Market power, growth, and wealth inequality

Giammario Impullitti, Pontus Rendhal

March 2025

Centre for Inclusive Trade Policy
Working Paper No.020

Centre for Inclusive Trade Policy
https://citp.ac.uk/
info@citp.ac.uk

Established in 2022, the Centre for Inclusive Trade Policy (CITP) is the first research centre dedicated to trade policy to be funded by the Economic and Social Research Council. As a centre of excellence for innovative research on trade policy and its inclusiveness, we aim to equip the UK with the capability to formulate and implement a trade policy tailored to the needs of the whole of the UK, while recognising the importance of the multilateral trading system and the UK's role within it. The CITP is funded by the Economic and Social Research Council [grant number ES/W002434/1]

Abstract

In recent decades, the United States has experienced a notable rise in markups, a slowdown in productivity growth, and an increase in wealth inequality. We present a framework that unifies these trends into a common driving force. In particular, increased barriers to entry raise markups and boost corporate profits. Rising profits elevates firm valuations, fuels the demand for capital, and drives up asset returns. At the same time, the reduction in competition stifles overall economic growth. Wealth inequality is shaped by the *return gap, r - g*, which represents the difference between asset returns and the economy's growth rate. The rise in capital demand together with a reduction in growth leads to a widening of the return gap, which amplifies inequality by affecting the saving patterns of households in different ways across the wealth distribution, deepening the divide between the rich and the poor. These trends result in substantial welfare losses for the majority of households, while only the top 1%, and especially the top 0:1%, experience gains.

Suggested citation

Impullitti, G; Rendahl, P (2025) Centre for Inclusive Trade Policy, Working Paper 020

Non-Technical Summary

From Washington to Brussels, policymakers are grappling with the dominance of a few companies in key industries, slowing productivity growth, and rising wealth inequality. These issues, highlighted by recent research, are central to today's economic challenges. In the U.S., the average markup—the difference between the price firms charge and production costs—have surged since the 1980s, rising from 20% to 55% by 2020, reflecting increased market power. Meanwhile, productivity growth has slowed, and wealth inequality has reached levels not seen since the Gilded Age. These trends are interconnected, and understanding their links is crucial for effective policymaking.

But what exactly is the link between market power, growth, and inequality? We provide a framework that ties these trends together, offering new insights into how market power shapes the economy. Piketty popularised the idea that wealth inequality is driven by the difference between the rate of return on assets and the growth rate of the economy, or the return gap. A higher return gap increases inequality as wealthier households, who own more assets, benefit from higher returns and save more, further increasing their wealth. Poorer households, who rely more on wages, see their incomes stagnate due to slower growth and higher markups. This dynamic deepens the divide between the rich and the poor, leading to a more unequal society.

How does market power affect the return on assets and the growth rate? To answer this question, we build a macroeconomic model where large firms invest in innovation to gain market shares and where aggregate innovation pushes overall productivity growth. Uninsurable income risk generates wealth diversity across households.

We engineer a rise in markups as a response to an external increase in the cost of entry to the market for firms. This reduces competition allowing incumbent firms to charge higher markups and boosting their profits. Higher profits, in turn, increase the value of these firms, driving up returns and asset prices.

Aggregate innovation contributes to the economy's overall stock of knowledge, which functions as a public good. Firms continuously learn from one another, fostering a cycle of innovation and progress. However, when competition declines, this knowledge-sharing process weakens, reducing the efficiency of innovation and ultimately slowing economic growth.

In addition, lower competition also conveys some bad news for labour income, both in the present and in the future. Higher markups create a wedge between the price of goods and the associated marginal costs: wages. As markups rise, real wages fall. Additionally, slower economic growth further exacerbates this outcome, dampening the prospects for future wage increases, which are closely tied to productivity growth.

Why does a rise in the return gap exacerbate wealth inequality? Our theory demonstrates that a widening return gap exacerbates inequality by affecting the saving behaviour of households in distinct ways across the wealth spectrum. Poorer households, driven by the need to buffer against income risk, predominantly save for precautionary reasons, whereas richer households – having attained a high level of self-insurance – primarily save for intertemporal reasons. An increase in asset returns enhances the incentives for intertemporal substitution but has little impact on the precautionary motive. As a result, wealthier households respond more strongly to rising returns, increasing their savings at a higher rate than asset-poor households and further accumulating wealth.

Finally, we find that the increase in markups and the slowdown in growth since 1980 have led to substantial welfare losses for most households. For the bottom 80% of the wealth distribution, these losses amount to roughly 34% of long-run consumption. In contrast, the top 1% of households have seen significant gains, with the top 0.1% experiencing a 30% increase in consumption.

Market power, growth, and wealth inequality

Giammario Impullitti
University of Nottingham and CEPR

Pontus Rendahl
Copenhagen Business School and CEPR

February 4, 2025

Abstract

In recent decades, the United States has experienced a notable rise in markups, a slowdown in productivity growth, and an increase in wealth inequality. We present a framework that unifies these trends into a common driving force. In particular, increased barriers to entry raise markups and boost corporate profits. Rising profits elevates firm valuations, fuels the demand for capital, and drives up asset returns. At the same time, the reduction in competition stifles overall economic growth. Wealth inequality is shaped by the *return gap*, $r - g$, which represents the difference between asset returns and the economy's growth rate. The rise in capital demand together with a reduction in growth leads to a widening of the return gap, which amplifies inequality by affecting the saving patterns of households in different ways across the wealth distribution, deepening the divide between the rich and the poor. These trends result in substantial welfare losses for the majority of households, while only the top 1%, and especially the top 0.1%, experience gains.

Keywords: Market Power, Growth, Heterogeneous Agents, Wealth Distribution.

Impullitti: giammario.impullitti@nottingham.ac.uk,; Rendahl: pontus.rendahl@gmail.com.

1 Introduction

Over the past four decades, the U.S. economy has seen a notable rise in wealth inequality. The share of wealth held by the top 1% increased from about 25% in 1980 to 35% in 2022, while the top 10% saw their share grow from roughly 65% to 71% during the same period. Additionally, the wealth-to-income ratio rose from 3.8 to 6, and the Gini coefficient observed a trough at 0.79 in 1984, and peaked at 0.84 in 2017 (World Inequality Database, 2024)[1].

Alongside this growing inequality, there has been a marked increase in various indicators of market power. Indeed, after remaining stable (or declining) between 1960 and 1980, aggregate markups rose by over 30 percentage points, climbing from 20% in 1980 to 55% in 2020 (De Loecker et al., 2020).[2] This rise in wealth inequality and market power has coincided with a slowdown in growth. Total factor productivity averaged an annual growth rate of 1.56% in 1960 to 1980, and then dropped to 0.77% in the following four decades (Fernald, 2014). Figure 1 reports the paths of markups, TFP, the Gini index and the top-10 wealth share in this period.

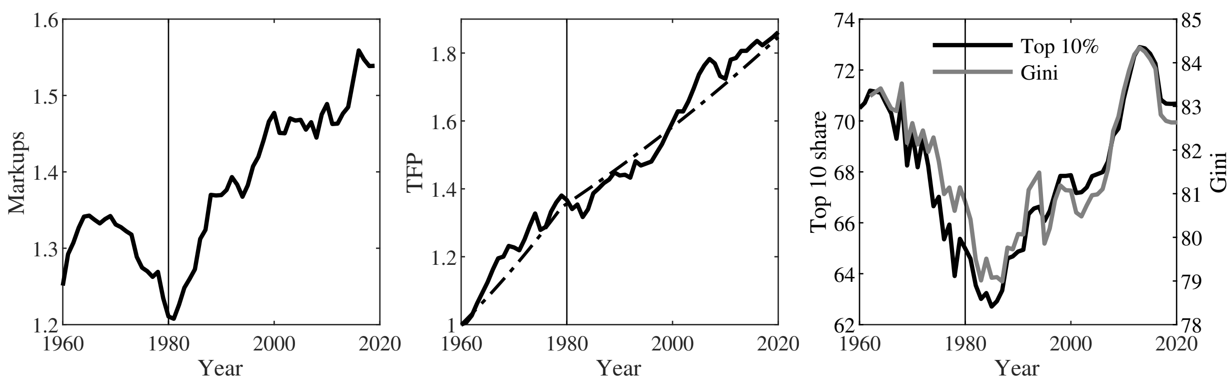

Figure 1: Market power, growth and wealth inequality in the US

Sources: De Loecker et al. (2020), Fernald (2014), and World Inequality Database (2024).

We propose a theory that ties these facts together, and use it to study the interconnectedness of market power, growth, and wealth concentration. Specifically, we consider a rise in the cost of entry, that reduces competition and raises markups. This reduction in competition leads to fewer active firms, and thereby less knowledge spillovers, which dampens growth. At the same time, higher markups benefit incumbent firms, increasing their profitability, and enhancing their stock market valuations. This surge in profitability leads to greater demand for financial capital, which drives up required returns on assets. As a result, higher asset returns and lower growth leads to a widening of the "return gap" ($r-g$) (Piketty, 2014; Jones, 2015; Moll et al., 2022). In a stable wealth distribution, wealth inequality is primarily governed by this return differential. In particular,

[1] There was also a significant rise in key dimensions of income inequality (Acemoglu and Autor, 2011; Autor, 2019)

[2] Similar trends have been documented in Bajgar et al. (2019) and Autor et al. (2020b) among others.

our theory demonstrates that a widening return gap exacerbates inequality by affecting the savings behaviors of households in distinct ways across the wealth spectrum, magnifying the divide between richer and poorer households. Taken together, we show that the observed trends in markups, growth, and wealth inequality, may well be connected, and follow a pattern that naturally emerges in a less competitive environment.

To formalize these arguments, we build an endogenous growth model with variable markups and heterogeneous households. Oligopolistic firms innovate to improve their productivity and compete strategically for market shares. Innovation drives long-run aggregate growth and free entry pins down product market competition and the markup. We first illustrate these mechanisms in a simple baseline *two-agent* growth model, with capitalists and workers, where only the former save and owns wealth, while the latter are hand-to-mouth workers. Later, we generalize this setting to an *incomplete markets* economy, with a rich exploration of wealth inequality.

Firms in our setting innovate to enhance their productivity, and the technological progress this generates drives long-run growth. The incentive to innovate is driven by firms' market size, as cost reductions from increased efficiency yield greater benefits when the scale of production is larger. This *market size* effect creates a negative relationship between competition and innovation, since firms tend to be smaller in more competitive markets. However, all innovations contribute to the economy's overall stock of knowledge, which acts as a public good. As a result, firms learn from one another. Reduced competition weakens this learning process, as fewer competitors limit opportunities for knowledge exchange, thereby lowering innovation efficiency and hampering growth. This endogenous *knowledge spillovers* channel establishes another link between market power and the efficiency of innovation. The trade-off embodied by these two channels shapes the relationship between competition and growth.

We engineer a rise in markups as a response to an exogenous increase in the cost of entry.[3] If the spillover channel prevails, an increase in markups leads to a reduction in growth, consistent with the patterns observed in the US economy. Notably, this slowdown in growth may happen even as firms invest more in innovation. This paradox occurs because increasing market power undermines knowledge spillovers, diminishing the effectiveness of individual innovation efforts.[4]

As markups rise, the wealth-to-income ratio – the two-agent economy's sole measure of inequality – also increases. Higher markups result in greater profits and elevated firm valuations, boosting the value of assets held by capitalists, thereby placing upward pressure on the return on assets. Capitalists take advantage of these higher returns by saving more, increasing asset supply. Over time, however, the additional supply of assets alleviates some of the upward pressure on the

[3] We also discuss other potential sources of change in market power.

[4] This aligns with key findings from Bloom et al. (2020), which indicate that falling R&D productivity may be a significant factor behind the growth slowdown in the US. Our theory posits that rising market power is a contributing factor to the decline in both innovation efficiency and productivity growth.

interest rate. Meanwhile, the reduced growth outlook encourages asset supply on its own due to intertemporal considerations (consumption smoothing).[5] The combination of increased asset supply and lower growth leads to a reduced real interest rate compared to the previous balanced growth path (BGP).

While the return gap initially widens in this narrative – as the real interest rate rises and growth declines – it eventually returns to its original level, as the Euler equation dictates that it must align with the discount factor on the BGP.[6] Yet, this *transitory* increase in the return gap represents the channel through which weakening competition generates *permanently* higher inequality. It should be noted that the return on assets only affects the income and wealth of capitalists. Workers, unable to access the asset market, derive income only from wages, which are now lower due to elevated markups, and grow at the economy's lower growth rate.

In the generalized framework, all agents work and have the ability to save. However, incomplete markets and (uninsurable) income risk, fosters precautionary savings and creates heterogeneity in asset holdings. This adds new dimensions to our previous analysis; in particular it allows us to investigate the model beyond the binary discrepancy in wealth between a poor worker, and a rich capitalist, and to make stronger contact with existing empirical evidence. The mechanisms through which an increase in markups affects growth remain consistent with the simpler model, as the production side of the economy is unchanged. However, interest rate determination differs in this framework, as the supply of assets is upward sloping, even in the long run (cf. Moll et al. (2022)). The mechanism through which an increase in the return gap affects inequality is similar to that of the simple model, but characterised by greater complexity and nuance, requiring a full numerical solution.

We parametrise the incomplete markets model to align with key cross-sectional and aggregate statistics of the U.S. economy, focusing our analysis on the BGP[7]. We feed it with an increase in the entry cost that allows us to match the increase in the markup observed in the data and study its impact on innovation, growth and wealth inequality. Similar to the simpler model, an increase in markups drives up the demand for assets, exerting upward pressure on asset returns. Weaker competition results in slower growth, which, combined with higher asset returns, leads to a wider return gap and consequently fuels wealth inequality. The slowdown in growth encourages saving and increases the supply of assets, ultimately resulting in a long-run equilibrium characterized by a lower real interest rate, alongside a *permanently* higher return gap due to the decline in the growth

[5] What matters to capitalists is the return gap, $r-g$, as slowdown in growth is akin to bad news about the future, which encourages saving *given* a certain real interest rate.

[6] Saving in our simple model is carried out by a representative household and thus the long-run supply of assets is infinitively elastic.

[7] Although our primary focus is on balanced growth paths, section 2.3.1 provides an insight into the transitional dynamics.

rate.

Why does a rise in the return gap exacerbate wealth inequality? After all, if all agents responded proportionally to a rise in the $r-g$ differential, wealth inequality would be unaffected.[8] In our economy, uninsurable income risk implies that there are two reasons for saving: intertemporal substitution, and a precautionary motive. Poorer households, motivated to mitigate income risk, predominantly save for precautionary reasons, whereas richer households, having attained a high level of self-insurance, primarily save for intertemporal reasons. An increase in asset returns affects the rewards associated with intertemporal subsitution, but does little to affect the precautionary motive. Hence, richer households respond stronger to an increase in the return gap than the asset poor. Simply put, the widening return gap increases wealth inequality by affecting the saving rates of poorer and richer households in different ways. In the capitalist-worker model, this occurs because workers are unable to save. In the incomplete markets setting, it arises from the differing motivations for saving between the rich and the poor.

