Heterogeneous Decline in Sectoral Business Dynamism*

Agnieszka Markiewicz† Riccardo Silvestrini‡

May 11, 2021

Abstract

The U.S. economy has witnessed a slowdown in business dynamism: firms' entry and exit rates have declined while price markups, market concentration, and profit margins have increased. Yet, different sectors have been affected by these changes to a different degree. This paper argues that the heterogeneity in the slowdown of business dynamism across U.S. sectors can be explained by the different trends in firms' mobility. We define firms' mobility as the likelihood of changing the productivity rank by the firms within a sector. In a model with oligopolistic competition, we show that an increase in firms' mobility leads to a reallocation of market shares towards more productive and larger firms, that charge higher markups. As a result, sectoral markups, market concentration, and profit margins increase.

Keywords: Markups, Business Dynamism, Superstar Firms, Oligopolistic Competition, Heterogeneous Firms

JEL: D21, D50, L13

*We thank Yao Chen, Andrea Colciago, Robert Dur, Ana Figueiredo, John Haltiwanger, Andreas Pick, Lorenzo Pozzi, John Van Reenen, Peter Sedláček, Bauke Visser, Casper De Vries, Felix Ward, our discussants Byoungchan Lee, Rutger Schilpzand and Ali Sen, and the participants of the WEAI 2021 Virtual International Conference, Nederlandse Economenweek 2020 at De Nederlandsche Bank, KVS New Paper Sessions at Leiden University, Brown Bag Virtual Seminar and Tinbergen Institute Macro Day at Erasmus School of Economics and NuCamp Virtual PhD Workshop at Oxford University for their helpful comments and suggestions.

†Erasmus University Rotterdam and Tinbergen Institute, markiewicz@ese.eur.nl

‡Erasmus University Rotterdam and Tinbergen Institute, silvestrini@ese.eur.nl

1 Introduction

In the past several decades, the U.S. economy has witnessed a slowdown in business dynamism: firms' entry and exit rates have declined and price markups, market concentration and profit margins have increased. Yet, the magnitude of these trends differ across sectors. Specifically, the growth of markups and market concentration has been more pronounced in sectors with higher exposure to the ICT or to digital technologies (e.g. Bessen (2017), Calligaris, Criscuolo, and Marcolin (2018) and Diez, Fan, and Villegas-Sánchez (2019)).

This paper argues that the heterogeneity in the dynamics of markups, market concentration and profits across the U.S. sectors can be explained by a specific technological shift that occurred in ICT-intensive sectors. The key insight is that this technological shift has increased the likelihood of changing the relative position of firms in the cross-sectional productivity distribution. We call this likelihood *firms' mobility*. An increase in firms' mobility leads to a reallocation of market shares towards more productive and larger firms that charge higher markups. As a result, sectoral markups, market concentration, and profit margins increase.

Our strategy is the following. We first show that sectors with different trends in firms' mobility display heterogeneous dynamics in terms of ICT exposure and evolution of markups. To this end, we use the entire population of firms in Compustat data to build firm and sectoral-level quantities. Despite the general slowdown in business dynamism, as much as 35% of the U.S. NAICS-3 sectors exhibits a higher mobility of firms in 2016 than five decades ago. These sectors also display the steepest increases in markups and, importantly, were the most exposed to the ICT revolution.

Second, we build a model of a sector to analyse the impact of an increase in the mobility of firms on the dynamics of markups, market concentration and profits. In order to describe an industry with a granular and oligopolistic structure, our model is populated by a finite number of firms. The competitors differ by their idiosyncratic productivity level and we call superstars the firms endowed with the highest productivity level. All firms are subject to sequential idiosyncratic exit, entry, and productivity shocks. The model structure implies endogenous and time-varying idiosyncratic markups, which increase in the market share. The intuition is straightforward: a firm endowed with highest productivity, and thus with the lowest marginal cost, is able to charge the lowest price and to expand its market share. Its cost leadership allows it to charge the highest markup.

Third, we contrast the business dynamics between sectors that differ in the mobility of firms, due to a different exposure to the technological shift. The initial steady state of both sectors is calibrated to reproduce key features of the pre-80s U.S. economy, a period of low and stable markups and market concentration. One of the leading explanation for the observed decline in the U.S. business dynamism is an increase in industries' entry costs.[1] We therefore assume as a baseline that both sectors exhibit a general and sharp exogenous increase in entry costs. However, only the sector experiencing the technology shift sees an increase in the mobility of firms in cross-sectional productivity distribution.

The transition towards the second steady state demonstrates how market concentration and markups arise endogenously in response to the technological shift, when the latter is paired with an entry costs increase. A boost to firms' mobility increases the turnover of leaders, namely, the likelihood of becoming an industry superstar as well as the chances of becoming less productive and losing the leadership position. If a superstar firm experiences a detrimental shock, additional market shares are captured by the surviving highly productive firms, which can further increase their markups. This reallocation of market shares leads to an increase in concentration and sectoral markups during the transition to the new steady state.

An increase in firms' mobility makes entry attractive for small, unproductive firms and translates into the growth of their relative number over the transition period. The high number of unproductive firms, and their disadvantage in terms of productivity, pushes them to charge the monopolistic competition markup. Any increase in the sectoral markup is driven by the firms at the top of the markup distribution, which reflects the empirical

[1] Gutiérrez, Callum, and Philippon (2019) and Goldschlag and Tabarrok (2018) argue that the observed increase in entry costs is a result of the growing complexity of regulation. Bessen (2017) provides evidence that the increase in federal regulation favors incumbent firms and leads to an increase in market valuations and operating margins.

evidence.[2] Further, the sectoral markup grows primarily via the reallocation of market shares towards the firms that charge already higher markups, a fact that again mirrors the findings of the empirical literature (see De Loecker et al. (2020) and Van Reenen (2018)).

These findings contrast with the transition dynamics of the industry where we assume higher entry costs but no increase in the mobility of firms. There, the increase in markups is modest and entirely driven by the change in the composition of the firms in the sector: since the entry costs are higher and the chances of becoming an industry leader are slim, few new firms find profitable to join. The progressive exit of small and unproductive incumbents results in an economy with a larger share of firms that charge high markups, hence the increase in average markups.

Fourth, the comparison of two model economies implies a number of predictions. First, the larger is the increase in firms' mobility, the more important is the reallocation of market shares. Second, the stronger is the reallocation of market shares towards high-markup firms, the higher is the revenue weighted sectoral markup. Third, the steeper is the increase in firms' mobility, the lower is the average productivity in the sector. We evaluate these predictions using the entire population of firms from Compustat. The units of observation are 3-digit NAICS sectors over the period between 1965 and 2016.

The first theoretical result is that an increase in firms' mobility leads to the reallocation of market shares. Results from cross-section regressions confirm this key prediction of the model. In sectors with larger increases in firms' mobility, the reallocation of market shares has been the main driver of the markups' change.

The second theoretical prediction from our model is about the reallocation of market shares leading to higher revenue weighted markups. This prediction is also confirmed by the data, as the 3-digit NAICS sectors that had the largest reallocation of market shares also experienced the steepest increase in markups.

The third prediction is a declining average productivity in the transition period. This stems from the fact that a higher intra-sectoral mobility with associated larger profit possibilities attracts a large number of small, unproductive firms, and that leads to a

[2] According to De Loecker, Eeckhout, and Unger (2020), for instance, the median markup has been constant over the last several decades, while the increase in average markup has been driven by the growing markups of the top firms.

decline in average productivity. Again, the data reflect this as the 3-digit NAICS sectors where the reallocation has been more prominent also saw the largest declines in the average productivity.

Overall, the empirical results confirm the predictions of our model and suggest that the differences in sectoral dynamics are driven by heterogeneous trends in firms' mobility over the last several decades.

Our paper is related to the growing literature exploring the rise in market concentration and markups in recent decades, see Grullon, Larkin, and Michaely (2019) and De Loecker et al. (2020). Autor, Dorn, Katz, Patterson, and Van Reenen (2020) attribute these developments to the rise of superstar firms, whose advantage in productivity allows them to increase their market shares, without necessarily compromising consumers' welfare and firms' investments. De Loecker et al. (2020) and Gutiérrez and Philippon (2019) argue, instead, that the observed increase in the market concentration and markups reflects an increase in the market power of large firms as well as a reduction in the competition within U.S. industries. In a way, this paper reconciles both hypotheses. Although the increase in firms' mobility, in high-ICT sector, is the key trigger of the reallocation of market shares, an increase in fixed entry cost is necessary to generate the increasing trend in markups. Additionally, relative to the previously mentioned studies, this paper is the first to offer a theoretical explanation for the heterogeneity in the observed trends in markups and market concentration.

Our contribution is related to several papers that link the technological change to the slowdown in productivity and business dynamism, such as Aghion, Bergeaud, Boppart, Klenow, and Li (2019) and De Ridder (2019). Our paper is different from these studies in two key dimensions. First, we use a different modelling framework where the key friction arises from oligopolistic competition, when consumers display love for variety. This allows us to represent a product market where multiple firms are active at the same time with non-zero market shares. Second, in contrast to Aghion et al. (2019) and De Ridder (2019) who mainly focus on aggregate productivity dynamics, our paper studies the role of the increase in firms' mobility in shaping the magnitude of the growth in markups and market concentration *across* sectors.

In our model, superstar firms determine market concentration and markups' dynamics. The proposed mechanism conceptually builds on Gabaix (2011) since, in our environment with a finite number of firms, idiosyncratic shocks propagate to the aggregate economy. The most recent contributions to the literature studying the role of large firms for the aggregate dynamics include Carvalho and Grassi (2019) and Burstein, Carvalho, and Grassi (2019). In contrast to these papers, which study business cycle fluctuations, we focus on the long run dynamics across sectors.

Our propagation mechanism relies on the firms' stochastic entry and exit dynamics in the spirit of Jovanovic (1982) and Hopenhayn (1992). Importantly, to capture more realistically the current U.S. sectors' oligopolistic structure, our framework departures from the competitive environment of a continuum of firms in Hopenhayn (1992). More recently, in Bilbiie, Ghironi, and Melitz (2012) entry and exit dynamics represent a crucial transmission channel for the propagation of aggregate exogenous shocks. Instead of a continuum of firms, we assume an oligopolistic environment populated by a finite number of competitors. This assumption is similar to Colciago (2016), who however considers only homogeneous firms operating in a continuum of homogeneous sectors, hence removing any markup dispersion, the key variable of our investigation.[3]

The remainder of this paper is organized as follows. Section 2 describes the heterogeneity in the degree of decline in the U.S. business dynamism and links it to the use of ICT. In section 3, we describe a tractable, oligopolistic competition model with a finite number of firms populating a single sector. In section 4, we expose the calibration strategy. Section 5 carries out the main quantitative exercise where we compare the transition dynamics of a sector that experienced an increase in firms' mobility with the one that is subject to an increase in entry cost only. In Section 6 we confront our model directly with the data and test model's predictions using the universe of firms in Compustat data. Section 7 concludes.

[3]A seminal contribution about the modelling structure comes from Atkeson and Burstein (2008).

2 Facts

We start by presenting a set of empirical facts related to the slowdown in the U.S. business dynamism. First, using the findings of the existing studies, we briefly describe the signs of this slowdown. We highlight the fact that the degree of the decline in business dynamism varies across sectors and can be linked to the exposure to ICT. Second, we explore the hypothesis that specific developments in ICT-intensive sectors have led to the more pronounced growth in markups. Here, we provide some new results which show that (i) there is a large heterogeneity across sectors in the dynamics of firms' mobility over the productivity distribution, and (ii) that ICT-intensive sectors have seen the steepest growth in this mobility. Finally, we briefly review one of the leading explanations for the weakening business dynamism, namely, the increase in entry costs. Both the increase in firms' mobility and the increase in entry costs are the key exogenous inputs into our model and they are designed to match their empirical counterparts. In contrast, heterogeneous growth in markups, profits and market concentration are the endogenous outcomes.

2.1 Heterogeneous decline in business dynamism

There is now a considerable body of empirical evidence showing the slowdown in U.S. business dynamism measured by the trends in markups, market concentration, job creation and destruction and firms' entry and exit rates. We focus our analysis on the growing trends in price markups and market concentration.

A wide range of U.S. industries has experienced an increase in the concentration of sales and employment over the last four decades. This has been well documented by Autor et al. (2020), De Loecker et al. (2020), Grullon et al. (2019) and Gutiérrez and Philippon (2017), among others. Another sign of weakening dynamism has been a large increase in price markups, e.g. De Loecker et al. (2020), Autor et al. (2020).

While most of the theoretical studies aim to provide a comprehensive model to explain these aggregate trends, the empirical literature documents large heterogeneities in the degree of this increase across sectors. Figure A.1 in Autor et al. (2020) shows, for instance, that there has been hardly any increase in concentration in the wholesale trade between

1980 and 2015. In contrast, in retail trade, the concentration increased by 6pp over the same period. Bessen (2017) shows that the higher market concentration across a wide range of sectors can be linked to the use of proprietary IT systems. Similarly, Calligaris et al. (2018) and Diez et al. (2019) argue that, in most developed economies, markups are higher in ICT-intensive sectors than in less ICT-intensive sectors and that the difference between the two has been increasing over time.

2.2 Sectors' heterogeneity and the ICT

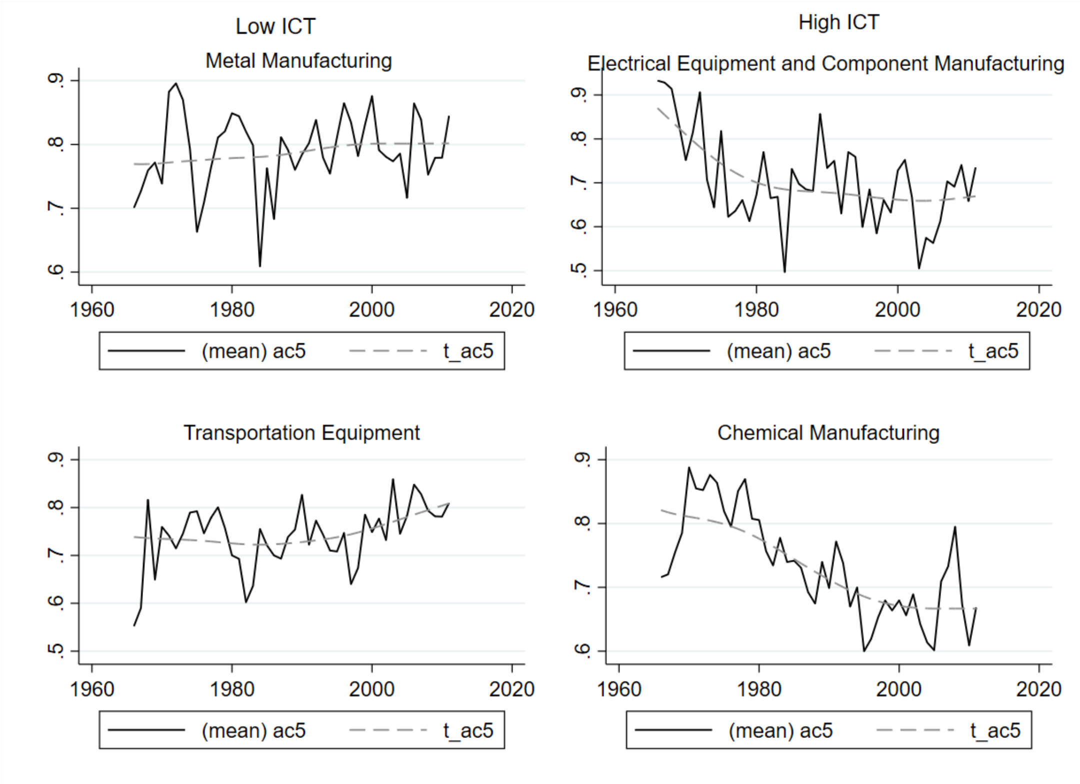
We explore the hypothesis that the differences in the degree of the decline in business dynamism across sectors can be linked to the exposure to ICT. Why would markups and market concentration grow faster in ICT-intensive sectors? When a sector is exposed to technological advances, a firm that is able to implement them becomes relatively more productive.[4] In contrast, an incumbent that fails to implement technological advances loses in terms of relative productivity and its market shares are reallocated towards more productive firm. The arrival of technological advances can therefore alternate the productivity rank of firms and, as a result, the industry leader can quickly emerge but can also easily lose its position.

We assume that the arrival of ICT exogenously shifts firms' mobility upwards. The reason is that we interpret the arrival of ICT, from firm's perspective, as an exogenous change in how technology operates. Specifically, it internalizes the faster decay and obsolescence rate of newer technologies or ideas (e.g. Koh, Santaeulàlia-Llopis, and Zheng (2020) and Bloom, Jones, Van Reenen, and Webb (2020)). In this context, an increase in firms' mobility implies a reduction in the average duration of a particular technology.

