price ratio $p_{n,t}/p_{m,t}$.

2.2 Households

Households are characterized by a representative agent who chooses sequences of consumption C_t and hours worked H_t to maximize discounted lifetime utility

$$\sum_{t=0}^{\infty} \beta^t \left(\ln C_t - \frac{\upsilon}{1+\chi} H_t^{1+\chi} \right) \qquad 0 < \beta < 1, \upsilon > 0, \chi \geq 0$$

in which β is the discount rate, υ denotes the disutility of working and χ relates to people's labour supply elasticity. The agent owns all firms and receives their profits Π_t. The period-budget is constrained by

$$w_t H_t + r_t K_t + \Pi_t \geq I_t + C_t$$

in which I_t is investment that adds to the aggregate capital stock K_t

$$K_{t+1} = (1-\delta) K_t + I_t \qquad 0 < \delta < 1.$$

The first-order conditions from the agent's maximization problem comprise of the labor supply

$$\upsilon H_t^{\chi} C_t = w_t,$$

the Euler equation

$$\frac{C_{t+1}}{C_t} = \beta \left(r_{t+1} + 1 - \delta \right)$$

along with the usual transversality condition.

3 Steady state and dynamics

To visualize the effect of scalable technology on firm dispersion in general equilibrium, we propose the following numerical steady state exercise. We fix the number of large firms at ten percent and vary both z and ω. We keep η at one, assume the CES-bundler subclass to abstract from the effects that arise from markup variations and set $\theta = 5$. Other parameters will be discussed further below. Figure 1 projects $m-$firm isoquants for $z-\omega-$combinations that yield market shares of these firms from 70 to 90 percent. You can see that higher market shares arise by means of higher z, larger ω or both. Going south to north in the figure illustrates that higher market share require overproportional increases in z. For example, if there were no scale effects of large firms, $\omega = 1$, going from 70 percent to 80 and to 90 percent market shares would require for z to increase by 14 percent and then another 23 percent. Keeping z fixed on the other hand, say at one, the required increases from ω would be 2.7 and then 3.4 percent. Thus, relatively modest changes in return

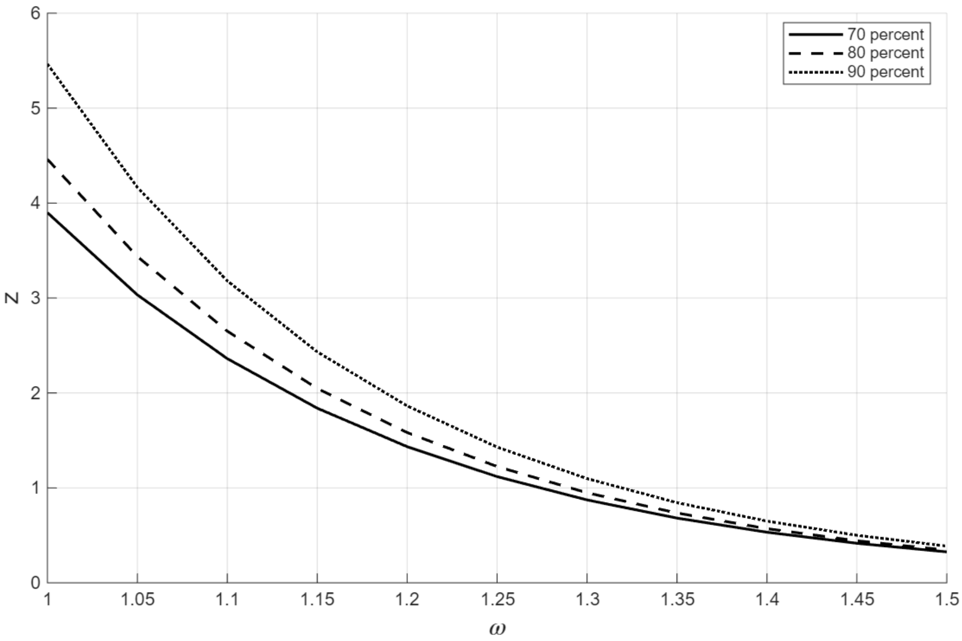

Figure 1: Isosales line in $z - \omega$-space.

to scales generate considerable size dispersions. We believe this aspect of relative importance of scalable technology is consistent with Hubmer et al. (2025) and as such it suggests support for our modelling of heterogeneity. As will be shown below, reasonable values for ω are compatible with realistic markup and firm dispersions, in particular for explaining the very top firms.

But first, we analyze the economy's local dynamics. The equilibrium conditions are log-linearized around the steady state and the dynamical system boils down to

$$\begin{bmatrix} \widehat{K}_{t+1} \\ \widehat{C}_{t+1} \end{bmatrix} = \mathbf{J} \begin{bmatrix} \widehat{K}_t \\ \widehat{C}_t \end{bmatrix}.$$

Hatted variables denote percentage deviations from their steady state values and $\mathbf{J}$ is the 2×2 Jacobian matrix of partial derivatives. Consumption is a non-predetermined variable and capital is predetermined, thus, a unique equilibrium requires one eigenvalue of $\mathbf{J}$ inside and one outside the unit circle. Equilibrium indeterminacy, and the potential presence of animal spirits, demands both roots of $\mathbf{J}$ within the unit circle. The current model structure does not lend itself well for studying the economy's existence and its dynamics algebraically, except perhaps for special cases. Thus, most of the study will be done numerically.

We solve for the steady state and local dynamics of the economy through a sequential calibration strategy. Given the two steady state markups, μ_m and μ_n, and the two Kimball parameters, θ and ϵ, the model jointly determines the mass of

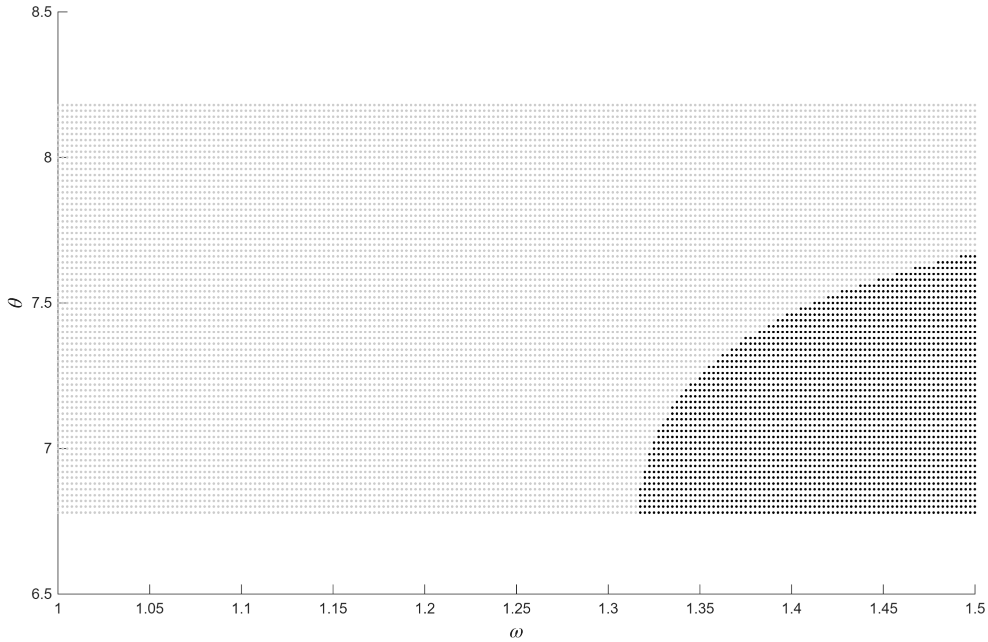

Figure 2: Indeterminacy in $\theta - \omega$-space. Dark area denotes indeterminacy.

large firms m as well as their market share in aggregate output $\frac{mp_m y_m}{Y}$. The Kimball bundler's θ must satisfy

$$\frac{\mu_m}{\mu_m - 1} = \theta^{\min} < \theta < \theta^{\max} = \frac{\mu_n}{\mu_n - 1}$$

to ensure an interior firm distribution. To facilitate comparison with the indeterminacy literature, particularly Farmer and Guo (1994), we set $a = \alpha = 0.3$, $\beta = 0.99$ and $\delta = 0.025$. Reflecting on Edmond et al. (2023) and Benkard et al. (2026), we fix markups of the top firms at $\mu_m = 1.5$ and those of the ordinary firms at $\mu_n = 1.05$. The market share of large firms is fixed at 70 percent (Ma et al., 2026). Then, the market-share-weighted average markup of the economy stands at 1.38 which is in between suggested values in De Loecker et al. (2020) and Edmond et al. (2023). Finally we set the fixed costs of n-firms such that their steady state profits are zero and $\omega > \eta = 1$ recreates the asymmetry in Hubmer et al. (2025). Lastly, after having calibrated markups and market shares, the choice of a feasible value of θ locks in a unique value of ϵ.