Our results underscore the significance of endogenous growth in understanding wealth inequality. Growth is not only a key factor influencing the return gap, shaping the distribution of wealth, but also interacts dynamically with interest rates. Without the feedback loop between growth and asset returns, an increase in market power would have resulted in a *higher* real interest rate in the long run, not *lower*. This interplay between interest rates and growth enables a richer set of predictions, where higher market power can either raise or lower interest rates.[9] Even in scenarios where the interest rate declines, as in ours, the return gap – and consequently inequality – can still widen due to slower growth rates. This is a key feature of our framework.

Finally, a welfare analysis, focused on comparing initial and final BGPs, reveals that the observed increase in markups between 1980 and 2020, alongside the reduction in growth and increase in wealth inequality, generate large welfare losses, equivalent to approximately 30% of long-run consumption. These losses are concentrated at the bottom 80% of the distribution, where they amount to roughly 34% of consumption. In contrast, only the top 1% of the distribution benefits, with the most significant gains accruing to the top 0.1%, where households instead record a 30% increase in compensated-variation consumption. The losses primarily stem from wages: rising markups erode real wages, and slower growth diminishes future wage prospects. Meanwhile, the gains are driven by wealth accumulation at the very top, enabling the richest households to significantly expand their consumption.

[8] Measures such as the Gini coefficient, the top-1 share, etc., are all invariant to scaling of the distribution. In the two-agent model, the effect of the return gap on inequality is quite mechanical, as households are differentially affected as workers do not have access to capital markets.

[9] In our period of analysis, the real interest rate has declined substantially in the US (Holston et al., 2017; Kroen et al., 2022).

Literature review. Our paper contributes to the literature on the determinants and dynamics of wealth inequality, with a particular focus on the evolution of wealth distribution in the United States in recent decades. Empirical studies have uncovered a substantial increase in wealth inequality in the US (e.g. Piketty, 2014; Piketty and Zucman, 2014; Kuhn and Rios-Rull, 2013; Jordá et al., 2019). This evolution has been attributed to changes in taxation and wage inequality; heterogeneity in portfolios and asset returns; as well as automation (e.g. Kaymak and Poschke, 2016; Benhabib et al., 2019; Straub, 2019; Hubmer et al., 2021; Moll et al., 2022; Brendler et al., 2024).

These studies typically employ frameworks where growth is either absent or exogenous, making a high real return on assets the primary driver of wealth inequality. A contribution of our's lies in highlighting a different potential source of wealth inequality: the nexus of market power and growth. By endogenizing both, we account for variations in both components of the return gap, $r-g$. From this, two novel insights emerge. First, we demonstrate that with endogenous growth, wealth inequality can increase even if asset returns decline. Second, we reveal that the divergence in saving motives between rich and poor – precautionary versus intertemporal – is the key mechanism through which an increase in the return gap exacerbates inequality.

That saving of the poor is less sensitive to movements in the interest rate than those of the rich is acknowledged, but not formally established, in the literature. Olivi (2018), for instance, shows in a continuous-time setting that a household far away from a period in which the liquidity constraint binds, is more sensitive to movements in the interest rate than those close by.[10] However, as the duration of time to when the constraint binds is endogenous, the result is an indication of this property, rather than an established fact. Relatedly, Achdou et al. (2022) illustrates numerically that the policy functions for saving pivot upwards around the liquidity constraint in response to a rise in the real interest rate, displaying a more pronounced change for the wealthy relative to the poor.[11] While this paper does not either formally establish a weaker interest-sensitivity of the poor, our numerical results align with this pattern, explaining the differential response of saving along the wealth distribution. Moreover, we provide an economic argument to as to why this feature is crucial for the return gap to affect wealth inequality.

The rise in markups in the US, alongside other indicators of market power such as increased market concentration, has been extensively documented in recent studies (e.g. De Loecker et al., 2020; Autor et al., 2020b; Covarrubias et al., 2020; Bajgar et al., 2019; Hall, 2018). Similarly, the slowdown in productivity growth in the US over recent decades has been the focus of numerous analyses (e.g. Gordon, 2012; Fernald, 2014; Cette et al., 2016).[12]

A recent stream of papers have used endogenous growth models to interpret this contemporary

[10] See also Achdou et al. (2022), proposition 5, for a simplification of this result relating to a constant real interest rate. Fahri et al. (2022) develops a similar result in discrete time.

[11] See figure 7 on page 66 in Achdou et al. (2022).

[12] Syverson (2017) shows that this is not due to mismesurement issues.

increase in market power and slowdown in productivity. The explanations highlight different potential sources driving both facts: falling overhead costs (Aghion et al., 2023), adoption of intangible inputs (De Ridder, 2023), declining population growth (Peters and Walsh, forthcoming). Akcigit and Ates (2023), Olmsted-Rumsey (2022), Cavenaile et al. (2020) have focused on the slowdown in knowledge diffusion which has hampered the possibility of technological laggard firms to catch up with the industry leaders, thereby increasing concentration and markups.

Our contribution to the literature is twofold. First, we introduce household heterogeneity to examine how the interaction between competition and growth shapes wealth inequality. Second, our endogenous knowledge spillovers channel provides a microfoundation for the key mechanisms in Akcigit and Ates (2023), Olmsted-Rumsey (2022), and Cavenaile et al. (2020). These studies attribute the dynamics of market power and productivity to an exogenous decline in knowledge diffusion. In contrast, we propose that the slowdown in knowledge diffusion arises endogenously as a by-product of increasing market power.[13]

2 Model

The economy is populated by a continuum of intermediate goods producing firms, a representative final good producer, and households. The households are heterogeneous, and either feature two types of agents (section 2.3), or exist on a continuum in the presence of uninsurable income risk (section 2.4). Aggregate labor supply will be exogenous and normalized to one. Time is continuous and denoted $t \in [0, \infty)$. We start by describing the production side of the economy, as this part does not change depending on the characteristics of the households. As there is (endogenous) growth, the model is stationarized in terms of wages, w_t.

2.1 Firms

2.1.1 Final good producers

The final good, Y_t, is produced under perfect competition using a continuum of intermediate goods, y_{jt}, aggregated by a constant elasticity of substitution (CES) production function

$$Y_t = \left(\int_0^1 y_{jt}^{\alpha} dj \right)^{\frac{1}{\alpha}}. \tag{1}$$

[13]Our approach to modeling competition and growth builds on the IO tradition of static oligopolistic models of innovation (e.g., Dasgupta and Stiglitz, 1980; Sutton, 1991; Impullitti et al., 2022), and on their extension to dynamic equilibrium frameworks (Peretto, 1996; Impullitti and Licandro, 2018).

The final good is the numéraire, so its price is set to one. The solution to the final good producer problem gives indirect demand for intermediate goods, y_{jt}, as

$$p_{jt} = \left(\frac{Y_t}{y_{jt}}\right)^{1-\alpha}.$$

Total sales of the final good producer are equal to total consumption expenditure from households, $Y_t = C_t$.

2.1.2 Intermediate goods producers.

Each intermediate good is produced within a sector, j, under oligopolistic competition. There are n_t firms competing in each intermediate goods sector, j, playing a dynamic Cournot game for market shares. Goods produced in the same sector are perfectly substitutable, and are produced using labor according to the technology

$$q_{ijt} = z_{ijt}^{\eta}\ell_{ijt}, \quad \eta > 0,$$

where ℓ_{ijt} is the amount of labor employed by firm i in sector j at time t, q_{ijt} is the quantity produced by this firm, and z_{ijt} is its productivity.

Firms may also allocate labor to innovation, h_{ijt}, in order to improve their productivity using the technology,

$$\dot{z}_{ijt} = A\kappa_{ijt}h_{ijt},, \quad A > 0. \tag{2}$$

The productivity of research, κ_{ijt}, is a combination of appropriable knowledge and public knowledge defined as

$$\kappa_{ijt} = z_{ijt}^{1-\beta}z_{jt}^{\beta}, \quad \beta \in [0,1], \tag{3}$$

where $z_{jt} = \int_i z_{ijt}di$, is the aggregate productivity in sector j. The parameter β determines the extent of firm i's learning from the knowledge of other firms in the sector.[14] When $\beta = 0$, knowledge is a private good, and firm i's R&D workers rely only on the firm's current state of technology to improve upon it. When instead $\beta = 1$, knowledge is a public good, and firm i's R&D workers have full access to the knowledge of all other firms in the economy when performing their innovation activity. Thus for any positive values of β, firms learn from each other via knowledge spillovers. The innovation technology exhibits constant returns to knowledge, which is the factor that accumulates. This feature is the source of endogenous growth in the model. In the symmetric equilibrium, on which we will focus, we have $\kappa_{ijt} = \kappa_t = z_t n_t^{\beta}$.

We characterize the open loop Nash equilibrium of this game. At time t a firm i in sector j

[14] This flexible representation of knowledge spillovers is similar to that in Peretto (1996).

choses quantity and innovation to solve

$$v_{ijt} = \max_{\{q_{ijs}, h_{ijs}, s \in [t,\infty)\}} \int_t^\infty e^{-\int_t^s (r_\tau + \delta) d\tau} [(p_{js} - z_s^{-\eta} w_s) q_{ijs} - h_{ijs} w_s] ds,$$

subject to

$$p_{jt} = \left(\frac{Y_t}{y_{jt}}\right)^{1-\alpha},$$
$$y_{jt} = \tilde{y}_{jt} + q_{ijt},$$
$$\dot{z}_{ijt} = A\kappa_{ijt} h_{ijt},$$

where $\hat{y}_{jt}$ is the quanitity produced by the firm's competitors in sector j, r_t is the real interest rate, and δ is a Poisson arrival rate of bankruptcy. Throughout we will let $R_t = r_t - \eta g_t$ denote the *return gap*.

The associated, stationarized, Hamilton-Jacobi-Bellman (HJB) equation for any firm within a sector is[15]

$$(R_t + \delta) v(\hat{z}_t, \hat{Z}_t, n_t) = \max_{q_t, h_t} \Big\{ (p\left(\frac{\hat{y}_t + \hat{q}_t}{\hat{Y}_t}\right) - \hat{z}_t^{-\eta}) \hat{q}_t - h_t + v_{\hat{z}}(\hat{z}_t, \hat{Z}_t, n_t)[A\hat{\kappa}_t h_t - g_t \hat{z}_t] \\ + v_{\hat{Z}}(\hat{z}_t, \hat{Z}_t, n_t)[A\hat{\kappa}_t H_t - g_t \hat{Z}_t] + v_n(\hat{z}_t, \hat{Z}_t, , n_t)\dot{n}_t \Big\}, \quad (4)$$

with $\hat{x}_t = x_t / w_t$, for $x_t \in \{q_t, \tilde{y}_t\}$, $\hat{z}_t = z_t / w_t^{1/\eta}$, and $\hat{\kappa}_t = \kappa_t / w_t^{1/\eta}$. Subscripts of the value function indicates a partial derivative. The first order condition with respect to h_t is

$$v_{\hat{z},t} A\hat{\kappa}_t = 1,$$

where the the dependence of $v_{\hat{z},t}$ on the state variables has been suppressed for clarity of exposition.

The first order condition with respect to $\hat{q}_t$ yields

$$\theta_t = \hat{z}_t^{-\eta}, \quad \text{with} \quad \theta_t = \frac{n_t - 1 + \alpha}{n_t}. \quad (5)$$

Since the final good is the numéraire, with a price normalized to one, the prices of the intermediate goods are one as well, and θ_t is the *inverse of the markup*. A higher number of firms and a higher elasticity of substitution across varieties lead to lower markups. Equation (5) implies that the wage

[15]For simplicity we omit the firm and sector indicators. The details of the derivation are in Appendix 7.1.1.

is given by

$$w_t = \theta_t z_t^{\eta}. \tag{6}$$

The wage is negatively affected by rising markups, as higher markups reduce θ_t, and positively influenced by productivity growth, $\dot{w}_t/w_t = \eta \dot{z}_t/z_t$, which impacts future wage prospects.

Combining the optimality conditions together with the envelope theorem (see appendix 7.2), gives the growth equation

$$g_t = \frac{\dot{z}_t}{z_t} = \frac{1}{\beta}\left[A\eta n_t^{\beta}\ell_t - (R_t + \delta)\right], \tag{7}$$

with, again, $R_t = r_t - \eta g_t$. Here we can clearly see that growth is positively linked to the size of the firms, ℓ, which is proportional to the size of the market, and to the knowledge spillovers, n^{β}. As we will see in details later, the effect of competition on growth operates via these two channels. Using the innovation technology we obtain the equilibrium innovation rate as $h_t = g_t/An_t^{\beta}$. Lastly, combining the optimality conditions with the HJB equation gives the stationarized value of the firm,

$$v_t = \frac{\frac{1-\theta_t}{\theta_t}\ell_t - \frac{g_t z_t}{A\kappa_t} + v_{n,t}\dot{n}_t}{R_t + \delta}, \tag{8}$$

where $\pi_t = ((1-\theta_t)/\theta_t)\ell_t - g_t z_t/(A\kappa_t)$ is the stationarised flow of profits. The value of the firm is given by the present discounted value of profits.

2.1.3 Free entry and labor market clearing

Intermediate goods entrepreneurs may freely enter the market by paying an upfront entry cost, ϕw_t. The value of setting up a firm is then $v_t w_t$.[16] Free entry implies that $v_t = \phi$, such that[17]

$$\phi = \frac{\frac{1-\theta_t}{\theta_t}\ell_t - \frac{g_t z_t}{A\kappa_t}}{R_t + \delta}. \tag{9}$$

The mass of firms evolves according to $\dot{n}_t = m_t - \delta n_t$, where m_t is the measure of entrants. For the labor market to clear, we require that the total amount of labor allocated to production, $n_t\ell_t$, innovation, $n_t h_t$, and the entry process of new firms, ϕm_t, equals that of labor supplied; namely one. Accordingly, $1 = n_t(\ell_t + h_t + m_t\phi)$. Thus, the supply side of the model, together with the (labor)

[16] Recall that v_t is stationarized using w_t.
[17] Notice that as $v_t = \phi$, we have $\dot{v}_t = v_{n,t}\dot{n}_t = 0$.

market clearing condition, is succinctly captured by the three equations

$$g_t = \frac{1}{\beta}\left(A\eta n_t^{\beta}\ell_t - (R_t+\delta)\right), \tag{10}$$

$$(R_t+\delta)\phi\theta_t = (1-\theta_t)\ell_t - \theta_t \frac{g_t}{n_t^{\beta}A}, \tag{11}$$

$$1 = n_t\left(\ell_t + \frac{g_t}{n_t^{\beta}A} + m_t\phi\right), \tag{12}$$

in the three unknowns g_t, ℓ_t, and m_t.