Similar to Comin and Philippon (2005), we construct a variable that measures the likelihood of the movements in the productivity rank of the firms within the sector: the firms' mobility. We use the data from Compustat and compute 5-year autocorrelation of the labour productivity rank for each firm, within its NAICS 3-digit sector. We then calculate the average autocorrelation for each sector and each year. A decline in the average

[4] ICT has been linked to the increase in U.S. firms' productivity in the 1990s, see e.g. Fernald (2015) and Bartelsman, Bassanini, Haltiwanger, Jarmin, and Scarpetta (2002).

Figure 1: Evolution of average autocorrelation in firms' productivity rank in several U.S. NAICS-3 sectors



Notes: The graph presents the 5-year average autocorrelation in firms' productivity rank in four U.S. NAICS-3 sectors: Metal Manufacturing (331), Electrical Equipment, Appliance, and Component Manufacturing (335), Transportation Equipment (336), and Chemical Manufacturing (325). Labour productivity is computed using firms' revenues and total hours worked in a year.

autocorrelation implies an increase in firms' mobility. The definition of high versus low ICT users is based on the industries' classification by Charles and Marye (2009).

Figure 1 plots the evolution of the productivity rank autocorrelation between 1965 and 2016 in 4 chosen manufacturing sectors. 2 low ICT users, (i) metal manufacturing and (ii) transportation equipment manufacturing, are plotted in the left panels while 2 high ICT users, (i) electrical equipment, appliance, and component manufacturing and (ii) chemical manufacturing, are plotted in the right panels. The black solid lines depict the autocorrelation coefficients themselves and the grey dashed lines plot the HP filtered trends.

In both low ICT-intensity sectors, firms' mobility (autocorrelation) remains essentially flat during the entire period, with a modest decrease (increase) in the second half of the sample. In contrast, high ICT-intensity industries experienced an increase in firms' mobility over the entire sample period.

We formally test whether industries with a higher exposure to the ICT have seen an increase in firms' mobility by regressing the change in the average 5-year autocorrelation in labor productivity rank between 2016 and 1965 on different measures of ICT intensity. Since the intangible capital is strongly associated with the ICT, we use its share in firms' total capital to proxy for the firms' exposure to ICT. We compute both the simple average across firms and a weighted average at the sectoral level. Although the share of intangibles changes over time and therefore allows us to investigate sectoral dynamics of ICT investment, it is only a rough proxy of ICT intensity. Therefore, we construct its direct measure, based on Charles and Marye (2009) classification of high ICT intensive industries and create a corresponding dummy.

In Table 1, we show the results of regressing the change in the autocorrelation of labor productivity ranks on the change in several sectoral characteristics over the same period. Because the sectoral quantities display large yearly fluctuations, we isolate the trends by using HP filters and compute the change between the last and first observation available.

Results documented in the first three columns of Table 1 detect a striking relationship between changes in firms' mobility and the ICT exposure. The sectors that have experienced the steepest increase in firms' mobility over the sample period (i.e. a decline in autocorrelation) are the ones that were exposed the most to the ICT revolution. This result is robust to all three measures of ICT intensity and various specifications including industry controls other than shifts in productivity and markups, as, for instance, profits and concentration. The second last column of Table 1 also shows that the revenue weighted markups have grown faster in sectors with increasing firms' mobility. This suggests that the latter may be the key to understand the steeper increase in markups in ICT-intensive sectors.

2.3 Entry Costs

A sharp increase in industries' entry costs has been proposed as one of the main reasons for the observed decline in the U.S. business dynamics, e.g. Gutiérrez et al. (2019) and

Table 1: **ICT exposure and firms' mobility**

$\Delta\rho_i = \alpha + \beta ict_i + \gamma\Delta X_i + \sigma_i$				
ICT	Δ IC-w	Δ IC-unw	Δ Markup	Δ Productivity
-0.169***				
(0.014)				
-0.202***			-0.093***	
(0.014)			(0.010)	
-0.213***			-0.101***	0.110
(0.014)			(0.010)	(0.09)
	-0.056***			
	(0.004)			
	-0.063***		-0.256***	
	(0.004)		(0.030)	
	-0.050***		-0.280***	-0.024***
	(0.004)		(0.030)	(0.006)
		-0.052***		
		(0.006)		
		-0.052***	-0.116***	
		(0.006)	(0.008)	
		-0.047***	-0.123***	-0.007
		(0.006)	(0.008)	(0.009)

Notes: This table shows results of regressions of the form: $\Delta\rho_i = \alpha + \beta ict_i + \gamma\Delta X_i + \sigma_i$. The unit of observation is a sector. Δ denotes the change between the last and the first value of the HP trend. Autocorrelation rank ρ_i is calculated as average correlation of firms' labour productivity rank within each sector between t and $t+5$, t being a year. Controls in X_i include revenue weighted-average labour productivity and revenue weighted markup in sector i. Values in brackets report White-corrected standard errors. Markups are constructed based on the following specification derived from the model: $\mu_t = \frac{\theta}{\theta-1}\frac{1}{1-\omega_t}$. Market shares, ω_t, are computed in terms of real revenues and elasticity of substitution θ is set to a standard value of 6. ICT is an ICT intensity dummy and IC stands for non-tangible capital proxy for ICT with $IC - unw$ being a simple average of firms' intangibles share of total capital stock in sector and $IC - w$ its weighted counterpart. X_i are sectoral level controls. Values in brackets report White-corrected standard errors.

Gutiérrez and Philippon (2019). According to Davis (2017) and Gutiérrez et al. (2019), the observed increase in entry costs is a result of the growing complexity of regulation. Similarly, Goldschlag and Tabarrok (2018) show that the increase in complexity in regulation is an economy-wide trend.

In what follows, we describe a tractable model of an oligopolistic sector. We use the model intuition to show how the reallocation of market shares and the change in composition of firms populating the sector affects markups and concentration. We then experiment with the model by introducing the empirical features described above: (i) increase in entry cost and (ii) boost in firms' mobility.

3 Model

To analyse the impact of increasing mobility of firms on the business dynamics we build a sectoral model in the spirit of Edmond, Midrigan, and Xu (2015). In order to describe an industry with a granular, oligopolistic structure, our model is populated by a finite number of firms. This is important because it implies that every incumbent possesses a non-atomistic mass, which translates into a strictly positive market share that depends on the idiosyncratic productivity level and on the number and type of active competitors. Motivated by Pugsley, Sedláček, and Sterk (2020), we allow for ex ante heterogeneities across firms; firms differ by their productivity level only. The industry is subject to endogenous sequential entry, exogenous exit and idiosyncratic productivity shocks. Since the core dynamics take place on the firms' side, a simple representative household is modelled on the demand side, in the spirit of Bilbiie et al. (2012).

3.1 Firms and Competition

3.1.1 Production

The economy features a single sector in which firms compete under oligopolistic competition á la Cournot, producing differentiated varieties $y(i)$, where the index represents firm type i [5]. After the production takes place, the individual goods are aggregated into the

[5] The case of Bertrand competition is described in Online Appendix 2.

bundle Y_t through a standard C.E.S. function. The aggregate output Y_t is used solely for consumption purposes. We assume that the economy is populated by a *finite* number of firms N_t and, as a result, every incumbent possesses a non-atomistic mass, which is reflected in a strictly positive market share that depends on the idiosyncratic productivity level and on the number and type of active competitors. When allowing for oligopolistic competition, the distribution of the market shares has a clear impact on the markup distribution: our focus justifies this assumption.

Firms are heterogeneous in their productivity level $x(i)$, which is drawn from a known discrete distribution function $f(x)$. Idiosyncratic productivity is assumed to be time variant and its dynamics can be summarized by a stationary and non-degenerating Markov process: in each period t, q_{ij} represents the probability of moving from productivity level $x(i)$ to $x(j)$ between period t and period $t+1$ and $\sum_j q_{ij} = 1$, $\forall i$. The number of distinct and active productivity levels, i.e. the number of firm types, is represented by S.[6]

Within each productivity type firms are identical and, thus, they are entirely identified by their productivity level. To simplify the notation, in the following, the variables related to a firm with a productivity, $x(i)$, are identified by the index (i). We call highly productive firms characterized by $x(S)$ superstar firms. Each firm type i produces an imperfectly substitutable good $y_t(i)$, which is aggregated into the bundle Y_t. The aggregator function is a standard C.E.S. function for discrete aggregation:

$$Y_t \equiv \left[\sum_{j=1}^{N_t} y_t(j)^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} = \left[\sum_{i=1}^{S} N_t(i) y_t(i)^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \tag{1}$$

where θ is the elasticity of substitution between varieties, with $\theta > 1$, $N_t(i)$ is the number of firms endowed with productivity level i, and N_t represents the total number of incumbents.

Production is linear in labor $l_t(i)$ and depends on the idiosyncratic productivity $x_t(i)$, which acts as a labor-augmenting technology:

$$y_t(i) = x_t(i) l_t(i) \tag{2}$$

[6]It is important to specify *active*. We model $S+1$ types: the type 0, i.e. the firm with productivity $x(0)$, mimics a productivity level that is not enough to guarantee firm survival and it is a proxy for fixed costs of production, which are not modelled explicitly. Thus, if a firm draws this productivity it is forced to leave the market immediately and this allows us to focus on the dynamics of S types only.

Firms compete under oligopolistic competition a la Cournot.[7] Firms maximize their per-period nominal profits by choosing the optimal quantity $y_t(i)$:

$$\max_{y_t(i)} p_t(i)y_t(i) - W_t l_t(i) \tag{3}$$

Subject to (2) and to the aggregate demand constraint:

$$y_t(i) = \left(\frac{p_t(i)}{P_t}\right)^{-\theta} Y_t$$

where W_t is the nominal wage, while P_t is the aggregate price index, defined as a function of the individual prices $p_t(i)$ as: $P_t \equiv \left[\sum_{j=1}^{N_t} p_t(j)^{1-\theta}\right]^{\frac{1}{1-\theta}} = \left[\sum_{i=1}^{S} N_t(i)p_t(i)^{1-\theta}\right]^{\frac{1}{1-\theta}}$.

In Online Appendix 2, we show that under this form of competition the optimal real price $\rho_t(i) = p_t(i)/P_t$ satisfies:

$$\rho_t(i) = \mu_t(i)\frac{w_t}{x_t(i)} \tag{4}$$

where $w_t = W_t/P_t$ is the real wage. As in Edmond et al. (2015), the markup $\mu_t(i)$ can be defined as a function of the market share $\omega_t(i)$, where $\omega_t(i) = \rho_t(i)^{1-\theta}$ given our structural assumptions about the sector(s). The markup $\mu_t(i)$ reads as:

$$\mu_t(i) = \left(\frac{\theta}{\theta-1}\right)\left(\frac{1}{1-\omega_t(i)}\right) \tag{5}$$

Note that equation (5) nests monopolistic competition case: when $\omega_t(i) \to 0$ the resulting markup is the standard $\theta/(\theta-1)$, independent from the number and type of competitors and from the idiosyncratic productivity level. Our markup captures both features through the market share and can be summarized as an extra idiosyncratic markup over the monopolistic benchmark which increases in the relative idiosyncratic productivity level. The markup in (5) also increases in the market share $\omega_t(i)$ and decreases in the real idiosyncratic price $\rho_t(i)$. Using equation (4) and the aggregate demand constraint, we can write the real profits for the firm with productivity $x(i)$ as:

$$d_t(i) = \left(1 - \frac{1}{\mu_t(i)}\right)\rho_t(i)^{1-\theta}Y_t \tag{6}$$

Profits $d_t(i)$ are increasing in the markup $\mu_t(i)$, with a lower bound on zero whenever $\mu_t(i) = 1$, as under perfect competition. Given that the markup is increasing in the market

[7] Under the chosen specification, incumbents internalize that the quantity they select affects the sectoral output Y_t, but not the total expenditure P_tY_t allocated to consumption.

share, profits are increasing in the market share as well. The intuition is straightforward: a superstar firm endowed with technology $x(S)$ and associated lower marginal cost is able to charge a lower relative price, with respect to the other incumbents. Because of its cost leadership, it gains larger market share and can charge, in turn, higher markups that generate higher profits.

3.1.2 Composition and Reallocation in the Average Price Markup

A simple average markup in a sector can change for two reasons: either the composition of the population of firms in the sector changes or the market shares are reallocated between the incumbents.[8] We call these two channels *composition* and *reallocation* channel and assess their individual impact on the average sectoral markup. For the sake of simplicity, we focus the following discussion on the simple average markup, although the results are similar for weighted markup. The average price markup in a sector $\bar{\mu}_t$ can be rewritten as, using the definition of idiosyncratic price markup in (5):

$$\bar{\mu}_t = \frac{\theta}{\theta - 1} \left[\sum_{i=1}^{S} \gamma_t(i) \varepsilon_t(i) \right] \tag{7}$$

where $\gamma_t(i)$ represents the period-t fraction of type i firms over the total number of incumbents, namely $\frac{N_t(i)}{N_t}$, and $\varepsilon_t(i)$ describes the inverse of the complement of the type-i idiosyncratic market share, i.e. $\frac{1}{1-\omega_t(i)}$. Note that $\varepsilon_t(i)$ is increasing in the market share $\omega_t(i)$ and that $\omega_t(S) > \omega_t(S-1) > ... > \omega_t(1)$, hence $\varepsilon_t(S) > \varepsilon_t(S-1) > ... > \varepsilon_t(1)$.

Composition channel *Ceteris paribus*, a change in the *composition* of the industry that results in relatively higher number of low productive firms, leads to a decrease in the average price markup $\bar{\mu}_t$. The opposite is true if the relative number of superstar firms increases.

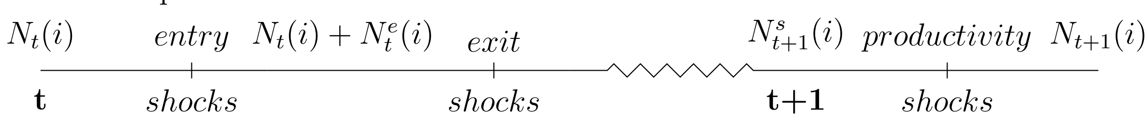
Reallocation channel Keeping the composition of the sector unaltered, a *reallocation* of market shares from small to large incumbents increases the average price markup, $\bar{\mu}_t$, e.g. a decrease in $\varepsilon_t(1)$ and an increase in $\varepsilon_t(S)$ increase $\bar{\mu}_t$ if the $\gamma_t(i)$ are kept constant. We derive the intuition of reallocation channel in Online Appendix 3.

[8]In this simple example we do not distinguish between within changes in the composition of the incumbents and changes in composition driven by new entrants/exiting firms.

In our general equilibrium framework, the two channels are not entirely independent of each other. Every change in the relative number of competitors alters not only the composition of the sector but also the competitive pressure and, as a result, the relative prices. Given our market structure where the market share directly depends on the relative price, $\omega_t(i) = \rho_t(i)^{1-\theta}$, this leads to a reallocation of market shares. In the model simulations, we will discuss in detail how each shock affects the price markup and through which channel primarily.

3.1.3 Idiosyncratic Exit, Entry and Productivity Shocks

Firms are subject to idiosyncratic entry, productivity and exit shocks. At the beginning of period t, surviving incumbents from period $t-1$, $N_t^s(i)$, are hit by productivity shocks, whose ex ante type-specific probabilities are known and determined by a Markov process. The potential entrants also draw a productivity level, $x(i)$, which determines whether the firms can successfully enter the market. At the very end of period t, each firm can be hit by an idiosyncratic exit shock. The timing of the shocks' occurrence is summarized in the timeline and explained in detail below.