Figure 2 plots the effect of increasing returns of large firms ω on the dynamics of the economy. The two white regions involving $\theta \notin \left[\theta^{\min}, \theta^{\max}\right]$ imply non-existence: for example if $\theta > \theta^{\max}$ the number of m-firms becomes strictly negative, m is zero at $\theta^{\max}$ and m equals one at $\theta^{\min}$. Conditional on existence, lighter regions denote determinate equilibria and darker regions denote indeterminacy. As you can see from the figure, indeterminacy requires $\omega > \omega^{\min} = 1.32$. While the mechanism for inde-

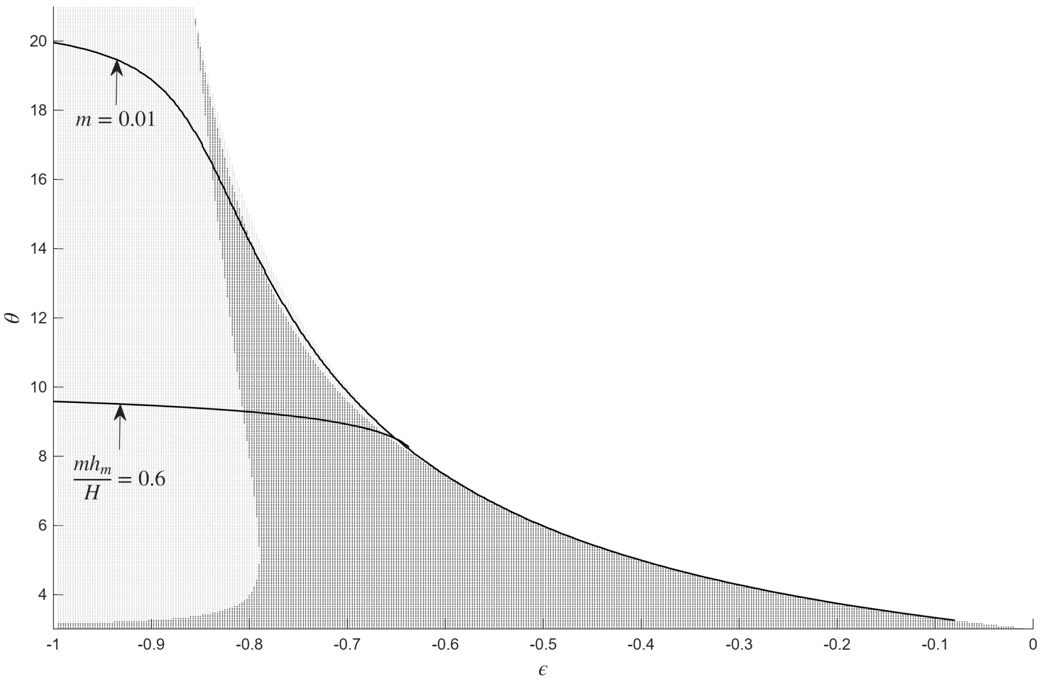

Figure 3: Indeterminacy in $\theta - \epsilon$-space.

terminacy is related to Farmer and Guo (1994), this minimum degree is significantly lower than their minimum returns to scale of 1.5 (corresponding to the limiting case $m = 1$ in the present model). At first glance, the finding appears difficult to reconcile with Herrendorf et al. (2000), who argue that heterogeneity eliminates indeterminacy. Our interpretation is instead that compositional effects amplify the conventional increasing-returns mechanism. In response to an expansionary impulse, resources are reallocated toward large firms, thereby endogenously increasing the economy's effective scale economies. As a consequence, the wage-hours locus becomes upwardly sloping at lower values of ω than in the representative-firm environment of Farmer and Guo (1994). The mechanism can be further appreciated when looking at the effects of the two Kimball parameters next.

The influence of θ and ϵ on indeterminacy is zoned in Figure 3. The calibration is as above but ω is now fixed at 1.5, which is its maximum possible value to ensure non-negative profits. The darker area indicates indeterminate equilibria and these arise for low values of θ and ϵ. The figure also reports two loci corresponding to parameter combinations that imply $m = 0.01$ and a 60 percent employment share for large firms, consistent with the evidence in Ma et al. (2026).

The indeterminacy pattern can be understood as follows. For parameter constellations along the horizontal axis, i.e., at the lower bound just above $\theta^{\min}$, the market share of m-firms approaches 100 percent. That is, as $\theta \to \frac{\mu_m}{\mu_m - 1} = \theta^{\min}$. The opposite situation arises at the upper bound of feasible θ. There, as $\theta \to \frac{\mu_n}{\mu_n - 1} = \theta^{\max}$ and because of the inconsequential market share of m-firms, the economy effectively operates under constant returns to scale. Phrased alternatively, at both $\theta^{\min}$ and $\theta^{\max}$, firm heterogeneity disappears and as such any repercussion of heterogeneity dissipates. Consequently, the Kimball parameter ϵ has no notable effect on the markup and model dynamics. For example, at the bottom range close to $\theta^{\min}$, the total economy effectively operates under increasing returns: it is indeterminate and morphs into the Farmer-Guo setup. The ϵ parameter therefore becomes asymptotically irrelevant in this region.

A more reasonable parametric position is denoted by the mostly-horizontal isoquant that keeps m-firms' employment share at 60 percent. Here, as $\epsilon < 0$, markups correlate with large firms' economic activity: $\mu_{m,t}$ and the employment-weighted average markup

$$\mu_t = \frac{m h_{m,t}}{H_t} \mu_{m,t} + \frac{n h_{n,t}}{H_t} \mu_{n,t}$$

are procyclical. On its own that counters indeterminacy. This mechanism is visible when moving leftward along the isoquant: the economy transitions from indeterminacy to determinacy as the stabilizing influence of procyclical markups on aggregate dynamics becomes progressively stronger. The isoquant sheds further light on the steady state properties of the economy: it is relatively flat which indicates that it is θ that primarily determines the employment and market shares. Bringing ϵ closer to zero lowers the number of big firms and increases their individual sales. At the boundary, as the isoquant approaches the non-feasibility zone, m-firms become very large but, at the same time, $m \to 0$. At this value of ϵ the markups of big firms, while not completely constant, are at their minimum elasticity.

Indeterminacy therefore emerges from the interaction between endogenous productivity effects and compositional reallocation induced by firm heterogeneity in technology, markups, and pricing behavior. To illustrate the mechanism, consider an economy initially in steady state that is hit by an extrinsic increase in optimism regarding future economic conditions. The resulting expansion reallocates resources toward more productive firms, namely the large firms in the present framework. This reallocation lowers the relative prices charged by large firms because their marginal costs decline as production expands, thereby validating the movement of inputs toward these firms. Consequently, an expansionary episode endogenously generates increasing returns beyond the mechanism emphasized in Farmer and Guo (1994).

As market shares adjust, the two markups move in opposite directions. Consistent with Burstein et al. (2025) and Callebaut and Peersman (2026), the markups of large firms rise during booms, whereas the markups of ordinary firms are countercyclical. Importantly, the model does not rely on aggregate countercyclical markups, in contrast to the mechanism in Galí (1994). Indeed, indeterminacy in the present environment is fully consistent with the evidence in Nekarda and Ramey (2020), since aggregate markups are procyclical overall. The dampening effects associated with rising average markups are dominated by the reallocation of resources from small to large firms, which amplifies aggregate productivity and sustains self-fulfilling fluctuations.

4 Variable capital utilization

With the aim of reducing requirements on scale economies and markups, we extend the model allowing for variable capital utilization, following Greenwood et al. (1988). Capital services are defined as the product of the capital stock and an endogenous utilization rate. Higher utilization raises effective capital input but increases depreciation. The household chooses utilization optimally, trading off higher current returns against future capital loss.

