Essays on Advertising, Growth, and Comparative Statics

Miquel Oliver-Vert

Thesis submitted to the University of Nottingham
for the Degree of Doctor of Philosophy

School of Economics
University of Nottingham

April 2025

Abstract

This dissertation presents three studies: the first on the macroeconomics of advertising, the second on growth and the role of competition, and the third on comparative statics.

The first study develops a new general equilibrium model to examine the aggregate effects of advertising. The model captures the two traditional views of advertising—informative and persuasive—and introduces a novel anticompetitive motive. We find that advertising has a significant positive aggregate effect through its role in spreading product awareness.

The second study explores the role of competition in shaping the long-run effect of fiscal policy on growth and the complementarity of fiscal and competition policies. To address these, we develop a step-by-step growth model with endogenous market structure, and we find that the market structure response amplifies the effect of fiscal policy.

The last study, which lies outside the field of macroeconomics, is motivated by the limited tools available to get analytical results in models frequently used in macroeconomics and other fields. It introduces a novel approach to comparative statics, building on Farkas' Lemma, and derives sufficient conditions under different assumptions. An application to an oligopoly model with differentiated goods under CES preferences illustrates the method.

Acknowledgements

I would like to thank my supervisors Giammario Impullitti, Omar Licandro, and Marit Hinnosaar for their guidance throughout this journey. I would like to thank Omar for believing in the potential of my job market paper idea from its early stages and for his patience during our long discussions about the model; I believe these helped a lot in my development as an economist. I would like to thank Giammario for his constant support and mentorship throughout the PhD. Working on a second project with him also provided a valuable safety net during more uncertain moments. I would like to thank Marit for her diligence and advice during the last stages of the PhD, especially in preparation for the job market.

I would also like to extend my thanks to the other professors and colleagues at the School of Economics with whom I had the pleasure of interacting over these years. I am also grateful for the financial support from the School of Economics, which enabled me to pursue my PhD at the University of Nottingham.

I would like to express my sincere gratitude to Murat Alp Celik for hosting me during my research visit at the University of Toronto. This experience was deeply enriching—both personally and academically—and gave me the opportunity to engage in stimulating discussions and connect with many wonderful people.

Last but not least, I am deeply grateful to my family and friends for their constant support, love, and encouragement throughout these years. Their presence, even from afar, has been a source of strength and motivation and made the journey more enjoyable.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation presents three studies: the first on the macroeconomics of advertising, the second on growth and the role of competition, and the third on comparative statics.

Chapter 2 develops a new general equilibrium model to examine the aggregate effects of advertising. The model captures the two traditional views of advertising—informative and persuasive—and introduces a novel anticompetitive motive. We find that advertising has an overall positive aggregate effect, not only through financing entertaining media goods, but also through expanding product awareness. Without advertising, firms remain smaller, which in turn reduces the incentives to create new products.

Chapter 3 explores the role of competition in shaping the long-run effect of fiscal policy on growth and the complementarity of fiscal and competition policies. To address these, we develop a step-by-step growth model with endogenous market structure, and we find that the market structure response amplifies the effect of fiscal policy. This is part of a joint project with my supervisor Giammario Impullitti and Antonin Bergeaud.

Chapter 4, whose applicability extends beyond macroeconomics, is motivated by the limited tools available to get analytical results in models frequently used in macroeconomics and other fields. It introduces a novel approach to comparative statics, building on Farkas' Lemma. This chapter shows that this approach can improve the sufficient conditions for the sign of the comparative static found in recent literature, as well as the bounds on the value. An application to an oligopoly model with differentiated goods under CES preferences illustrates the method.

Chapter 2

Advertising Motives and Firm Life-Cycle Dynamics in a General Equilibrium Model

2.1 Introduction

Firms advertise for different motives. First, firms have an incentive to advertise to make consumers aware of them and, in doing so, increase their customer base (i.e., the *informative motive*). By raising consumer awareness, advertising allows consumers to enjoy more variety and fosters competition. Second, firms may advertise to persuade current customers to spend more on their products (i.e., the *persuasive motive*). By enhancing consumer preferences for goods, advertising may positively affect utility from consumption but can also increase market power by reinforcing product differentiation. An extensive literature has examined whether advertising is informative or persuasive, with supporting evidence for both views.[1] However, research on the macroeconomic implications of advertising remains limited, and a framework accommodating both views is missing. This paper aims to fill this gap.

In addition to these traditional views, consumers' limited attention capacity implies that firms need to compete for the attention of consumers to make their way into their consumption sets. This introduces a novel effect of advertising: advertising by one firm diverts consumers' atten-

[1]See Bagwell (2007) for a review.

tion away from competitors. Thus, firms have an incentive to advertise to avoid competition by hindering competitors' ability to expand their customer base, which I refer to as the *anti-competitive motive*. This seems particularly relevant in settings like Google search or Amazon advertising, where firms compete to be placed in the top positions within a keyword, as these receive most of the attention.

What are the aggregate implications of the different advertising motives, and how do these motives evolve over the firm life cycle? This paper develops a novel model that is able to speak about these questions. Additionally, it is the first general equilibrium model of informative advertising where firms play strategically in a dynamic game, and I believe the methodology developed here will be valuable for subsequent research.

In the model, each industry is composed of a generic good produced under perfect competition by a fringe and an endogenous discrete set of oligopolistic firms each producing a differentiated good. A key feature of the model is that, within an industry, consumers are characterized by the set of goods they are aware of, which I refer to as awareness sets. All consumers know the generic good of each industry, but may not know all of the differentiated goods. The awareness sets evolve stochastically, affected by firms' advertising decisions. Thus, firms face a dynamic problem, as building a customer base takes time, which is motivated by the empirical evidence from Foster et al. (2016).[2] There is one type of advertising that has two effects. First, it increases the probability a consumer becomes aware of the good, and second, it increases the demand shifter for those consumers already aware of the good, inducing them to spend more on the advertised good. Given that the advertising space of each industry is limited, firms internalize that by increasing their advertising expenditure they increase the price of the advertising space in their industry, which reduces the advertising space acquired by competitors, reducing their visibility (the anticompetitive motive). This aligns with Google search advertising, where firms bid on specific keywords in a cost-per-click (CPC) auction.[3] So, heterogeneity in the demand for advertising in a specific keyword translate into heterogeneity

[2]In particular, Foster et al. (2016) take advantage of data on physical quantities in industries that are plausibly little subject to quality differentiation and find that the fact that older firms are bigger than younger firms cannot be explained by differences in productivity, and then they find support for the hypothesis that firms play an active role, not just a passive effect from aging. Einav et al. (2022), focusing on the retail sector, find that most of the variability in sales is accounted by the number of clients.