Before entering the market, each potential entrant draws a productivity level $x(i)$ from a discrete distribution function $f(x)$, the same as the incumbents'. With probability Ω_0, the potential entrant is successful, i.e. with probability $1 - \Omega_0$ the firm draws the null productivity and cannot join the market. Given the number of potential entrants M_t, the number of successful entrants N_t^e follows a binomial distribution with success probability Ω_0 and M_t trials:

$$Pr\left[N_t^e = x\right] = \binom{M_t}{x} \left(\Omega_0\right)^x \left(1 - \Omega_0\right)^{M_t - x} \tag{8}$$

where $0 \leq x \leq M_t$. Conditional on successful entry, there is a probability Ω_i of drawing the productivity level $x(i)$. Again, given N_t^e, the type of successful entrants follows a multinomial distribution with N_t^e trials, in formula:

$$Pr\left[N_t^e(i) = x, N_t^e(j) = y_j\right] = \frac{N_t^e!}{x! \prod_{j=1, j\neq i}^{S} y_j!} \Omega_i^x \prod_{j=1, j\neq i}^{S} \Omega_j^{y_j} \tag{9}$$

where j represents all the types j for which $j \neq i$ and x_j all the possible discrete values for all the j types such that $x + \sum_j y_j = N_t^e$. Moreover, note that $\sum_{i=1}^{S} \Omega(i) = 1$. Given the realization of the stochastic process, the following must hold:

$$\sum_{i=1}^{S} N_t^e(i) = N_t^e \tag{10}$$

After the successful draw, and before knowing the exact productivity level assigned for period t, entry cost must be payed by each entrant in order to join the market. Real entry costs are measured in terms of units of labor and they are equal to $f_{e,t} w_t$.[9] The entry fee $f_{e,t} w_t$ is payed conditional on successful entry only. Firms enter the market up to the point where their expected value is at least equal to the cost of entry. The free entry condition is:

$$(1 - \Omega_0)\, 0 + \Omega_0 \left[\sum_{i=1}^{S} \Omega_i e_{i,t}(i) - f_{e,t} w_t \right] \geq 0 \tag{11}$$

where $e_{i,t}(i)$ is the value of the potential entrant when the productivity $x(i)$ is drawn. The value of a potential entrant can be written recursively:

$$e_{i,t}(i) = E_t \Lambda_{t+1,t} \left(1 - \delta(i)\right) \left[\sum_{j=1}^{S} q_{ij} \left(d_{i,t+1}(j) + e_{i,t+1}(j) \right) \right] \tag{12}$$

where $\Lambda_{t+1,t}$ is the stochastic discount factor between period t and period $t+1$, $\delta(i)$ is the type-specific exit probability and $d_{i,t+1}(j)$ is the period $t+1$ profit for the marginal entrant that joins with productivity $x(i)$ in period t and it is endowed with productivity $x(j)$ in period $t+1$. When considering entry, the marginal entrant internalizes that its action affects the (expected) number of operating firms in the following periods, hence the subscript $_i$ in the variables, to distinguish them from the incumbents' ones. This effect in known in the literature as the *business stealing dynamics*. In particular, the value of a potential entrant is different from the value of an incumbent since the first considers that, by entering in the market with a productivity $x(i)$, the expected number of competitors

[9] Alternatively, we could have specified entry cost in terms of consumption. However, under this specification, profits are increasing in entry, if entry is low enough.

of type j in the following period is

$$E_t N_{t+1}(j) = q_{ij}\left(1-\delta(i)\right)\left(N_t^e(i)+N_t(i)+1\right) + \sum_{k=1,k\neq i}^{S}\left[q_{kj}\left(1-\delta(k)\right)\left(N_t^e(k)+N_t(k)\right)\right] \tag{13}$$

for $j \in [1, S]$. This has an impact on the expected sectoral price and, hence, on the expected profits and firm value and it is a direct consequence of the assumption of the finite number of firms.

At the very end of period t, each firm i can be hit by an idiosyncratic exit shock with a time invariant exogenous probability $\delta(i)$. The exit shocks realize after the entry of new firms has occurred, and they can hit potentially every firm, not only the incumbents already operating in the market. Since the entrants can start producing only in the period that follows their entry, as in Bilbiie et al. (2012), new firms may be forced to leave the sector even before being active. This assumption simplifies the formalization of the budget constraint of the household, in particular of the investment decisions.

After exit has occurred at the end of period t, the economy enters period $t+1$ with a given number of surviving firms of type (i) $N_{t+1}^s(i)$. Conditional on the number of survivors, the number of incumbents of each type in period $t+1$ is determined by the realization of idiosyncratic productivity shocks.[10] In particular, the number of firms endowed with the productivity $x(i)$ in period t which survives in period $t+1$, i.e. $N_{t+1}^s(i)$, follows a binomial distribution with success probability $1-\delta(i)$ and $N_t^e(i)+N_t(i)$ trials. Formally:

$$Pr\left[N_{t+1}^s(i) = x\right] = \binom{N_t^e(i)+N_t(i)}{x}\left[1-\delta(i)\right]^x \left[\delta(i)\right]^{N_t^e(i)+N_t(i)-x} \tag{14}$$

for $i \in [1, S]$ and $0 \leq x \leq N_t^e(i)+N_t(i)$ where x is an integer. To conclude, a description of the productivity shocks follows. Given the number of survivors of each type, the realization of multinomial distributions determines the fraction of survivors endowed with productivity $x(i)$ that keeps their own productivity level, i.e. $N_{t+1}^i(i)$, against the number

[10] The law of large numbers cannot be used in our framework due to the finite number of firms. As a result, it is not possible to reduce the idiosyncratic stochastic processes to their expected values, making the aggregate exit, entry and productivity dynamics deterministic laws. Due to this feature, the law of motion of firms evolves according to (the realization of) binomial distributions. This is the reason why we continue to talk about exit, entry and productivity dynamics as shocks.

of survivors that switch to any of the remaining types j, $N^i_{t+1}(j)$.[11]

$$Pr\left[N^i_{t+1}(i)=x, N^i_{t+1}(j)=y_j\right]=\frac{N^s_{t+1}(i)!}{x!\prod_{j=1,j\neq i}^{S}y_j!}q_{ii}^{x}\prod_{j=1,j\neq i}^{S}q_{ij}^{y_j} \tag{15}$$

The following holds after the realization of the productivity shocks:

$$N^s_{t+1}(i)=\sum_{j=1}^{S}N^i_{t+1}(j) \tag{16}$$

After the realizations of these processes, the number of firms for each type in period $t+1$ can be computed as:

$$N_{t+1}(i)=\sum_{j=1}^{S}N^j_{t+1}(i) \tag{17}$$

These specifications determine the number of firms in period $t+1$ and their types, which are considered as given by the households in period $t+1$.

3.2 Households

The household side of the economy is kept as simple as possible. The economy is populated by a continuum of identical households of unitary mass. The representative household consumes an aggregate consumption bundle c_t and supplies labor L_t. The quantity of labor supplied has two purposes: a fraction of the aggregate labor supply is employed in the production process and the remaining part is used to invest in new firms.[12] The

[11]The same result can be obtained by using a chain of conditional binomial distribution, as shown in Online Appendix 5.

[12]In the standard model with monopolistic competition, the marginal entrant has atomistic mass. Hence, in equilibrium: $\sum_{i=1}^{S}e_t(i)N^e_t(i)=\sum_{i=1}^{S}e_{i,t}(i)N^e_t(i)=N^e_t w_t f_{e,t}=w_t L^e_t$, where L^e_t is the fraction of labor supplied used to repay entry costs. However, in this setting, $e_t(i)\neq e_{i,t}(i)$ and profits possibilities are not exploited completely due to the integer nature of the number of competitors and entrants. Due to these issues, we must assume the existence of an investment fund. The fund creates new firms at their costs and sell them at their higher value to the households. Given that profit possibilities are not exhausted, the fund makes positive profits, and these rents are distributed as lump sum transfers to the households, closing the budget constraint. With the introduction of the fund, in equilibrium we still have that the labor supply that is used to invest in new firms, i.e. L^e_t, is equal to $N^e_t f_{e,t}$.

aggregate consumption bundle can be defined through a discrete C.E.S. aggregator as:

$$c_t \equiv \left(\sum_{j=1}^{N_t} c_t(j)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}} = \left(\sum_{i=1}^{S} N_t(i) c_t(i)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}} \tag{18}$$

where $c_t(i)$ is the consumption of the variety produced by a firm with productivity $x(i)$. The minimization of the total expenditure $\sum_{i=1}^{N_t} p_t(i)c_t(i)$ subject to equation (18) delivers the following demand function: $c_t(i) = \left(\frac{p_t(i)}{P_t}\right)^{-\theta} c_t$.

The household can invest in a portfolio that represents the ownership of the firms, by purchasing shares x_{t+1}. Finally, the households receive the rents F_t from the investment fund which pays the entry costs for every successful entrant. The rents equal, thus, the difference between the total value of the entrants and the total entry costs paid.

The household maximizes her lifetime utility in real terms, U:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln c_t - \chi \frac{L_t^{1+\frac{1}{\phi}}}{1+\frac{1}{\phi}} \right) \tag{19}$$

where c_t and L_t are aggregate consumption and labor supply, as defined above, $0 \leq \beta \leq 1$ is the discount factor, $\chi \geq 0$ is a scale parameter for the disutility of labor, useful for calibration, and $\phi > 1$ represents the elasticity of labor supply. The maximization is subject to the budget constraint:

$$c_t + x_{t+1}V_{t+1,t} = L_t w_t + x_t V_t + F_t \tag{20}$$

For the sake of clarity, the period t value of the entire portfolio, gross of dividends, is represented here by V_t, while $V_{t+1,t}$ describes the period t value of the new portfolio purchased in period t to be carried to period $t+1$. Given the definition above, $w_t L_t$ describes the total labor income.

The value of the portfolio is the sum of the net-of-dividend value of the portfolio in period t, A_t, and dividends payments, D_t:

$$V_t = A_t + D_t = \left(\sum_{i=1}^{S} \left[e_t(i) + d_t(i) \right] N_t(i) \right) \tag{21}$$

Entry occurs in the beginning of period t, when the representative household purchases the portfolio to be brought to period $t+1$. Because this happens before the exit shock occurs (end of period t), the household does not know the number of surviving firms and finances them all. The value of the purchased portfolio is:

$$V_{t+1,t} = \sum_{i=1}^{S} \left(e_t(i) \left[N_t(i) + N_t^e(i) \right] \right) \tag{22}$$

Finally, for completeness, the rents received by the intermediary are equal to the following (note that they do not affect the maximization of the household being a lump sum transfer):

$$F_t = \sum_{i=1}^{S} e_t(i) N_t^e(i) - N_t^e w_t f_{e,t}$$

The F.O.C. with respect to c_t and to L_t are:

$$\lambda_t = \frac{1}{c_t} \tag{23}$$

and

$$\chi L_t^{\frac{1}{\phi}} c_t = w_t \tag{24}$$

By combining the F.O.C.s with respect to c_t and x_{t+1}, the following equation can be written:

$$V_{t+1,t} = \beta E_t \left[\left(\frac{c_t}{c_{t+1}} \right) V_{t+1} \right] \tag{25}$$

This condition is equivalent, in expectation, to the following Euler equations for assets value that derive from the definition of $e_t(i)$, similar to the one for $e_{i,t}(i)$ provided above, and of the stochastic discount factor $\Lambda_{t+1,t}$ as $\beta \frac{c_t}{c_{t+1}}$:[13]

$$e_t(i) = \beta \left(1 - \delta(i) \right) E_t \frac{c_t}{c_{t+1}} \left[\sum_{j=1}^{S} q_{ij} \left(d_{t+1}(j) + e_{t+1}(j) \right) \right] \tag{26}$$

for $i = 1, 2, ..., S$.

3.3 Aggregation

In equilibrium, the representative household holds the entire portfolio of firms, i.e. $x_{t+1} = x_t = 1$. Using the definition of F_t the following resource constraint can be obtained:

$$c_t + N_t^e w_t f_{e,t} = L_t w_t + \sum_{i=1}^{S} d_t(i) N_t(i) \tag{27}$$

[13] This is true provided that $E_t \beta \frac{c_t}{c_{t+1}} (e_{t+1}(i) + d_{t+1}(i)) E_t N_{t+1}(i) = E_t \beta \frac{c_t}{c_{t+1}} (e_{t+1}(i) + d_{t+1}(i)) N_{t+1}(i)$

From the definition of the aggregate output $Y_t = c_t$, we obtain aggregate labor supply:

$$L_t = L_t^p + L_t^e = N_t^e f_{e,t} + \sum_{i=1}^{S} l_t(i) N_t(i)$$

where L_t^p represents the labor employed for production, L_t^e the one used to finance entry. Finally, from the definition of the aggregate price:

$$1 = \sum_{i=1}^{S} N_t(i) \rho_t(i)^{1-\theta} \tag{28}$$

3.4 Risky Steady State

Due to the assumption regarding the finite number of firms, paired with stochastic entry, exit and productivity dynamics, the standard definition of the deterministic steady state, where entry perfectly balances exit, cannot be applied. Building on Coeurdacier, Rey, and Winant (2011) and Juillard (2011), we derive instead a version of the *Risky Steady State*. The proposed equilibrium with risk-neutral agents describes a state in which no shocks occur but economic agents take into consideration the possibility that the shocks might happen in the future. In the steady state, households invest in entry up to the point where it compensates the *expected* exit, thus keeping the number of incumbents and entrants constant over time in expectation.[14] As in the calibration below, we consider a particular case in which the number of active types S is equal to 3.

The definition of the *Risky Steady State* follows. Given the steady state value for the exogenous entry costs f_e, and given the calibration of the exogenous parameters, the steady state is the set $\{\rho(i), w, d(i), Y, N(i), N^e(i), N^e, M, e(i), L, c\}$ with $i = 1, 2, 3$ that solves the system of equations described in Appendix A. Note that this deterministic equilibrium holds in expectation, since we consider the expected realizations of the stochastic processes. Because of that, no restriction is put on the integer nature of the variables relative to the number of firms (incumbents and entrants). This means that a marginal entrant can be of infinitesimal size, as the free entry condition closes: the effect of its entry on the mass

[14] The steady state we present here is the one that would also emerge under the assumption of oligopolistic competition within sectors and continuum of sectors in aggregate, in the spirit of Atkeson and Burstein (2008).

of competitor is negligible. Given that, marginally, entry does not affect the sectoral price and the profits here, the value of an incumbent or of a potential entrant of the same type is equivalent.

4 Calibration

The initial steady state is calibrated to reproduce key features of the U.S. industries in the year 1980. The values of the internally and externally calibrated parameters are presented in Table 2. Since the two permanent shocks are the key drivers of model dynamics, we pay particular attention to their calibration and present their values before and after the occurrence of the shock in a separate table.

To keep the analysis tractable, the discrete distribution function of productivities consists of four mass points: $x(0)$, $x(1)$, $x(2)$ or $x(3)$, with $x(3) > x(2) > x(1) > x(0)$. Thus, we keep track of four types of firms and three active types, i.e. $S = 3$. We believe that the classification of firms into 3 types does not compromise our calibration exercise. Several studies show how little firms differ from each other even if in different percentiles of the size distribution. The main differences can be found between very top firms and the rest, see for instance Crouzet, Mehrotra, et al. (2017). The null productivity mimics the presence of fixed costs of production, which are not introduced explicitly in the model.

Each period t represents a quarter and the discount factor β is fixed at the value of 0.99, implying an annual interest rate of approximately 4%. The parameter that governs the elasticity of substitution between goods, θ, is set to 6. The value is standard in the literature and gives a baseline monopolistic competition markup of 1.2. This implies that all the incumbents charge a markup equal or higher than 1.2, depending on the size of their market shares. The elasticity of labor supply ϕ is equal to 3. We normalize the multiplier for the disutility of labor, χ, to a value such that the labor supply in the second steady state, i.e. year 2005, is equal to one. Under this specific calibration, $\chi = 0.8576$. The remaining internally calibrated parameters are set to match the key U.S. economy variables.

4.1 Productivity levels, entry and exit probabilities

Empirically, the firm level productivity is distributed under a power law. To accommodate this fact, we assume that the function $f(x)$ is the discrete counterpart of a continuous Pareto distribution with a minimum at 1 and a tail parameter $\kappa = 1.05$.[15] This results in three productivity levels $x(1) = 1.315$, $x(2) = 4.631$ and $x(3) = 80.309$ reflecting the underlying assumption about firms' productivity: they represent, respectively, the 25^{th} percentile, the 80^{th} percentile and 99^{th} percentile of the Pareto distribution. Similarly, the ex-ante probability of being a successful entrant, i.e. Ω_0, is equal to 0.75 reflecting the assumption that an entrant is successful whenever a productivity higher than $x(1)$ is drawn from the continuous Pareto distribution. This implies that $\Omega_0 = Pr[x \geq x(1)]$, which gives $\Omega_0 = 0.75$. Given that Ω_2 and Ω_3 represent, respectively, the conditional probability of drawing productivity level $x(2)$ and $x(3)$ for a successful entrant, they are set to 0.2533 and 0.0133. Note that $\Omega_2 = Pr[x \geq x(2) \wedge x \leq x(3)|x \geq x(1)] = \frac{0.19}{\Omega_0} = 0.2533$ and that $\Omega_3 = Pr[x \geq x(3)|x \geq x(1)] = \frac{0.01}{\Omega_0} = 0.0133$. Given this calibration, $x(3)$ describes large and highly productive superstar firms, $x(2)$ medium-big incumbents while $x(1)$ unproductive small to medium firms.