Capital services that are used to produce intermediate goods are denoted by $\kappa_{i,t} \equiv U_t k_{i,t}$ in which U_t is the utilization rate that determines the intensity of operation for a given stock. Thus, m-firms produce with technology

$$y_{i,t} = z\left(\kappa_{i,t}^{\alpha} h_{i,t}^{1-\alpha}\right)^{\omega} - \phi_m$$

and similar adjustments are made for ordinary firms. The intensity of capital use affects capital's depreciation as

$$\delta_t = \frac{1}{\vartheta} U_t^{\vartheta} \qquad \vartheta > 1$$

and the rate of utilization is determined by the representative agent, for whom the intertemporal budget constraint is now

$$K_{t+1} = (1 - \delta_t)K_t + w_t H_t + r_t U_t K_t + \Pi_t - C_t.$$

The utility maximization problem delivers the Euler equation

$$\frac{C_{t+1}}{C_t} = \beta\left(r_{t+1} U_{t+1} + 1 - \delta_{t+1}\right)$$

and the optimal rate of capital utilization

$$r_t = U_t^{\vartheta - 1}$$

in which $\vartheta = (1/\beta - 1 + \delta)/\delta$ comes from the the steady state conditions.

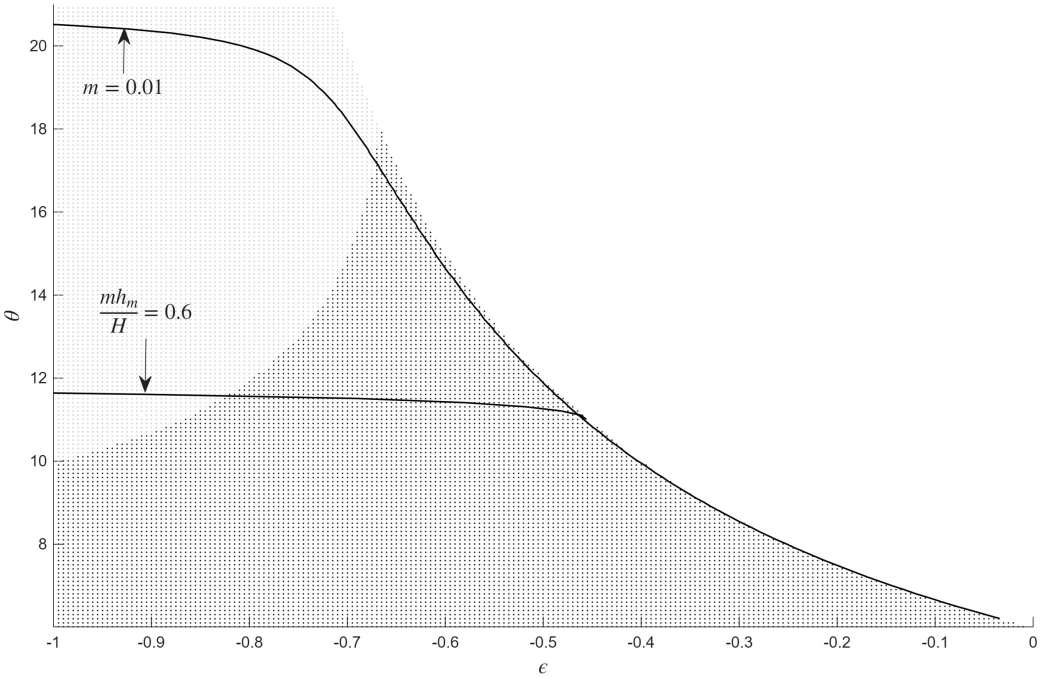

Figure 4: Indeterminacy in $\theta - \epsilon$-space (variable capital utilization)

We retain the calibration of the previous section but now lower the steady state markup of large firms to $\mu_m = 1.2$ and reduce their scalability parameter to $\omega = 1.15$. The latter value is consistent with the estimated gap in returns to scale between top and bottom firms reported by Hubmer et al. (2025). Quantitatively, this degree of scalability is well aligned with their evidence for firms in the top one percent of the size distribution, firms operating large establishment networks, high markup firms and also if incorporating tangible capital.[3] These characteristics closely correspond to the class of dominant firms to which we calibrate the model economy. Figure 4 shows that the qualitative role of the Kimball parameters remains similar to the benchmark case with constant capital utilization. However, the degree of increasing returns and the level of markups required to generate indeterminacy are now substantially lower and comfortably within or even below empirically plausible ranges.[4]

[3] See Figures A.8 and A.9 in Hubmer et al. (2025).

[4] In fact, if we further reduce small firms' markups to $\mu_n = 1.01$, essentially towards perfect competition and therefore increasing the markup gap between the two markups, $\omega^{\min}$ is close to 1.07 which is significantly lower than Wen (1998).

5 Estimation

Thus far, we have established that scalability among large firms, and their interaction with smaller producers, may generate macroeconomic instability. This environment admits a role for non-fundamental fluctuations—so-called animal spirits—in driving business cycle dynamics. We now assess their quantitative relevance in conjunction with a range of fundamental disturbances. To this end, the baseline framework is extended to incorporate exogenous growth, aggregate supply and demand shocks, and external habit formation in consumption.

5.1 Shocks

We augment the baseline environment with a set of aggregate disturbances to both supply and demand. On the supply side, we introduce technological progress A_t

$$\ln A_t = \ln A_{t-1} + \ln \zeta_t$$

in which the component ζ_t evolves according

$$\ln \zeta_t = (1 - \psi_A) \ln \zeta + \psi_A \ln \zeta_{t-1} + \varepsilon_t^A$$

where $0 \leq \psi_A < 1$ and ε_t^A is an i.i.d. innovation with variance σ_A^2. The other fundamental shocks follow similar AR(1) processes. Thus, m-firm technology becomes

$$y_{m,t} = zA_t \left(\kappa_{i,t}^{\alpha} h_{i,t}^{1-\alpha}\right)^{\omega} - \phi_{m,t}$$

and n-firm technology is

$$y_{n,t} = A_t^{\tau} \left(\kappa_{i,t}^{a} h_{i,t}^{1-a}\right)^{\eta} - \phi_{n,t}.$$

Here $\tau = 1/(1 - \omega\alpha)$ to allow for a balanced growth path under different returns to scale technologies. Growing variables are detrended by A_t^{τ}, for example, fixed costs grow with technological progress, and the detrended version is $\tilde{\phi}_m \equiv \phi_{m,t}/A_t^{\tau}$. Investment-specific technological change comes about by marginal efficiency of investment ξ_t shocks which enter through capital accumulation

$$K_{t+1} = (1 - \delta_t)K_t + \xi_t I_t.$$

Justiniano et al. (2011) interpret such disturbances as fluctuations in financial conditions. On the demand side, we add a preference shock Δ_t that shifts the marginal utility of consumption as in Christiano (1988). Household period-preferences are then given by

$$\Delta_t \ln(C_t - bC_{t-1}) - \frac{\nu}{1+\chi} H_t^{1+\chi}$$

in which $0 \leq b < 1$ captures external habit formation. Beyond its literal interpretation as a taste shifter, Δ_t can be viewed more broadly as capturing movements in the labor wedge. Another demand disturbance arises from government expenditure shocks financed through lump-sum taxation. Government spending evolves around a stochastic trend defined by

$$A_{G,t} = A_{G,t-1}^{\psi_{AG}} A_{t-1}^{1-\psi_{AG}}$$

in which ψ_{AG} governs the smoothness of the trend relative to the trend in output. Detrended government spending is then $\tilde{G}_t \equiv G_t / A_{G,t}^{\tau}$. In addition to these fundamental disturbances, we introduce a non-fundamental "animal spirits" shock. This shock is modeled as an expectational error in output that is orthogonal to fundamentals. Under indeterminacy, equilibrium outcomes are not uniquely pinned down by fundamentals alone. We therefore represent output dynamics as

$$\widehat{Y}_t = E_{t-1}\widehat{Y}_t + \Omega_A \varepsilon_t^A + \Omega_\xi \varepsilon_t^\xi + \Omega_\Delta \varepsilon_t^\Delta + \Omega_G \varepsilon_t^G + \varepsilon_t^S$$

in which the coefficients Ω_i capture the transmission of fundamental shocks, and ε_t^S is an i.i.d. disturbance with variance σ_S^2. This formulation captures fluctuations driven by extrinsic shifts in expectations that are unrelated to any underlying economic fundamentals.[5]

5.2 Bayesian estimation

We estimate the model using full-information Bayesian methods applied to U.S. quarterly data covering 1990:I to 2025:IV, as this period aligns with the documented rise in market power. However, we omit the 2020:II-III quarters as our small-scale and linear model is not suitable to deal with the COVID-19 lockdowns. The observable variables consist of real per capita growth rates of output, consumption, investment, and government spending, together with the logarithm of per capita hours worked and a measure of credit spreads.