This part of the model – summarized by equations (10)-(12) – also implies the demand for capital, $n_t\phi$. The supply of capital comes instead from the households. The real interest rate, r_t, will link these two parts of the economy, ensuring overall market clearing, and determining the general equilibrium.

2.2 General equilibrium

Before specifying the behavior of households, it is instructive to observe how the interest rate clears overall markets. Suppose the (stationary) *aggregated* budget constraint of the households is given by

$$a_t R_t + 1 - c_t = \dot{a}_t,$$

where a_t represents aggregate assets.[18] Market clearing implies that

$$c_t = \frac{Y_t}{w_t} = n_t \frac{\ell_t}{\theta_t}.$$

The free entry condition in equation (11) can be rewritten as

$$\begin{aligned} n_t \frac{\ell_t}{\theta_t} &= n_t(\ell_t + h_t) + n_t\phi(r_t - \eta g_t + \delta), \\ &= 1 - m_t\phi + n_t\phi(r_t - \eta g_t + \delta), \\ &= 1 - \dot{n}_t\phi + n_t\phi(r_t - \eta g_t), \end{aligned}$$

[18]Two things ought to be noted here. First, recall that the model is stationarized using wages, and that aggregate labor supply is equal to one. Second, this budget constraint is general and holds under any type of heterogeneity.

where we have used the market clearing condition for labor, and the law of motion for n_t. Combining this with the budget constraint gives

$$(a_t - n_t\phi)R_t = (\dot{a}_t - \dot{n}_t\phi).$$

Thus, the equilibrium real interest rate must be such that $a_t = \phi n_t$, and that aggregate asset supply, a_t, driven by households' saving decisions, equals asset demand, $n_t\phi$, determined by the market value of firms. However, to obtain this interest rate, we must have a theory of the determination of a_t, which is the topic of the next section.

2.3 Households: Capitalists and workers

Capitalist solve the dynamic consumption problem

$$\max_{C_t^c} \int_0^\infty \ln C_t^c e^{-\rho t} dt$$
$$s.t.\ \dot{A}_t = r_t A_t - C_t^c$$

where C_t^c denotes total consumption expenditures of capitalists, and A_t is the their asset holdings. Let x_t denote a stationarized variable, such that $x_t = \frac{X_t}{w_t}$, with $\frac{\dot{w}_t}{w_t} = \eta g_t$. The associated, stationarized, HJB equation is given by

$$\rho v(a_t) = \max_{c_t} \{\ln(c_t^c) + v'(a_t)\dot{a}_t\}, \quad \text{with} \quad \dot{a}_t = R_t a_t - c_t^c$$

The first order condition to this problem, together with the envelope condition, gives the Euler equation

$$\frac{\dot{c}_t^c}{c_t^c} = R_t - \rho. \tag{13}$$

Workers' consumption expenditures are $C_t^w = w_t$, such that $c_t^w = 1$. The price of the final good one and total expenditures is therefore $c_t = c_t^c + 1$. Workers are hand-to-mouth, and endowed with one unit of labor that they supply inelastically.

Balanced Growth Path (BGP). The Euler equation implies that on the BGP $R = \rho$. Thus, the BGP of the model is characterized by the three equations

$$g = \frac{1}{\beta}\left(A\eta n^{\beta}\ell - \rho - \delta\right), \tag{14}$$

$$(\rho+\delta)\phi\theta = (1-\theta)\ell - \theta\frac{g}{n^{\beta}A}, \tag{15}$$

$$1 = n\left(\ell + \frac{g}{n^{\beta}A} + \delta\phi\right), \tag{16}$$

determining the three endogenous variables ℓ, n and g. Notice that in this case, capitalists' asset supply, a, is indeterminate, and uniquely pinned down via capital demand, ϕn.

2.3.1 Exploring the mechanism

In this section, we provide insights on the mechanisms at work, using both numerical illustrations and analytical results on the link between competition, growth and inequality. For the numerical results, we consider an increase in the entry cost that replicates the rise in the markup observed in the post-1980 period shown in figure 1. While we primarily view these results as indicative, and a tool of gaining insights, the precise calibration used is outlined in section 3.1.

The BGP. As stated above, we generate the increase in the markup via an increase in the entry cost. The reason a higher entry cost leads to a higher markup is straightforward and operates via the free entry condition: when the cost of entry increases, the value of the firm needs to be higher. This only occurs when there are less firms in the market and, consequently, firm size and profitability is higher. This relationship is depicted in the top right panel of figure 2, and is formalized in proposition 1.

Proposition 1. *If the markup is sufficiently large, $1/\theta > 1+\eta/\beta$, an increase in the entry cost, ϕ, reduces the number of firms and increases the markup.*

Proof. See appendix 7.3.1. □

The condition used in proposition 1 is not necessary, but sufficient, and holds in our numerical simulation.[19] Moreover, the condition is also sufficient to ensure positive equilibrium profits.

Figure 2 reveals additional important features of the economy. First, growth has an inverted U-shape, but is for the most part declining along with the rise in entry costs. Second, innovation efforts, that is labor allocated to R&D, also follows an inverted U-relationship, displaying a more pronounced relationship than in the case of growth. Third, firm size, measured as labor in variable

[19] Following the calibration in section 3.1, $\eta/\beta \approx 0.215$.

production, ℓ, increases. Fourth, the number of firms, n, declines. And lastly, the wealth-to-income ratio rises. What economic forces underpin these patterns?

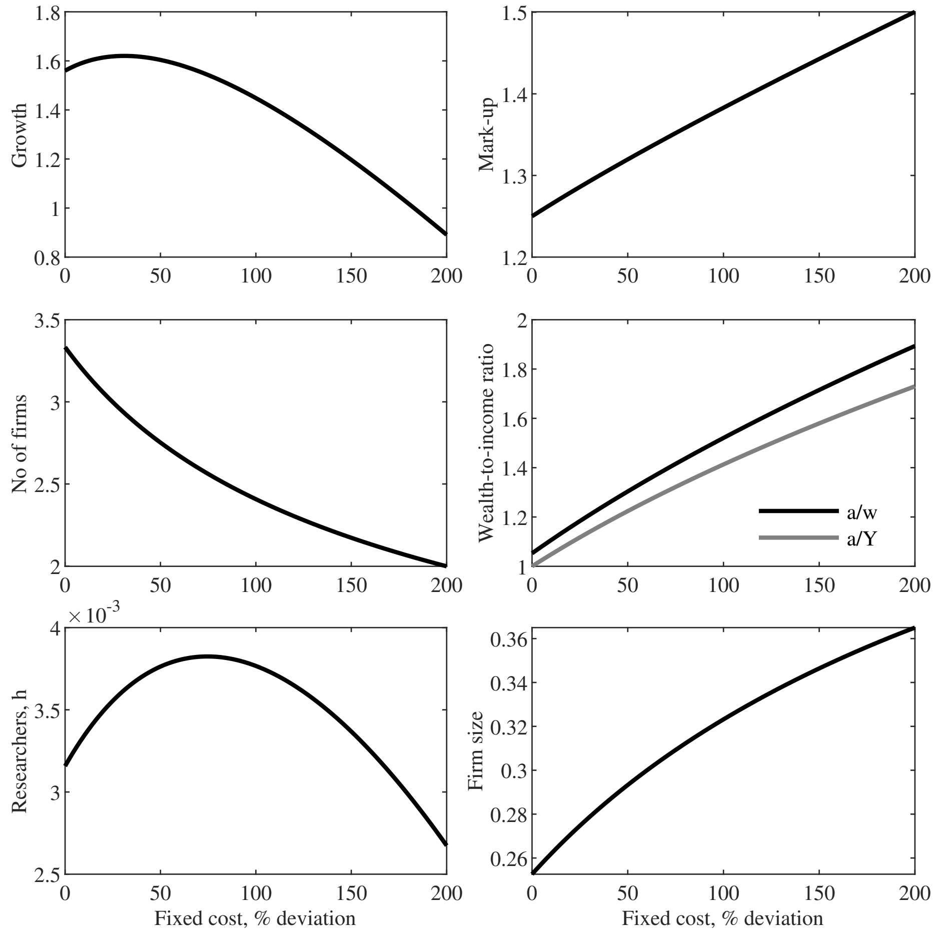

Figure 2: Capitalist-worker model – comparative statics.

Notes. The relationship between entry costs (x-axis) and equilibrium quantities (y-axis) in the capitalist-worker model of section 2.3. Entry costs are illustrated as log point deviations from baseline. The calibration follows that of section 3.1. The increase in the entry cost replicates the rise in the markup observed in the post-1980 period.

While the effect of a higher entry cost on the number of firms is straightforward and intuitive, its impact on growth is more complex. The reason is that the growth rate, g, the number of firms, n, and the firm size, ℓ, are jointly determined, and their relationship is shaped by several forces. Thus, to gain insight it is useful to differentiate the growth equation in (10) with respect to the entry cost

$$\frac{\partial g}{\partial \phi} = \frac{A\eta}{\beta} n^{\beta} \left(\underbrace{\frac{\partial \ell}{\partial \phi}}_{\text{firm size}} + \underbrace{\beta \frac{1}{n} \frac{\partial n}{\partial \phi} \ell}_{\text{spillovers}} \right). \tag{17}$$

Equation (17) shows that the effect of a change in the entry cost on growth depends on two forces: the *firm size* which depends on market size, and the *knowledge spillovers*. The first of these arises as there is a complementarity between innovation and firm size in production; larger firms produce more, and any incremental increase in productivity, z, has therefore a larger impact on (total) profits. That is, innovation concentrated to a *smaller number* of *larger firms*, leads, ceteris paribus, to higher growth.[20] The latter force, spillovers, however, arises because firms learn from each other in their innovation activity. In particular, with $\beta > 0$, a larger number of firms expands the body of public knowledge on which each firm draws, thereby increasing research efficiency. Via this knowledge spillovers channel, a *larger number* of firms increases innovation and growth. As a rise in the entry cost reduces the number of firms, but leaves the remaining firms bigger, the firm-size effect tends to be positive, while the spillover effect is negative. And since these two channels predict an opposite effect of competition on growth, this relationship is potentially non-monotonic and depends on the channels' relative strength.[21]

This tradeoff is clearly visible in figure 2, and is (partly) formalized in proposition 2.

Proposition 2. *The effect of changes in the entry cost on growth, and the relationship between competition and growth is ambiguous. When knowledge is a full public good* $\beta = 1$*, if the elasticity of the number of firms to the entry cost* $|\varepsilon_{n,\phi}| = |\frac{\partial n}{\partial \phi}\frac{\phi}{n}| < 1$*, higher entry costs are associated with lower growth.*

Proposition 2 states that if the elasticity of the number of firms to the entry cost is below one, the model predicts a positive relationship between competition and growth. This prediction is in line with what we observe in the period of interest in the US, as well as in the top left panel of figure 2.

What does the (sufficient) condition $|\varepsilon_{n,\phi}| < 1$ entail? We know that n decreases alongside an increase in ϕ. The condition then states that n cannot decrease "too fast". The reason is that while a reduction in n leads to a ceteris paribus decline in spillovers (and thereby growth), it also reduces the amount of labor allocated to the entry cost, freeing up labor being used for other activities, including innovation. The condition is sufficient to ensure the latter cannot happen.[22] Indeed, $\varepsilon_{n,\phi} > -1$ is

[20] A decrease in the number of firms, n, increases the size of each remaining firm, ℓ, thereby increasing innovation and growth. In particular, combining equations (14) and (16) leads to $1/n + (\rho+\delta)/(A\beta n^{\beta}) = (1+\eta)\ell$, which shows that lowering the number of firms increases the size of each firm on the market ℓ. Growth is directly tied to ℓ via the growth equation in (14).

[21] The non-monotonic relationship between competition and growth can be also obtained in 'step by step' Schumpeterian growth models (Aghion et al., 2001, 2005). Our framework is simpler but more general has it does not impose a duopolistic market structure and naturally accommodates entry. The non-monotonicity is important because there is no empirical consensus on the relationship between competition and growth. Bloom et al. (2016) argue that Chinese import competition induced innovative activity in exposed domestic sectors in Europe [see also Coelli et al. (2022), Gorodnichenko et al. (2010), and Iacovone (2012)], whereas Autor et al. (2020a) argue the opposite, using data on U.S. firms and sectors [see also Hashmi (2013) and Hombert and Matray (2015)], and yet a third set of papers including Aghion et al. (2017) finds ambiguous results.

[22] Notice that labor allocated to entry costs are given by $\delta n\phi$. The derivative of the logarithm of this expression is

ensures that the amount of labor allocated to entry costs increases, which is a sufficient condition for the result.

Figure 2 illustrates that the firm size effect is more pronounced at lower levels of entry costs, while the spillover channel becomes dominant at higher levels. This dynamic creates a mild hump-shaped relationship between entry costs and growth. The relationship between entry costs and innovation follows a similar hump-shaped pattern, albeit more distinctly. Equilibrium innovation, defined as $h = g/An^{\beta}$, highlights how the negative effects of market power on knowledge spillovers manifest both directly, by reducing R&D productivity, and indirectly, by impacting growth. These combined effects result in a more pronounced inverted-U relationship between innovation and entry costs. Our model can therefore accommodate scenarios where increasing innovation effort leads to lower growth due to worsening of R&D productivity.[23] This aligns with recent empirical finding of declining research productivity in the US in recent decades (Bloom et al., 2020). Our findings offer another insight: the rise in market power may play a pivotal role in diminishing innovation efficiency by weakening knowledge spillovers.

Lastly, as seen in the middle right panel of figure 2, the wealth-to-labor-income ratio, a, increases. To understand why, recall that the stationarized value of the capitalists' assets, a, is equal to the value of the firms, nv, which, in turn, equals total entry cost expenditures, $n\phi$ (see section 2.2). Thus, $a = n\phi$, and as a consequence

$$\frac{\partial(n\phi)}{\partial\phi} = n\left(\varepsilon_{n,\phi} + 1\right), \tag{18}$$

where $\varepsilon_{n,\phi}$, again, is the elasticity of the number of firms with respect to the entry cost. This is positive provided that $\varepsilon_{n,\phi} > -1$, which is the same condition as in proposition 2. The reason is not far-fetched: when $\varepsilon_{n,\phi} > -1$, the number of firms declines when ϕ rises, but its decline does not outpace the rise in ϕ. As a consequence, the product $n\phi$ increases, leading to a rise in capital demand, and ultimately the capitalists' wealth. Thus, the same condition that ensures that a rise in ϕ does not lead to less labor allocated to entry costs, and that a rising entry cost hampers growth, also dictates that wealth-to-income increases.

Transitional dynamics. While the above analysis makes clear that the wealth-to-income ratio *must* rise as entry costs increase, it does not address the economic forces as to why and how this happens. And to this end, a comparative statics analysis falls short; on any balanced growth path, (stationarize) consumption is constant, and the return gap is equal to ρ. Asset supply is therefore indeterminate, and set to equal asset demand. But how do capitalist acquire this new-found wealth?