The parameter that determines the likelihood of an exit shock for type 2 firms, $\delta(2)$, is normalized to 0.0025, which implies an approximate one percent yearly destruction rate of type 2 businesses. In order to identify all the entries of Markov transition matrix, we need to impose several restrictions. Specifically, our identification strategy relies on the assumption that superstar firms are a subset of highly-productive type 2 firms and, therefore, they share some characteristics with $x(2)$ incumbents. Specifically, we assume that exit probabilities are the same, i.e. $\delta(2) = \delta(3) = 0.0025$, and that the probabilities of receiving a detrimental productivity shock that lowers the productivity level to $x(1)$ are equal for both types as well, i.e. $q_{31} = q_{21}$.[16] $\delta(1)$ is set to 0.035. This calibration delivers a yearly business destruction rate of 10%, as in Colciago (2016). It also implies that the probability that a new entrant is still active in the market after three years following

[15] Axtell (2001) estimates a tail parameter of 1.059, approximately.

[16] A different approach would be to set $q_{31} = 0$, as we do later for q_{13}, and to iterate the Markov process to pin down q_{21}, given $\delta(2)$ and $\delta(3)$.

Table 2: **Calibration of fixed exogenous parameters**

Parameter	Calibration	Target
S	3	Four types of firms, three active
β	0.99	$\approx 4\%$ yearly interest rate
θ	6	Monopolistic competition markup = 1.2
ϕ	3	Elasticity of labor supply close to King and Rebelo (1999)
χ	0.8576	Aggregate labor supply = 1 in post-ICT steady state
$x(1)$	1.3155	25^{th} percentile in a Pareto distribution with $\kappa = 1.05$
$x(2)$	4.631	80^{th} percentile in a Pareto distribution with $\kappa = 1.05$
$x(3)$	80.309	99^{th} percentile in a Pareto distribution with $\kappa = 1.05$
Ω_0	0.75	$Pr[x > x(1)]$ under Pareto with $\kappa = 1.05$
Ω_2	0.2533	$Pr[x \geq x(2) \wedge x \leq x(3) \| x \geq x(1)]$ under Pareto with $\kappa = 1.05$
Ω_3	0.0133	$Pr[x \geq x(3) \| x \geq x(1)]$ under Pareto with $\kappa = 1.05$
$\delta(1)$	0.035	Yearly business destruction rate $\approx 10\%$, Colciago (2016)
$\delta(2)$	0.0025	Normalization, yearly business destruction rate for type 2 $\approx 1\%$
$\delta(3)$	0.0025	$= \delta(2)$ for identification strategy

Notes: The table presents the calibration of the exogenous parameters. The second column describes the value assigned to the parameters. The third column describes the targets of the calibration. These parameters are kept fixed along the entire transition in every simulation.

its entry, is equal to 0.75 (0.67 if restricted to small entrants only). Despite being non-targeted, this value is in line with the estimation of 0.61 in Hyytinen, Pajarinen, and Rouvinen (2015).

4.2 A shift in firms' mobility and in entry cost

We model two simultaneous, structural shifts that took place in the U.S. economy in the 1980s. The first one corresponds to an increase in firms' mobility within sectoral productivity distribution, initially documented by Comin and Philippon (2005). To capture this change, we modify the Markov transition matrix entries to match the increase in the firms' turnover probabilities between 1980 and 2005. Accordingly, under our identifying restrictions, the ex-ante probability that a top-20% firm never leaves its leadership position for 5 years is given by $[(1-\delta(2))(1-q_{21})]^{20}$. We choose $q_{21} = q_{31} = 0.00006152$ for the pre-shift steady state and $q_{21} = q_{31} = 0.0145$ for the post-shift steady state, to obtain a 5-year top firms turnover of 0.05 and 0.29, the turnover rates observed in data for 1980 and 2005, respectively.

Using the values for q_{21} and q_{31} calculated above, we can calibrate the remaining elements of the Markov matrix by matching the 5-year productivity autocorrelations, also documented by Comin and Philippon (2005). The detailed procedure is presented in the Appendix B. The correlation coefficients between the productivity of a firm at time t and at time $t+20$ are set to 0.78 before the shift and 0.7 after, their empirical counterparts in 1980 and 2000, respectively. With this procedure we are able to pin down the Markov process for both the pre and the post-shift model economy and the corresponding values are reported in Table 3.

The entry costs are set to 0.05 before the shift allowing us to generate sectors with the number of units corresponding to current small 4-digit or 6-digit NAICS industries. The entry costs after the shift are set to 0.075.[17] One way of interpreting the entry cost in our model is as the foregone labor force required to introduce a new variety in the economy. Bloom et al. (2020) estimate a sharp increase in the number of effective workers employed for R&D in the U.S. economy, which are required to sustain a stable technological growth when the research productivity is decreasing. They show that the number of researchers doubled between 1980 and 2000, consistent with a 100% increase in entry costs.

5 Model Simulations

In our main quantitative exercise, we study the transition dynamics between two steady states of two different economies. In both economies, the initial steady state corresponds to the period pre-80s, when the markups and other measures of U.S. market concentration were low. We then introduce two permanent shifts into the ICT-intensive sector: a boost in firms' mobility over the cross-sectional productivity distribution, and an increase in entry costs. In a low ICT sector, only entry costs go up. We focus our attention on the resulting transition dynamics before the model economy settles around the second steady state. Specifically, we study the trends in markups, market concentration, profits,

[17]Relatively to other papers, as, for instance, De Loecker, Eeckhout, and Mongey (2019) who need three or four times increase to match the markup growth, we impose a conservative 50% increase.

Table 3: **Calibration of the key exogenous parameters pre and post-shift**

Parameter	Pre-shock	Post-shock	Target
f_e	0.05	0.075	50% increase in barriers, decrease in entry rates $\approx$ 15%
q_{11}	0.9876	0.9794	5-year correlation = 0.78 pre-shock and = 0.7 post
q_{12}	0.0124	0.0206	$1 - q_{11}$
q_{13}	0	0	Assumption for computational purposes
q_{21}	6.15e-05	0.0145	5-year top firms turnover = 0.05 pre-shock and = 0.29 post
q_{22}	0.9857	0.9792	5-year correlation = 0.78 pre-shock and = 0.7 post
q_{23}	0.0142	0.0063	$1 - q_{21} - q_{22}$
q_{31}	6.15e-05	0.0145	$= q_{21}$ for identification strategy
q_{32}	0.0142	0.0035	$1 - q_{31} - q_{33}$
q_{33}	0.9857	0.9820	5-year correlation = 0.78 pre-shock and = 0.7 post

Notes: The table presents the calibration of the remaining exogenous parameters. The numerical targets are taken from 1980 (pre-shock) and 2005 (post-shock) estimates in Comin and Philippon (2005).

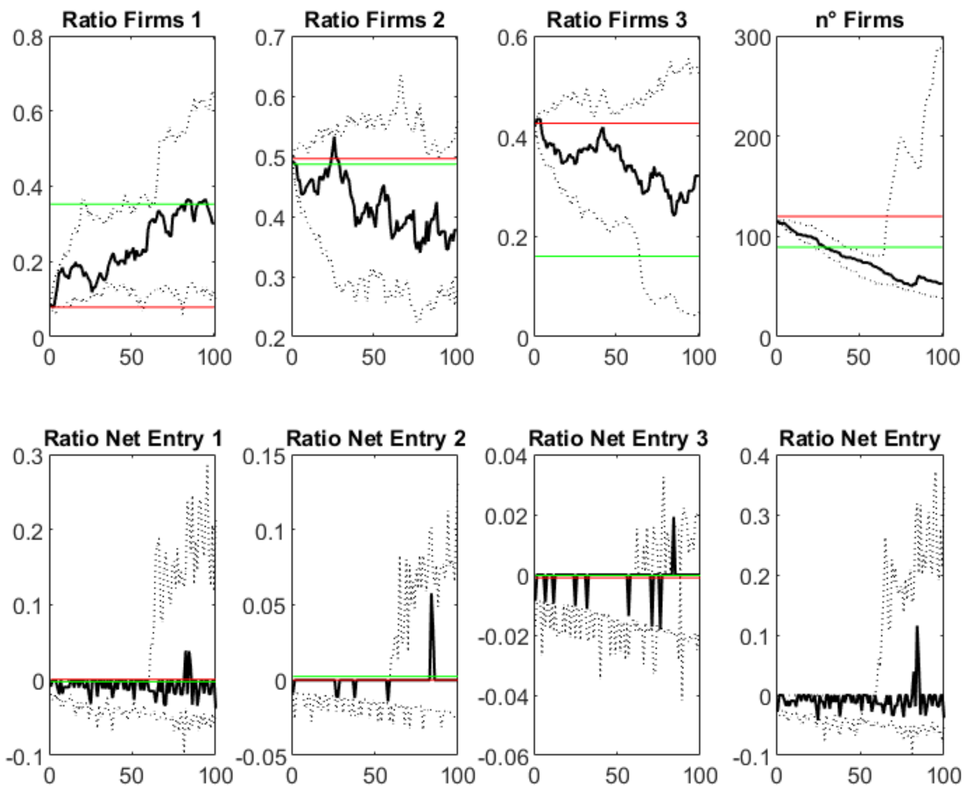
average productivity, and the number of incumbents. We compare the behaviour of these quantities between two sectors.

5.1 Transition dynamics in ICT-intensive sector

We solve the model using the original algorithm described in detail in Online Appendix 4. We simulate the model economy over 100 periods (25 years) and repeat the experiment 100 times to build confidence intervals (dotted lines). We present the results of one simulation instead of averaging over 100 to preserve the properties deriving from our assumption about the finite number of firms. In period 0, we make it more costly for all the potential firms to enter the economy and we boost the firms' mobility. Figures 2 and 3 show the resulting patterns and clearly point towards the decline in the business dynamism. The results of an alternative simulation, where both entry cost and firms' mobility shifts are progressive, are presented in Figures 8 and 9 of Online Appendix 7.

The red and green lines in each of the panels represent the initial and the final steady state, respectively. The panels in the first row of Figure 2 plot the relative share of each type of firms. An initial increase in firms' mobility makes the entry particularly attractive for type 1 firms and translates into the growth in their relative number over the transition

Figure 2: High-ICT sector: firms' dynamics in response to the permanent shifts.

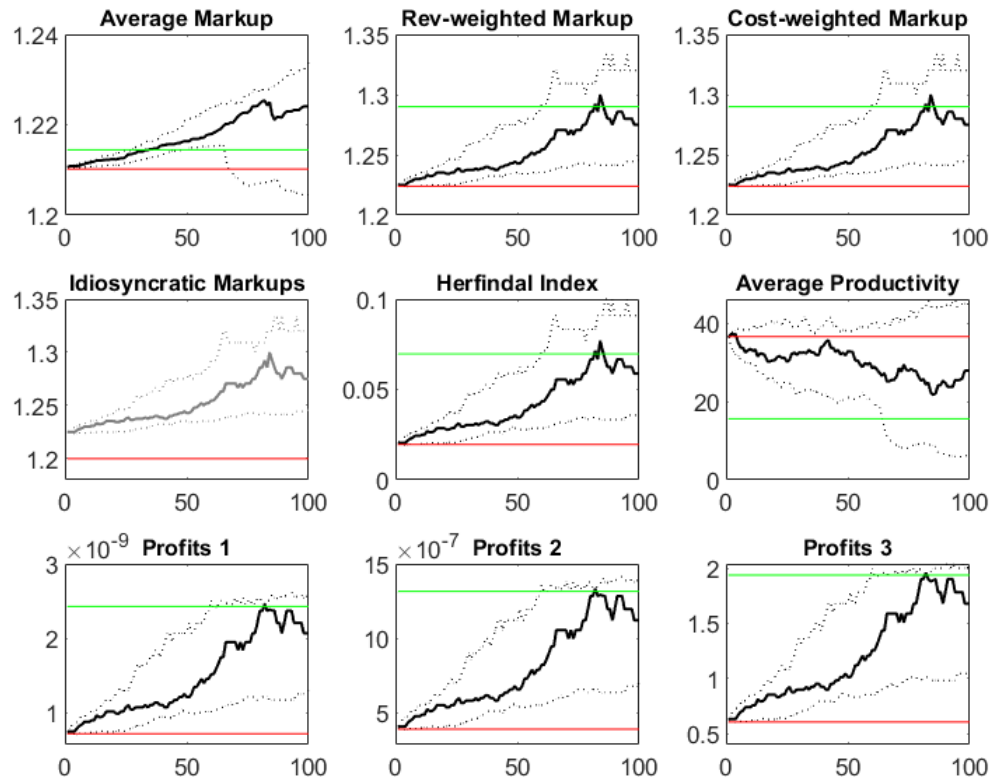


Notes: The figure plots dynamics of the firms' quantities in response to a 50% increase in entry costs and in the firms' mobility. The black solid lines describe the dynamics of the variables over time, and the dotted black lines the confidence intervals. The red and green lines represent, respectively, the initial and the final steady state.

period. In contrast, the relative number of type 2 and superstar firms declines over the transition. In fact, the absolute and relative number of superstar firms is always lower while the number of low productive firms is always higher in the second steady state. This is driven by the heterogeneous impact of the firms' mobility shock depending on the firms' type. Superstar firms can only be hit by a detrimental productivity shock, by definition. An increase in firms' mobility over productivity distribution translates therefore for them into a higher likelihood of becoming less productive or disappearing from the market. If this happens, a large available market share is captured by other superstar firms.

The relative number of type 1 firms is always higher in the second steady state because

Figure 3: High-ICT sector: evolution of aggregate variables in response to the shifts in entry cost and in the firms' mobility.



Notes: The graph presents the response of the economy to a 50% increase in entry costs and in the firms' mobility. The black solid lines describe the dynamics of the variables over time, and the dotted black lines the confidence intervals. The red and green lines represent, respectively, the initial and the final steady state.

mobility shock additionally increases the likelihood of types 2 and 3 to become type 1 firm. Small and low productivity firms are also more likely to exit the market than the other two types. Combined with high calibrated likelihood of entering the market, this model feature captures high turnover of small and unproductive firms and can be detected in the frequent net exits, displayed in the left panel in the second row of Figure 2.[18]

Variations in the number of different firm types during the transition translate into the fluctuations of the market concentration presented in Figure 3. The first row of the

[18] Bartelsman et al. (2002) show that the U.S. entering firms are smaller relative to industry average and have a lower level of labour productivity relative to the average incumbent.

figure plots the evolution of sectoral markup computed in three different ways: simple average markup, revenues weighted markup and costs weighted markup. The aggregate markups in the model economy increase during the transition no matter how they are defined. Yet, individual markup dynamics largely differ, conditional on the productivity level. The first panel in the second row of Figure 3 documents that type 1 and type 2 firms always charge the monopolistic competition markup of 1.2 due to the high number of same type competitors and resulting negligible market shares. The entire increase in the average markup is therefore driven by the firms at the top of the markup distribution. This finding has been largely documented by empirical literature. According to De Loecker et al. (2020), for instance, the median markup has been constant over the last several decades around 1.2, while the increase in average markup has been driven by growing markups of the top firms.

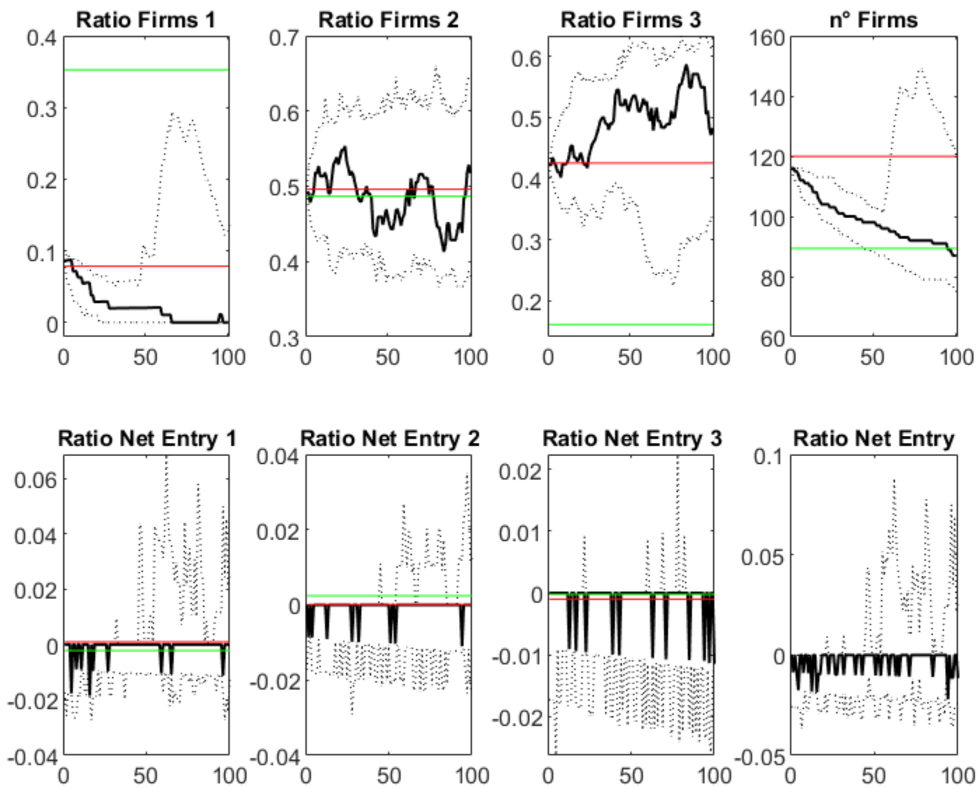
A declining total number of firms (last panel in the first row of Figure 2), translates into a growing Herfindahl index (HHI) over the transition to the second steady state (the middle panel of Figure 3). The average productivity plotted in the last panel of the second row in Figure 3 shows an interesting pattern. Because of the growing relative number of small, unproductive firms the average productivity declines over the transition.