To identify investment-specific shocks, we follow the empirical strategy of relating the marginal efficiency of investment to observed credit spreads. Specifically, we use the spread between BAA corporate bond yields and long-term Treasury yields and impose the measurement equation

$$\text{spread}_t = \varkappa \widehat{\xi}_t \qquad \varkappa < 0.$$

[5] Our results are not sensitive to the choice of expectation error and remain robust with Bianchi and Nicolò's (2021) approach of solving and estimating linear rational expectations models under indeterminacy.

so that improvements in investment efficiency correspond to tighter spreads. The measurement system links observed growth rates to their model counterparts as

$$\begin{bmatrix} 100\ln(Y_t/Y_{t-1}) \\ 100\ln(C_t/C_{t-1}) \\ 100\ln(I_t/I_{t-1}) \\ 100\ln(G_t/G_{t-1}) \\ 100\,(\ln H_t/H) \\ spread_t \end{bmatrix} = \begin{bmatrix} \widehat{Y}_t - \widehat{Y}_{t-1} + \tau\widehat{\zeta}_t \\ \widehat{C}_t - \widehat{C}_{t-1} + \tau\widehat{\zeta}_t \\ \widehat{I}_t - \widehat{I}_{t-1} + \tau\widehat{\zeta}_t \\ \widehat{G}_t - \widehat{G}_{t-1} + \tau\left(\hat{g}_t - \hat{g}_{t-1} + \widehat{\zeta}_t\right) \\ \widehat{H}_t \\ \varkappa\widehat{\xi}_t \end{bmatrix} + \begin{bmatrix} \tau\overline{\zeta} \\ \tau\overline{\zeta} \\ \tau\overline{\zeta} \\ \tau\overline{\zeta} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_t^{m.e.} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

in which $\overline{\zeta} = 100(\zeta - 1)$, $\tau\overline{\zeta} = 0.36$ is equal to the quarterly growth rate of real GDP in the data, $g_t = A_{G,t}/A_t = g_{t-1}{}^{\psi_{AG}}/\zeta_t$, $\varepsilon_t^{m.e.}$ is a measurement error restricted to account for not more than ten percent of output growth and H stands for the average hours worked over the sample period.

We consider an economy subject to six exogenous disturbances: four of a fundamental nature, together with a non-fundamental (sunspot) disturbance and a measurement error. This matches the number of observables and, consequently, avoids excess shocks that would limit the economic interpretation of the estimated shocks (Pagan and Robinson, 2022).

Several parameters are calibrated prior to estimation to target post-1990 regularities, including the government expenditure share at 19 percent. We capture heterogeneity as follows. Following Ma et al. (2026) we assume that the top one percent of firms, $m = 0.01$, account for 60 percent of aggregate employment. We assign a markup of 1.5 to the top firms and 1.1 for the rest of firms (De Loecker and Eeckhout, 2026). Then, the employment-weighted average markup equals 1.34. Thus, despite its simple two-firm structure, the model aligns with Edmond et al. (2023). Together, this calibration implies Kimball parameters at $\theta = 5.06$ and $\epsilon = -0.41$, as well as a superelasticity $-\theta\epsilon$ of 2.06, which pairs well with Beck and Lein (2020). The labor (capital) share is 63 (27) percent and the steady state profit share is 10 percent, a value that sits within the range suggested by Barkai (2020) and Hasenzagl and Pérez (2026). The market share of the top firms at 64 percent is at the lower end of Ma et al. (2025). Yet, given the simple model of dispersions, it illustrates well how large firms are quantitatively important not merely because they employ a disproportionate quantity of labor, but also because they operate at a larger scale conditional on inputs.

Table 1 reports the posterior estimates. Of prime interest to us, the parameter ω underpins the key amplification mechanism in the model operating through endogenous increasing returns to scale, and it plays a central role in making way for endogenous fluctuations. This scalability degree is estimated with a posterior mean

Table 1: Prior and posterior distributions

Prior					Posterior	
Name	Range	Density	Mean	Std. Dev.	Mean	90% Interval
ω	R^{+}	Normal	1.15	0.03	1.24	[1.23 1.26]
b	[0,1)	Beta	0.5	0.1	0.50	[0.42,0.58]
ψ_A	[0,1)	Beta	0.5	0.2	0.00	[0.00,0.01]
ψ_ξ	[0,1)	Beta	0.5	0.2	0.83	[0.77,0.89]
ψ_Δ	[0,1)	Beta	0.5	0.2	0.94	[0.91,0.96]
ψ_G	[0,1)	Beta	0.5	0.2	0.99	[0.99,0.99]
ψ_{AG}	[0,1)	Beta	0.5	0.2	0.99	[0.99,0.99]
σ_S	R^{+}	Inverse Gamma	0.1	Inf	0.28	[0.25,0.32]
σ_A	R^{+}	Inverse Gamma	0.1	Inf	0.55	[0.50,0.61]
σ_ξ	R^{+}	Inverse Gamma	0.1	Inf	0.07	[0.04,0.10]
σ_Δ	R^{+}	Inverse Gamma	0.1	Inf	1.05	[0.89,1.22]
σ_G	R^{+}	Inverse Gamma	0.1	Inf	0.82	[0.74,0.90]
$\sigma^{m.e.}$	$[0, 0.19]$	Uniform	0.1	0.05	0.19	[0.19,0.19]
Ω_A	[-3,3]	Uniform	0	1.73	-1.12	[-1.27,-0.98]
Ω_ξ	[-3,3]	Uniform	0	1.73	1.92	[1.15,2.92]
Ω_Δ	[-3,3]	Uniform	0	1.73	0.40	[0.30,0.50]
Ω_G	[-3,3]	Uniform	0	1.73	0.20	[0.14,0.26]
$\varkappa$	[-20,0]	Uniform	-10	5.77	-4.66	[-6.59,-2.71]

The table presents the prior and posterior distributions for model parameters and shocks under indeterminacy. Standard deviations are in percent terms.

of 1.24.[6] This value implies a returns to scale higher than the baseline in Hubmer et al. (2025). Yet, such value is in line for the kind of firms we consider here. Hubmer et al. (2025) find values around 1.20 for firms with characteristics typical of the very largest firms.

Table 2 compares second moments in the data with those implied by the model evaluated at the posterior mean. Despite its parsimonious structure, the model provides a reasonable account of U.S. business cycle dynamics. It overstates the volatility of investment, while understating fluctuations in hours worked. Correlations with output are well replicated. In addition, the model generates realistic autocorrelation patterns, owing to its internal propagation mechanism, without relying on the range of frictions typically introduced in medium-scale models.

Table 3 presents the forecast error variance decomposition. Supply-side disturbances—technology and investment shocks—together account for around a third of aggregate fluctuations. While investment shocks explain a large share of investment dynamics, their overall importance is smaller than in some earlier studies such as Justiniano

[6]Our model features only two productivity levels, yet, it still produces significant productivity dispersion. Cunningham et al. (2023) present micro evidence on dispersion in establishment-level productivity with a standard deviation of the log total factor productivity of establishments across detailed US manufacturing industries is 0.46. This standard deviation is 0.57 in our model (sales weighted).

Table 2: Business cycle dynamics

	Data			Model		
x	σ_x	$\rho(x, \ln(Y_t/Y_{t-1}))$	ACF	σ_x	$\rho(x, \ln(Y_t/Y_{t-1}))$	ACF
$\ln(Y_t/Y_{t-1})$	0.62	1	0.26	0.83	1	0.44
$\ln(C_t/C_{t-1})$	0.55	0.72	0.41	0.74	0.64	0.48
$\ln(I_t/I_{t-1})$	1.61	0.66	0.59	2.99	0.78	0.61
$\ln(G_t/G_{t-1})$	0.79	0.24	0.22	0.83	0.20	0.00
$\ln(H_t/H)$	5.89	0.13	0.99	4.81	0.10	0.99
spread_t	0.57	-0.53	0.86	0.57	-0.23	0.83

Business cycle statistics for the artificial economy are calculated at the posterior mean. σ_x denotes the standard deviation of variable x, $\rho(x, \ln(Y_t/Y_{t-1}))$ is the correlation of variable x and output growth, and ACF is the first order autocorrelation coefficient.

Table 3: Unconditional variance decomposition (in percent)

	$\ln\left(\frac{Y_t}{Y_{t-1}}\right)$	$\ln\left(\frac{C_t}{C_{t-1}}\right)$	$\ln\left(\frac{I_t}{I_{t-1}}\right)$	$\ln\left(\frac{G_t}{G_{t-1}}\right)$	$\ln\left(\frac{H_t}{H}\right)$	$spread_t$
ε_t^S	13.00	0.15	23.46	0.00	2.52	0.00
ε_t^A	17.97	27.21	14.89	0.38	15.67	0.00
ε_t^ξ	14.41	0.62	23.92	0.00	15.45	100
ε_t^Δ	44.50	71.92	36.15	0.00	56.19	0.00
ε_t^G	4.95	0.09	1.58	99.62	10.17	0.00
$\varepsilon_t^{m.e.}$	5.17	0.00	0.00	0.00	0.00	0.00

Variance decompositions are performed at the posterior mean.

et al. (2011). On the demand side, preference shocks dominate and account for the bulk of variation in aggregate output, echoing Galí and Rabanal (2005). Animal spirits disturbances make a further non-trivial contribution, explaining a modest share of output fluctuations and about a quarter of investment variability.