[3]Google doesn't charge firms just to be placed at the top, instead, the CPC is the amount the firm will be charged for each click their ad receives. Therefore, Google doesn't necessarily place the highest bidding firm at the top; it also considers the relevance of the ad.

in prices.[4] Additionally, there is a media sector that transforms the advertising expenditure into media goods using labor. These media goods are supplied to consumers at a zero monetary price.[5] Consumers choose the time they spend on them based on their entertainment value, and during this time, they are exposed to advertising.

The paper studies how the three motives change over the firm life cycle. Intuitively, younger firms have more potential customers to acquire, and so they tend to have a stronger informative motive. In contrast, the anticompetitive motive, which is about retaining market power over the existing customers by reducing the probability they learn about competitors, is stronger for older firms, as they have more to lose due to their larger customer base. The persuasive motive also tends to be stronger in older firms. Intuitively, if advertising persuades existing customers to spend more, the revenue increase will be larger the bigger the customer base.

The model is estimated by simulated method of moments for the U.S. economy to fit key empirical patterns regarding (i) the evolution of average firm growth by age, which is important to discipline the customer base building process in the model; (ii) the relationship between advertising expenditure and sales, which, given the previous results, is informative of the strength of the motives; and (iii) macroeconomic aggregates, specifically the sales-weighted average markup and its standard deviation, aggregate advertising expenditure as a share of GDP, the fraction of time spent on media, and the labor share. The model does well in matching the targets. In addition, the model also features an inverted-U relationship between advertising and sales as documented in previous literature.

To assess the aggregate effects of the motives and advertising as a whole, the calibrated model is compared to counterfactual economies where some or all of the motives are shut down from the firms' first-order conditions. The counterfactual where all motives are shut down (i.e., there is no advertising) reveals that advertising has a significant positive effect on the aggregate, not only through its role in financing media goods that provide entertainment but also through its role in spreading product awareness. In this counterfactual, the probability consumers learn about goods is lower, as they only learn through an exogenous probability; consequently, firms

[4]This heterogeneity is shown in Figure 1.A1. The same applies if we look at CPC in Google shopping ads across industries, although these are considerably cheaper, rarely more than one dollar per click.

[5]The presence of media goods that get their revenue from selling advertising space is pervasive, spanning traditional outlets like radio and TV as well as digital platforms such as YouTube, Instagram, Facebook, and Google. Although these media goods are barely reflected in GDP (see Greenwood et al. (2024)), the time consumers spend on them suggests they have a significant impact on welfare.

tend to be smaller. This lowers firms' growth prospects and discourages new firms from entering the market. As a result, consumption decreases, both due to consumers enjoying less variety and because a larger share of consumption comes from a fringe of small, unproductive firms. Overall, shutting down advertising would decrease consumption by 16.68%. The counterfactuals also show that while the persuasive motive increases markups and reduces entry, it has a net positive effect by increasing consumers' taste for the advertised good. The anticompetitive motive is detrimental to output, resulting in higher markups and lower entry. However, it can still have a positive aggregate effect, due to its contribution to the provision of media goods. In fact, a result of the calibrated model is that both the anticompetitive and persuasive motives matter mostly through the entertainment value of media goods, as their effects on consumption are modest. This is despite the fact that these motives significantly influence firms' advertising decisions: shutting down the anticompetitive motive reduces total advertising expenditure by 12%, and additionally shutting down the persuasive motive reduces it by a further 42%. A complementary decomposition analysis based on the firms' first-order condition shows that 8.45% of the (marginal) incentives to advertise are attributable to the anticompetitive motive, 33.74% to the persuasive motive, and 57.81% to the informative motive.

To quantify the inefficiencies, the decentralized equilibrium is compared to the one resulting from solving the social planner problem, while maintaining the consumers' information frictions. A novel feature of the model is that the social planner values media goods not only because they entertain consumers, as in existing literature, but also because, through the advertising in media, consumers get information that allows them to improve consumption. In other words, media serves as a vehicle for product awareness. Unsurprisingly, as the entertainment value of media goods increases, the social planner reallocates more labor from the production sector to the media sector. After a certain point, consumption under the planner's allocation becomes lower than under the decentralized one. More interestingly, when the entertainment value of media is negligible, the optimal quantity of media is lower than in the decentralized equilibrium, suggesting excessive advertising expenditure. However, this conclusion would be inaccurate as one must also consider how the advertising space is allocated among firms. In other words, the 'overprovision' of media, through its effect on learning, may help mitigating the inefficiencies arising from the misallocation in the advertising space. In this direction, the exercise examining the optimal uniform tax on advertising reveals that advertising should be subsidized.[6] Finally, given that the informative motive declines with firm age, while the persuasive and anticompetitive motives increase, a natural question to ask is what the welfare gains from allowing the advertising tax to be age-dependent would be. However, I find that the gains from such a policy are negligible.

The paper is organized as follows: Section 2.2 introduces the model and characterize the equilibrium. Section 2.3 estimates the model, studies the evolution of the motives, their contribution to total advertising expenditure, and their aggregate effect. Section 2.4 discusses the inefficiencies of the model, compares the informationally-constrained social optimal equilibrium to the decentralized one, and examines the gains from taxing advertising. Finally, section 2.5 concludes.

Related literature. This paper relates to the literature that studies the implications of customer capital for firm, industry, and macroeconomic dynamics (e.g. Dinlersoz and Yorukoglu (2012), Gourio and Rudanko (2014), Molinari and Turino (2018), Argente et al. (2023), Einav et al. (2022), Ignaszak and Sedlacek (2023), Greenwood et al. (2024)). In these models, firms grow via increasing their idiosyncratic demand (customer capital). Together with Cavenaile et al. (n.d.), we contribute to this literature by showing that it is not just the quantity of customers that matters, but also the degree of information the customers have about alternative goods. Relative to Cavenaile et al. (n.d.), this paper allows for strategic advertising decisions as well as the interaction between firms of different sizes and ages. In Cavenaile et al. (n.d.), advertising also serves to expand product awareness, but the advertising choices are coordinated at the industry level, made once and for all at industry inception, and firms are assumed to be symmetric. Their focus is on how the improvements in targeted advertising may lead to increased market power through market segmentation.