The last row of Figure 3 plots the evolution of profits and it shows a clear upward trend for all types of incumbents. In our model, increasing profits are a natural counterpart of decreasing total number of incumbents, presented in Figure 2 and have been documented by Furman (2015), Grullon et al. (2019), Gutiérrez and Philippon (2017) and Barkai (2020).

5.2 Transition dynamics in low-ICT sector

We now contrast the transition dynamics of a low-ICT user sector, which has not experienced the increase in firm's mobility, with the ICT-intensive sector. To do it, we simulate a model economy subject only to an increase in entry costs. The magnitude of the shock and all the other parameters of the model are the same as in the ICT-intensive sector simulation. Figure 4 presents the firms' dynamics in response to a shift in entry costs only. The red horizontal lines display the initial steady state. To contrast the results of

Figure 4: Low-ICT sector: firms' dynamics in response to entry costs increase.

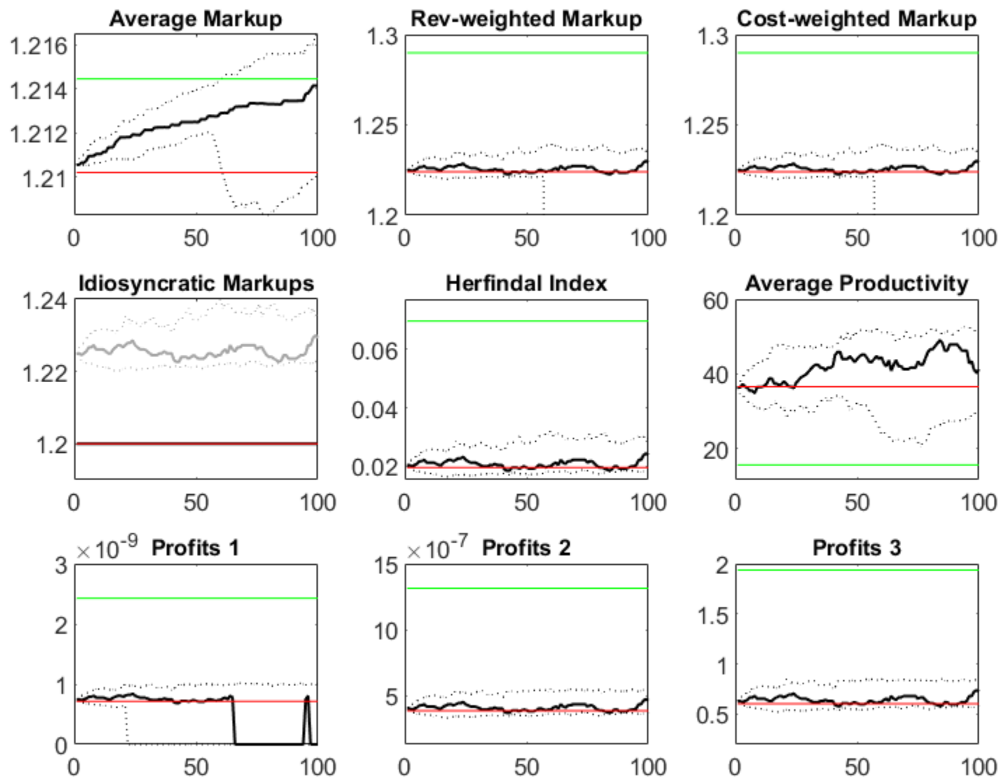


Notes: The graph presents the dynamics of the firms' variables in response to a one time shock in entry costs (50% increase). The black solid lines describe the dynamics of the variables over time, and the dotted black lines the confidence intervals. The red and green lines represent, respectively, the initial and the final steady state.

low-ICT with high-ICT sectors' simulations, we plot the final steady state of the latter (green horizontal lines).

The first row of Figure 4 plots relative number of each type of firms and, in the last panel, the total number of incumbents. Similar to the benchmark case and the data pattern, the total number of incumbents decreases during the transition. In contrast to the benchmark simulation, this trend is almost entirely driven by the exit of type 1 firms (the first panel of Figure 4). In fact, in this model economy the entry is almost always nil during the entire simulation. This occurs because, in the absence of firms' mobility boost, a substantial increase in entry costs prohibits new firms from entry. Without higher firms' mobility, in the environment of nil entry, the idiosyncratic exit shocks are the key driver

Figure 5: Low-ICT sector: evolution of aggregate variables in response to the increase in entry costs only.



Notes: The graph presents the response of the economy to a 50% increase in entry costs, only. The black solid lines describe the dynamics of the variables over time, and the dotted black lines the confidence intervals. The red and green lines represent, respectively, the initial and the final steady state.

of the firms' dynamics (last row of Figure 4).

Figure 5 shows how various market concentration measures respond to a sharp increase in the entry cost. The average markup, plotted in the first panel of the figure, displays an increasing trend as does its empirical counterpart. However, the trend is mainly driven by a change in the composition of the sector and, in particular, by an increase in the relative number of high-markup superstar firms at the expense of the number of type-1 firms. The absence of a reallocation channel in this model economy also results in stable revenue weighted and costs weighted markups and in the Herfindahl index. Individual profits of type 2 and type 3 firms also display no trend and they drop to zero for type 1 firms as soon as they exit the market.

6 Reallocation of Market Shares in the Model and the Data

The responses of sectoral level variables largely differ depending on whether a sector has experienced a boost in firms' mobility. We exploit these differences to develop a set of cross-sectional predictions that we then test on U.S. sectors.

First, we highlight the importance of the reallocation mechanism for the sectoral dynamics. For this purpose, we decompose the theoretical markup's change by its driver (reallocation versus composition) in a sector with and without the shift in firms' mobility. We then test theoretical predictions on 3 digits NAICS sectors in Compustat.[19]

6.1 Reallocation of Market Shares in the Model

We focus on the simple average markup's decomposition, instead of weighted average, because it provides a much clearer contribution of reallocation and composition to the markup's growth. In Online Appendix 6, we also provide a decomposition of the revenue weighted markup and the reasons for why the simple average specification is preferred.

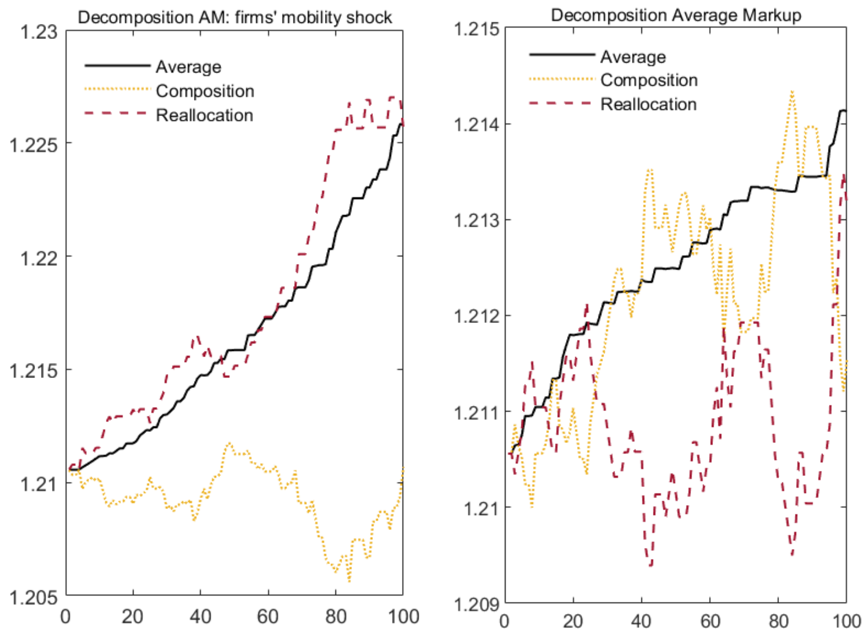
Growth of the simple average markup between period t and $t-1$, $\Delta\bar{\mu}_t$, can be decomposed as:

$$\begin{aligned}\Delta\bar{\mu}_t &= \sum_{i=1}^{3}\frac{\theta}{\theta-1}\frac{1}{1-\omega_t(i)}\left(\frac{N_t(i)}{N_t}-\frac{N_{t-1}(i)}{N_{t-1}}\right)+ \\ &+ \sum_{i=1}^{3}\frac{N_{t-1}(i)}{N_{t-1}}\frac{\theta}{\theta-1}\left(\frac{1}{1-\omega_t(i)}-\frac{1}{1-\omega_{t-1}(i)}\right)\end{aligned} \tag{29}$$

where the composition is captured by the first line of (29) and reallocation of market shares by the second line. Composition changes are driven by changes in the relative number of each type of firms, $\frac{N_t(i)}{N_t}$, while keeping their markups constant. The reallocation channel

[19] To clean the database we used the codes from De Loecker et al. (2020)

Figure 6: Evolution of the average markup and its decomposition in high-ICT and low-ICT sectors.



Notes: The left panel presents the decomposition of the simple average markups in economy hit by increase in firms' mobility shock. The right panel plots the decomposition of markup's growth in an economy without this shock. The black solid lines represent the simple average markup, the red dashed lines describe the counterfactual series considering pure changes in reallocation only and the yellow dotted line in composition only.

describes the changes in the average markup driven by variations in the market shares of the incumbents, $\omega_t(i)$, while keeping the relative number of each type constant at the previous period level. The reallocation has an impact on the simple average markup in our model because the individual markups are a function of market weights, $\omega_t(i)$. Intuitively, changes in market shares directly translate into changes in individual markups and hence, the simple average markup.

To perform a markup increase decomposition, we directly draw on the model simulations described in sections 5.1 and 5.2. Figure 6 plots the contribution of the composition and reallocation channels to the simple average markup's growth in an economy that has experienced an increase in firms' mobility (left panel) and in an economy that has not experienced it (right panel). In the first one, the entire increase in markup is explained by the reallocation of market shares towards superstar firms (red dashed line). Without the composition effects, the changes in reallocation would often predict an average markup that grows even faster than the actual simulated series. This potential growth is mitigated by the change in the composition of firms populating the economy. The relative number of type 1 low markup firms increases over the transition, a result demonstrated in Figure 2. In fact, a shift in the composition of the firms in the economy would, on its own, predict a modest decline in the average markup.

The right panel of Figure 6 shows that, in the absence of firms' mobility increase, the growth in markup is of smaller magnitude. Importantly, this trend is principally driven by a change in the composition of the firms that we documented in Figure 4 of Section 5.2. An increase in entry cost, on its own, drives small unproductive firms out of the market and the relative number of superstars increases boosting the average markup. In the absence of higher firms' mobility, there is little reallocation dynamics which alone would predict rather a stable markup.

6.2 Reallocation of Market Shares in the Data

We exploit the differences in dynamics implied by the model for low ICT and high ICT sectors to develop a set of cross-sectional predictions, with a 3 digits NAICS sector being a unit of observation. To compute sectoral characteristics, we use the entire population of Compustat firms, between 1965 and 2016. When possible, we construct the variables of interest directly from the model's specifications. We compute markups from equation (5), with market shares computed in terms of real revenues. The elasticity of substitution θ is set to 6, the value used in the model simulations. Given that we observe all the listed firms, we do not need to impose types although the results are robust to the classification of firms into three size-based categories.[20] Labour productivity is computed from equation

[20]We tested a specification in which we classify firms into the three types used in our theoretical framework and results do not change. The methodology was as follows: by exploiting the one-to-one correspondence between size and productivity in the model, we used the actual distribution of sizes, in terms of employment EMP, to estimate a Pareto distribution of sizes. Once we pin down the scale parameter κ, this value can be used to compute cut-offs that describe the exact same three percentiles we used to classify our types in the model. Given the cut-offs, we can reduce the panel to a 3 by t matrix, where in every year we have just one observation for each type, as in the simulation. The observation is

(2).

Tables 4 and 5 report results of cross-section regressions of the following form: $ra_i = \alpha + \beta X_i + \sigma_i$, where ra_i denotes the share of reallocation in the revenue weighted markup increase in sector i, over the considered period. In the tables, Δ denotes the change between the last and the first value of the HP filtered trend for the variable of interest. Table 4 displays results of regressions with average reallocation share as an independent variable, computed as the average contribution of the reallocation component to the markup change. The decomposition of the markup change between the two channels is summarized by equation (29).

Table 5 describes results of regressions with cumulative reallocation share as an independent variable, calculated as the cumulative, over the entire period, contribution of the reallocation component to the markup change. ICT stands for ICT intensity dummy and IC for nontangible capital share of total capital stock, used as a proxy for investment in ICT. IC-unw stands for a simple average of firms' intangibles share in a sector and IC-w its weighted couterpart. X_i includes additional regressors.

Our key theoretical result is that increasing firms' mobility has led to the reallocation of market shares. Firms' mobility is measured here as 5-year autocorrelation of productivity rank within its sector. The sectoral mobility is computed as an average across autocorrelations of all firms in the sector. A decline in autocorrelation implies an increase in firms' mobility and a negative sign in regressions reported in Tables 4 and 5. Results documented in the first column of both tables robustly confirm this prediction. In all but one considered specifications, the sectors where reallocation of market shares has been the main driver of the markups' change, have also experienced increasing firms' mobility.

In the model, the reallocation of market shares towards high-markup firms leads to higher revenue weighted sectoral markup. This result is strongly supported by positive and significant estimates in the second columns of Tables 4 and 5. 3 digits NAICS sectors, where reallocation of market shares has been most prevalent, have displayed the steepest markups' growth.

obtained by computing the mean for all of the relevant variables, evaluated in a given period within every active incumbent assigned to that specific type.

Table 4: **Reallocation and sectoral characteristics**

			Average reallocation share		
$-\Delta$ Mobility	Δ Markup	Δ Productivity	ICT	Δ IC-*unw*	Δ IC-*w*
-0.162***					
(0.017)					
-0.088***	0.097***				
(0.016)	(0.008)				
-0.206***	0.080***	-0.146***			
(0.021)	(0.008)	(0.008)			
-0.194***	0.086***	-0.146***	0.086***		
(0.021)	(0.007)	(0.008)	(0.018)		
			0.133***		
			(0.018)		
-0.064**	0.057***	-0.103**		0.105***	
(0.020)	(0.008)	(0.007)		(0.008)	
				0.099***	
				(0.007)	
-0.191***	0.021***	-0.180***			0.110***
(0.028)	(0.005)	(0.001)			(0.009)
					0.090***
					(0.005)

Notes: This table shows results of cross-section regressions specified by: $\Delta ra_i = \alpha + \beta X_i + \sigma_i$. The unit of observation i is a sector. Δ denotes the change between the last and the first value of the HP trend. The top panel displays results of regressions with average reallocation share as an independent variable. The average reallocation share is computed as the average contribution of the reallocation component to the markup change. The bottom panel shows results of regressions with cumulative reallocation share as an independent variable. The Cumulative reallocation share is calculated as the cumulative, over the entire period, contribution of the reallocation component to the markup change. ICT is an ICT intensity dummy and IC stands for nontangible capital proxy for ICT with $IC - unw$ being a simple average of firms' intangibles share of total capital stock in sector and $IC - w$ its weighted couterpart. X_i are sectoral level controls. Values in brackets report White-corrected standard errors.

Table 5: **Reallocation and sectoral characteristics (*cont.*)**

			Cumulative reallocation share		
$-\Delta$ Mobility	Δ Markup	Δ Productivity	ICT	Δ IC-*unw*	Δ IC-*w*
-4.566***					
(0.563)					
-1.256***	4.603***				
(0.506)	(0.363)				
-3.648***	4.211***	-4.354***			
(0.788)	(0.340)	(0.336)			
-2.564***	4.767***	-4.312***	7.598***		
(0.770)	(0.341)	(0.327)	(0.977)		
			8.922***		
			(0.989)		
-4.125***	3.145***	-5.195**		6.102***	
(0.934)	(0.387)	(0.362)		(0.427)	
				5.831***	
				(0.384)	
-0.479	6.909**	-5.532***			5.889***
(1.258)	(2.396)	(0.507)			(0.427)
					4.803***
					(0.298)

Our model predicts that higher intra sectoral mobility and resulting increasing profit possibilities attract a growing number of small, unproductive firms. As a result, the model economy exhibits declining average productivity. The third columns of the tables shows that sectors where reallocation has been the most important have also become the least productive.