Since the shocks are identified within the model, it is useful to assess whether they correspond to observable empirical counterparts. As neither total factor productivity nor sentiment data are used in estimation, we conduct an external validation exercise. For technology, we compare the estimated series to Fernald's (2014) measure of utilization-adjusted TFP. After transforming both series into levels and HP-filtering them, we find a correlation of 0.62, lending support to the interpretation of the estimated disturbance as a technology shock. Furthermore, the correlation between output and TFP is -0.24 in the data and -0.47 in the model. We conduct a similar exercise for a level index of animal spirits that we construct from the estimated non-fundamental shocks. While we find positive co-movement of this index with survey-based measures of sentiment, we remain sceptical about the interpretation since such sentiment measures likely combine both intrinsic and extrinsic elements. Another concept of judging the plausibility of extrinsic expectation shocks maps their coherence with aggregate output over time and in particular at business

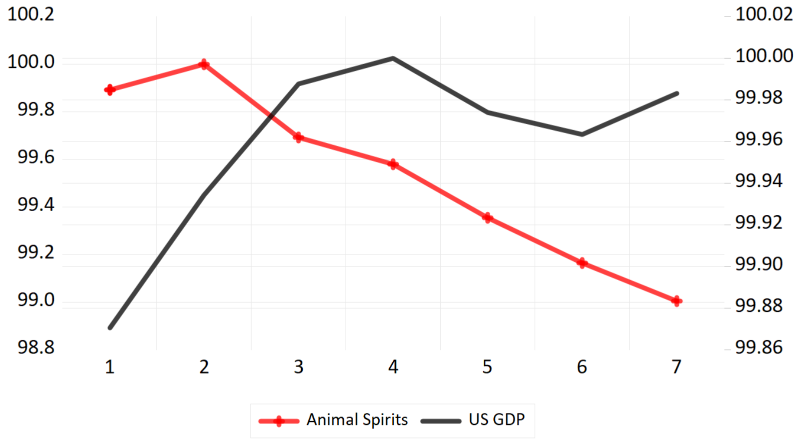

Figure 5: Burns and Mitchell meet animal spirits.

cycle turning points. For being of relevance and to fall into the extrinsic affiliation, one would expect at minimum that animal spirits are not only procyclical but also lead aggregate economic activity. We explore this pattern in Figure 5. Adopting a graphical device originally proposed by Burns and Mitchell (1946), the figure places side by side the sequences of output and animal spirits centered around business cycles' upper turning points.[7] In particular, we plot the average cycles of logged real U.S. GDP and the animal spirits index around three NBER business cycle peaks since 1990. We leave out the COVID-slump for obvious reasons. The GDP series can be understood as the reference of aggregate economic activity with stage four the peak quarter which we set equal to 100. Likewise, we normalize animal spirits' peak quarter to 100 as well. As you can see, animal spirits, by construction orthogonal to fundamental shocks, lead aggregate economic activity by at least two quarters: on average, a decline in optimism pre-dates recessions. What is more, while the figure plots the average spirits around three peaks, the characteristic pattern is largely driven by the 1990-91 recession and the Great Recession. Animal spirits remain essentially flat around 2001. As per the recession of 1990-91, our result parallels Blanchard's (1993) interpretation that animal spirits were one of the culprits for that particular economic slump.

[7] The business cycle series consists of a sequence of reference cycles, each spanning seven quarters centered on a business cycle peak. Each observation within a reference cycle is expressed as a percentage of the cycle mean, referred to as a cycle relative. Cycle relatives are then averaged across all reference cycles to produce a graphical summary of the average business cycle, shown in the seven-quarter plots of Figure 5.

6 Conclusion

This paper has set out a general equilibrium framework in which top firms are distinguished by their ability to operate scalable technologies. Such firms are thereby able to sustain higher markups and to command a larger share of the market than their peers. The model gives rise to heterogeneous firm behavior, as well as to reallocations which bear importantly upon aggregate outcomes. Foremost, the framework can simultaneously account for observed employment shares and markups of the top one percent of firms. The coexistence of large and small firms expands the set of parameter values consistent with multiple equilibria, creating scope for endogenous fluctuations driven by self-fulfilling beliefs. The model has been taken to the data by means of Bayesian estimations. The results point to the quantitative significance of the technology dispersion channel as a mechanism of amplification. A number of external checks suggest that the estimated disturbances—notably those associated with technology and sentiment—as well as the implied firm-level dynamics, accord reasonably well with the available empirical evidence. Overall, the findings suggest that non-fundamental shocks contribute meaningfully, although not dominantly, to business-cycle fluctuations. Their effects are particularly pronounced for investment dynamics, where shifts in extrinsic expectations account for a substantial share of observed variation alongside conventional fundamental forces.

References

[1] Autor, D., D. Dorn, L. Katz, C. Patterson and J. Van Reenen (2020): "The Fall of the Labor Share and the Rise of Superstar Firms", *Quarterly Journal of Economics* **135**, 645–709.

[2] Barkai, S. (2020): "Declining Labor and Capital Shares", *Journal of Finance* **75**, 2421–2463.

[3] Beck, G. and S. Lein (2020): "Price Elasticities and Demand-side Real Rigidities in Micro Data and in Macro Models", *Journal of Monetary Economics* **115**, 200–212.

[4] Benkard, C., N. Miller and A. Yurukoglu (2026): "The Rise of Market Power and the Macroeconomic Implications: Comment", Stanford University, mimeo.

[5] Bianchi, F. and G. Nicolò (2021): "A Generalized Approach to Indeterminacy in Linear Rational Expectations Models", *Quantitative Economics* **12**, 843–868.

[6] Blanchard, O. (1993): "Consumption and the Recession of 1990-1991", *American Economic Review* **83**, 270-274.

[7] Burns, A. and W. Mitchell (1946): *Measuring Business Cycles*, National Bureau of Economic Research, New York.

[8] Burstein, A., V. Carvalho and B. Grassi (2025): "Bottom-up Markup Fluctuations", *Quarterly Journal of Economics* **140**, 2619–2684.

[9] Callebaut, A. and G. Peersman (2026): "Markup Cyclicality? Firm and Sector Heterogeneity Matters", Ghent University, mimeo.

[10] Christiano, L. (1988): "Why Does Inventory Investment Fluctuate So Much?", *Journal of Monetary Economics* **21**, 247–280.

[11] Cunningham, C., L. Foster, C. Grim, J. Haltiwanger, S. Pabilonia, J. Stewart and Z. Wolf (2023): "Dispersion in Dispersion: Measuring Establishment-Level Differences in Productivity", *Review of Income and Wealth* **69**, 999–1032.

[12] De Loecker, J. and J. Eeckhout (2026): "The Macroeconomics of Market Power", *Annual Review of Economics*, in press.

[13] De Loecker, J., J. Eeckhout and G. Unger (2020): "The Rise of Market Power and the Macroeconomic Implications", *Quarterly Journal of Economics* **135**, 561–644.

[14] Edmond, C., V. Midrigan and D. Xu (2023): "How Costly Are Markups?", *Journal of Political Economy* **131**, 1619–1675.

[15] Farmer R. and J.-T. Guo (1994): "Real Business Cycles and the Animal Spirits Hypothesis", *Journal of Economic Theory* **63**, 42–72.

[16] Fernald, J. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity", Federal Reserve Bank of San Francisco, mimeo.

[17] Galí, J. (1994): "Monopolistic Competition, Business Cycles, and the Composition of Aggregate Demand", *Journal of Economic Theory* **63**, 73–96.

[18] Galí, J. and P. Rabanal (2005): "Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar U.S. Data?", *NBER Macroeconomics Annual 2004*, 225–288.

[19] Greenwood, J., Z. Hercowitz and G. Huffman (1988): “Investment, Capacity Utilization, and the Real Business Cycle“, *American Economic Review* **78**, 402–17.

[20] Hasenzagl, T. and L. Pérez (2026): “The Micro-Aggregated Profit Share“, Federal Reserve Bank of Richmond, mimeo.