$\varepsilon_{n,\phi} + 1$. Hence, if $\varepsilon_{n,\phi} > -1$, *more* labor are allocated to entry costs, and not less.

[23] If, for instance, entry costs double, growth declines, yet innovation efforts increase. This is visible in figure 2.

Figure 3 illustrates the transitional dynamics of the capitalist/worker economy.[24] For this example we use a subset of parameters derived from the calibration of the more general model in section 2.4. The top left panel shows that the return gap, R_t, starts at its BGP value of ρ, rises, and then asymptotes back. That is, its movements are transitory. The real interest rate itself, however, displays a similar pattern, but returns instead to a new permanently *lower* value.[25] The reason for this is the growth rate; on the new BGP, growth is permanently lower. Hence, for the return gap to asymptote to ρ, the real rate must settle down at a new reduced value. Lastly, the wealth-to-income ratio permanently rises, and eventually reaches its new BGP value.

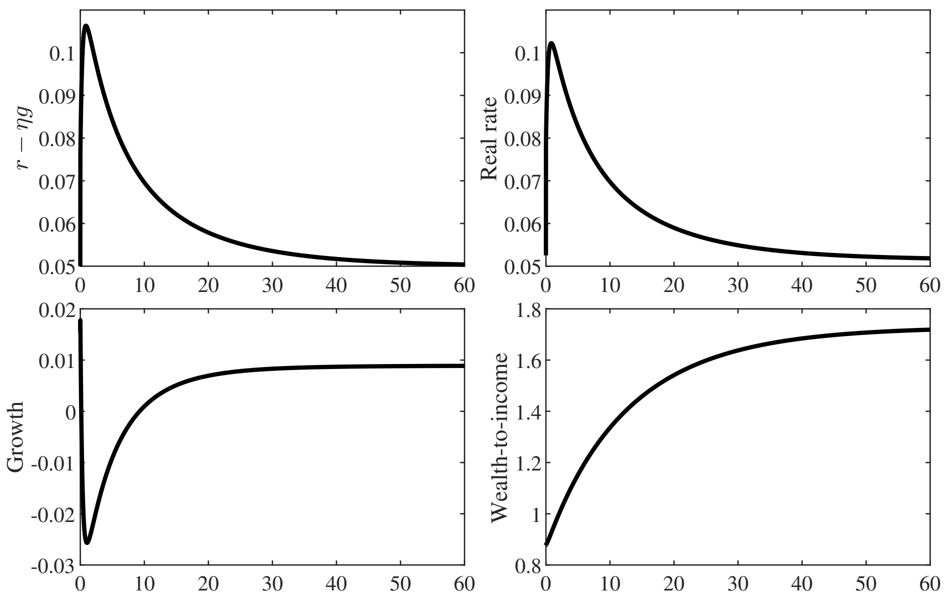

Figure 3: Capitalist-worker model – transitional dynamics.

Notes. The figure illustrates the transitional dynamics arising from an increase in the entry cost that renders a new markup of 1.5. Time in years is on the x-axis. The calibration follows that of section 3.1.

The transitional pattern of the return gap is the culprit explaining this rise in wealth. To see this, conjecture that capitalists' consumption is a constant share of wealth, $c_t^c = \gamma a_t$. Then the budget constraint suggests that $\dot{a}_t = (R_t - \gamma)a_t$. The conjecture is consistent with the Euler equation if $\gamma = \rho$. Hence,

$$a_t = a_0 e^{\int_0^t (R_s - \rho) ds}.$$

Thus, for any initial BGP at time $t = t_0$, and assuming the economy has reached its new BGP in

[24] We consider an event where the entry cost incrementally increases to a value that renders a markup of 1.5 on the new BGP. See appendix 7.4.2 for details.

[25] The real interest rate starts at 5.25% and ends at 5.14%.

$t = t_1$, we have

$$a_{t_1} = a_{t_0} e^{\int_{t_0}^{t_1} (R_s - \rho) ds}.$$

That is, the *permanent* rise in wealth materializes as the return gap *temporarily* exceeds ρ.

We now have the full set of ingredients to present an internally consistent story of the link between competition, growth, and wealth inequality. Higher entry costs lead to less entry, higher markups and increased capital demand. This exerts upward pressure on the real interest rate, increasing the return gap, and encourages asset supply. With less firms, markups increase, and, due to spillovers, growth subdues. Over time, the reduction in growth leads to increased desire to save for intertemporal reasons, and the pressure on the real interest rate alleviates. Thus in the long run, the real interest rate is left permanently lower, assets and wealth inequality permanently higher, growth subdued, and competition hampered.

2.4 Incomplete markets

To further explore the impact of reduced competition on inequality, we extend the household sector to accommodate market incompleteness. In this framework, all households are ex ante identical, but ex post heterogenous, due to idiosyncratic, uninsurable, risk. This implies that all households are "workers" to some extent, but that no household is entirely a capitalist; indeed the households exist on a continuum of wealth, with a wide disparity of affluence. This extension allows the model captures a more nuanced representation of wealth dynamics, and aligns more closely with empirical evidence on wealth inequality. Beyond offering predictions for the evolution of inequality, the primary distinction from the simpler framework lies in the determination of the interest rate, which is the price connecting firms and households. In the presence of incomplete markets and uncertainty, households save both for intertemporal consumption smoothing and as a precautionary measure against unforeseen risks.

With incomplete markets, the stationarized HJB equations for the households are now

$$\rho_s v(a,s) = \max_c \{\ln c + v'(a,s)\dot{a} - \sum_{s' \in S} \lambda_{s',s}(v(a,s) - v(a,s'))\}, \tag{19}$$

$$\text{with} \quad \dot{a} = y_s + aR - c, \quad \text{and subject to} \quad v'(0,s) \geq \frac{1}{y_s}, \quad \forall s \in S,$$

where S is a set of possible exogenous states, and $\lambda_{s',s}$ are the associated Poisson arrival rates. That is, a household with wealth a, belonging to the exogenous state s, can consume, c, or save, $\dot{a}$, taking into account that she may transition to state s' with (conditional) arrival rate $\lambda_{s',s}$. The final restriction is a boundary condition that rules out borrowing (see Achdou et al. (2022)).

We consider six states, $S = \{(y_i, \rho_j) : i \in \{e,u\}, j \in \{l,m,h\}\}$, where l, m, h are abbreviations for low, medium, and high values of the discount factor; and e, u, for employment and unemployment. That is, households may be unemployed and have high patience; employed with low patience, and so on. The difference in patience is included to obtain an empirically relevant wealth distribution (cf. Krusell and Smith (1998)).

While the model does not have a BGP in the sense that individual (stationarized) quantities are constant, it does feature a stationary cross-sectional distribution, $f(a,s)$, which implies that aggregate quantities are constant. In particular, solving equation (19) gives rise to policy functions, $c = g(a,s)$ and $\dot{a} = h(a,s)$, that together with the Poisson arrival rates maps out the Kolmogorov forward equation (see Achdou et al. (2022)),

$$\dot{f}_t(a,s) = -\frac{\partial[f_t(a,s)h(a,s)]}{\partial a} - \sum_{s' \in S} \lambda_{s',s}(f_t(a,s) - f_t(a,s')),$$

determining the law of motion of the cross sectional distribution, f_t. The stationary cross-sectional distribution then satisfies

$$0 = -\frac{\partial f(a,s)h(a,s)}{\partial a} - \sum_{s' \in S} \lambda_{s',s}(f(a,s) - f(a,s')). \tag{20}$$

Given this distribution, we may obtain aggregate asset supply as

$$A^s = \sum_{s \in S} \int_a a f(a,s) da,$$

and the equilibrium return gap, R, must satisfy

$$A^s = n\phi.$$

Thus, in contrast to the capitalist-worker framework, there is no immediately available equilibrium R. Instead, we must solve equation (19), use the associated policy function together with equation (20) to find the cross-sectional distribution, and integrate it to retrieve asset supply. The equilibrium return gap is then determined as the rate that ensures that asset supply equals asset demand. The algorithm used to achieve this is outlined in appendix 7.4.

3 Numerical analysis

3.1 Calibration

The model is calibrated to the US economy in 1980 at an annual frequency. The discount rate, ρ, is set at 0.05. This implies a real return of 5% in the capitalist/worker model, and is in between common estimates of the real risk free rate and stock market returns.[26] The Poisson arrival rate of bankruptcy, δ, is set to 0.14, giving rise to a BGP entry rate of 14% matching that observed in Business Dynamics Statistics. We target an elasticity of substitution of 1.5, which implies a value for α of 0.394, consistent with the macro estimates in Feenstra et al. (2018). Bloom et al. (2013) provides estimates for knowledge spillovers finding that in highly innovative sectors the impact of other firms R&D on firms' citation-weighted patents ranges from 0.58 to 1.4. Bloom et al. (2020) reports a spillover value of 0.8 for semiconductors and Peters (2022) finds a value of 0.71. We select a value within this narrower interval, setting $\beta = 0.77$, which is roughly the average.[27]

The parameters A, η, and ϕ are jointly set to match a BGP growth rate of 1.56%, corresponding to the 1980 estimate in Fernald (2014); a markup of 25%, as found in De Loecker et al. (2020); and a share of R&D to GDP at 1% matching the share of GDP accounted for by business R&D expenditure from the National Science Foundation Scientists and Engineering Indicators.

For the incomplete markets model, we further need to calibrate the arrival rates, $\lambda_{s',s}$, the income process, y_s, and the values of ρ_s. To do this, we assume that the arrival rate of employment/unemployment is independent of current patience. Similarly, the arrival rate of patience is independent of current employment status. For the arrival rates for employment/unemployment, we target a BGP unemployment rate of 6% (Bureau of Labor Statistics), and an unemployment duration of 12 weeks (Westcott and Bednarzik (1981)). These two moments discipline the parametrization of the arrival rate of employment, λ_{eu}, and of unemployment, λ_{ue}. The income process, y_s, is such that a household that becomes unemployed loses 50% of its income for the duration of an average unemployment spell, implying an annual income loss of 12.5%, relative of the employed. The income of the employed is set such that expected labor income, stationarized by aggregate wages, is equal to one.

For patience, there are a total of six arrival rates, three for the employed and three for the unemployed. We partly follow Krusell and Smith (1998) and assume that agents cannot transition from high to low patience, nor from low to high, without first visiting the intermediate state, and that transitioning to the high or the low state are equally likely. This leaves only two arrival rates to be parametrized; the transition probabilities into the intermediate state, $\lambda_{hm} = \lambda_{lm}$, and the transition

[26]In the model, all assets are stocks. But since we abstract from aggregate risk, there is no risk premium. Hence we consider a real return of 5% to be a reasonable compromise.

[27]Although Peters (2022) uses historical German data, it proves valuable in narrowing the range.

out of the intermediate state, $\lambda_{mh} = \lambda_{ml}$. Following Krusell and Smith (1998) we discipline the first parameter to insure that the mass of high- and low-patient agents are 10% respectively. This leaves one additional arrival rate to calibrate.

Table 1: Calibration summary

External parameters	Value	Source
CES parameter (α)	0.394	Feenstra et al. (2018)
Discount factor (ρ)	0.05	Annual real return
Spillover parameter (β)	0.77	Bloom et al. (2013)
Bankruptcy rate (δ)	0.14	Census (BDS)
Calibrated parameters	Value	
R&D productivity (A)	0.33	
R&D tech. curvature (η)	0.40	
Entry cost (ϕ)	0.40	
Arrival rate of employment (λ_{eu})	0.8125	
Arrival rate of unemployment (λ_{ue})	0.0519	
Arrival rate of h cond. m (λ_{hm})	See Section 7.5.1	
Arrival rate of m cond. h (λ_{mh})	See Section 7.5.1	
Patience gap (ε)	3.4e(-4)	
Moments	Data (Model)	Source
Markup	25%	De Loecker et al. (2020)
TFP growth rate	1.56%	Fernald (2014)
R&D/GDP	1%	NSF S&E Indicators
Unemployment rate	6%	Bureau Labor Statistics
Unemployment duration	12 weeks	Westcott and Bednarzik (1981)
Mass of medium patient	80%	Krusell and Smith (1998)
Top-10% wealth share	63%	World Inequality Database (2024)
Elasticity of current wealth to wealth 30 years ago	0.71	Clark and Cummins (2015)

Notes. The table summarizes the main calibration. For the “Calibrated parameters” see the main text.

For the values of ρ_s, we assume symmetry, such that

$$\rho_h = \rho + \varepsilon, \quad \text{and} \quad \rho_l = \rho - \varepsilon. \tag{21}$$

Thus, we need to calibrate one arrival rate, as well as ε. To do so, we target a top-10% wealth share of 63%, which is roughly its value in 1980 (see figure 1). Moreover, interpreting 30 years as one generation, we target an elasticity of wealth of a “child” to “father’s” wealth of 0.71 (Clark and Cummins, 2015). Appendix 7.5.1 provides further details. A summary of the calibrated parameters and moments targeted is provided in table 1.

3.2 Results

The left panel of figure 4 shows the distributions of households along the asset dimension for the three different groups of patience. For comparison reasons, all distributions have been normalized to integrate to one. As can be seen, assets for the impatient households are highly concentrated towards the left tail, with a substantial share of households having no assets at all. A similar, albeit less pronounced, pattern emerges for the intermediate group, while the asset holdings of the patient are instead skewed towards the right (middle panel). Indeed, despite only comprising 10% of the population, the patient group holds over 47% of total assets, while the impatient group holds only 2%. Relatedly, almost 6% of the impatient households hold no assets at all, a figure that drops to 2.5% of the intermediate group, and to 0.1% of the patient.

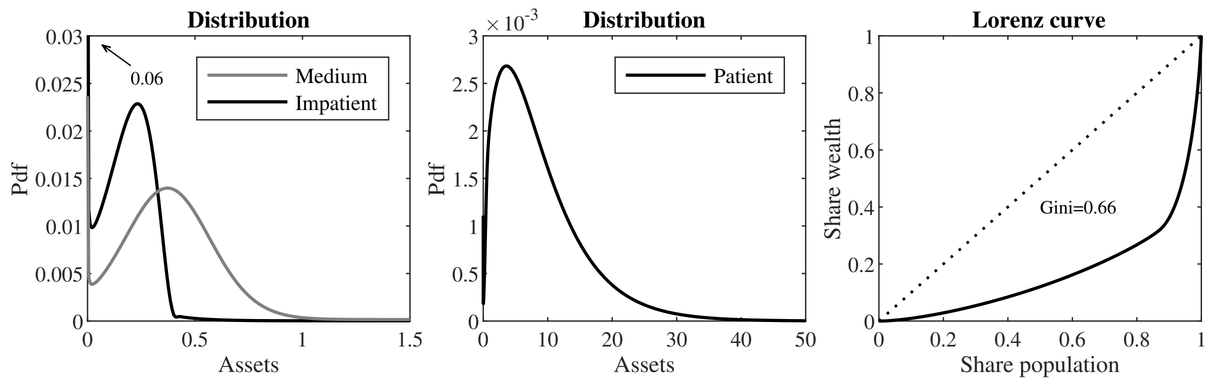

Figure 4: Wealth distribution and the Gini index.