Finally, we explore the hypothesis that the observed heterogeneities across sectors are driven by differences in implemented technologies. The last three columns of Tables 4 and 5 report a striking relationship between the importance of reallocation of market shares and the ICT exposure. The sectors with higher ICT exposure have exhibited higher reallocation of the market shares across firms, no matter how the measure is defined. This result is also in line with our initial intuition that the more ICT intensive sectors have become more dynamic over the last five decades.

In essence, results of our empirical exercise strongly support model predictions and suggest that the differences in the degree of decline in business dynamism across sectors are linked to the degree of exposure to ICT over the last several decades.

7 Conclusion

In this paper, we argue that an increase in firms' mobility within cross-sectional productivity distribution of a sector can explain a more pronounced decline in business dynamism in ICT-intensive sectors. We study the impact of higher firms' mobility on business dynamics in an oligopolistic model economy, populated by a finite number of firms. The firms differ by their productivity level and their markups increase in the market share and productivity level.

The initial steady state is calibrated to reproduce key features of the U.S. industries in the year 1980. We then simulate an economy that approximates an ICT-intensive sector and is subject to joint increase in firms' mobility and in entry costs. We contrast the resulting dynamics with the simulation of a low ICT sector which only experiences an increase in entry costs. An increase in firms' mobility triggers reallocation of market shares towards more productive, larger firms that charge higher markups. As a result, markups, market concentration and profits increase in an ICT-intensive sector. In contrast, in a low ICT sector, the increase in markup is modest and mainly driven by the change in the composition of firms populating the sector.

The comparison of two sectors implies a number of predictions that we test across U.S. industries. First, we show that the larger is the increase in firms' mobility, the more important is the reallocation of market shares. Second, the more reallocation of market shares towards high-markup firms, the higher is the revenue weighted sectoral markup. Third, the steeper is the increase in firms' mobility, the lower is the average productivity in the sector. These results support model predictions and suggest that the differences in the degree of decline in business dynamism across sectors are associated with their heterogeneous exposure to technologies.

References

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., & Li, H. (2019). *A theory of falling growth and rising rents* (Tech. Rep.). National Bureau of Economic Research.

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

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

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

Barkai, S. (2020). Declining labor and capital shares. *Journal of Finance*.

Bartelsman, E., Bassanini, A., Haltiwanger, J., Jarmin, R., & Scarpetta, S. (2002). Spread of ict and productivity growth: Is Europe really lagging behind in the new economy? *CEPN Working Paper*.

Bessen, J. (2017). Information technology and industry concentration.

Bilbiie, F. O., Ghironi, F., & Melitz, M. J. (2012). Endogenous entry, product variety, and business cycles. *Journal of Political Economy*, *120*(2), 304–345.

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–44.

Burstein, A., Carvalho, V., & Grassi, B. (2019). Bottom-up markup fluctuations. In *2019 meeting papers* (Vol. 505).

Calligaris, S., Criscuolo, C., & Marcolin, L. (2018). *Mark-ups in the digital era* (Tech. Rep.). OECD Science, Technology and Industry Working Papers 2018/10.

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

Charles, K., & Marye, P. (2009). Using industry classification codes to sample high-technology firms: Analysis and recommendations. *Journal of Accounting, Auditing and Finance*, *24*, 35–58.

Coeurdacier, N., Rey, H., & Winant, P. (2011). The risky steady state. *American Economic Review*, *101*(3), 398–401.

Colciago, A. (2016). Endogenous market structures and optimal taxation. *Economic*

Journal, *126*(594), 1441–1483.

Comin, D., & Philippon, T. (2005). The rise in firm-level volatility: Causes and consequences. *NBER macroeconomics annual*, *20*, 167–201.

Crouzet, N., Mehrotra, N., et al. (2017). Small and large firms over the business cycle. *Unpublished working paper*, *101*.

Davis, S. J. (2017). Regulatory complexity and policy uncertainty: headwinds of our own making. *Becker Friedman Institute for Research in economics working paper*(2723980).

De Loecker, J., Eeckhout, J., & Mongey, S. (2019). *Quantifying market power* (Tech. Rep.). Working Paper.

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, *135*(2), 561–644.

De Ridder, M. (2019). Market power and innovation in the intangible economy.

Diez, M. F. J., Fan, J., & Villegas-Sánchez, C. (2019). *Global declining competition.* International Monetary Fund.

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

Etro, F., & Colciago, A. (2010). Endogenous market structures and the business cycle. *Economic Journal*, *120*(549), 1201–1233.

Fernald, J. G. (2015). Productivity and potential output before, during, and after the great recession. *NBER macroeconomics annual*, *29*(1), 1–51.

Furman, J. (2015). Business investment in the united states: Facts, explanations, puzzles, and policies. *Reviving Private Investment" Remarks at the Progressive Policy Institute Washington, DC*.

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

Goldschlag, N., & Tabarrok, A. (2018). Is regulation to blame for the decline in american entrepreneurship? *Economic Policy*, *33*(93), 5–44.

Grullon, G., Larkin, Y., & Michaely, R. (2019). Are US industries becoming more concentrated? *Review of Finance*, *23*(4), 697–743.

Gutiérrez, G., Callum, J., & Philippon, T. (2019). Entry costs and the macroeconomy. *IMF Working Paper*.

Gutiérrez, G., & Philippon, T. (2017). Declining competition and investment in the US. *NBER Working Paper*.

Gutiérrez, G., & Philippon, T. (2019). *The failure of free entry* (Tech. Rep.). National Bureau of Economic Research.

Haltiwanger, J. C. (1997). Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence. *Federal Reserve Bank of St. Louis Review*, *79*(3), 55.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, *60*(5), 1127–1150.

Hyytinen, A., Pajarinen, M., & Rouvinen, P. (2015). Does innovativeness reduce startup survival rates? *Journal of Business Venturing*, *30*(4), 564–581.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, *50*(3), 649–670.

Juillard, M. (2011). Local approximation of dsge models around the risky steady state. *Working Paper*, *87*.

King, R. G., & Rebelo, S. T. (1999). Resuscitating real business cycles. *Handbook of macroeconomics*, *1*, 927–1007.

Koh, D., Santaeulàlia-Llopis, R., & Zheng, Y. (2020). Labor share decline and intellectual property products capital. *Econometrica*, *88*(6), 2609–2628.

Krusell, P., & Smith, A. A., Jr. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, *106*(5), 867–896.

Pugsley, B., Sedláček, P., & Sterk, V. (2020). The nature of firm growth. *American Economic Review*.

Van Reenen, J. (2018). Increasing differences between firms: Market power and the macro-economy. *Centre for Economic Performance, London School of Economics and Political*.

Appendix A: Risky Steady State

The equilibrium conditions for the *Risky Steady State* presented in the main text are:

Production

$$\rho(i) = \left[\left(\frac{\theta-1}{\theta}\right)\left(1-\rho(i)^{1-\theta}\right)\right]^{-1}\frac{w}{x(i)} \quad for \quad i=\{1,2,3\}$$

$$d(i) = \left(\frac{1}{\theta}+\left(\frac{\theta-1}{\theta}\right)\rho(i)^{1-\theta}\right)\rho(i)^{1-\theta}Y \quad for \quad i=\{1,2,3\}$$

Entry and Exit

$$N(i) = \sum_{j=1}^{3} q_{ji}\left[1-\delta(j)\right]\left[N(j)+N^e(j)\right] \quad for \quad i=\{1,2,3\}$$

$$N^e = \Omega_0 M$$

$$N^e(i) = \Omega_i N^e \quad for \quad i=\{2,3\}$$

$$N^e = N^e(1)+N^e(2)+N^e(3)$$

$$f_e w = \left(1-\Omega_2-\Omega_3\right)e(1)+\Omega_2 e(2)+\Omega_3 e(3)$$

Households

$$\chi L^{\frac{1}{\phi}} c = w$$

$$e(i) = \beta\left[1-\delta(i)\right]\sum_{j=1}^{3}\left[q_{ij}\left(d(j)+e(j)\right)\right] \quad for \quad i=\{1,2,3\}$$

Aggregation

$$Y = c$$

$$c + N^e f_e w = wL + N(1)d(1)+N(2)d(2)+N(3)d(3)$$

$$1 = N(1)\rho(1)^{1-\theta}+N(2)\rho(2)^{1-\theta}+N(3)\rho(3)^{1-\theta}$$

Appendix B: Calibration of Markov process

Given that we will obtain a system of three equations, we start by restricting the probability to switch from productivity $x(1)$ to $x(3)$ to be equal to zero, i.e. $q_{13} = 0$. The Markov process is then iterated for 20 periods while keeping the three unknown values:

$$\begin{bmatrix} x & 1-x & 0 \\ q_{21} & y & 1-y-q_{21} \\ q_{31} & 1-z-q_{31} & z \end{bmatrix}$$

This means that, abstracting from exit shocks, the probability $p(1|1)_5$ that a type 1 firm keeps its own productivity level after a 5-year period (i.e. 20 periods in our model) is given by the first element of the row vector resulting from:

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x & 1-x & 0 \\ q_{21} & y & 1-y-q_{21} \\ q_{31} & 1-z-q_{31} & z \end{bmatrix}^{20} = \begin{bmatrix} p(1|1)_5 & p(2|1)_5 & p(3|1)_5 \end{bmatrix}$$

The same can be computed for the probabilities of keeping productivity $x(2)$ or $x(3)$ after 5 years conditional on starting on that given productivity level. Given those highly non-linear equations in terms of x, y and z, we restrict the solution for the three unknowns to be in the interval $[0, 1]$. Our goal is to have $p(1|1)_5 = p(2|2)_5 = p(3|3)_5 = 0.78$ pre-shock and $p(1|1)_5 = p(2|2)_5 = p(3|3)_5 = 0.7$ post-shock. In this way, no matter the initial condition, the ex-ante correlation between the productivity of a firm at time t and at time $t + 20$ equals 0.78 before and 0.7 after the shift, values presented in Comin and Philippon (2005). The resulting values are $q_{11} = 0.9876$, $q_{22} = 0.9857$ and $q_{33} = 0.9857$ pre-shock and $q_{11} = 0.9794$, $q_{22} = 0.9792$ and $q_{33} = 0.9820$ post-shock.

Online Appendix

Online Appendix 1: Derivation of the aggregate demand constraint

In this appendix, the aggregate demand constraint is derived. There are two ways in which the constraint can be derived: the first assumes a continuum of final good producers competing under perfect competition, which purchase the individual firms' production. The final good producers use the individual outputs as inputs to produce the aggregate bundle Y_t, which is sold to the households at a price P_t. The second method, which exploits the fact that the aggregate production is entirely consumed by the households, is based on the minimization of the total aggregate expenditure. In the following, we present both methods, primarily because they complement each other and, together, they provide a consistent definition of the aggregate price P_t as a function of the individual prices $p_t(i)$. Note that the time index t is dropped in the following since firms maximize their per-period profits (no frictions regarding re-optimization are present).

The first method implies aggregate/sectoral producers. Each individual good $y(i)$ is aggregated into a final output Y, purchased by the households at a price P. The maximization of the final good producers is:

$$\max_{y(i)} PY - \sum_{i=1}^{N} p(i)y(i)$$

subject to:

$$Y = \left[\sum_{i=1}^{N} y(i)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}$$

The aggregator function through which the individual goods and the final good can be linked follows a standard C.E.S. function, as presented in the constraint of the maximization (note that N is the number of firms in the economy). Note that $\theta > 1$ represents the elasticity of substitution between intermediate goods. The F.O.C. with respect to $y(i)$ is:

$$P\left(\frac{\theta}{\theta-1}\right)\left[\sum_{i=1}^{N} y(i)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}-1}\left(\frac{\theta-1}{\theta}\right)y(i)^{\frac{\theta-1}{\theta}-1} = p(i)$$

This can be rewritten as:

$$PY^{\frac{1}{\theta}}y(i)^{-\frac{1}{\theta}} = p(i)$$

From this F.O.C., we obtain the demand for the individual output as:

$$y(i) = \left(\frac{p(i)}{P}\right)^{-\theta} Y$$

Alternatively, given that the entire production is consumed by the households, i.e. $c(i) = y(i)$ and $C = Y$, we can obtain the same condition, and a definition for the aggregate price P, from the minimization of the households' consumption expenditure. Households choose the optimal mixture of varieties $c(i)$ to minimize the aggregate expenditure, given an aggregate level of consumption C, by purchasing each good directly from the firms. Formally:

$$\min_{c(i)} \sum_{i=1}^{N} p(i)c(i)$$

subject to:

$$C = \left[\sum_{i=1}^{N} c(i)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}$$

The Lagrangian is:

$$\mathcal{L} = \sum_{i=1}^{N} p(i)c(i) + \lambda \left[C - \left[\sum_{i=1}^{N} c(i)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}\right]$$

The F.O.C. with respect to $c(i)$ is:

$$p(i) = \lambda \left(\frac{\theta}{\theta-1}\right) \left[\sum_{i=1}^{N} c(i)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}-1} \left(\frac{\theta-1}{\theta}\right) c(i)^{\frac{\theta-1}{\theta}-1}$$

which is equal to:

$$\lambda C^{\frac{1}{\theta}} c(i)^{-\frac{1}{\theta}} = p(i)$$

By raising each side to the power of $1-\theta$ and summing from 1 to N, we can write:

$$\lambda^{1-\theta} C^{\frac{1-\theta}{\theta}} \sum_{i=1}^{N} c(i)^{\frac{\theta-1}{\theta}} = \sum_{i=1}^{N} p(i)^{1-\theta}$$

Using the definition of aggregate consumption C provided above:

$$\lambda = \left[\sum_{i=1}^{N} p(i)^{1-\theta}\right]^{\frac{1}{1-\theta}} \equiv P$$

Finally, we can plug back the expression for the Lagrange multiplier λ in the F.O.C. and write:

$$c(i) = \left(\frac{p(i)}{P}\right)^{-\theta} C$$

which is consistent with the aggregate demand constraint presented above.

Online Appendix 2: Maximization of the intermediate goods' producers under Cournot and Bertrand

Cournot In this appendix we present the maximization for the intermediate goods producers. The notation follows the one introduced in the Online Appendix 1 and the time index is dropped for convenience. Each variety $y(i)$ is produced by employing one factor of production only. In particular, production is linear in labor $l(i)$ and depends on the idiosyncratic productivity level $x(i)$:

$$y(i) = x(i)l(i)$$

Inside the economy/industry, firms compete oligopolistically á la Cournot. In the next appendix, we present an alternative market structure (Bertrand competition). The incumbents internalize that their optimal quantity affects the sectoral output Y. However, firms cannot alter the total consumption expenditure $EXP = PY$ that is allocated to the production side of the economy. Firms maximize their per-period nominal profits by choosing the optimal quantity $y(i)$:

$$\max_{y(i)} p(i)y(i) - Wl(i) = y(i)\left(p(i) - \frac{W}{x(i)}\right)$$

subject to the following aggregate demand constraint:

$$y(i) = \left(\frac{p(i)}{P}\right)^{-\theta} Y$$

where W is the nominal wage. Substituting the idiosyncratic price $p(i)$ using the demand constraint, we can write the Lagrangian as:

$$\mathcal{L} = y(i)^{-\frac{1}{\theta}+1} P Y^{\frac{1}{\theta}} - W\frac{y(i)}{x(i)} = EXP y(i)^{-\frac{1}{\theta}+1} Y^{\frac{1}{\theta}-1} - W\frac{y(i)}{x(i)}$$

where the second equality comes from the definition of aggregate expenditure EXP provided above. The F.O.C. with respect to $y(i)$ is:

$$EXP\left(\frac{\theta-1}{\theta}\right)y(i)^{-\frac{1}{\theta}}Y^{\frac{1}{\theta}-1}-EXP\left(\frac{\theta-1}{\theta}\right)y(i)^{-\frac{1}{\theta}+1}Y^{\frac{1}{\theta}-2}\left(\frac{y(i)}{Y}\right)^{-\frac{1}{\theta}}-\frac{W}{x(i)}=0$$

Plugging back the definition of total expenditure EXP, this can be written as:

$$\left[\left(\frac{\theta-1}{\theta}\right)-\left(\frac{\theta-1}{\theta}\right)\left(\frac{y(i)}{Y}\right)^{1-\frac{1}{\theta}}\right]PY^{\frac{1}{\theta}}y(i)^{-\frac{1}{\theta}}=\frac{W}{x(i)}$$

Note that the market share $\omega(i)$ can be defined as the ratio between the individual revenues and the total revenues, i.e. $\frac{p(i)y(i)}{PY}$. Hence, using the aggregate demand constraint for $y(i)$:

$$\omega(i)=\frac{p(i)y(i)}{PY}=\left(\frac{y(i)}{Y}\right)^{1-\frac{1}{\theta}}=\left(\frac{p(i)}{P}\right)^{1-\theta}$$

It is worth to mention that the market share reasonably reduces to $1/N$ if we assume homogeneity across firms. This would provide the same markup function as the one presented in Etro and Colciago (2010). Using the aggregate demand constraint again, we can simplify the LHS of the F.O.C. and write:

$$\left[\left(\frac{\theta-1}{\theta}\right)(1-\omega(i))\right]p(i)=\frac{W}{x(i)}$$

In the main text, we directly present the implicit formula for the relative price $\rho(i) = p(i)/P$, which can be easily computed from the previous equation:

$$\rho(i)=\left[\left(\frac{\theta-1}{\theta}\right)\left(1-\rho(i)^{1-\theta}\right)\right]^{-1}\frac{w}{x(i)}=\mu(i)\frac{w}{x(i)}$$

where $w = W/P$ is the real wage.