[21] Herrendorf, B., A. Valentinyi and R. Waldmann (2000): “Ruling Out Multiplicity and Indeterminacy: The Role of Heterogeneity“, *Review of Economic Studies* **67**, 295–307.

[22] Hirose, Y., T. Kurozumi and W. Van Zandweghe (2023): “Inflation Gap Persistence, Indeterminacy, and Monetary Policy“, *Review of Economic Dynamics* **51**, 867–887..

[23] Hopenhayn, H. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium“, *Econometrica* **60**, 1127–1150.

[24] Hubmer, J., M. Chan, S. Ozkan and S. Salgado (2025): “Scalable versus Productive Technologies“, University of Pennsylvania, mimeo.

[25] Justiniano A., G. Primiceri and A. Tambalotti (2011): “Investment Shocks and the Relative Price of Investment“, *Review of Economic Dynamics* **14**, 102–121.

[26] Kimball, M. (1995): “The Quantitative Analytics of the Basic Neomonetarist Model”, *Journal of Money, Credit and Banking* **27**, 1241–1277.

[27] Kurozumi, T. and W. Van Zandweghe (2022): “Macroeconomic Changes With Declining Trend Inflation: Complementarity With the Superstar Firm Hypothesis”, *European Economic Review* **141**, 103998.

[28] Lubik, T. and F. Schorfheide (2004): “Testing for Indeterminacy: An Application to U.S. Monetary Policy“, *American Economic Review* **94**, 190–217.

[29] Lucas, R. (1978): “On the Size Distribution of Business Firms“, *Bell Journal of Economics* **8**, 508–523.

[30] Ma, Y., B. Pugsley, H. Qin and K. Zimmermann (2025): “Superstar Firms through the Generations“, University of Chicago, mimeo.

[31] Ma, Y., M. Zhang and K. Zimmermann (2026): “Business Concentration around the World: 1900–2020“, University of Chicago, mimeo.

[32] Nekarda, C. and V. Ramey (2020): "The Cyclical Behavior of the Price-Cost Markup", *Journal of Money, Credit and Banking* **52**, 319–353.

[33] Pagan, A. and T. Robinson (2022): "Excess Shocks Can Limit the Economic Interpretation", *European Economic Review* **145**, 104120.

[34] Pintus, P., Y. Wen and X. Xing (2022): "The Inverted Leading Indicator Property and Redistribution Effect of the Interest Rate", *European Economic Review* **148**, 104219.

[35] Wen, Y. (1998): "Capacity Utilization under Increasing Returns to Scale", *Journal of Economic Theory* **81**, 7-36.

A Appendix

A.1 Data sources and construction

This Appendix details the source and construction of the U.S. data used in Section 5. All data is quarterly and for the period 1990:I-2025:IV.

1. Gross Domestic Product. Seasonally adjusted at annual rates, billions of chained (2017) dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.6.

2. Gross Domestic Product. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

3. Personal Consumption Expenditures, Nondurable Goods. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

4. Personal Consumption Expenditures, Services. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

5. Personal Consumption Expenditures, Durable Goods. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

6. Gross Private Domestic Investment, Fixed Investment, Residential. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

7. Gross Private Domestic Investment, Fixed Investment, Nonresidential. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

8. Government consumption expenditures and gross investment. Seasonally adjusted at annual rates, billions of dollars. Source: Bureau of Economic Analysis, NIPA Table 1.1.5.

9. Nonfarm Business Hours. Seasonally adjusted, index. Source: Bureau of Labor Statistics, Series Id: PRS85006033.

10. Civilian Noninstitutional Population. 16 years and over, thousands. Source: Bureau of Labor Statistics, Series Id: LNU00000000Q.

11. GDP Deflator = (2)/(1).

12. Real Per Capita Output, $Y_t = (1)/(10)$.

13. Real Per Capita Consumption, $C_t = [(3) + (4)]/(11)/(10)$.

14. Real Per Capita Investment, $I_t = [(5) + (6) + (7)]/(11)/(10)$.

15. Real Per Capita Government Expenditures, $G_t = (8)/(11)/(10)$.

16. Per Capita Hours Worked, $H_t = (9)/(10)$.

17. Moody's Seasoned Baa Corporate Bond Yield [DBAA], retrieved from FRED,

Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DBAA.

18. Market Yield on U.S. Treasury Securities at 30-Year Constant Maturity, Quoted on an Investment Basis [DGS30], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DGS30.

19. Credit spread, (17) − (18).

20. Deviation from average credit spread, $spread_t$ = (19)− average of (19).

21. University of Michigan. Consumer Sentiment [UMCSENT], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/UMCSENT.

22. Total Factor Productivity. "A Quarterly Utilization-Adjusted Series on Total Factor Productivity", retrieved from https://www.johnfernald.net/TFP.

A.2 Model appendix

A.2.1 Markups and marginal costs

Final good producers' cost minimization

$$\min_{y_{i,t}/Y_t} 1 - \int_0^1 \frac{p_{i,t}y_{i,t}}{P_tY_t}di$$

subject to (1). Optimization together with (2) gives

$$\frac{p_{i,t}}{P_t} = v_t\left((1+\epsilon)\frac{y_{i,t}}{Y_t} - \epsilon\right)^{\frac{-1}{\gamma}}$$

where v_t is the multiplier on the constraint. This rearranges for (3) and substituting it into (2) and some algebra gives the price index (4). Zero profits imply

$$1 = \int_0^1 \frac{p_{i,t}y_{i,t}}{P_tY_t}di$$

Then substituting (3) into the above and more algebra gives us

$$v_t = 1 + \epsilon - \int_0^1 \frac{p_{i,t}}{P_t}\epsilon di$$

Intermediate good producers maximize profits

$$\pi_{i,t} = p_{i,t}y_{i,t} - w_th_{i,t} - r_tk_{i,t}$$

and for the big firm this is subject to the production function (6). The first-order conditions are

$$w_t = \lambda_{i,t}\omega(1-\alpha)zk_{i,t}^{\alpha\omega}h_{i,t}^{(1-\alpha)\omega-1}$$

$$r_t = \lambda_{i,t}\omega\alpha zk_{i,t}^{\alpha\omega-1}h_{i,t}^{(1-\alpha)\omega}$$

$$y_{i,t} + (p_{i,t} - \lambda_{i,t}) \frac{\partial y_{i,t}}{\partial p_{i,t}} = 0$$

where the multiplier $\lambda_{i,t}$ amounts to marginal cost $\mathbf{MC}_{i,t}$. Assuming firms take P_t and v_t as given yields

$$\frac{\partial y_{i,t}}{\partial p_{i,t}} = -\gamma p_{i,t}^{-1} \left(y_{i,t} - \frac{\epsilon}{1+\epsilon} Y_t \right)$$

and substituting it into the above first-order condition and rearranging gives us the markup (8). The small firm's markup is found the same way and the wage and rental rates are

$$w_t = \lambda_{i,t} \eta (1-a) k_{i,t}^{a\eta} h_{i,t}^{(1-a)\eta - 1}$$
$$r_t = \lambda_{i,t} \eta a k_{i,t}^{a\eta - 1} h_{i,t}^{(1-a)\eta}$$

Minimizing the cost function

$$\min_{h_{i,t}, k_{i,t}} \mathbf{C}_{i,t} = w_t h_{i,t} + r_t k_{i,t}$$

subject to their production technology provides the conditional demands for the big firm

$$h_{i,t} = \left(\frac{r_t}{w_t} \frac{1-\alpha}{\alpha} \right)^{\alpha} \left(\frac{y_{i,t} + \phi_m}{z} \right)^{\frac{1}{\omega}}$$
$$k_{i,t} = \left(\frac{r_t}{w_t} \frac{1-\alpha}{\alpha} \right)^{-(1-\alpha)} \left(\frac{y_{i,t} + \phi_m}{z} \right)^{\frac{1}{\omega}}.$$

Inserting these back into the cost function yields

$$\mathbf{C}_{i,t} = \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} w_t^{1-\alpha} r_t^{a} \left(\frac{y_{i,t} + \phi_m}{z} \right)^{\frac{1}{\omega}}$$

and marginal costs

$$\mathbf{MC}_{i,t} = \frac{1}{\omega} \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} w_t^{1-\alpha} r_t^{a} \left(y_{i,t} + \phi_m \right)^{\frac{1}{\omega} - 1} z^{\frac{-1}{\omega}}.$$

Similarly, we can find marginal costs for the small firm

$$\mathbf{MC}_{i,t} = \frac{1}{\eta} a^{-a} (1-a)^{-(1-a)} w_t^{1-a} r_t^{a} \left(y_{i,t} + \phi_n \right)^{\frac{1}{\eta} - 1}.$$