Notes. The left panel displays the equilibrium distribution for the impatient and medium group; the middle panel the distribution for the patient group; and the right panel the overall Lorenz curve.

The rightmost panel of figure 4 shows the overall Lorenz curve of the economy. Assets are quite dispersed, and the associated Gini measure is 0.66. A Gini coefficient of 0.66 is below that commonly reported in US data of 0.79 (see figure 1). However, given that we do not model any income inequality except through unemployment, we believe that the framework captures the main features of wealth inequality that arises for other reasons than those of income dispersion.

In similarly to figure 2 for the capitalist/worker economy, figure 5 illustrates how several BGP measures of the incomplete markets economy change as the entry cost increases. The two top panels show that the two frameworks display remarkable similarities for the key outcomes pertaining to markups and growth.[28] However, while the simpler framework permits a coarse evaluation of

[28] In fact, the frameworks deliver such similar outcomes for all equilibrium quantities, that we refrain from illustrating these, almost duplicate, results.

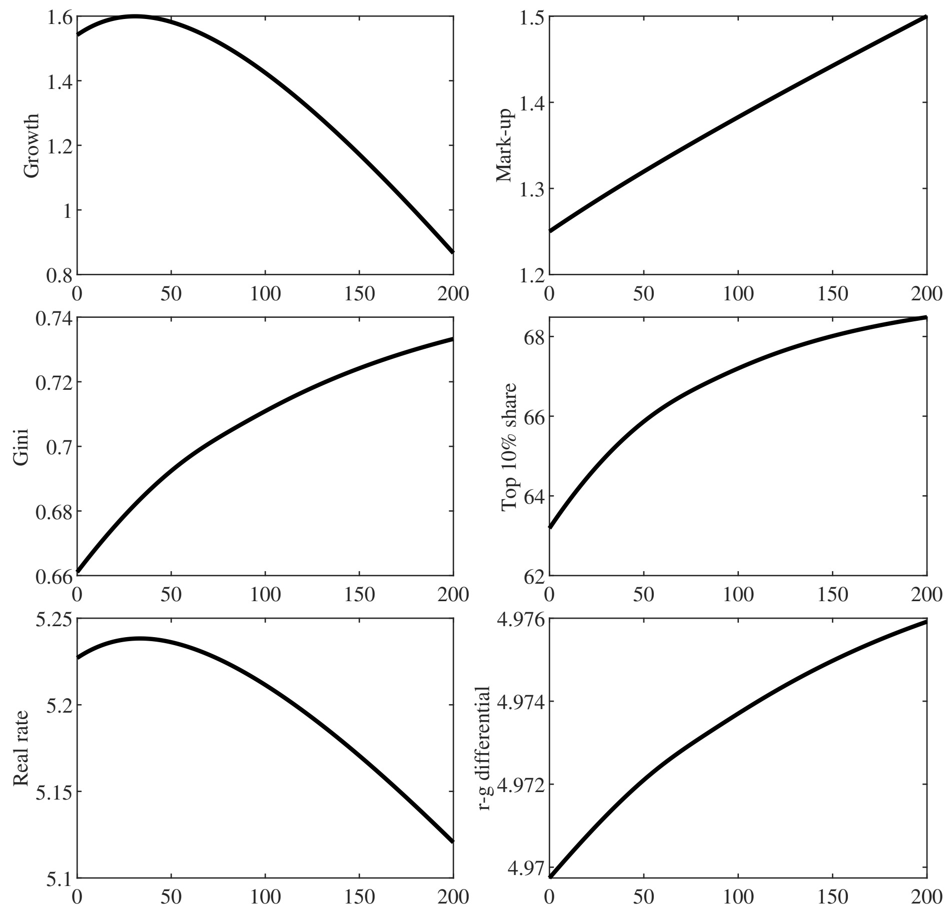

Figure 5: Incomplete markets model – comparative statics.

Notes. The relationship between entry costs (*x*-axis) and equilibrium quantities (*y*-axis) in the heterogenous agents model of section 2.4. Entry costs are illustrated as log point deviations from baseline. The calibration follows that of section 3.1. The increase in the entry cost replicates the markup rise observed in the post-1980 period.

inequality – by comparing the wealth difference between the two types of agents – the incomplete markets setting allows us to make full contact with the empirical evidence on inequality, as it generates a more realistic endogenous distribution of wealth. To this end, the two bottom panels show the evolution of the Gini coefficient, as well as the top 10% share. As can be seen, the model delivers a substantial increase in inequality, with the Gini coefficient rising from 0.66 to 0.73, and the top 10% share from 64% to 70%. The next section will explore the underlying reasons behind these results.

Finally, the bottom two panels display the evolution of the real interest rate, as well as the return gap. As can be seen, the real interest rate declines, while the return gap increases. The reason is that the rise in the entry cost increases capital demand. Further capital demand puts upward pressure on the interest rate, lowers entry, and leads to a slowdown in growth. Over time, the increasing demand for assets peters out, which alleviates some of the pressure on the interest rate. At the same time, the worsened outlook for growth reinforces asset supply for intertemporal reasons.[29] An increased demand for assets, combined with a stronger supply of assets driven by intertemporal considerations, results in a lower real interest rate compared to the previous BGP.

This outcome arises due to an endogenous decline in the growth rate, g. If productivity had remained constant, or followed an exogenous growth trajectory, an increase in market power would have instead resulted in a *higher* real interest rate. Our finding of a lower real rate is particularly compelling given the substantial decline in the real interest rate observed in the United States during the period analysed (see, e.g., Holston et al., 2017). The feedback loop between growth and saving in an endogenous growth framework illustrates how wealth inequality can rise even as returns on asset accumulation diminish.

3.3 Inspecting the mechanism

Figure 6 shows the economy's distributions when the entry cost is increased to render a markup of 1.5. To facilitate comparison, the baseline results from figure 4 are left as transparent lines in the background. As is visible, all distributions are stretched towards the right, leading to an increase in average wealth holding for *all* groups. Yet, however, inequality increases. This section explores why.

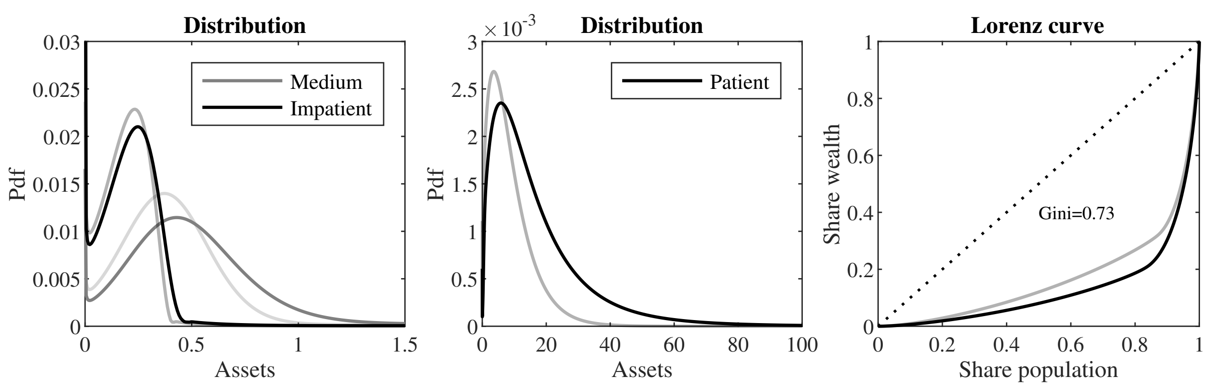

Figure 6: Market power and the distribution of wealth.

Notes. The figure shows the impact of the observed increase in market power on the wealth distribution by age type and on the Gini index.

The left panel of figure 7 displays the *expected* policy function for $\dot{a}(a,s)$ relative to income. That is, saving as a fraction of income, integrated across the six different types of households (employed and patient, unemployed and impatient, and so on). Each marker is adjusted in size to reflect the density of that particular asset value in the baseline distribution. The black markers refer to the baseline distribution, and the grey of those in the new equilibrium.

[29] Lower growth leads to a more bleak outlook for future consumption, which encourages saving on its own.

The saving functions are downward sloping and convex, reflecting the concavity of the consumption function that emerges due to the borrowing constraint alongside the precautionary motive (Carroll and Kimball, 1996). But more importantly, the increase in the return gap pivots the saving function upwards, and more so for the asset-rich. That is, richer agents appear more affected by the rise in the return gap than the poor (cf. Achdou et al. (2022), figure 7, p. 66).

This feature is highlighted in the right panel, which instead illustrates the *change* in the policy function. That is, how saving behavior change when the $r - g$ differential increases. The overall distribution is left in the background, to give an idea of where changes primarily take place. As is visible, the change in the policy function is upward sloping and concave, suggesting that – as in the left panel – saving change more for the asset-rich, than for households closer to the constraint.[30] That rich and poor households are responding asymmetrically to the return gap is key to our result that inequality increases.

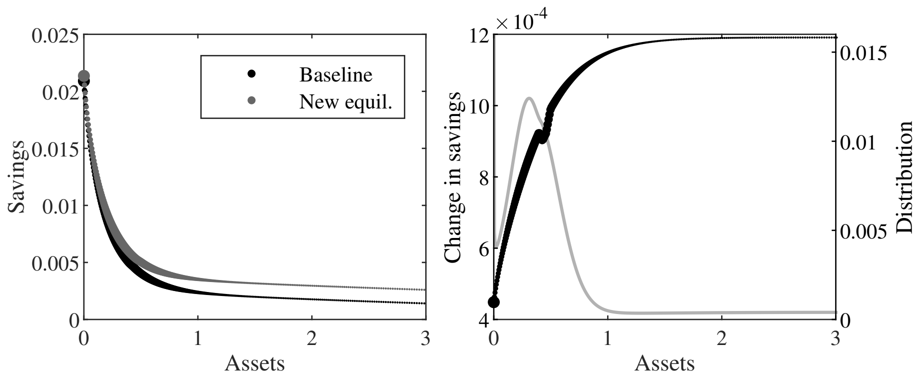

Figure 7: Market power and saving rates

Notes. The left panel show the expected saving function at the two different return gaps. The right panel illustrates the change (left axis) together with the overall distribution (right axis).

To appreciate this, notice that if all households increased their saving by the same proportion, inequality would be left unchanged; indeed, measures of inequality – the Gini, or the top shares – are invariant to scaling of the distribution. If, however, rich households respond stronger to an increases in the return gap relative to poorer households, inequality instead rises. In our framework, poor households save for precautionary reasons, while richer households save for the purpose of intertemporal substitution. A change in the return gap alters the reward for intertemporal substitution, but does little to affect the incentives involved in precautionary decision-making. Hence, the rich respond more strongly than the poor, and a rise in inequality ensues.

It ought to be noted that an analogous logic applies to the capitalist/worker model as well. The capitalists respond strongly to the rise in the return gap for reasons of intertemporal substitution, while workers do not. However, in that setting, workers do not respond as they are precluded to

[30] The small wiggle around assets of 0.5 arises because this is where the policy function for the impatient agents crosses the zero line.

participate in any asset markets. Here, most households do indeed change their saving behavior, but some less than others, as their motive for saving is driven by precautionary concerns, and not due to intertemporal substitution.

3.3.1 Inequality between and within groups

Table 2 outlines some key properties of within- and between-group inequality, with the first six rows pertaining to the baseline ("1980"), and the last six to the situation with a markup of 1.5 ("2020"). Focussing on the two top rows in each cluster – average and median wealth holdings –

Table 2: Evolution of inequality

	Total	Patient	Intermed.	Impatient	
1980					
Average wealth	1.05	4.993	0.604	0.224	
Median wealth	0.36	3.580	0.363	0.184	
Wealth share		0.473	0.459	0.021	
Share in top10	0.64[a]	0.070	0.030	0.000	*Between:*
Gini	0.66	0.467	0.556	0.455	0.407
Theil	1.02	0.364	0.852	0.653	0.448
2020					
Average wealth	1.90	8.296	1.217	0.415	
Median wealth	0.47	5.531	0.466	0.206	
Wealth share		0.438	0.514	0.022	
Share in top10	0.70[a]	0.055	0.044	0.002	*Between:*
Gini	0.73	0.499	0.691	0.656	0.374
Theil	1.25	0.428	1.255	1.528	0.385

Notes. The table shows key ine quality measures for ϕ at the benchmark ("1980"), and at the level that generates a markup of 1.5 ("2020"). *Low*, *Med.*, *High* are referring to ρ_s, and is, for the case of *Gini* and *Theil*, measuring within group wealth inequality. *Across* is referring to wealth inequality across groups.

[a]The total share in top10 shows the overall top 10% share.

three key patterns emerge. First, both average and median wealth holdings increases for all groups. Second, average wealth holdings increase relatively more for the two less patient groups compared to the patient. And three, average wealth increases by more than median wealth, and this pattern is stronger the less patient a household is. This latter result further reinforces the view that inequality increases.[31]

Next, the two middle rows in each cluster – the wealth shares and the shares in the top 10% – provides some additional supportive evidence for these patterns. First, the wealth shares of the less patient households increase, while that of the patient decreases. This echos the aforementioned

[31] Average wealth increases by about 65% for patient households, and almost doubles for less patient households. Median wealth increases by about 50% for patient households, 28% for the intermediate group, and 12% for the patient.

fact that average wealth increases more for these former groups than for the patient. Second, and similarly, the shares in the top 10% increase for both the intermediate and impatient households, but decreases for the patient. Together, the top four rows suggest that between-group inequality actually decreases, but that inequality within groups might have increased.

Turning to the latter two rows, we consider two measures of within-group inequality; the Gini coefficient as well as the Theil index. We augment the analysis with the latter measure, since it is decomposable into within- and between-group indices (Bourguignon, 2020).[32] The Gini coefficient reveals that within-group inequality increases for all three types of households, with the impatient group displaying the most pronounced result. The Theil index reinforces this view, as it paints a similar picture. However, both the Gini coefficient and the Theil index suggest that between-group inequality *decreases*.

To summarize, *within-group wealth inequality* as well as *average wealth* increase for all groups, but more so for the less patient households. In contrast, across-group wealth inequality decreases. These results seem to contradict the idea that the rise in inequality is driven by the top of the distribution, and instead suggest that wealth inequality primarily arises somewhere in the "middle class". This is not true; indeed, the top is the driving force of this change.