Bertrand Assume that intermediate goods producers compete oligopolistically on prices under Bertrand competition. Production still is linear in labor $l(i)$ and depends on the idiosyncratic productivity level $x(i)$:

$$y(i) = x(i)l(i)$$

Inside the economy/industry, firms compete oligopolistically á la Bertrand. Thus, the incumbents internalize that their optimally chosen price affects the sectoral price P, defined in Online Appendix 1. However, firms cannot alter the total consumption expenditure $EXP = PY$ that is allocated to the production side of the economy. Firms maximize their per-period nominal profits by choosing the optimal nominal price $p(i)$:

$$\max_{p(i)} p(i)y(i) - Wl(i) = y(i)\left(p(i) - \frac{W}{x(i)}\right)$$

subject to the aggregate demand constraint:

$$y(i) = \left(\frac{p(i)}{P}\right)^{-\theta} Y$$

where W is the nominal wage. Substituting the individual quantity $y(i)$ using the demand constraint, we can write the Lagrangian as:

$$\mathcal{L} = p(i)^{1-\theta}P^{\theta}Y - \frac{W}{x(i)}p(i)^{-\theta}P^{\theta}Y = p(i)^{1-\theta}P^{\theta-1}EXP - \frac{W}{x(i)}p(i)^{-\theta}P^{\theta-1}EXP$$

where the second equality comes from the definition of aggregate expenditure EXP, provided above.

The F.O.C. with respect to $p(i)$ is:

$$(1-\theta)\,p(i)^{-\theta}P^{\theta-1}EXP + (\theta-1)\,p(i)^{1-\theta}P^{\theta-2}EXP\left(\frac{p(i)}{P}\right)^{-\theta} +$$

$$+\theta\frac{W}{x(i)}p(i)^{-\theta-1}P^{\theta-1}EXP - \frac{W}{x(i)}(\theta-1)\,p(i)^{-\theta}P^{\theta-2}EXP\left(\frac{p(i)}{P}\right)^{-\theta} = 0$$

Multiplying by $\frac{p(i)^{1+\theta}}{-\theta P^{\theta-1}EXP}$, this can be written as:

$$p(i)\left(\frac{\theta-1}{\theta}\right)\left[1-\left(\frac{p(i)}{P}\right)^{1-\theta}\right] = \frac{W}{x(i)}\left[1-\left(\frac{\theta-1}{\theta}\right)\left(\frac{p(i)}{P}\right)^{1-\theta}\right]$$

As previously, the market share $\omega(i)$ can be defined as the ratio between the individual revenues and the total revenues, i.e. $\frac{p(i)y(i)}{PY}$. Hence, using the aggregate demand constraint for $y(i)$:

$$\omega(i) = \frac{p(i)y(i)}{PY} = \left(\frac{p(i)}{P}\right)^{1-\theta}$$

The implicit formula for the relative price $\rho(i) = p(i)/P$, can be easily computed from the previous equation:

$$\rho(i) = \frac{1-\left(\frac{\theta-1}{\theta}\right)\rho(i)^{1-\theta}}{\left(\frac{\theta-1}{\theta}\right)\left(1-\rho(i)^{1-\theta}\right)}\frac{w}{x(i)} = \mu(i)\frac{w}{x(i)}$$

where $w = W/P$ is the real wage.

It is possible to compare the markup under both market structures. The markup under Bertrand competition $\mu(i)^B$ is:

$$\mu(i)^B = \frac{1-\left(\frac{\theta-1}{\theta}\right)\rho(i)^{1-\theta}}{\left(\frac{\theta-1}{\theta}\right)(1-\rho(i)^{1-\theta})} = \frac{1}{\left(\frac{\theta-1}{\theta}\right)(1-\rho(i)^{1-\theta})} - \frac{\left(\frac{\theta-1}{\theta}\right)\rho(i)^{1-\theta}}{\left(\frac{\theta-1}{\theta}\right)(1-\rho(i)^{1-\theta})} = \mu(i)^C - \frac{\omega(i)}{1-\omega(i)}$$

where $\mu(i)^C$ is the markup under Cournot competition.

Whenever the market share goes to zero, the markups are the same as they both converge to the monopolistic competition markup $\frac{\theta}{\theta-1}$. If the market share is non-zero, the markup is always lower under Bertrand competition. In particular as:

$$\frac{\partial(\mu(i)^C - \mu(i)^B)}{\partial\omega(i)} = \frac{1}{(1-\omega(i))^2}$$

the difference in the markups is increasing in the market share. Given that high market share firms charge higher markups, under Bertrand competition the markup dispersion is lower.

Online Appendix 3: Impact of reallocation of market shares on markups.

Case of simple average markup The proof of the proposition follows.

Assume 3 types of active firms: $S = 3$. The average markup is:

$$\bar{\mu}_t = \frac{\theta}{\theta-1}\left[\gamma_t(1)\frac{1}{1-\omega_t(1)} + \gamma_t(2)\frac{1}{1-\omega_t(2)} + \gamma_t(3)\frac{1}{1-\omega_t(3)}\right]$$

Given that $1 = N_t(1)\omega_t(1) + N_t(2)\omega_t(2) + N_t(3)\omega_t(3)$, this can be rewritten as:

$$\bar{\mu}_t = \frac{\theta}{\theta-1}\left[\gamma_t(1)\frac{1}{1-\frac{1-N_t(2)\omega_t(2)-N_t(3)\omega_t(3)}{N_t(1)}} + \gamma_t(2)\frac{1}{1-\omega_t(2)} + \gamma_t(3)\frac{1}{1-\omega_t(3)}\right]$$

Note that an increase in $\omega_t(3)$, while keeping $N_t(i)$ and $\omega_t(2)$ constant, decreases $\omega_t(1)$. The derivative with respect to $\omega_t(3)$ is:

$$\frac{\partial \bar{\mu}_t}{\partial \omega_t(3)} = \frac{\theta}{\theta-1}\left[-\gamma_t(1)N_t(1)\frac{N_t(3)}{(N_t(1)-1+N_t(2)\omega_t(2)+N_t(3)\omega_t(3))^2}+\gamma_t(3)\frac{1}{(1-\omega_t(3))^2}\right]$$

Given that $\gamma_t(i) = \frac{N_t(i)}{N_t(1)+N_t(2)+N_t(3)}$ we get:

$$\frac{N_t(3)}{N_t(1)+N_t(2)+N_t(3)}\frac{1}{(1-\omega_t(3))^2} > \frac{N_t(1)^2}{N_t(1)+N_t(2)+N_t(3)}\frac{N_t(3)}{(N_t(1)-1+N_t(2)\omega_t(2)+N_t(3)\omega_t(3))^2}$$

Since $N_t(1)-1+N_t(2)\omega_t(2)+N_t(3)\omega_t(3) = N_t(1)(1-\omega_t(1))$:

$$\frac{1}{(1-\omega_t(3))^2} > \frac{1}{(1-\omega_t(1))^2}$$

this simplifies to $\varepsilon_t(3) > \varepsilon_t(1)$, which is always true.

Case of revenue weighted markup The proof of the proposition follows.

Assume 3 types of active firms: $S = 3$. The revenue weighted average markup is:

$$\bar{\mu}_t^R = \frac{\theta}{\theta-1}\left[N_t(1)\frac{\omega_t(1)}{1-\omega_t(1)}+N_t(2)\frac{\omega_t(2)}{1-\omega_t(2)}+N_t(3)\frac{\omega_t(3)}{1-\omega_t(3)}\right]$$

Given that $1 = N_t(1)\omega_t(1)+N_t(2)\omega_t(2)+N_t(3)\omega_t(3)$, this can be rewritten as:

$$\bar{\mu}_t^R = \frac{\theta}{\theta-1}\left[N_t(1)\frac{\frac{1-N_t(2)\omega_t(2)-N_t(3)\omega_t(3)}{N_t(1)}}{1-\frac{1-N_t(2)\omega_t(2)-N_t(3)\omega_t(3)}{N_t(1)}}+N_t(2)\frac{\omega_t(2)}{1-\omega_t(2)}+N_t(3)\frac{\omega_t(3)}{1-\omega_t(3)}\right]$$

Note that an increase in $\omega_t(3)$, while keeping $N_t(i)$ and $\omega_t(2)$ constant, decreases $\omega_t(1)$. The derivative with respect to $\omega_t(3)$ is:

$$\frac{\partial \bar{\mu}_t^R}{\partial \omega_t(3)} = \frac{\theta}{\theta-1}\left[-N_t(1)\frac{N_t(3)N_t(1)}{(N_t(1)-1+N_t(2)\omega_t(2)+N_t(3)\omega_t(3))^2}+N_t(3)\frac{1}{(1-\omega_t(3))^2}\right]$$

$$\frac{1}{(1-\omega_t(3))^2} > \frac{N_t(1)^2}{(N_t(1)-1+N_t(2)\omega_t(2)+N_t(3)\omega_t(3))^2}$$

Since $N_t(1)-1+N_t(2)\omega_t(2)+N_t(3)\omega_t(3) = N_t(1)(1-\omega_t(1))$:

$$\frac{1}{(1-\omega_t(3))^2} > \frac{1}{(1-\omega_t(1))^2}$$

this simplifies to $\omega_t(3) > \omega_t(1)$, which is always true.

Online Appendix 4: Algorithm and Theoretical Challenges.

In this appendix, we present the original algorithm developed to simulate the dynamic behavior of the economy. For tractability reasons, some further assumptions have been added with respect to the baseline theoretical model: the discussion about their impact on the equilibrium outcomes is presented below.[21]

In order to introduce the algorithm and its main mechanisms, we describe here the dynamics of the economy for a given period t. When the entry decisions are formed during period t, the current number of incumbents and their types, i.e. $N_t(1)$, $N_t(2)$ and $N_t(3)$, is known. Indeed, note that they come from the realization of the stochastic processes that regulate exit and productivity shocks, which have already occurred between period $t-1$ and period t and at the very beginning of period t. Given those three quantities, the first step in our solution technique is to pin down the number of potential entrants, M_t, by using a sequential approach.

Starting from a given stock of zero entrants, namely from $M_t = 0$, we evaluate the free entry condition for the first entrant, in formula:

$$(1 - \Omega_2 - \Omega_3)\, e_{1,t}(1) + \Omega_2 e_{2,t}(2) + \Omega_3 e_{3,t}(3) - f_{e,t} w_t$$

Whenever the condition is positive, i.e. when the expected value for the new firm is higher than the entry costs, the marginal entrant joins the market. If this happens, the algorithm continues by re-evaluating the free entry condition. This time, however, the condition is computed conditional on the number of potential entrants being one. The idea behind the procedure is that the sector may present some unexploited profit possibilities. Whenever those expected revenues are higher than the entry costs, the entrant joins the market. However, its entry increases future competition and, thus, it lowers firms' value, making harder for new competitors to enter. The algorithm continues with this mechanism until

[21] The assumptions are imposed on how firms form their expectations. The constraints simplify the computation of the stream of expected future profits, which is required to pin down the value of the firms. This type of restriction is not new to the literature and it is often assumed in order to solve similar dynamic problems. See, for instance, Krusell and Smith (1998).

the free entry condition turns negative because entry is not profitable anymore.

Note that, in the computation of the competitive equilibrium, the number of successful entrants N_t^e, and, hence, $N_t^e(1)$, $N_t^e(2)$ and $N_t^e(3)$ as well, is not known with certainty since it depends on the realization of stochastic processes, given the equilibrium M_t. The same uncertainty holds for period $t+1$ competitors $N_{t+1}(1)$, $N_{t+1}(2)$ and $N_{t+1}(3)$, since they depend on the realization of idiosyncratic exit and productivity shocks. The households solve this uncertainty by considering the expected values of these quantities in their maximization process.[22] In formulas:

$$N_{t+1}(i) = \sum_{j=1}^{3} \left[q_{ji} \left(1 - \delta(j)\right) \left(N_t^e(j) + N_t(j)\right) \right]$$

where

$$N_t^e = \Omega_0 M_t$$

$$N_t^e(3) = \Omega_3 N_t^e$$

$$N_t^e(2) = \Omega_2 N_t^e$$

$$N_t^e = N_t^e(1) + N_t^e(2) + N_t^e(3)$$

Furthermore, note that the equations above hold for the incumbents: from the perspective of a potential entrant, the expected number of competitor is different since the latter internalizes that, by entering the market with a productivity level $x(i)$, the number of type-i firms is $N_t(i) + N_t^e(i) + 1$. As explained in the modelling section, this means that the entrants internalize the effects of their entry on the future price index, hence creating the wedge between their value and the value of the incumbents. These dynamics are taken into account in the algorithm.

In order to be able to compute the first part of the algorithm, i.e. to pin down M_t, some simplifying assumptions have been included. The constraints regard how firms form their expectations. First of all, potential entrants are imposed to expect that period $t+1$ output

[22] One alternative solution is to consider separately every possible state of the world for $N_t^e(i)$ and $N_{t+1}(i)$ given $N_t(i)$ and a specific M_t. In every state of the world it is possible to pin down the value for the marginal entrant and, only after this, obtain $e_{i,t}(i)$ as the average of the values computed, weighted by the probability that a given state occurs. In this way, every state of the world conserves a finite number of firms. However, this solution method gets computationally unsolvable quite easily.

Y_{t+1}, and, hence, consumption c_{t+1}, does not change between period t and period $t+1$. The assumption affects the computation of the expected profits for period $t+1$, which are necessary to pin down the value of the incumbents and of the (potential) entrants. In other words, potential entrants do not consider the effect of their entry on the sectoral production of the following periods and they assume that the economy is on a stable path. This assumption is quite restrictive: in the per-period profits maximization, we assume Cournot competition, which entails that the producers internalize that their optimally chosen quantity has an impact on the contemporaneous aggregate output.

Nevertheless, the assumption may still be justified. When the number of active incumbents in the sector grows, the marginal effect of entry on output growth is negligible, since the competition is already tight. This is true in our case, given that the steady state present hundreds of incumbents, and the effect of entry on output is particularly irrelevant if the entrant does not realizes as a superstar firm with productivity $x(3)$. Furthermore, the effect of entry on future competition and prices, i.e. the increase in the expected number of competitors, is clear from the firm's perspective (this is why we assumed that the entrant can internalize these dynamics). On the other hand, the aggregate effects on consumption and output are significantly harder to be predicted *ex ante*, given that they rely also on households' response. Thus, it is not unreasonable to assume that firms are partially myopic and that their period t expectation of Y_{t+1} is simply Y_t.[23] Finally, note that this assumption affects also the stochastic discount factor, which reduces to β.

A second assumption regards the computation of the firms' value itself, and it is in line with the previous constraint. When firms evaluate their stream of conditional expected profits, which pins down their own value, they anticipate correctly the number of competitors in period $t+1$, required to estimate period $t+1$ profits. However, it is imposed that those profits are assumed to stay constant from period $t+2$ onward, conditional of being active in the market. Again, firms are myopic, since they consider that similar entry and exit dynamics occur every period. Given these assumptions, the value of incumbents and

[23]Alternatively, one could say that firms are myopic simply because they perceive the economy as if it was in the steady state. Hence, incumbents and entrants expect no variations in the aggregate quantities over time, although they internalize idiosyncratic exit and entry dynamics.

entrants can be easily defined as, respectively:

$$e_t(i) = \frac{\beta\left(1-\delta(i)\right)}{1-\beta\left(1-\delta(i)\right)} d_{t+1}(i)$$

and

$$e_{i,t}(i) = \frac{\beta\left(1-\delta(i)\right)}{1-\beta\left(1-\delta(i)\right)} d_{i,t+1}(i)$$

Note that an alternative approach to the above restrictions is to directly assume that firms are myopic and render the entry choice static.[24] Results do not vary significantly since the value function takes a similar form. However, we think that it is worth to keep the entry choice dynamic, even at the costs of some restrictive assumptions.