A.2.2 Symmetric equilibrium

Since both firms set different prices, (5) can be written as

$$v_t = 1 + \epsilon - \epsilon \int_0^n \frac{p_{i,t}}{P_t} di - \epsilon \int_n^1 \frac{p_{i,t}}{P_t} di$$

In symmetric equilibrium with $P_t = 1$ as the numeraire

$$v_t = 1 + \epsilon - \epsilon n p_{n,t} - \epsilon m p_{m,t} \tag{A.1}$$

Demand for firms' goods

$$\frac{y_{m,t}}{Y_t} = \frac{p_{m,t}^{-\gamma} v_t^{\gamma} + \epsilon}{1+\epsilon} \tag{A.2}$$

$$\frac{y_{n,t}}{Y_t} = \frac{p_{n,t}^{-\gamma} v_t^{\gamma} + \epsilon}{1+\epsilon} \tag{A.3}$$

Production functions

$$y_{m,t} = k_{m,t}^{\alpha\omega} h_{m,t}^{(1-\alpha)\omega} - \phi_m \tag{A.4}$$

$$y_{n,t} = k_{n,t}^{a\eta} h_{n,t}^{(1-a)\eta} - \phi_n \tag{A.5}$$

Markups

$$\mu_{m,t} = \frac{\gamma\left(\frac{y_{m,t}}{Y_t} - \frac{\epsilon}{1+\epsilon}\right)}{(\gamma-1)\frac{y_{m,t}}{Y_t} - \gamma\frac{\epsilon}{1+\epsilon}} \tag{A.6}$$

$$\mu_{n,t} = \frac{\gamma\left(\frac{y_{n,t}}{Y_t} - \frac{\epsilon}{1+\epsilon}\right)}{(\gamma-1)\frac{y_{n,t}}{Y_t} - \gamma\frac{\epsilon}{1+\epsilon}} \tag{A.7}$$

The employment weighted average markup is

$$\mu_t = \frac{m h_{m,t}}{H_t}\mu_{m,t} + \frac{n h_{n,t}}{H_t}\mu_{n,t} \tag{A.8}$$

Wage and rental rates

$$w_t = \omega(1-\alpha)\frac{p_{m,t}}{\mu_{m,t}} z k_{m,t}^{\alpha\omega} h_{m,t}^{(1-\alpha)\omega-1} = \eta(1-a)\frac{p_{n,t}}{\mu_{n,t}} k_{n,t}^{a\eta} h_{n,t}^{(1-a)\eta-1} \tag{A.9}$$

$$r_t = \omega\alpha\frac{p_{m,t}}{\mu_{m,t}} z k_{m,t}^{\alpha\omega-1} h_{m,t}^{(1-\alpha)\omega} = \eta a\frac{p_{n,t}}{\mu_{n,t}} k_{n,t}^{a\eta-1} h_{n,t}^{(1-a)\eta} \tag{A.10}$$

From the zero profit equation of final good firms we have

$$P_t Y_t = \int_0^1 p_{i,t} y_{i,t} di = n p_{n,t} y_{n,t} + m p_{m,t} y_{m,t}$$

and so we get

$$n\frac{p_{n,t} y_{n,t}}{Y_t} + m\frac{p_{m,t} y_{m,t}}{Y_t} = 1 \tag{A.11}$$

Aggregate capital and labour

$$K_t = m k_{m,t} + n k_{n,t} \tag{A.12}$$

$$H_t = m h_{m,t} + n h_{n,t} \tag{A.13}$$

A.2.3 Calibration and the steady state

Calibrating steady state markups μ_m and μ_n, and the Kimball parameters θ and ϵ allows us to find the steady state terms on the firms' side. First, this pin down $\frac{y_n}{Y}$ and $\frac{y_m}{Y}$ from the markup equations (A.6) and (A.7). Next, the symmetric equilibrium version of (1) is

$$n\left((1+\epsilon)\frac{y_n}{Y}-\epsilon\right)^{\frac{\gamma-1}{\gamma}}+m\left((1+\epsilon)\frac{y_m}{Y}-\epsilon\right)^{\frac{\gamma-1}{\gamma}}=1$$

and since $m+n=1$, we obtain m and n. Next, the symmetric equilibrium version of (4) is

$$v=\left(np_n^{1-\gamma}+mp_m^{1-\gamma}\right)^{\frac{1}{1-\gamma}} \tag{A.14}$$

and substituting it into (A.2) gives

$$\frac{y_m}{Y}=\frac{\left(m+n\left(\frac{p_n}{p_m}\right)^{1-\gamma}\right)^{\frac{\gamma}{1-\gamma}}+\epsilon}{1+\epsilon}$$

which pins down the price ratio $\frac{p_n}{p_m}$. Combining (A.14) with (A.1), then rearranging for p_m gives

$$p_m=\frac{1+\epsilon}{\left(m+n\left(\frac{p_n}{p_m}\right)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}+\epsilon m+\epsilon n\frac{p_n}{p_m}}$$

Since we have $\frac{p_n}{p_m}$, we can then determine p_m and p_n. Together, this pins down the market shares $n\frac{p_n y_n}{Y}$ and $m\frac{p_m y_m}{Y}$.

Substituting the cost function into the profits of big firms we get

$$\pi_m=p_m y_m-\omega \mathbf{MC}_m\left(y_m+\phi_m\right)$$

which can be re-written as

$$\frac{m\pi_m}{Y}=\left(1-\frac{\omega}{\mu_m}\right)\frac{mp_m y_m}{Y}-\frac{\omega}{\mu_m}\frac{mp_m\phi_m}{Y}$$

where $\omega<\mu_m$ creates positive pure profits. If there were no fixed costs, $\phi_m=0$, the first term on the right hand side gives the maximum profits that can be set at the steady state. Lowering steady state profits from there would then entail $\phi_m>0$. Therefore, the calibration of big firms' profits in aggregate output must satisfy

$$0\leq\frac{m\pi_m}{Y}\leq\left(1-\frac{\omega}{\mu_m}\right)\frac{mp_m y_m}{Y}$$

and this pins down $\frac{\phi_m}{Y}$. We assume that small firms have zero profits i.e., $\pi_n=0$, and their profits can be simplified to

$$0=1-\frac{\eta}{\mu_n}\left(1+\frac{Y}{y_n}\frac{\phi_n}{Y}\right)$$

which determines $\frac{\phi_n}{Y}$.

Using the intermediate good firms' production functions and the wage equations allows us to pin down

$$\frac{h_n}{h_m} = \frac{\eta(1-a)\frac{p_n}{\mu_\eta}\frac{y_n+\phi_n}{Y}}{\omega(1-\alpha)\frac{p_m}{\mu_m}\frac{y_m+\phi_m}{Y}}$$

and dividing (A.9) by (A.10) we get

$$\frac{k_n}{k_m} = \frac{h_n}{h_m}\frac{1-\alpha}{\alpha}\frac{a}{1-a}.$$

Since we now have $\frac{h_n}{h_m}$ and $\frac{k_n}{k_m}$, using (A.12) and (A.13) we can determine

$$\begin{aligned}\frac{k_m}{K} &= \frac{1}{m+(1-m)\frac{k_n}{k_m}}\\ \frac{k_n}{K} &= \frac{1}{m\frac{k_m}{k_n}+1-m}\\ \frac{h_m}{H} &= \frac{1}{m+(1-m)\frac{h_n}{h_m}}\\ \frac{h_n}{H} &= \frac{1}{m\frac{h_m}{h_n}+1-m}\end{aligned}$$

Using the rental rates and intermediate good firms' production functions we get

$$\frac{rk_m}{Y} = \omega\alpha\frac{p_m}{\mu_m}\left(\frac{y_m}{Y}+\frac{\phi_m}{Y}\right)$$

$$\frac{rk_n}{Y} = \eta a\frac{p_n}{\mu_n}\left(\frac{y_n}{Y}+\frac{\phi_n}{Y}\right)$$

Therefore

$$\frac{rK}{Y} = m\frac{rk_m}{Y}+n\frac{rk_n}{Y}$$

Similarly, using wages and production functions we obtain

$$\frac{wh_m}{Y} = \omega(1-\alpha)\frac{p_m}{\mu_m}\left(\frac{y_m}{Y}+\frac{\phi_m}{Y}\right)$$

$$\frac{wh_n}{Y} = \eta(1-a)\frac{p_n}{\mu_n}\left(\frac{y_n}{Y}+\frac{\phi_n}{Y}\right)$$

$$\frac{wH}{Y} = m\frac{wh_m}{Y}+n\frac{wh_n}{Y}$$

To complete the firms' side, relative total factor productivity z is endogenously determined by our calibration. Rewriting relative marginal costs $\mathbf{MC}_n/\mathbf{MC}_m$ we obtain

$$z^{\frac{-1}{\omega}} = \left(\frac{w}{r}\right)^{\alpha-a} Y^{\frac{1}{\eta}-\frac{1}{\omega}}\frac{\omega}{\eta}\frac{p_m}{p_n}\frac{\mu_n}{\mu_m}\frac{(1-\alpha)^{1-\alpha}\alpha^\alpha}{(1-a)^{1-a}a^a}\frac{\left(\frac{y_n}{Y}+\frac{\phi_n}{Y}\right)^{\frac{1}{\eta}-1}}{\left(\frac{y_m}{Y}+\frac{\phi_m}{Y}\right)^{\frac{1}{\omega}-1}}$$

which shows that z is a complicated function of the model's parameters and steady state when $\eta \neq \omega$ or $\alpha \neq a$.