This paper is also related to Greenwood et al. (2024), who present a static model to study the inefficiencies from advertising. Here, the contribution is to add to the analysis the inefficiencies arising from markups typical of oligopoly frameworks, as well as identifying three novel sources of inefficiency from advertising. The first two are reminiscent of growth models, namely: (i)

[6] In this exercise, I compare the stationary equilibria resulting from the different tax levels, without considering the transition.

lack of full appropriability, as the producers cannot extract the entire surplus; and (ii) business-stealing, as firms don't consider the losses from the reduction of consumption from competitors. The third source of inefficiency arises from the anticompetitive motive particular to this model, as firms try to avoid suffering from the business-stealing effect. Note that the sources of inefficiency push in different directions, so it is not clear whether there is too much or too little advertising, and requires a quantitative answer. Finally, there is inefficient entry, again due to lack of full appropriability and business-stealing.

For the persuasive aspect of advertising, this paper builds on the literature that adopts the persuasive view, e.g. Cavenaile et al. (2024a), Rachel (2024), Molinari and Turino (2018). These papers model advertising as a static demand shifter. A novel contribution of the current paper is the combination of the persuasive and informative views of advertising within a single framework. In particular, this paper relates to Cavenaile et al. (2024a), as they also develop an oligopolistic model with endogenous market structure. In their paper, they study the interaction between R&D and advertising. As in my model, advertising in Rachel (2024) and Greenwood et al. (2024) also finances the provision of media goods that improve utility.

Finally, this paper uses the concept of consideration sets introduced by Manzini and Mariotti (2014) that are widely used in other fields.[7] In macroeconomics, this concept (using the term *awareness set*) is introduced by Cavenaile et al. (n.d.). Relative to them, this paper explores how the evolution of awareness sets both influences and is influenced by firms' advertising decisions. The presence of awareness sets complicates the firm problem, as firms need to keep track of the distribution of consumers across these sets. In their model, they abstract from this by assuming the evolution of the awareness sets is determined at industry inception, so the only state variable is industry age.

[7]Manzini and Mariotti (2014) model choice as a two-stage process. In the first stage, some of the available alternatives are selected into a *consideration set*, with a probability that is linked to attention. In the second stage, the agent maximizes utility restricted to the consideration set.

2.2 Model

2.2.1 Environment

Market structure and the production sector. There is a continuum of mass 1 of industries indexed by i. In each industry, there is a generic good and a discrete set $\mathcal{J}_{i,t}$ of firms, indexed by j, each one producing a single differentiated good with the production function $y_{j,i,t} = N_{j,i,t}$, where $N_{j,i,t}$ is the labor employed by firm j. The generic good is produced under perfect competition by many small firms with the production function $y_{0,i,t} = A_0 N_{0,i,t}$, where $N_{0,i,t}$ is the total labor employed by these small firms in industry i at period t.

Advertising and the media sector. There is a media sector populated by media firms that employ labor to produce media goods, which are supplied to consumers at zero monetary price, and generate revenue by selling advertising space to production firms. There is free entry. Each media firm produces a differentiated variety of media good of equal quality, so consumers will allocate the time they spend on media equally among the different media goods. The aggregate quality of media is given by

$$Q = AN_m^{\frac{1}{2}}, \tag{2.1}$$

where N_m is total labor employed in media. In order to rule out an equilibrium with no advertising expenditure and no media produced, I assume the government employs $\bar{N}_m$ units of labor in media, which is financed by a lump-sump tax to consumers.

Within the media sector, each industry of the production sector has α units of advertising space.[8] The process whereby firms acquire advertising space follows a kind of auction, where media firms post a price per unit of ad space in industry i, $p_{a,i,t}$, which is the minimum bid accepted, and supply at most α units of ad space. Letting $e_{j,i,t}$ be the advertising expenditure of firm j in industry i, then the final price per unit of ad space in industry i will be equal to $\max\left\{p_{a,i,t}, \sum_{j\in\mathcal{J}_{i,t}} e_{j,i,t}/\alpha\right\}$. Therefore, the advertising space acquired by firm j, $\alpha_{j,i,t}$, will be:

[8] This is a reasonable assumption for search advertising: there is one top search position for a specific keyword, so higher demand only leads to higher price, as suggested in Figure 1.A1. More generally, you could think of this α as some measure of attention.

$$\alpha_{j,i,t} = \min\left\{\frac{e_{j,i,t}}{p_{a,i,t}}, e_{j,i,t}\frac{\alpha}{\sum\limits_{k\in\mathcal{J}_{i,t}} e_{k,i,t}}\right\}. \tag{2.2}$$

Entry and Exit of firms. A firm is hit by a death shock with probability κ, independent of whether other firms are affected (so, the probability n firms exit is κ^n). Regarding entry, there is a measure one of entrepreneurs that employs $N_{e,i,t}$ units of labor to create a new differentiated good in industry i with probability $z_{e,i,t} = \phi_s N_{e,i,t}^{\frac{1}{2}}$. Upon successfully creating a new good (and, for computational purposes, provided the number of firms in the industry is below $\bar{J}$), a new firm enters the market, and initially no consumer is aware of the new firm. Entry and exit occur simultaneously right at the start of $t+1$.

Consumers. There is a unit mass of individuals indexed by ℓ who maximize lifetime utility, where the instantaneous utility is a function of her consumption (C_ℓ) and entertainment (L_ℓ) goods. Individuals die with an exogenous probability δ, in which case they are replaced with an offspring who inherits the assets $a_{\ell,t}$, and individuals discount the offspring's utility with the same discount rate; thus, we can write utility as if they were infinitely lived:[9]

$$U_\ell = \sum_{t=0}^{\infty} \beta^t \left[\mathbb{E} \ln C_{\ell t} + L_{\ell t}\right]. \tag{2.3}$$

Each individual supplies inelastically one unit of labor and chooses how much time to allocate to media goods, $T_{\ell,t}$, in order to maximise her entertainment good $L_{\ell,t}$, which is defined as follows:[10]

$$L_{\ell t} = \upsilon \left(Q_t T_{\ell,t} - \frac{T_{\ell,t}^2}{2}\right), \tag{2.4}$$

where Q_t is an output of the media sector production function. Anticipating that all individuals choose the same $T_{\ell,t}$, in what follows I drop the subindex ℓ from T_t and L_t. Individual ℓ gets her

[9]Note that the only source of uncertainty on $C_{\ell,t}$ comes from the probability of dying, but not from awareness. There is uncertainty at the industry level due to the stochastic evolution of the awareness sets (see next section), but the law of large numbers over the continuum of industries removes the uncertainty at the aggregate level.