To appreciate this, consider the following counterfactual experiment: take the new equilibrium's policy functions for saving, but replace those that belong to the top 20% of the patient group with their corresponding saving functions at the baseline (these comprise approximately the top 2% of wealth holders in the economy). That is, assume that this small, but wealthy, group's decision-making is unaffected by the rise in the return gap. By doing so, and recalculating the distribution, as well as the metrics in table 2, we find that the overall Gini-coefficient declines to 0.68, and that the within-group Gini-coefficients decline to 0.43, 0.60, and 0.53, for the patient, intermediate, and impatient group, respectively. That is, almost 70% of the rise in the Gini coefficient can be accounted for by the saving decisions made at the very top. Moreover, the increase in average wealth is halved for all three groups, including the less patient, despite the fact that these latter two groups are not directly affected by this counterfactual exercise.[33] The behavior at the top indeed trickles through the rest of the economy.

These results, however, raise some questions. If inequality is driven by the top of the distribution, how come inequality does not significantly increase in the patient group? Relatedly, how can within-group inequality rise in the less patient groups, when the overall rise in inequality can be

[32]That is, the total Theil index is equal to the sum of the wealth-share weighted within-group indecies, added to the across-group index. Thus, for instance the Theil index 1.02 (row six) is given by $0.473 \times 0.364 + 0.459 \times 0.852 + 0.021 \times 0.653 + 0.448$. The non-decomposability of the Gini is explored in Mookherjee and Shorrocks (1982). Between group inequality is the inequality that would arise if each household within a group held assets equal to that group's average.

[33]Average wealth is now 6.7, 0.9,, and 0.3, for the three respective groups.

attributed to the behavior of the top wealth holders in the most patient group? And lastly, how can overall inequality increase, when the less patient groups are catching up to the patient?

The first of these questions has a simple answer: as the vast majority of households in the patient groups has sufficient wealth to be unbothered by the borrowing constraint (as well as income risk), these households are largely *symmetrically* affected by the rise in the return gap. Hence, their distribution is largely scaled towards the right, leading to a moderate rise in within-group inequality. The second and third question have more intricate answers, pertaining to the fact that agents transition across patience-groups. First, notice that even if wealth only increases by about 65% for the patient group, the rise in absolute terms of this group is five times larger than the rise for the intermediate group, and 15 times larger than the rise in the impatient group. This matters as inflows to a group follows the neighboring groups' endogenous distribution. Thus, inflows from, say, the patient group to the intermediate group, will on average bring large volumes of wealth, which will increase both within-group and overall inequality, while simultaneous raise average wealth holdings in this group.

A simple numerical example can help to clarify this point. Suppose there are two, equally sized, groups: rich and poor. The rich are in possession of 5 units of wealth, while the poor only of 0.5 units.[34] This leads to a within-group Gini of zero, and between-group inequality that coincides with the overall Gini coefficient

$$\text{Gini}_0 = \frac{1}{2.75} 0.5 \times 0.5(5 - 0.5) \approx 0.41.$$

Now, suppose that wealth of the rich increases by a factor of 2, but that the wealth of the poor remains the same. After this event, 10% of the rich group transition to the poor group, and 10% of the poor group transition to the rich group. This leads to a within-group Gini of 0.1 for the rich, and to 0.59 for the poor. Moreover, the overall Gini coefficient rises to 0.45, and average wealth increases only by 80% for the rich, but by 190% for the poor. Lastly, the between-group Gini coefficient declines to 0.36. Thus, within-group inequality increases the most for the poor group, between-group inequality declines, the overall Gini coefficient rises, and average wealth increases the least for the rich.

Taking stock. The patient group, which includes the wealthiest households, does not experience a significant rise in inequality. This is because their savings are primarily driven by intertemporal considerations, making the effects of an increase in the return gap relatively balanced within the group. In contrast, inequality within the less patient group, which includes the poorest households, grows as households in this groups respond differentially to the rise in the return gap, and as wealthy

[34] The numbers are chosen to largely mimic those of table 2.

households transition into this category. Although inequality increases within the less patient group and remains largely unchanged within the patient group, the rise in aggregate inequality is largely driven by the top-wealth households in the patient group. Their focus on intertemporal saving makes them particularly sensitive to shifts in the return gap.

4 Alternative sources of rising market power

A natural question that arises is to what extent increased entry costs are the primary drivers of growing market power. While directly measuring entry costs can be challenging, the evolution of regulatory burdens in the product market serves as a useful proxy for both increased entry costs and broader barriers to entry. Substantial evidence indicating that the regulatory burden has increased over the study period can be found in the works of Kalmenovitz (2023), Dawson and Seater (2013), Akcigit and Ates (2023) and Trebbi et al. (2023), among others.

However, it is important to consider that the U.S. economy has undergone a wide array of transformations over the past forty years, which could have influenced the development of market power in addition to the rise in entry barriers. To further elaborate on this, we follow Akcigit and Ates (2023) and consider two alternative key changes; the role of corporate taxes as well as R&D subsidies. The corporate income tax rate is 20% in 1980 and declines to 21% in 2020. In the same period, the effective R&D subsidy rate increases from 5% to 20% (Akcigit et al., forthcoming). As we will see, in the context of our model, taxes and/or subsidies are not able to replicate the key patterns we observe in the data.

Introducing taxes, τ, and subsidies, s, alters the BGP growth equation and the free entry condition according to

$$g = \frac{1}{\beta}\left[\frac{A\eta n^{\beta}\ell}{(1-s)} - (R+\delta)\right],$$

$$\phi = \frac{(1-\tau)\frac{1-\theta}{\theta}\ell - (1-s)\frac{g}{An^{\beta}}}{R+\delta}.$$

It is useful to develop an intuition of how these policy instruments are likely to affect markups and growth. Any policy that increases markups, must also reduce the number of active firms, n. With fewer firms, there is more labor available for each existing firm. This leads to an expansion in both ℓ and h, with a positive effect on growth. As a consequence, an increase in the tax rate, τ, will have a direct negative impact on profits, reduce n, and increase growth. The effect of the subsidy, s, is more intricate, as it affects both relations above; it directly increases profits and growth, but as growth negatively affects profits, the net effect is somewhat unclear. Thus, inserting the first equation into the latter gives

$$\phi = \frac{[(1-\tau)\frac{(1-\theta)}{\theta} - \frac{\eta}{\beta}]\ell + \frac{(1-s)}{\beta A n^{\beta}}(R+\delta)}{R+\delta},$$

revealing that s has a negative effect on profits. Thus, in similarity to taxes, an increase in s leads to both a rise in markups and in growth.

To gauge the role of these policy instruments, we start by holding ϕ constant, and vary the tax, τ, or the subsidy, s, rate to obtain an empirically relevant rise in the markup. By increasing taxes from zero in the baseline, to about 65% raises the markup from 1.25 to 1.5. At the same time, however, growth increases from about 1.6% to 5.9%. The reason is that higher taxes reduces profits and entry, which renders existing firms larger, leading to more innovation. Alternatively, raising the R&D subsidy to 0.89 gives an identical effect on markups, but raises growth to 91%. Thus, neither taxes nor subsidies can, in isolation, lead to an empirically relevant increase in markups alongside a decline in growth.

If, however, taxes were to increase to 72%, and subsidies decreased to -24.5%, markups would rise to 1.5, and growth fall to 0.8%. Yet, in this case, average wealth declines from the baseline of 1.05 to 0.63, the Gini coefficient falls from 0.66 to 0.56, and the top10 share from 0.64 to 0.52. Again, even jointly, taxes and subsidies are unable to replicate the empirical patterns of market power, growth, and inequality.

Akcigit and Ates (2023) also consider exogenous changes in knowledge diffusion, demonstrating that a slowdown in this process can account not only for the observed rise in markups and market concentration but also for other key trends they investigate. To support the hypothesis of a slowdown in knowledge diffusion in recent decades, they provide evidence related to patent concentration and the strategic use of patents. Our findings offer an alternative perspective: the slowdown in knowledge diffusion is not an exogenous factor but rather a consequence of rising market power. In our model, the extent of knowledge spillovers, which drive diffusion, depends on the number of firms from which each firm can learn when innovating. Increased entry costs reduce the number of firms, weakening spillovers, and ultimately leading to slower economic growth and greater wealth inequality.

Liu et al. (2020) argue that the decline in the U.S. real interest rate over recent decades could be a key driver of rising market power. Using a Schumpeterian growth model, they demonstrate that an exogenous decrease in the interest rate leads to increased market concentration and reduced long-run growth. Our findings align with the same set of stylized facts but differ in their mechanism. We endogenize the decline in the interest rate, showing it as a consequence of higher market power and slower growth. This perspective provides a novel insight: the observed decline in the interest rate can in part be attributed to a general equilibrium feedback loop initiated by the rise in market power.

Recent studies have argued that the decline in population growth observed in recent decades could be at the root of increasing market power and productivity slowdown (e.g. Hopenhayn et al., 2022; Peters and Walsh, forthcoming). The impact of slower population growth is conceptualised via a reduction in firm entry. Although we do not have population growth in the model, it is possible to broadly interpret the effect of these changes through the lenses of our model assuming that population growth operate similarly to a barrier to entry. In line with Peters and Walsh (forthcoming) we show that a reduction in the creation of new firms leads to higher concentration, higher markups and slower growth.

5 Welfare on the balanced growth path

The main analysis of this paper has centred around two BGPs: the first with high growth, low markups, and little capital; and the second with the opposite features. This creates some tension with respect to welfare, as both lower growth and higher markups depress labor income, in the present and the future, while more capital expands consumption opportunities. From (6) we can infer that along the balanced growth path wages evolve according to $w_t = \theta z_t^{\eta}$. An increase in the markup has a direct negative impact on wages via θ and an indirect impact via depressing the growth rate of productivity. This section explores to which extent these tensions play out, and how welfare is affected on these two BGPs, and for which households. As we will see, only the very rich are better off in the new equilibrium.

The model is stationarized by subtracting the expected net present value of wages, w_t, from the expected net present value of utility. Thus, to compute overall welfare we need to add this former component to the stationarized value function. That is, if $v(a,s)$ is the stationarized value function, net present value utility is given by

$$V_t(\phi) = v(a,s) + E_t\left[\int_t^{\infty} e^{\int_t^{\tau}\rho_s di} w_\tau d\tau\right], \tag{22}$$

where $V_t(\phi)$ denotes overall utility conditional on a certain value of the entry cost, ϕ.

On the BGP of the capitalist/worker economy, this has a simple expression

$$V_t(\phi) = \frac{1}{\rho}\left[\ln c + \ln w_t + \frac{g}{\rho}\right].$$

That is, welfare depends on consumption relative to wages, c, the level of wages at time t, w_t, and the evolution of wages, captured by the growth rate, g. To obtain a measure of the welfare gains or losses associated with a change in ϕ, we focus on the percentage change in consumption that would be necessary to make a household indifferent between the old and the new value of ϕ. If ϕ denotes the baseline entry cost, and ϕ' the new, the welfare gain is then given by

$$\rho[V_t(\phi') - V_t(\phi)].$$

For the baseline BGP, we use the values $w_t = 1$, $g = 0.0156$, and $c = \rho \times 1.1$ for the capitalists.[35] At the new BGP, we use the values $w_t = 0.833$, $g = 0.008$, and $c = \rho \times 1.9$ for the capitalists, reflecting a rise in markups from 1.25 to 1.5, halving of the growth rate, and a rise in a in accordance with the results in figure 2. For the workers, this implies a welfare loss of 33.4% of consumption. For the capitalists, it instead implies a welfare gain of 21.2%. Thus, workers are substantially worse off on the new BGP, as the rise in markups erodes real wages, and the decline in growth reduces futures wage prospects. Capitalists, however, gain, as they are unaffected by wages, both in the present and in the future, and acquire more capital. The additional capital expand their consumption, and thereby welfare.

In the incomplete markets economy, the expression for overall utility in equation (22) remains the same, but its value varies depending on a household's level of wealth, a, as well as the exogenous state, s. That is, there is a distribution of welfare gains/losses. Moreover, as no household is purely a capitalist nor a worker, welfare will fall in between the results of the capitalist/worker above, with outcomes endogenously determined by the household's level of assets, as well as the exogenous state. Lastly, as both wages as well as the discount factor stochastically fluctuate over time, the expectation term in (22) is numerically computed.

Table 3: Welfare

Percentile	0-50	50-80	80-90	90-95	95-99	99-99.5	top 1%	top 0.1%
Welfare gain	-34.4	-33.3	-27.8	-19.9	-10.3	2.4	8.9	30.5

Notes. The table lists the welfare gains and losses for the incomplete markets model at different parts of the distribution.

Table 3 displays the results of this exercise. The welfare gains/losses exhibits three distinct patterns across the distribution. First, losses are almost uniform for agents at the bottom 80% of the distribution. This group experience a welfare loss on the new BGP equivalent to a drop in consumption of about 34%, echoing the results of the workers above. Second, for agent in the range of the $80th$-$99th$ percentile, losses are linearly reduced from about 33% at the $80th$ percentile, to close to 0% at the $99th$, averaging a loss of about 22 percent. Lastly, these losses then turns to gains for the top 1%. In particular, gains range from 0% for the 99th percentile, and exponentially increase to 70% at the extreme top. The average gain for the top 1% is 9%, and 30% for the top

[35] Recall from page 2.3.1 that capitalist consumption is given by $c = \rho \times a$, and from figure 2 that $a \approx 1.1$. For the workers it is always the case that $c = 1$ by definition.

0.1%. On average, welfare losses are about 30 percent, with gains exclusively accumulating to the extremely wealthy.

6 Concluding remarks

This paper proposed a theory linking market power to growth and wealth inequality, to help interpret a set of empirical regularities observed in the US economy in recent decades. It starts from two key ideas: first, there is a natural link between market power and wealth inequality, as profits shape the value and the return on assets. Second, since the key return driving wealth inequality is the difference between the asset return and the growth rate of the economy, a comprehensive analysis of the wealth distribution needs to account for the evolution of growth.

Building on these concepts, we develop an endogenous growth model featuring variable markups and household heterogeneity. Our theory posits that an increase in markups, driven by higher entry barriers, enhances corporate profitability and raises the demand for assets, thereby exerting upward pressure on asset returns. At the same time, weaker competition hampers growth by limiting the opportunities for firms to learn from one another's technologies.

The rise in markups erodes real wages, while slower growth diminishes future wage prospects, making it harder for households reliant on labor income – primarily poorer households – to accumulate assets. In contrast, wealthier households respond more strongly to rising asset returns, increasing their saving rates significantly. This divergence arises because affluent households predominantly save for the purpose of intertemporal substitution (consumption smoothing), while less wealthy households, constrained by income risk and less able to borrow – they are closer to the borrowing constraint – save primarily for precautionary reasons. Precautionary savings, however, are relatively insensitive to changes in interest rates.

Together, the increase in asset returns and the slowdown in growth widen the return gap, $r - g$, thereby intensifying wealth inequality. Lower growth prospects induces more saving for intertemporal reasons, thereby increasing the supply of assets and putting downward pressure on the interest rate which ultimately falls below its initial level. Even with a lower interest rate the return gap, and therefore inequality, rises due to lower growth.

Although confined to the balanced growth path, the welfare analysis suggests that the increase in market power observed in the data since 1980, along with the slowdown in growth and widening wealth inequality generate large welfare losses for most of the population except the very rich, the top 1% of the wealth distribution.