Finally, from the household side, the steady state capital accumulation equation determines

$$\frac{I}{K} = \delta$$

and the Euler equation gives

$$r = \frac{1}{\beta} - 1 + \delta.$$

These then allow us to find the investment and consumption shares

$$\frac{I}{Y} = \frac{I}{rK}\frac{rK}{Y} = \frac{\delta}{\frac{1}{\beta} - 1 + \delta}\frac{rK}{Y} = 1 - \frac{C}{Y}.$$

A.2.4 Linearized equations

The symmetric equilibrium equations listed in Appendix A.2.2 are linearized:

$$v\hat{v}_t = -\epsilon n p_n \hat{p}_{n,t} - \epsilon(1-n)p_m\hat{p}_{m,t}$$

$$\widehat{Y}_t = \widehat{y}_{n,t} + \frac{\gamma p_n^{-\gamma} v^{\gamma}}{p_n^{-\gamma} v^{\gamma} + \epsilon}\widehat{p}_{n,t} - \frac{\gamma p_n^{-\gamma} v^{\gamma}}{p_n^{-\gamma} v^{\gamma} + \epsilon}\widehat{v}_t$$

$$\widehat{Y}_t = \widehat{y}_{m,t} + \frac{\gamma p_m^{-\gamma} v^{\gamma}}{p_m^{-\gamma} v^{\gamma} + \epsilon}\widehat{p}_{m,t} - \frac{\gamma p_m^{-\gamma} v^{\gamma}}{p_m^{-\gamma} v^{\gamma} + \epsilon}\widehat{v}_t$$

$$\widehat{y}_{n,t} = \alpha\mu_n\widehat{k}_{n,t} + (1-\alpha)\mu_n\widehat{h}_{n,t}$$

$$\widehat{y}_{m,t} = \alpha\mu_m\left(1 - \frac{\pi_m}{p_m y_m}\right)\widehat{k}_{m,t} + (1-\alpha)\mu_m\left(1 - \frac{\pi_m}{p_m y_m}\right)\widehat{h}_{m,t}$$

$$\widehat{\mu}_{m,t} = \left(\frac{1}{\mu_m}\frac{\gamma\frac{y_m}{Y}}{(\gamma-1)\frac{y_m}{Y} - \gamma\frac{\epsilon}{1+\epsilon}} - \mu_m\frac{(\gamma-1)\frac{y_m}{Y}}{\gamma\left(\frac{y_m}{Y} - \frac{\epsilon}{1+\epsilon}\right)}\right)\left(\widehat{y}_{m,t} - \widehat{Y}_t\right)$$

$$\widehat{\mu}_{n,t} = \left(\frac{1}{\mu_n}\frac{\gamma\frac{y_n}{Y}}{(\gamma-1)\frac{y_n}{Y} - \gamma\frac{\epsilon}{1+\epsilon}} - \mu_n\frac{(\gamma-1)\frac{y_n}{Y}}{\gamma\left(\frac{y_n}{Y} - \frac{\epsilon}{1+\epsilon}\right)}\right)\left(\widehat{y}_{n,t} - \widehat{Y}_t\right)$$

$$\hat{\mu}_t = \frac{mh_m}{H}\frac{\mu_m}{\mu}\left(\hat{\mu}_{m,t} + \widehat{h}_{m,t}\right) + \frac{nh_n}{H}\frac{\mu_n}{\mu}\left(\hat{\mu}_{n,t} + \widehat{h}_{n,t}\right) - \hat{H}_t$$

$$\widehat{w}_t = \widehat{p}_{m,t} - \widehat{\mu}_{m,t} + \alpha\omega\widehat{k}_{m,t} + ((1-\alpha)\omega - 1)\widehat{h}_{m,t}$$

$$\widehat{w}_t = \widehat{p}_{n,t} - \widehat{\mu}_{n,t} + a\eta\widehat{k}_{n,t} + ((1-a)\eta - 1)\widehat{h}_{n,t}$$

$$\widehat{r}_t = \widehat{p}_{m,t} - \widehat{\mu}_{m,t} + (\alpha\omega - 1)\widehat{k}_{m,t} + (1-\alpha)\omega\widehat{h}_{m,t}$$

$$\widehat{r}_t = \widehat{p}_{n,t} - \widehat{\mu}_{n,t} + (a\eta - 1)\widehat{k}_{n,t} + (1-a)\eta\widehat{h}_{n,t}$$

$$np_n x_n\left(\widehat{p}_{n,t} + \widehat{y}_{n,t} - \widehat{Y}_t\right) + mp_m x_m\left(\widehat{p}_{m,t} + \widehat{y}_{m,t} - \widehat{Y}_t\right) = 0$$

$$\widehat{K}_t = m\frac{k_m}{K}\widehat{k}_{m,t} + (1-m)\frac{k_n}{K}\widehat{k}_{n,t}$$

$$\widehat{H}_t = m\frac{h_m}{H}\widehat{h}_{m,t} + (1-m)\frac{h_n}{H}\widehat{h}_{n,t}$$

Furthermore, from the household side we have

$$\hat{K}_{t+1} = (1-\delta)\hat{K}_t + \delta\hat{I}_t$$

$$\hat{w}_t = \hat{C}_t + \chi\hat{H}_t$$

$$-\hat{C}_t = -\hat{C}_{t+1} + \beta\left(\frac{1}{\beta} - 1 + \delta\right)\hat{r}_{t+1}$$

$$\widehat{Y}_t = \left(1 - \frac{C}{Y}\right)\widehat{I}_t + \frac{C}{Y}\widehat{C}_t$$

where the last comes from the resource constraint $Y_t = C_t + I_t$.

A.2.5 CES version of the model

If $\epsilon = 0$, the model collapses to CES with constant markups. The aggregator is

$$\int_0^1 \left(\frac{y_{i,t}}{Y_t}\right)^{\frac{\theta-1}{\theta}} di = 1$$

with demand

$$\frac{y_{i,t}}{Y_t} = \left(\frac{p_{i,t}}{P_t}\right)^{-\theta}$$

and both firm types have the usual Dixit-Stiglitz markups of $\theta/(\theta-1)$. In symmetric equilibrium, the aggregator becomes

$$n\left(\frac{y_{n,t}}{Y_t}\right)^{\frac{\theta-1}{\theta}} + m\left(\frac{y_{m,t}}{Y_t}\right)^{\frac{\theta-1}{\theta}} = 1.$$

With the price index as the numeraire ($P = 1$) we get

$$1 = np_{n,t}^{1-\theta} + mp_{m,t}^{1-\theta}$$

and substituting in the demand equations yields the market share

$$1 = \frac{np_n y_n}{Y} + \frac{mp_m y_m}{Y}.$$

The modified linearized equations are

$$\widehat{y}_{m,t} - \widehat{Y}_t = -\theta\widehat{p}_{m,t}$$

$$\widehat{y}_{n,t} - \widehat{Y}_t = -\theta\widehat{p}_{n,t}$$

$$\frac{np_n y_n}{Y}\left(\widehat{p}_{n,t}+\widehat{y}_{n,t}-\widehat{Y}_t\right)+\frac{mp_m y_m}{Y}\left(\widehat{p}_{m,t}+\widehat{y}_{m,t}-\widehat{Y}_t\right)=0$$

Calibrating the market share $\frac{mp_m y_m}{Y}$ and the number of ordinary firms n, then gives us $\frac{p_n y_n}{Y}$ and $\frac{p_m y_m}{Y}$. These are then used together with the demand functions to pin down the relative price

$$\left(\frac{p_n}{p_m}\right)^{1-\theta}=\frac{p_n y_n}{Y}\frac{Y}{p_m y_m}.$$