[10]Note that $T_{\ell,t}$ is not restricted to be below 1; this is consistent with the way media time is measured in the data where multitasking is counted separately, see Appendix 2.7.1.

C_ℓ following a Cobb-Douglas aggregator of her consumption over the continuum of industries of mass 1

$$\ln C_{\ell,t} = \int_0^1 \ln C_{\ell,i,t} di \tag{2.5}$$

where the industry i consumption good of individual ℓ is a CES aggregator of her consumption on the generic good and each of the differentiated goods she is aware of:

$$C_{\ell,i,t} = \left(c_{\ell,0,i,t}^{\frac{\sigma-1}{\sigma}} + \sum_{j\in\mathcal{I}_{\ell,i,t}} \omega_{j,i,t} c_{\ell,j,i,t}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \tag{2.6}$$

where $c_{\ell,j,i,t}$ is the quantity of good j consumed by ℓ at t; $\mathcal{I}_{\ell,i,t}$ will be referred to as the awareness set of individual ℓ in industry i at period t, as it is the subset of the differentiated goods $\mathcal{J}_{i,t}$ the individual is aware of at period t (in the next section, I describe the evolution of this object); and $\omega_{j,i,t}$ is a demand shifter that depends on the exposure of individuals to the ad of good j. In particular:

$$\omega_{j,i,t} = 1 + \nu_s(\alpha_{j,i,t}T_t)^{\nu_c}, \quad \nu_c \in (0,1). \tag{2.7}$$

Note that the more time consumers spend on media, the larger the effect of advertising on the demand shifter.

Individual ℓ's budget constraint is given by:

$$w_t N_{\ell,t} + r_t a_{\ell,t} = \int_0^1 \sum_{j\in\mathcal{I}_{\ell,i,t}\cup\{0\}} c_{\ell,j,i,t} p_{j,i,t} di + a_{\ell,t+1} - a_{\ell,t} + \tau_t, \tag{2.8}$$

where w_t is the wage, $a_{\ell t}$ is the asset holding of individual ℓ at period t, r_t is the return on each unit of asset in period t, and τ_t is the lump-sum tax the government uses to employ $\bar{N}_m$ units of labor in the media sector. At time 0, all individuals are assumed to have the same level of assets a_0.

Product learning and the evolution of the awareness sets. I assume that the probability an individual gets aware of a product thanks to advertising is an increasing and concave function of the exposure to the ad of that good. In particular, assume an individual will get aware of product j in industry i with the following probability

$$\rho_{j,i,t} = \min\{1, \hat{\rho} + \psi_s(\alpha_{j,i,t}T_t)^{\psi_c}\}, \quad \psi_c \in (0,1). \tag{2.9}$$

Note that consumers can learn about multiple goods in the same period, and the events are independent. Although I focus on advertising as an active way through which firms can increase their customer base, consumers can get to know a firm in other ways (word-of-mouth, seeing the product in a shop...), and these are captured by $\hat{\rho}$. The inclusion of T_t is to capture the idea that the more time consumers spend on media, the more they are exposed to ads, and so the more effective advertising is, just like in the demand shifter $\omega_{j,i,t}$.

In addition, when a consumer dies, they are replaced by a newborn individual who starts knowing only the generic good of each industry (i.e. $\mathcal{I}_{\ell,i} = \emptyset$ for all i). This is equivalent to say that individuals forget all the differentiated goods they know with an exogenous probability δ. This assumption is not crucial for the results, and its only implication is that, even if a firm lived forever, there would always be some consumers that are not aware of it.

Then, we have all the information needed to find the probability of moving between any pair of awareness sets. Let $\Theta_{(\mathcal{I}\to\mathcal{I}')}$ be the probability of moving from $\mathcal{I}$ to $\mathcal{I}'$. Given that, conditional on not dying, the awareness set can only expand, then if $\mathcal{I}'$ doesn't contain $\mathcal{I}$, the transition is only possible (i.e. $\Theta_{(\mathcal{I}\to\mathcal{I}')} > 0$) if $\mathcal{I}' = \emptyset$, which happens with the probability of dying δ. Conversely, if $\mathcal{I}'$ contains $\mathcal{I}$, then the probability this transition takes place is the probability an individual doesn't die, $(1-\delta)$, times the probability of learning all the goods that are in $\mathcal{I}'$ but not in $\mathcal{I}$, $\prod\limits_{j\in\mathcal{I}'\setminus\mathcal{I}} \rho_{j,i}$, times the probability of not learning any of the goods that are not in

$\mathcal{I}'$, $\prod_{j \notin \mathcal{I}'} (1 - \rho_{j,i})$. Formally:

$$\Theta_{(\mathcal{I} \to \mathcal{I}')} = \begin{cases} 0, & \text{if } \mathcal{I} \nsubseteq \mathcal{I}' \neq \emptyset \\ \delta, & \text{if } \mathcal{I} \nsubseteq \mathcal{I}' = \emptyset \\ (1-\delta) \prod_{j \in \mathcal{I}' \setminus \mathcal{I}} \rho_{j,i} \cdot \prod_{j \notin \mathcal{I}'} (1 - \rho_{j,i}), & \text{if } \mathcal{I} \subseteq \mathcal{I}' \neq \emptyset \\ (1-\delta) \prod_{j \in \mathcal{J}_{i,t}} (1 - \rho_{j,i}) + \delta, & \text{if } \mathcal{I} = \mathcal{I}' = \emptyset. \end{cases} \tag{2.10}$$

2.2.2 Equilibrium

In this section, I characterize the pure strategy Markov perfect stationary equilibrium, that is such that the time spent in media T_t and the relative wage $\hat{w}_t = \frac{w_t}{E_t}$ are constant.