7 Appendix

7.1 Derivation of HJB equations

7.1.1 Intermediate firms

In discrete time, the Bellman equation for the intermediate firms is given by

$$V(z_t, Z_t, n_t) = \max_{\ell_t, h_t} \{(\Delta p\left(\frac{\tilde{y}_t + z_t^{\eta}\ell_t}{Y_t}\right) z_t^{\eta} - w_t)\ell_t - \Delta h_t w_t + e^{-\Delta(r_t+\delta)} V(z_{t+\Delta}, Z_{t+\Delta}, n_{t+\Delta})\},$$
$$z_{t+\Delta} - z_t = \Delta A \kappa_t h_t,$$
$$Z_{t+\Delta} - Z_t = \Delta A \kappa_t H_t,$$
$$n_{t+\Delta} - n_t = \Delta[m_t - \delta n_t].$$

where z_t is the firm's productivity level, and Z_t is sector or aggregate productivity, as sectors are symmetric. m_t is the mass of new firms entering the market at t. Dividing by w_t on both sides, and defining $\hat{z}_t$ as

$$\hat{z}_t = \frac{z_t}{w_t^{1/\eta}} \quad \Leftrightarrow \quad \hat{z}_t^{\eta} = \frac{z_t^{\eta}}{w_t}$$

gives

$$v(\hat{z}_t, \hat{Z}_t, n_t) = \max_{\ell_t, h_t} \{(\Delta p\left(\frac{\hat{y}_t + \hat{q}_t}{\hat{Y}_t}\right) - \hat{z}_t^{-\eta})\hat{q}_t - \Delta h_t + e^{-\Delta(r_t - \eta g_t + \delta)} v(\hat{z}_{t+\Delta}, \hat{Z}_{t+\Delta}, n_{t+\Delta})\}, \tag{23}$$

with $\hat{y}_t = \tilde{y}_t / w_t$, $\hat{q}_t = q_t / w_t$, and $e^{\Delta \eta g_t} = w_{t+\Delta}/w_t$.

In continuous time this is

$$(R_t + \delta) v(\hat{z}_t, \hat{Z}_t, n_t) = \max_{q_t, h_t} \Big\{ (p\left(\frac{\hat{y}_t + \hat{q}_t}{\hat{Y}_t}\right) - \hat{z}_t^{-\eta})\hat{q}_t - h_t + v_{\hat{z}}(\hat{z}_t, \hat{Z}_t, n_t)[A\hat{\kappa}_t h_t - g_t \hat{z}_t]$$
$$+ v_{\hat{Z}}(\hat{z}_t, \hat{Z}_t, n_t)[A\hat{\kappa}_t H_t - g_t \hat{Z}_t] + v_n(\hat{z}_t, \hat{Z}_t, , n_t)\dot{n}_t \Big\}, \tag{24}$$

with $\hat{\kappa}_t = \kappa_t / w_t^{1/\eta}$

7.1.2 Incomplete markets

In the incomplete market setting, the model is stationarized using average wages. Thus, at constant prices, r, and growth rate, g, the Bellman equation in Δ units of time is

$$v(a,s) = \Delta \ln c + e^{-\Delta\rho_s}\left[\left(1 - \sum_{s'\neq s}(1-e^{-\Delta\lambda_{s',s}})\right)v(a',s) + \sum_{s'\neq s}(1-e^{-\Delta\lambda_{s',s}})v(a',s')\right],$$

with

$$a' - a = a\frac{1-e^{\Delta g}}{e^{\Delta g}} + \Delta(ra + y_s - c).$$

Subtracting $v(a,s)$, dividing with Δ, and rearranging gives

$$\begin{aligned} 0 = \ln c + e^{-\Delta\rho_s}\sum_{s'\neq s}\frac{(1-e^{-\Delta\lambda_{s',s}})}{\Delta}v(a',s') + \frac{(e^{-\Delta\rho_s}-1)}{\Delta}v(a',s) \\ - e^{-\Delta\rho_s}\left(\sum_{s'\neq s}\frac{(1-e^{-\Delta\lambda_{s',s}})}{\Delta}\right)v(a',s) + \frac{v(a',s)-v(a,s)}{\Delta}. \end{aligned}$$

Taking the limit $\Delta \to 0$

$$0 = \ln c + \sum_{s'\neq s}\lambda_{s',s}v(a,s') - \rho_s v(a,s) - \sum_{s'\neq s}\lambda_{s',s}v(a,s) + v'(a,s)\dot{a}.$$

or

$$\rho_s v(a,s) = \ln c + v'(a,s)\dot{a} - \sum_{s'\in S}\lambda_{s',s}[v(a,s) - v(a,s')].$$

with

$$\dot{a} = (r-g)a + y_s - c.$$

7.2 Derivation of optimality conditions

Suppressing notation such that $v_{\hat{z}} = v_{\hat{z}}(\hat{z}_t, \hat{Z}_t, n_t)$, the first order conditions to the intermediate goods producers problem are

$$\begin{aligned} v_{\hat{z}} A\hat{\kappa}_t &= 1, \\ \hat{z}_t^{-\eta} &= \theta_t, \end{aligned}$$

with $\theta_t = (n_t - 1 + \alpha)/n_t$. The envelope condition gives

$$
\begin{aligned}
(R_t+\delta)v_{\hat{z}} = \eta \hat{z}_t^{-\eta-1}\hat{q}_t + v_{\hat{z}\hat{z}}[A\hat{\kappa}_t h_t - g_t\hat{z}_t] \\
+ v_{\hat{z}}[A(1-\beta)n_t^{\beta}h_t - g_t] + v_{\hat{Z}\hat{z}}[A\hat{\kappa}_t H_t - g_t\hat{Z}_t] + v_{n\hat{z}}\dot{n}_t.
\end{aligned}
$$

From the first order condition we have

$$
\begin{aligned}
v_{\hat{z}\hat{z}}A\hat{\kappa}_t + v_{\hat{z}}A(1-\beta)n_t^{\beta} &= 0,\\
v_{\hat{z}\hat{Z}}A\hat{\kappa}_t + v_{\hat{z}}A\beta n_t^{\beta-1} &= 0,\\
v_{\hat{z}n}A\hat{\kappa}_t &= 0,
\end{aligned}
$$

Inserting the first and the last gives

$$
\begin{aligned}
(R_t+\delta)v_{\hat{z}} &= \eta \hat{z}_t^{-\eta-1}\hat{q}_t - v_{\hat{z}}(1-\beta)\frac{n_t^{\beta}}{\hat{\kappa}_t}[A\hat{\kappa}_t h_t - g_t\hat{z}_t] \\
&\quad + v_{\hat{z}}[A(1-\beta)n_t^{\beta}h_t - g_t] + v_{\hat{Z}\hat{z}}[A\hat{\kappa}_t H_t - g_t\hat{Z}_t],\\
&= \eta \hat{z}_t^{-\eta-1}\hat{q}_t + v_{\hat{z}}(1-\beta)\frac{n_t^{\beta}}{\hat{\kappa}_t}g_t\hat{z}_t - v_{\hat{z}}g_t + v_{\hat{Z}\hat{z}}[A\hat{\kappa}_t H_t - g_t\hat{Z}_t],\\
&= \eta \hat{z}_t^{-\eta-1}\hat{q}_t - v_{\hat{z}}g_t\beta + v_{\hat{Z}\hat{z}}[A\hat{\kappa}_t H_t - g_t\hat{Z}_t].
\end{aligned}
$$

Inserting the second gives

$$
\begin{aligned}
(R_t+\delta)v_{\hat{z}} &= \eta \hat{z}_t^{-\eta-1}\hat{q}_t - v_{\hat{z}}g_t\beta - v_{\hat{z}}\frac{\beta}{\hat{\kappa}_t}n_t^{\beta-1}[A\hat{\kappa}_t H_t - g_t\hat{Z}_t],\\
&= \eta \hat{z}_t^{-\eta-1}\hat{q}_t - v_{\hat{z}}g_t\beta - v_{\hat{z}}\beta n_t^{\beta-1}[AH_t - \frac{g_t Z_t}{\kappa_t}],\\
&= \eta \hat{z}_t^{-\eta-1}\hat{q}_t - v_{\hat{z}}g_t\beta - v_{\hat{z}}\beta n_t^{\beta-1}[\frac{\dot{Z}_t - g_t Z_t}{\kappa_t}],\\
&= \eta \hat{z}_t^{-\eta-1}\hat{q}_t - v_{\hat{z}}g_t\beta.
\end{aligned}
$$

Thus

$$
\begin{aligned}
g_t &= \frac{1}{\beta}\left[A\hat{\kappa}_t\eta\hat{z}_t^{-\eta-1}\hat{q}_t - (R_t+\delta)\right],\\
&= \frac{1}{\beta}\left[A\eta n_t^{\beta}\ell_t - (R_t+\delta)\right].
\end{aligned}
$$

7.3 Proofs

7.3.1 Proof of proposition 1

The equilibrium equations are:

$$g = \frac{1}{\beta}\left(A\eta\kappa(n)\ell - \rho - \delta\right), \tag{25}$$

$$(\rho+\delta)\phi\theta = (1-\theta)\ell - \theta\frac{g}{\kappa(n)A}, \tag{26}$$

$$1 = n\left(\ell + \frac{g}{\kappa(n)A} + \delta\phi\right). \tag{27}$$

where $\kappa(n) = n^{\beta}$.

Proposition 1. *Condition $1-\theta(1+\eta/\beta) > 0$ is sufficient for profits to be positive and for an increase in the fixed entry cost ϕ to reduce the number of firms and increase the markup*

Proof. Combining (25) and (26) we get

$$\ell = \frac{\tilde{\rho}\theta\left(\phi - \frac{1}{Ak(n)\beta}\right)}{1-\theta\hat{\eta}} \tag{28}$$

where $\hat{\eta} = 1+\eta/\beta$ and $\tilde{\rho} = \rho+\delta$. Combining (25) and (27) we get

$$\ell = \frac{\frac{1}{n} + \frac{\tilde{\rho}}{Ak(n)\beta} - \delta\phi}{\hat{\eta}}. \tag{29}$$

Combining the two expressions for ℓ we get

$$F = \frac{\frac{1}{n} + \frac{\tilde{\rho}}{Ak(n)\beta} - \phi\delta}{\hat{\eta}} - \frac{\tilde{\rho}\theta\left(\phi - \frac{1}{Ak(n)\beta}\right)}{1-\hat{\eta}\theta}.$$

Using the implicit function theorem we have

$$\begin{aligned}\frac{\partial n}{\partial \phi} &= -\frac{\partial F/\partial \phi}{\partial F/\partial n} \qquad (30)\\ &= \frac{\frac{\delta}{\hat{\eta}} + \frac{\tilde{\rho}\theta}{1-\theta\hat{\eta}}}{\frac{1}{\hat{\eta}}\left(-\frac{1}{n^2} - \frac{Ak'(n)\beta\tilde{\rho}}{(Ak(n)\beta)^2}\right) - \frac{1}{(1-\theta\hat{\eta})^2}\left[\tilde{\rho}\theta\frac{Ak'(n)\beta}{(Ak\beta)^2}(1-\theta\hat{\eta}) + \tilde{\rho}\theta'\left(\phi - \frac{1}{Ak(n)\beta}\right)\right]}.\end{aligned}$$

Condition $1-\theta\hat{\eta} > 0$ implies that $(\phi - 1/(Ak(n)\beta)) > 0$ so that firm size ℓ is positive and so are profits. Moreover, $\theta' = \partial\theta/\partial n = (1-\alpha)/n^2 > 0$ and $\partial\kappa(n)/\partial n > 0$ imply that $\frac{\partial n}{\partial \phi} < 0$.

□

7.3.2 Proof of proposition 2

Proposition 2. *The effect of changes in the entry cost on growth and the relationship between competition and growth is ambiguous.If $\beta = 1$, if the elasticity of the number of firms to the entry cost, $|\varepsilon_{n,\phi}| = |\frac{\partial n}{\partial \phi}\frac{\phi}{n}| < 1$ for any ϕ, higher entry cost are associated with lower growth.*

Proof. From the growth equation we get

$$\frac{\partial g}{\partial \phi} = \frac{A\eta}{\beta}\left(\frac{\partial k(n)}{\partial \phi}\ell + \kappa(n)\frac{\partial \ell}{\partial \phi}\right)$$

From (29) we get

$$\frac{\partial \ell}{\partial \phi} = \frac{1}{\hat{\eta}}\left[\left(-\frac{1}{n^2}\right)\frac{\partial n}{\partial \phi} - \delta - \frac{\tilde{\rho}A\beta}{(A\kappa(n)\beta)^2}\frac{\partial \kappa(n)}{\partial \phi}\right]$$

and then

$$\frac{\partial g}{\partial \phi} = \frac{A\eta}{\beta}\left[-\frac{\partial n}{\partial \phi}\frac{\kappa(n)}{\hat{\eta}n^2} - \frac{\delta k(n)}{\hat{\eta}} + \frac{\partial \kappa(n)}{\partial \phi}\left(\ell - \frac{\tilde{\rho}}{\hat{\eta}A\kappa(n)\beta}\right)\right]. \tag{31}$$

When $\beta = 1$, $\kappa(n) = n$, expression (31) becomes,

$$\frac{\partial g}{\partial \phi} = \frac{nA\eta\delta}{\hat{\eta}}\left(-\frac{\partial n}{\partial \phi}\frac{\phi}{n} - 1\right), \tag{32}$$

where I have used (29) to simplify. If the elasticity of the number of firms to the entry cost is in absolute value smaller than one for any ϕ, $|\varepsilon_{n,\phi}| = |\frac{\partial n}{\partial \phi}\frac{\phi}{n}| < 1$, then higher entry cost are associated with lower growth and therefore there is a positive relationship between competition and growth, which is what we observe in the data.

□

7.4 Computational appendix and calibrated values

7.4.1 Stationary firm problem

Given a constant interest rate, r, solving the firm problem is straightforward, as it amounts to solving a system of three equations in three unknowns. We solve this using Newton's method with a tolerance criterion of $1e(-9)$ for the maximum absolute error of the three equations. This applies both for the BGPs of the capitalist-worker economy, and the incomplete markets setting, as well as the transitional dynamics of the former.

7.4.2 Transitional dynamics

To solve for the transitional dynamics, we assume that ϕ increases from its BGP benchmark, ϕ_0, to it's new value, ϕ_1, according to

$$\dot{\phi}_t = -\gamma(\phi_t - \phi_1),$$

with γ set to 0.07, leading to convergence in approximately 40 years. Recall that the solution is given by

$$a_t = a_0 e^{\int_0^t (R_s - \rho) ds}.$$

Thus, we guess for a path of capital supply, $\{a_t\}$, for $t = 0, dt, 2dt, \ldots T$, with $dt = 0.01$ and $T = 180$, and use the above equation to obtain

$$\int_0^t R_s ds = \ln(a_t) - \ln(a_0) + \rho t.$$

We subsequently differentiate this expression to obtain R_t, for all t, and solve the problem in equations (10)-(12) for all t. This leads to a sequence of capital demand, $\{n_t \phi_t\}$. We then update the guess for $\{a_t\}$ according to

$$\{0.9a_t\} + \{0.1n_t \phi_t\},$$

until $a_t \approx n_t \phi_t$ for all t.