Once M_t is pinned down, the algorithm proceeds with the simulation of the stochastic realization of N_t^e from the given M_t and of $N_t^e(1)$, $N_t^e(2)$ and $N_t^e(3)$ from the realized N_t^e. Conditional on those variables, and recalling that $N_t(1)$, $N_t(2)$ and $N_t(3)$ are predetermined, we can compute the competitive equilibrium of the economy by solving the household's maximization problem, constrained by the previous assumptions. It is worth to mention that, when clearing the market, $N_t^e(1)$, $N_t^e(2)$ and $N_t^e(3)$ are now known by the households, differently from the information set through which M_t is determined. Finally, having solved for the market equilibrium, the stochastic realizations of the idiosyncratic exit and productivity processes are computed. In this way, we can obtain the realized $N_{t+1}(1)$, $N_{t+1}(2)$ and $N_{t+1}(3)$ from the previous quantities, which serve as a basis for the algorithm in the following period.

Online Appendix 5: Alternative formalization of productivity shocks

In this section, we present an alternative formalization of the productivity shocks for $S = 3$. Given the number of survivors, the realization of binomial distributions determines the fraction of survivors endowed with productivity $x(i)$ that keeps their own productivity level $N_{t+1}^i(i)$ versus the number of switchers $N_{t+1}^i(j,k)$. In formulas:

$$Pr\left[N_{t+1}^i(i) = x\right] = \binom{N_{t+1}^s(i)}{x} \left[q_{ii}\right]^x \left[q_{ij}+q_{ik}\right]^{N_{t+1}^s(i)-x} \tag{30}$$

[24] Similar to the approach in De Loecker et al. (2019).

For $i = \{1,2,3\}$ and $j,k = \{1,2,3\}, j,k \neq i$.

Of course, note that the following holds after the realization of the productivity shocks:

$$N^s_{t+1}(i) = N^i_{t+1}(i) + N^i_{t+1}(j,k) \tag{31}$$

For $i = \{1,2,3\}$ and $j,k = \{1,2,3\}, j,k \neq i$. Given the number of switcher for each type i, a further realization of a binomial distribution determines the number of switcher with type j versus the number of switcher with type k:

$$Pr\left[N^i_{t+1}(j) = x\right] = \binom{N^i_{t+1}(j,k)}{x}\left[\frac{q_{ij}}{q_{ij}+q_{ik}}\right]^x\left[\frac{q_{ik}}{q_{ij}+q_{ik}}\right]^{N^i_{t+1}(j,k)-x} \tag{32}$$

and

$$Pr\left[N^i_{t+1}(k) = x\right] = \binom{N^i_{t+1}(j,k)}{x}\left[\frac{q_{ij}}{q_{ij}+q_{ik}}\right]^x\left[\frac{q_{ik}}{q_{ij}+q_{ik}}\right]^{N^i_{t+1}(j,k)-x} \tag{33}$$

For $i = \{1,2,3\}$ and $j,k = \{1,2,3\}, j,k \neq i$. After the realizations of these processes, the number of type 1, 2 and 3 firms in period $t+1$ can be easily computed as:

$$N_{t+1}(i) = N^i_{t+1}(i) + N^j_{t+1}(i) + N^k_{t+1}(i) \tag{34}$$

For $i = \{1,2,3\}$ and $j,k = \{1,2,3\}, j,k \neq i$.

Online Appendix 6: Revenue weighted markup decomposition

In addition to the benchmark simple average markup decomposition, described in Section 6.1, we provide an alternative adjusted version of the revenue weighted markup decomposition presented in De Loecker et al. (2020) and originally coming from Haltiwanger (1997). The growth of the revenue weighted markup between period t and $t-1$, $\Delta\bar{\mu}^R_t$, can be decomposed as:

$$
\begin{aligned}
\Delta\bar{\mu}_t^R &= \sum_{i=1}^{3} N_t(i)\omega_{t-1}(i)\frac{\theta}{\theta-1}\left(\frac{1}{1-\omega_t(i)}-\frac{1}{1-\omega_{t-1}(i)}\right)+ \\
&+ \sum_{i=1}^{3} N_t(i)\left(\frac{\theta}{\theta-1}\frac{1}{1-\omega_{t-1}(i)}-\bar{\mu}_{t-1}^R\right)(\omega_t(i)-\omega_{t-1}(i))+ \\
&+ \sum_{i=1}^{3} N_t(i)\frac{\theta}{\theta-1}\left(\frac{1}{1-\omega_t(i)}-\frac{1}{1-\omega_{t-1}(i)}\right)(\omega_t(i)-\omega_{t-1}(i))+ \\
&+ \sum_{i=1}^{3} N_t^e(i)\omega_t(i)\left(\frac{\theta}{\theta-1}\frac{1}{1-\omega_t(i)}-\bar{\mu}_{t-1}^R\right)-\sum_{i=i}^{3} N_{t-1}^{ex}(i)\omega_{t-1}(i)\left(\frac{\theta}{\theta-1}\frac{1}{1-\omega_{t-1}(i)}-\bar{\mu}_{t-1}^R\right)
\end{aligned}
$$

Note that N_t^{ex} represents the number of incumbents that exits the market in period t.

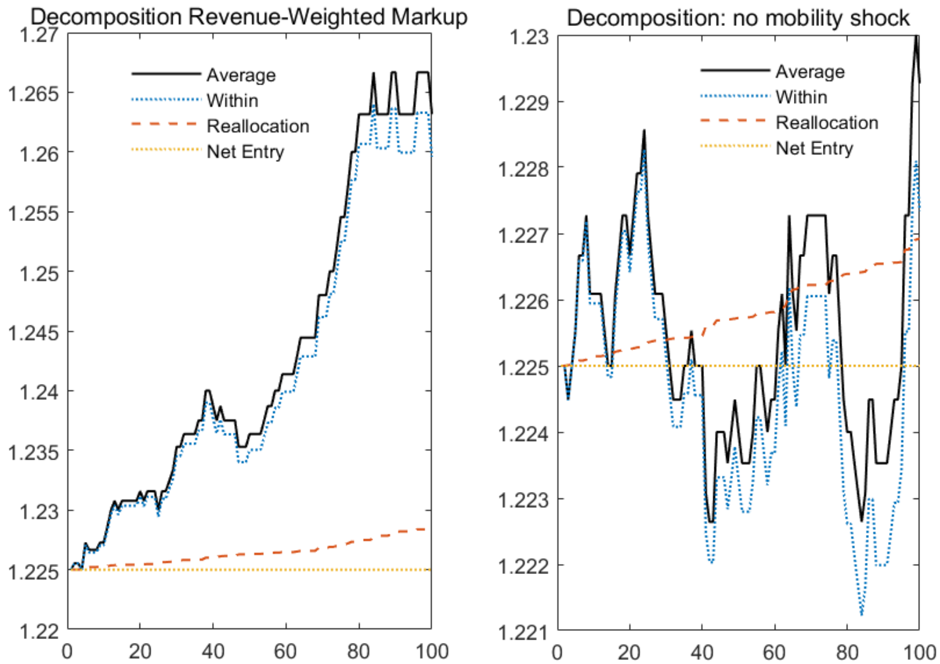
The within changes (first line of 35) represent variations in the average markup driven by changes in the firm-level markups, while keeping the market shares constant at the previous period level. The last line of 35 describes changes driven by net entry: whenever a firm joins or leaves the market with an idiosyncratic markup that is different from the average, the event has a non-zero impact on the average markup which is captured by this channel.

Between changes, i.e. the reallocation of market power, are captured by the second and third lines of 35, which represent, respectively, changes in average markup induced by pure changes in market shares, while keeping markups constant, and by cross terms changes, i.e. changes in covariances.

This decomposition is particularly useful for the empirical analysis. First of all, by disentangling the net entry channel, we are able to partially mitigate the misrepresentation of entry and exit dynamics in the Compustat.[25] Second, by isolating the pure reallocation of market shares toward high-markup firm, we can test our predictions: that sectors with a higher increase in firms' mobility experienced a stronger reallocation and, thus, a stronger increase in markups. Finally, by being performed on the weighted average, this decomposition is more robust to outliers than the unweighted average decomposition.

[25] A misalignment exists between entry in our model, i.e. the introduction of an extra variety in the market from a completely new firm, and entry in the Compustat, i.e. the act of becoming a listed company, even if the firm was already an incumbent in the market.

Figure 7: Evolution of revenue weighted average markup and its decomposition in economy with and without increase in firms mobility, simulation over 100 periods.



Notes: The left panel presents the decomposition of the revenue weighted average markups in economy hit by firms' mobility shock. The right panel plots the decomposition of the revenue weighted markup's growth in an economy without the shock. The black solid line represents the markup, the red dashed line describes the counterfactual series considering changes in reallocation, blue dotted line within and the yellow dotted line in net entry.

In spite of these useful properties, the reason why we chose to present the decomposition of the unweighted average for the theoretical part in the main text is the suitability of the latter for our model. In our model, individual markups are a function of the idiosyncratic market shares only. As a result, changes in market shares directly affect both the idiosyncratic markups and the weights of the average, i.e. the market shares themselves, affecting both reallocation and composition channels. Consequently, the within channel often captures by construction the largest part of the transmission. In contrast, the simple average decomposition allows to isolate the composition and the reallocation channels in a straightforward way.

Figure 7 shows the decomposition of the revenues weighted markups. Due to our modelling assumption it is not possible to disentangle the reallocation channel. Yet, Figure

7 demonstrates that the implied increase in markup is larger when the firms' mobility shock is present.

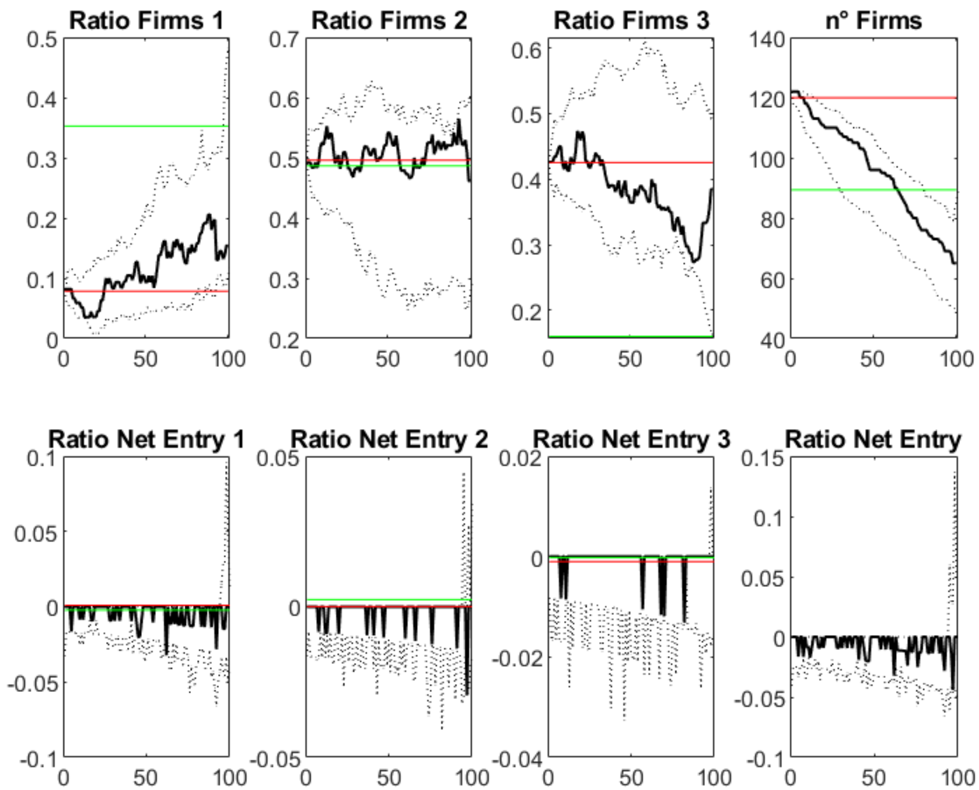
Online Appendix 7: Step-increase in entry costs and mobility

We assume that the increase in entry cost and firms' mobility has been gradual instead of a one-time structural shift, presented in the main text. We therefore implement a more realistic exercise where both entry cost and firms' mobility shifts are progressive. The cumulative magnitude of entry cost shift is assumed to be the same as in initial simulation, but it now accumulates in steps at constant intervals in a way to match empirical entry rates in U.S. sectors. To obtain the final cumulative Markov chain probabilities, we increase the diagonals in constant steps every 10 periods.

The results of this alternative exercise largely confirm initial findings and are plotted in Figures 8 and 9. The number of incumbents decays over time and, again, the decline is driven by the decreasing number of superstar firms while the relative number of small incumbents is increasing.

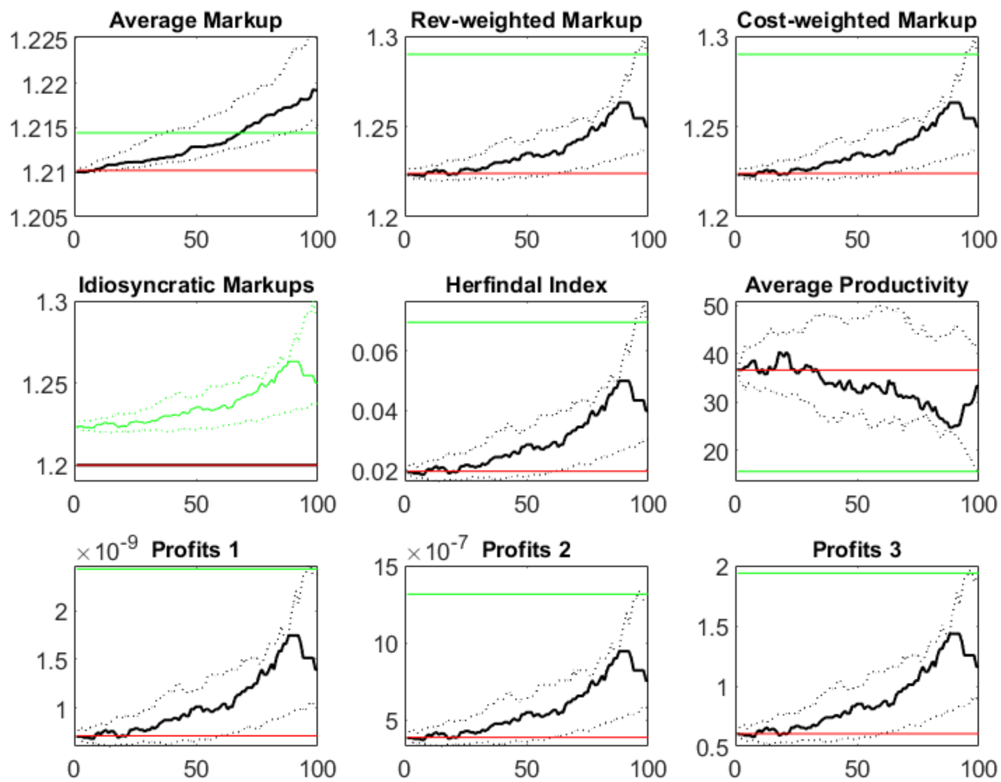
Figure 9 shows that the response of the model economy to gradual changes in entry cost and firms' mobility is also very similar to the baseline scenario. In particular, the increase in concentration and in the average price markup display a similar trend and magnitude. The main difference is the sluggish initial response of markups and concentration. The initial increase in price markup is driven by the composition channel with the growing relative number of type 2 and type 3 firms in the first 30 periods of the simulation (Figure 8). In the second part of the simulation, the surviving superstar firms capture the market shares of the exiting and damaged firms leading to market shares' reallocation. Finally, idiosyncratic profits rise over time, similar to the baseline simulation and the trend in the U.S. data.

Figure 8: Firms' dynamics in response to two gradual changes in entry cost and firms' mobility, simulation over 100 periods.



Notes: The graph presents the response of the firms to two shifts: firms' mobility and entry costs increases. The increase is imposed in steps every period. The blue line describes the dynamics of the variables over time, and the dotted blue line the confidence intervals, while the red and green lines represent, respectively, the initial and the final steady state presented in the previous section.

Figure 9: Evolution of economy in response to two gradual changes in entry cost and firms' mobility, simulation over 100 periods.



Notes: The graph presents the response of the economy to two permanent shocks: firms' mobility and entry costs. The increase in imposed in steps every period. The black solid lines describe the dynamics of the variables over time, and the dotted grey lines the confidence intervals. The red and green lines represent, respectively, the initial and the final steady state presented in the previous section.