2.2.2.1 Consumption.

On the one hand, logarithmic preferences on $C_{\ell,t}$, together with $a_{\ell,0} = a_0$, imply that all consumers choose the same expenditure at all t: $E_{\ell,t} = E_t$.[11] On the other hand, CES preferences over the varieties within an industry implies that consumer's spending in an industry is independent of her industry price index: $E_{\ell,i,t} = E_t$. Therefore, the awareness sets $\mathcal{I}_{\ell,i,t}$ only affect the allocation of the expenditure within each industry. That is, in order to characterize consumer ℓ's consumption choices in industry i, we only need to know her awareness set in i, $\mathcal{I}_{\ell,i,t}$. In other words, within industry i, there are as many types of consumers as subsets $\mathcal{I} \subseteq \mathcal{J}_{i,t}$. So, the set of consumer types in industry i is identified by the power set of $\mathcal{J}_{i,t}$, $\mathcal{P}(\mathcal{J}_{i,t})$, and, within an industry, I'll use subindex $\mathcal{I}$ to denote the choice of an individual with awareness set $\mathcal{I}$.

[11]Equation 2.31 in the Appendix gives the analytical expression for E_t. In the stationary equilibrium, consumers spend all their income.

The optimal choices satisfy:[12]

$$c_{\mathcal{I},j,i,t} = E_t P_{\mathcal{I},i,t}^{\sigma-1} p_{j,i,t}^{-\sigma} \omega_{j,i,t}^{\sigma}, \quad j \in \mathcal{I}, \tag{2.11}$$

$$\frac{E_{t+1}}{E_t} = \beta(1 + r_{t+1}), \tag{2.12}$$

$$T_t = Q_t. \tag{2.13}$$

where $P_{\mathcal{I},i,t} = \left(p_{0,i,t}^{1-\sigma} + \sum_{j\in\mathcal{I}} \omega_{j,i,t}^{\sigma} p_{j,i,t}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$.[13] Note that consumers consume a positive amount of all the goods they are aware of, and the particular quantity consumed follows equation (2.11). Equation (2.12) is the Euler equation, and (2.13) states that the aggregate quality of media determines the time spent on media. From $E_t = P_{\mathcal{I},i,t} C_{\mathcal{I},i,t}$ we see one of the channels through which (the informative) advertising will increase welfare: advertising will increase the amount of products the consumer is aware of, which reduces the price $P_{\mathcal{I},i,t}$ of her industry composite good. This is a standard preference for variety effect.

2.2.2.2 The industry state and its evolution

Given that firms have the same production technology, all the heterogeneity comes from the consumer side. The relevant state of the industry is characterised by the triple $(\mathcal{J}_{i,t}, \mathcal{P}(\mathcal{J}_{i,t}), \vec{M}_{i,t})$, where $\vec{M}_{i,t} = (M_{i,t}(\mathcal{I}))_{\mathcal{I}\in\mathcal{P}(\mathcal{J}_{i,t})}$ is the mass of consumers in each awareness set (that is, the distribution of consumers over the awareness sets). Note that since there is a mapping from $\mathcal{J}_{i,t}$ to $\mathcal{P}(\mathcal{J}_{i,t})$, I may write the state simply as $(\mathcal{J}_{i,t}, \vec{M}_{i,t})$. There are two processes that shape the evolution of the industry state.

On the one hand, the industry state changes as consumers' awareness sets evolve due to learning and death, which, by law of large numbers, is a deterministic process at the industry level. Calling Θ_t the transition matrix, where the element in row r and column s indicates the probability of going from subset $\mathcal{I}_r$ to $\mathcal{I}_s$ at time t (i.e. $\Theta_{t,(\mathcal{I}_r\to\mathcal{I}_s)}$), and calling $\vec{M}_{i,t}$ the $2^{\#\mathcal{J}_{i,t}}$-dimensional row vector (where $\#\mathcal{J}_{i,t}$ is the cardinalilty of $\mathcal{J}_{i,t}$) containing the masses of consumers in each awareness set at time t; then, by the law of large numbers, the distribution of consumers in

[12] Together with the No-Ponzi condition $\lim_{\tau\to\infty} \frac{a_{t+\tau}}{\prod_{s=0}^{\tau}(1+r_{t+s})} = 0$.

[13] Note that consumers may not only have different industry price index, $P_{\ell,i,t}$, but also a different aggregate price index $P_{\ell,t} = exp\left(\int_0^1 \ln P_{\ell,i,t} di\right)$. In particular, as explained in section 2.2.2.8, individuals with the same age have the same aggregate price index, and the numeraire of the economy is the geometric mean of $P_{\ell,t}$.

$t+1$ in the absence of entry and exit of goods, which I denote by $\vec{\hat{M}}_{i,t+1}$, would be:

$$\vec{\hat{M}}_{i,t+1} = \vec{M}_{i,t}\Theta_t. \tag{2.14}$$

On the other hand, the industry state changes stochastically due to entry and exit of firms. If the realization of exit and entry changes the set of firms in industry i from $\mathcal{J}$ to $\mathcal{J}'$, then the next period industry state is obtained using the application $(\mathcal{J}, \vec{\hat{M}}, \mathcal{J}') \longmapsto (\mathcal{J}', \vec{M}')$ defined as follows.

$$\text{For } \mathcal{I}' \in \mathcal{P}(\mathcal{J}'), \qquad M'(\mathcal{I}') = \begin{cases} \sum\limits_{\{\mathcal{I}\in\mathcal{P}(\mathcal{J}):\mathcal{I}\cap\mathcal{J}'=\mathcal{I}'\}} \hat{M}(\mathcal{I}) & \text{, if } \mathcal{I}' \subseteq \mathcal{J} \\ 0 & \text{, if } \mathcal{I}' \nsubseteq \mathcal{J}, \end{cases} \tag{2.15}$$

where the first case says that two consumers become identical in industry i if all the firms in which they differed exit, whereas the second case says that there are no consumers who are aware of a newborn firm. The last piece of information needed to compute expected values is the probabilities that the set of differentiated goods moves from $\mathcal{J}$ to $\mathcal{J}' \subseteq \mathcal{J} \cup \{e\}$, where e denotes an entrant. These probabilities are given by:

$$\text{For } \mathcal{J}' \in \mathcal{P}(\mathcal{J} \cup e)), \qquad Prob\{\mathcal{J} \to \mathcal{J}'\} = \begin{cases} (1 - z_{e,i,t}) \prod\limits_{j\in\mathcal{J}\cap\mathcal{J}'} (1-\kappa) \prod\limits_{j\in\mathcal{J}\setminus\mathcal{J}'} \kappa & \text{, if } e \notin \mathcal{J}' \\ z_{e,i,t} \prod\limits_{j\in\mathcal{J}\cap\mathcal{J}'} (1-\kappa) \prod\limits_{j\in\mathcal{J}\setminus\mathcal{J}'} \kappa & \text{, if } e \in \mathcal{J}', \end{cases} \tag{2.16}$$

where $z_{e,i,t}$ is the probability of an entrant, $\prod\limits_{j\in\mathcal{J}\cap\mathcal{J}'} (1-\kappa)$ is the probability that all the firms in $\mathcal{J} \cap \mathcal{J}'$ survive, and $\prod\limits_{j\in\mathcal{J}\setminus\mathcal{J}'} \kappa$ is the probability that all the firms in $\mathcal{J} \setminus \mathcal{J}'$ exit.

2.2.2.3 Production firms' problem

Given the large number of small firms producing a homogeneous product, the price of the generic good is equal to its marginal cost, $p_{0,i,t} = \frac{w_t}{A_0}$. The differentiated firms compete in prices a la Bertrand and in advertising expenditures for the attention of consumers. Both decisions are made simultaneously.

Profits. Using the production function $y_{j,i,t} = N_{j,i,t}$, we can express profits decomposed as

$$\pi_{j,i,t} = \underbrace{M_{j,i,t}}_{\substack{\text{Customer} \\ \text{Base}}} \cdot \underbrace{(1 - \mathcal{M}_{j,i,t}^{-1}) \underbrace{\sum_{\mathcal{I} \in \mathcal{P}_j(\mathcal{J}_{i,t})} \frac{M_{i,t}(\mathcal{I})}{M_{j,i,t}} s_{\mathcal{I},j,i,t} E_t}_{\substack{\text{Average spending} \\ \text{by customers}}}}_{\text{Average rents from customers}}, \tag{2.17}$$

where $\mathcal{M}_{j,i,t} = \frac{p_{j,i,t}}{w_t}$ is the markup of firm j, $s_{\mathcal{I},j,i,t} = \frac{p_{j,i,t} c_{\mathcal{I},j,i,t}}{E_t}$ is type $\mathcal{I}$ individual's share of expenditure in good j, $\mathcal{P}_j(\mathcal{J}_{i,t}) = \{\mathcal{I} \in \mathcal{P}(\mathcal{J}_{i,t}) \mid j \in \mathcal{I}\}$ is the family of awareness sets containing good j, and $M_{j,i,t} = \sum_{\mathcal{I} \in \mathcal{P}_j(\mathcal{J}_{i,t})} M_{i,t}(\mathcal{I})$ is the customer base of firm j.

This expression offers a first intuition of the motives driving firms to advertise. First, they want to advertise to increase their customer base. I refer to this as the informative motive. Second, as shown in the Appendix 2.7.3.1, all else equal, firms prefer to have customers that know as fewer competitors as possible. Intuitively, the fewer alternative goods they know, the more they will spend in j (i.e. higher $s_{\mathcal{I},j,i,t}$) and the lower their demand elasticity (so, the firm is able to extract more rents by rising the markup). So, given that by increasing the advertising space they occupy, firms reduce the attention of consumers to the competitors' goods and so the probability they will add them to their awareness sets; then, firms may have the incentive to do advertising for the mere purpose of reducing the mass of customers who learn about competitors. I refer to this as the anticompetitive motive, as under this motive the firm is doing advertising to avoid competition by precluding competitors to expand their customer base. Finally, given that the demand shifter $\omega_{j,i,t}$ increases with the the advertising space, firms want to do advertising to persuade current consumers to buy more. This is the persuasive motive.

Price setting. I assume pure strategy Markov perfect equilibrium where policy functions only depend on the current industry state $(\mathcal{J}_{i,t}, \vec{M}_{i,t})$. Given that the price has no direct effect on the evolution of the industry state and that advertising and price choices are made simultaneously, then the price setting problem is static. The optimal markup $\mathcal{M}_{j,i,t}$ is such that profits (2.17) are maximised, given its own demand-shifter $\omega_{j,i,t}$, the markups and demand-shifters of the competitors $\{\mathcal{M}_{k,i,t}, \omega_{k,i,t}\}_{k \in \mathcal{J}_{i,t}, k \neq j}$, and the distribution of consumers over the awareness sets $\vec{M}_{i,t} = (M_{i,t}(\mathcal{I}))_{\mathcal{I} \in \mathcal{P}(\mathcal{J}_{i,t})}$, and taking into account that individuals' spending shares are given by

$$s_{\mathcal{I},j,i,t} = \left[(A_0 \mathcal{M}_{j,i,t})^{\sigma-1} \omega_{j,i,t}^{-\sigma} + \sum_{k \in \mathcal{I}} \left(\frac{\omega_{k,i,t}}{\omega_{j,i,t}} \right)^{\sigma} \left(\frac{\mathcal{M}_{j,i,t}}{\mathcal{M}_{k,i,t}} \right)^{\sigma-1} \right]^{-1}. \tag{2.18}$$

The equilibrium markups are given by:

$$\mathcal{M}_{j,i,t} = \frac{\frac{\sigma}{\sigma-1} - \bar{s}_{j,i,t}}{1 - \bar{s}_{j,i,t}} \text{ , with } \bar{s}_{j,i,t} = \sum_{\mathcal{I} \in \mathcal{P}_j(\mathcal{J}_{i,t})} \frac{M_{i,t}(\mathcal{I}) p_{j,i,t} c_{\mathcal{I},j,i,t}}{p_{j,i,t} y_{j,i,t}} s_{\mathcal{I},j,i,t}, \tag{2.19}$$

where $\bar{s}_{j,i,t}$ is the sales-weighted average of firm j customers' share of expenditure in industry i allocated to good j.