7.4.3 Heterogeneous agents

Given an $r - g$ differential, it is possible to solve the HJB equations in (19), letting us obtain the policy functions $\dot{a} = h(a,s)$. These policy function, together with the transition probabilities, $\lambda(s', s)$, gives rise to the Kolmogorov forward equation.

$$\dot{p}_t(a,s) = -\frac{\partial p_t(a,s)h(a,s)}{\partial a} - \sum_{s' \neq s} \lambda(s', s)(p_t(a,s) - p_t(a,s')),$$

determining the law of motion of the cross sectional distribution, p_t. At a stationary equilibrium, this satisfies

$$0 = -\frac{\partial p(a,s)h(a,s)}{\partial a} - \sum_{s' \neq s} \lambda(s', s)(p(a,s) - p(a,s')). \tag{33}$$

We then obtain asset supply

$$\hat{a} = \int_{(s,a)} ap(a,s)dads,$$

and the equilibrium $r - g$ differential, R, must satisfy

$$\hat{a} = n\phi.$$

The algorithm is now straightforward

(i) Guess for an interest rate, R_n.

(ii) Given the interest rate, solve equations (10)-(12) for g and n (and ℓ).

(iii) Given R, solve the HJB equations in (19), and obtain the policy function $\dot{a} = h(a,s)$.

(iv) Find the stationary cross-sectional distribution, p, via the Kolmogorov forward equation (33), and integrate it to obtain $\hat{a}$.

(v) If $\hat{a} > \phi n$, set $R_{n+1} < R_n$. Otherwise set $R_{n+1} > R_n$.

(vi) Return to step 2. and iterate until convergence, $\hat{a} \approx \phi n$.

7.5 Details

The firm's problem is always solved using Newton's method, with a tolerance criterion of $1e(-9)$ for the maximum error of the three equations. The incomplete markets model is solved using 800 grid points in wealth, resulting in a total of 4,800 grid points. We use the upwind implicit method to solve for the HJB equations, and a bisection method for the equilibrium interest rate.

7.5.1 Calibration

Since the arrival rates for employment are independent of those for patience, we discuss them separately, starting with the former. Let $\lambda_{e,u}$ denote the arrival rate for employment (conditional on unemployment). If we wish to target an expected duration of 12 weeks, this would require an arrival rate of 4.33. There are two issues associated with such a choice. First, it would imply that unemployed agents have better job prospects than employed agents. Second, the model becomes numerically unstable. To circumvent this problem, we will use a trick from the discrete time literture that is commonly deployed when the frequency of the model falls short of the duration of a shock. To this end, we assume that employed households that are hit by an unemployment shock has a strictly positive probability of finding a new job *immediately*, and that this probability is equal to the arrival rate of employment.[36] The duration of unemployment is then

$$E[\text{Unemp. duration}] = \frac{(1-\lambda_{e,u})}{\lambda_{e,u}},$$

requiring $\lambda_{e,u} = 0.8125$. This implies that agents have a 80% probability of having a zero duration, and a 20% probability of an expected duration of 1.2 years. The arrival rate for unemployment (conditional on employment) is set to ensure a BGP unemployment rate of 6%, leading to $\lambda_{u,e} = 0.0519$.

For patience, we require $\lambda_{h,l} = \lambda_{l,h} = 0$. Moreover, symmetry implies $\lambda_{h,m} = \lambda_{l,m}$, and $\lambda_{m,l} = \lambda_{m,h}$. Thus, we need to calibrate two parameters: the first of these, $\lambda_{h,m}$, is set to ensure 80% of mass in the medium patience regime (Krusell and Smith, 1998). The other, $\lambda_{m,l}$, is jointly determined alongside ε to ensure at top-10% wealth share of 63%, and an average elasticity of current wealth to wealth 30 years ago of 0.71 (Clark and Cummins, 2015). We achieve this via indirect inference. The parameters are combined, and presented below

$$\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \end{array} \begin{array}{c} \begin{array}{cccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \end{array} \\ \begin{bmatrix} & 0.0518 & 0.0006 & 0 & 0 & 0 \\ 0.8120 & & 0.0005 & 0.0001 & 0 & 0 \\ 0.0001 & 0 & & 0.0519 & 0.0001 & 0 \\ 0.0001 & 0 & 0.8124 & & 0.0001 & 0 \\ 0 & 0 & 0.0006 & 0 & & 0.0518 \\ 0 & 0 & 0.0005 & 0.0001 & 0.8120 & \end{bmatrix} \end{array}$$

with $s_1 = (e,l)$, $s_2 = (u,l)$, $s_3 = (e,m)$, $s_4 = (u,m)$, $s_5 = (e,h)$, and $s_6 = (u,h)$, and where e represents *employment* (high income), and u *unemployment* (low income), and h, m, l are referring to *high*, *medium*, and *low* patience.

[36] This is commonly done in discrete time models when the frequency of the model – say, quarterly – falls short of the expected duration of an unemployment spell (see, for instance, Challe et al. (2017)).

References

Acemoglu, Daron and David Autor, "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," in David Card and Orley Ashenfelter, eds., *David Card and Orley Ashenfelter, eds.*, Vol. 4 of *Handbook of Labor Economics*, Elsevier, 2011, pp. 1043–1171.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 2022, *89* (1), 45–86.

Aghion, Philippe, Antonin Bergeaud, Mathieu Lequien, and Marc Melitz, "The Impact of Exports on Innovation: Theory and Evidence," 2017. Manuscript.

_ , _ , Timo Boppart, Peter J Klenow, and Huiyu Li, "A Theory of Falling Growth and Rising Rents," *The Review of Economic Studies*, 2023, *90* (6), 2675–2702.

_ , Christopher Harris, Peter Howitt, and John Vickers, "Competition, Imitation and Growth with Step-by-step Innovation," *Review of Economic Studies*, 2001, *68* (3), 467–492.

_ , Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Akcigit, Ufuk and Sina T. Ates, "What Happened to U.S. Business Dynamism?," *Journal of Political Economy*, 2023, *131* (8), 2059–2124.

_ , Sina Ates, and Giammario Impullitti, "Innovation and Trade Policy in a Globalized World," *Journal of Political Economy Macroeconomics*, forthcoming.

Autor, David, David Dorn, Gordon H. Hanson, Gary Pisano, and Pian Shu, "Foreign Competition and Domestic Innovation: Evidence from US Patents," *American Economic Review: Insights*, September 2020, *2* (3), 357–74.

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

Autor, David H., "Work of the Past, Work of the Future," *AEA Papers and Proceedings*, May 2019, *109*, 1–32.

Bajgar, Matei, Giuseppe Berlinghieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis, "Industry Concentration in Europe and North America," *OECD Productivity Working Papers*, 2019.

Benhabib, Jess, Alberto Bisin, and Ming Luo, "Wealth Distribution and Social Mobility in the US: A Quantitative Approach," *American Economic Review*, 2019, *109*, 1623–1647.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," *American Economic Review*, April 2020, *110* (4), 1104–44.

_ , Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

_ , Mirko Draca, and John Van Reenen, "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity," *Review of Economic Studies*, 2016, *83* (1), 87–117.

Bourguignon, Francois, "Decomposable Income Inequality Measures," *Econometrica*, July 2020, *47* (4), 901–920.

Brendler, Pavel, Moritz Kuhn, and Ulrike Stein, "To Have or Not to Have: Understanding Wealth Inequality," CEPR Discussion Paper DP19412, CEPR 2024.

Carroll, Christopher D. and Miles S. Kimball, "On the Concavity of the Consumption Function," *Econometrica*, 1996, *64* (4), 981–992.

Cavenaile, Laurent, Murat Alp Celik, and Xu Tian, "Are Markups Too High? Competition, Strategic Innovation, and Industry Dynamics," *mimeo*, 2020.

Cette, Gilbert, John Fernald, and Benoît Mojon, "The pre-Great Recession slowdown in productivity," *European Economic Review*, 2016, *88*, 3–20. SI: The Post-Crisis Slump.

Challe, Edouard, Julien Matheron, Xavier Ragot, and Juan F. Rubio-Ramirez, "Precautionary saving and aggregate demand," *Quantitative Economics*, 2017, *8* (2), 435–478.

Clark, Gregory and Neil Cummins, "Intergenerational Wealth Mobility in England, 1858-2012: Surnames and Social Mobility," *The Economic Journal*, 2015, *125*, 61–85. SI: The Post-Crisis Slump.

Coelli, Federica, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "Better, Faster, Stronger: Global Innovation and Trade Liberalization," *The Review of Economics and Statistics*, 03 2022, *104* (2), 205–216.

Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon, "From Good to Bad Concentration? US Industries over the Past 30 Years," *NBER Macroeconomics Annual*, 2020, *34*, 1–46.

Dasgupta, Partha and Joseph E. Stiglitz, "Industrial Structure and the Nature of Innovation Activity," *Economic Journal*, 1980, *57* (3), 267–308.

Dawson, John W. and John J. Seater, "Federal Regulation and Aggregate Economic Growth," *Journal of Economic Growth*, 2013, *18* (2), 137–177.

Fahri, Emmanuel, Alan Olivi, and Iván Werning, "Price Theory for Incomplete Markets," Working Paper, MIT 2022.

Feenstra, Robert C., Philip Luck, Maurice Obstfeld, and Katheryn N. Russ, "In Search of the Armington Elasticity," *The Review of Economics and Statistics*, 2018, *100* (1), pp. 135–150.

Fernald, John, "A quarterly, utilization-adjusted series on total factor productivity," in "in" Federal Reserve Bank of San Francisco 2014.

Gordon, Robert J, "Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds," Working Paper 18315, National Bureau of Economic Research August 2012.

Gorodnichenko, Yuriy, Jan Svejnar, and Katherine Terrell, "Globalization and Innovation in Emerging Markets," *American Economic Journal: Macroeconomics*, April 2010, *2* (2), 194–226.

Hall, Robert E, "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy," Working Paper 24574, National Bureau of Economic Research May 2018.

Hashmi, Aamir Rafique, "Competition and Innovation: The Inverted-U Relationship Revisited," *Review of Economics and Statistics*, 2013, *95* (5), 1653–1668.

Holston, Kathryn, Thomas Laubach, and John C. Williams, "Measuring the natural rate of interest: International trends and determinants," *Journal of International Economics*, 2017, *108*, S59–S75.

Hombert, Johan and Adrien Matray, "Can Innovation Help US Manufacturing Firms Escape Import Competition from China?," 2015. CEPR Discussion Papers 10666.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania, "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," *Econometrica*, 2022, *90* (4), 1879–1914.

Hubmer, Joachim, Per Krusell, and Anthony Smith, "Sources of U.S. Wealth Inequality: Past, Present, and Future," *NBER Macroeconomics Annual*, 2021, *35*, 391–455.

Iacovone, Leonardo, "The Better You Are the Stronger It Makes You: Evidence on the Asymmetric Impact of Liberalization," *Journal of Development Economics*, 2012, *99* (2), 474–485.

Impullitti, Giammario and Omar Licandro, "Trade, Firm Selection, and Innovation: The Competition Channel," *Economic Journal*, 2018, *128* (608), 189–229.

_, _, and Pontus Rendahl, "Technology, market structure and the gains from trade," *Journal of International Economics*, 2022, *135*, 1035–57.

Jones, Charles I., "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality," *Journal of Economic Perspectives*, February 2015, *29* (1), 29–46.

Kalmenovitz, Joseph, "Regulatory Intensity and Firm-Specific Exposure," *The Review of Financial Studies*, 01 2023, *36* (8), 3311–3347.

Kaymak, Baris and Markus Poschke, "The Evolution of Wealth Inequality Over Half a Century: The Role of Taxes, Transfers, and Technology," *Journal of Monetary Economics*, 2016, *77*, 1–25.

Kroen, Thomas, Ernest Liu, Atif Mian, and Amir Sufi, "Low Interest Rates, Market Power and Productivity Growth," *NBER Working Paper*, 2022, (29368).

Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Kuhn, Moritz and Victor Rios-Rull, "2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics," *Quarterly Review*, 2013. Federal Reserve Bank of Minneapolis.

Liu, Ernest, Atif Mian, and Amir Sufi, "Low Interest Rates, Market Power and Productivity Growth," *Econometrica*, 2020, *90*, 193–221.

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.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo, "Uneven Growth: Automation's Impact on Income and Wealth Inequality," *Econometrica*, 2022, *90* (6), 2645–2683.

Mookherjee, Dilip and Anthony Shorrocks, "A Decomposition Analysis of the Trend in UK Income Inequality," *The Economic Journal*, December 1982, *92* (368), 886–902.

Olivi, Alan, "Sufficient Statistics for Heterogeneous Agent Models," Working Paper, University College London 2018.

Olmsted-Rumsey, Jane, "Market Concentration and the Productivity Slowdown," *mimeo*, 2022.

Óscar Jordá, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor, "The Rate of Return on Everything, 1870–2015*," *The Quarterly Journal of Economics*, 2019, *134* (3), 1225–1298.

Peretto, Pietro, "Sunk Costs, Market Structure and Growth," *International Economic Review*, 1996, *37* (4), 895–923.

Peters, Michael, "Market Size and Spatial Growth - Evidence From Germany's Post-War Population Expulsions," *Econometrica*, 2022, *90*, 2357–2396.

— **and Conor Walsh**, "Population Growth and Firm Dynamics," forthcoming.

Piketty, Thomas, *Capital in the Twenty-First Century*, Harvard University Press, 2014.

— **and Gabriel Zucman**, "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010," *Quarterly Journal of Economics*, 2014, *129* (3), 1255–1310.

Ridder, Maarten De, "Market Power and Innovation in the Intangible Economy," *American Economic Review*, 2023, *113* (1), 33–66.

Straub, Ludwig, "Consumption, Saving and the Distribution of Permanent Income," *Mimeo Harvard*, 2019.

Sutton, John, *Sunk Costs and Market Structure*, MIT University Press, 1991.

Syverson, Chad, "Challenges to Mismeasurement Explanations for the US Productivity Slowdown," *Journal of Economic Perspectives*, May 2017, *31* (2), 165–86.

Trebbi, Francesco, Miao Ben Zhang, and Michael Simkovic, "The Cost of Regulatory Compliance in the United States," 2023. mimeo, Berkeley.

Westcott, Diane and Robert Bednarzik, "Employment and unemployment: a report on 1980," *Monthly Labor Review*, February 1981, *104* (4).

World Inequality Database, 2024. http://data.worldbank.org/indicator/SP.DYN.LE00.FE.IN.