Note that in a standard oligopoly model with Bertrand competition, the optimal markup is given be the expression in (2.19) but with the market share $s_{j,i,t}$ instead of $\bar{s}_{j,i,t}$. So, while the optimal markup in a standard oligopoly model with Bertrand is increasing with size, this is not necessarily the case here. Here, the markup depends on the composition of the customers, not in the size: a smaller firm can have a higher markup if a larger fraction of its customers spend a larger share of expenditure on it. However, the model will still predict that, within an industry, larger firms have higher markups. The intuition is as follows: a firm that entered earlier had more time to accumulate customers (so older firms will be larger); but also, since as time passes consumers get aware of more goods and advertising is undirected, then a firm that enters later will get consumers that, on average, know more goods (and we have seen that customers with more alternative goods spend a smaller share). So, within an industry, larger firms will have customers that on average spend a larger share of expenditure, and thus they set higher markups.

Advertising choice. Each firm chooses dynamically its advertising expenditure $e_{j,i,t}$, taking into account (i) the advertising expenditure choices of its competitors $\{e_{k,i,t}\}_{k \in \mathcal{J}_{i,t}, k \neq j}$; (ii) markups $\{\mathcal{M}_{k,i,t}\}_{k \in \mathcal{J}_{i,t}}$; (iii) the time consumers spend on media, T_t; (iv) the law of motion of the industry state; and (v) that the actual advertising space purchased by each firm is given by (2.2). In practice, given that in equilibrium $p_{a,i,t}$ is such that total advertising expenditure in industry i exactly purchases α units of ad space, then, in all industries with more than one differentiated firms, $\alpha_{j,i,t}$ will be given by the second argument in (2.2), and so there will be an anticompetitive motive to advertise: by increasing $e_{j,i,t}$, firm j will achieve to increase the actual price for advertising space and thus reduce the advertising space of competitors, which will reduce the probability consumers learn about competitors. In industry states with only one differentiated firm, there is trivially no anticompetitive motive because there is no competitor and so the unique firm has no incentive to spend more than $p_{a,i,t}\alpha$, and so in such industry states $\alpha_{j,i,t}$ will be given by the first argument in (2.2).

Given that I focus on Markov perfect equilibrium, then the firm problem can be expressed in recursive form, with the value of the firm being a function of the state. Given that profits are linear on E_t, by guess and verify, the value of the firm is also linear in E_t. Therefore, defining $V_j(\mathcal{J}_{i,t}, \vec{M}_{i,t}) = \frac{V_{j,i,t}}{E_t}$, $\hat{e}_{j,i,t} = \frac{e_{j,i,t}}{E_t}$, $\hat{p}_{a,i,t} = \frac{p_{a,i,t}}{E_t}$ and $\pi_j(\omega_{j,i,t}, \mathcal{J}_{i,t}, \vec{M}_{i,t}) = \frac{\pi_{j,i,t}}{E_t}$ and using the Euler equation and that in the stationary equilibrium it will be $T_t = T$, we can write the dynamic firm problem recursively as

$$V_j\Big(\mathcal{J}, \vec{M}\Big) = \max_{\hat{e}_j} \Big\{\pi_j\Big(\omega_j, \mathcal{J}, \vec{M}\Big) - \hat{e}_j + \beta \mathbb{E} V_j\Big(\mathcal{J}', \vec{M}'\Big)\Big\}$$

$$\text{s.t. } \{\hat{e}_k\}_{k\in\mathcal{J}\setminus\{j\}},\ \{\mathcal{M}_k\}_{k\in\mathcal{J}},\ T,\ (2.9),\ (2.10),\ (2.14),\ (2.35),\ (2.36),\ (2.2).$$

We can decompose the first-order condition into the three motives to advertise: the informative motive (increase ρ_j), the anticompetitive motive (decrease $\rho_{j'}$, $j' \neq j$), and the persuasive motive (increase $s_{\mathcal{I},j,i}$ for $\mathcal{I} \in \mathcal{P}_j$):

$$1 = \underbrace{\frac{\partial \pi_{j,i}}{\partial e_j}}_{\substack{\text{Persuasive}\\ \text{motive}}} + \underbrace{\frac{\partial V_j}{\partial \rho_j}\frac{\partial \rho_j}{\partial \alpha_j}\frac{\partial \alpha_j}{\partial e_j}}_{\substack{\text{Informative}\\ \text{motive}}} + \underbrace{\sum_{j'\neq j}\left(-\frac{\partial V_j}{\partial \rho_{j'}}\right)\frac{\partial \rho_{j'}}{\partial \alpha_{j'}}\left(-\frac{\partial \alpha_{j'}}{\partial e_j}\right)}_{\text{Anticompetitive motive}}. \tag{2.20}$$

In section 2.3.2 I show how the intensity of the three motives evolve with firm age.

2.2.2.4 Entrepreneurs' problem

The entrepreneurs in an industry $(\mathcal{J}, \vec{M})$ choose N_e to maximise their expected value:

$$v^e(\mathcal{J}, \vec{M}) = \max_{N_e} \Big\{-N_e\hat{w} + \beta z_e \mathbb{E}_e V_e\Big(\mathcal{J}' \cup \{e\}, \vec{M}'\Big)\Big\},\ \text{s.t. } z_e = \phi_s N_e^{\frac{1}{2}}, \tag{2.21}$$

where $\mathbb{E}_e V_e\Big(\mathcal{J}' \cup \{e\}, \vec{M}'\Big)$ is the expected value of being a new firm conditional on successfully creating a new differentiated good (so, the expectation comes from the uncertainty on which of the $\mathcal{J}$ incumbents will survive). Then, the equilibrium labor employed in entry in an industry

$(\mathcal{J}, \vec{M})$ will be:

$$N_{e,(\mathcal{J},\vec{M})} = \left(\frac{\phi_s}{2\hat{w}} \beta \mathbb{E}_e V_e \Big(\mathcal{J}' \cup \{e\}, \vec{M}' \Big) \right)^2. \tag{2.22}$$

2.2.2.5 Stationary distribution

In the Appendix 2.7.7 I prove that, for any aggregates $\hat{w}$ and T given, with their associated solutions of the firms and entrepreneurs problems $\{\alpha_{j,(\mathcal{J},\vec{M})}, N_{e,(\mathcal{J},\vec{M})}\}$, the probability that an industry is at a given state $(\mathcal{J}, \vec{M})$ converges to an ergodic distribution (existence), which is independent of the initial state (uniqueness), and satisfies that the set of different states realised, call it Ω, is at most countably infinite.[14]

By the strong law of large numbers, this implies that the economy converges to a stationary distribution associated to the aggregates $\hat{w}, T$. Let $\mu_{(\mathcal{J},\vec{M})}$ be the mass of industries in state $(\mathcal{J}, \vec{M}) \in \Omega$ in this stationary distribution. If $\hat{w}, T$ are consistent with this stationary distribution, then we are in the stationary equilibrium.

The stationary distribution is computed using the method described in Appendix 2.7.8. To the best of my knowledge, no other paper computes a stationary distribution with a countably infinite number of states in its support, making this a methodological contribution of the paper.

2.2.2.6 Media sector problem

Given the identical media production by media firms, consumers allocate their media time T equally among the media firms, and production firms allocate their advertising expenditure equally among the media firms. Therefore, all media firms have the same profits, and so each media firm has positive profits if and only if the overall profits in the media sector are positive. Then, since there is free entry into the media sector, profits in the media sector must be zero in equilibrium; so, the equilibrium Q_t satisfies:

$$\int_0^1 \sum_{j \in \mathcal{J}_{i,t}} \hat{e}_{j,i,t} E_t di + w_t \bar{N}_m - w_t \left(\frac{Q_t}{A} \right)^2 = 0, \tag{2.23}$$

[14]Note that I have not formally proved whether the solution of the firms' and entrepreneurs' problems is unique; and so the stationary equilibrium is not necessarily unique.

where recall that $\bar{N}_m$ is the labor in media employed by the public sector. In the stationary equilibrium, $Q_t = Q$ is constant.

2.2.2.7 Labor market clearing

The labor market must clear, that is, the amount of labor supplied has to be equal to the labor demanded by the production firms, media firms and entrepreneurs. Without any loss of generality (just a change in the units we measure labor), I can normalize labor supply N to 1.

$$1 = N = \int_0^1 \left(\sum_{j \in \{0\} \cup \mathcal{J}_{i,t}} N_{j,i,t} + N_{e,i,t} \right) di + N_{m,t}, \tag{2.24}$$

where $N_{j,i,t} = \frac{s_{j,i,t}}{\mathcal{M}_{j,i,t}} \hat{w}^{-1}$, and $N_{e,i,t}$ and $N_{m,t}$ are given by (2.22) and (2.23), respectively. This pins down the equilibrium relative wage $\hat{w}_t$, and verifies that it is constant in the stationary equilibrium.

2.2.2.8 Aggregate output and representative consumer conditional on age

I define the aggregate consumption good as the geometric mean of the individuals' aggregate consumption goods; that is $\ln C_t = \int_0^1 \ln C_{\ell,t} d\ell$. Using the definitions of $C_{\ell,t}$ and $C_{\ell,i,t}$, together with $c_{\ell,j,i,t} = \frac{s_{\ell,j,i,t}}{\mathcal{M}_{j,i,t}} \hat{w}^{-1}$ and $c_{\ell,0,i,t} = s_{\ell,0,i,t} A_0 \hat{w}^{-1}$, and interchanging the integrals over ℓ and i, we obtain the level of the consumption good:

$$\ln C = -\ln \hat{w} + \sum_{(\mathcal{J},\vec{M}) \in \Omega} \mu(\mathcal{J},\vec{M}) \sum_{\mathcal{I} \in \mathcal{P}(\mathcal{J})} M(\mathcal{I}) \frac{\sigma}{\sigma - 1} \ln \left(\left(s_{\mathcal{I},0,(\mathcal{J},\vec{M})} A_0\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j \in \mathcal{I}} \omega_{j,(\mathcal{J},\vec{M})} \left(\frac{s_{\mathcal{I},j,(\mathcal{J},\vec{M})}}{\mathcal{M}_{j,(\mathcal{J},\vec{M})}} \right)^{\frac{\sigma-1}{\sigma}} \right). \tag{2.25}$$

The aggregate price index of the economy is P_t such that $P_t C_t = E_t$, and it is the numeraire (i.e. $P_t = 1$). GDP in the economy is given by $Y = E + \sum_{(\mathcal{J},\vec{M}) \in \Omega} \mu_{(\mathcal{J},\vec{M})} N_{e,(\mathcal{J},\vec{M})} \hat{w} E$.

Finally, note that applying a law of large numbers to the continuum of industries, two consumers with the same age will have the same level of aggregate consumption good. That is, although they will differ on their awareness sets for particular industries, at the aggregate level they will have the same distribution of awareness sets.

2.3 Quantitative analysis

2.3.1 Calibration

In this section, we describe the calibration of the model, and the details of the data sources and how the moments are computed are provided in the Appendix 2.7.1. One of the main components of the model is firms' customer base accumulation, which has a strong relationship with firm size, both in the model and in the data, as pointed by the empirical literature cited in the introduction. Therefore, it is important that the model reproduces the evolution of the average firm sales growth by age, in order to calibrate this customer base building process. In particular, we target the constant and the linear coefficient from the fitted line of the plot of average firm relative sales growth by age. Also, as shown in section 2.3.2, the intensity of the different motives to advertise varies with firm size, so the coefficient from a regression of advertising expenditure and sales is a good candidate to discipline the model.

To compute these three moments I use Compustat data. Given that firms typically enter Compustat a few years after their foundation (and certainly not with zero customers as it is assumed in the model for new firms), for the computation of the model-implied moments of these three targets, I assume that firms in the model are unobserved until they are at least five years old.

We estimate the model for the US at an annual frequency and set the consumer discount rate to $\beta = 0.98$. We also set (i) $\delta = 0.01$ corresponding to the mortality rate of 1% in the data, (ii) the concavity parameter for the persuasive advertising $\nu_c = 0.2972$ is taken from (the inverse) Cavenaile et al. (2024a), (iii) given that public sector spending on media represents roughly 0.008% of US GDP, dividing this by the (capital-adjusted) labor share, we set $\bar{N}_m = \frac{8 \cdot 10^{-5}}{0.8359}$, and (iv) we set $\kappa = 0.1151$ corresponding to the entry rate in the data. Acknowledging the difficulty to find good proxies for the utility value of media goods, we leave the weight of the entertainment good on the utility function, υ, uncalibrated and all the exercises involving welfare are made for a range of values of υ. This leaves 8 parameters to estimate: the elasticity of substitution parameter, σ; the relative productivity of the small firms producing the generic goods, A_0; the scale parameter for the persuasive effect of advertising, ν_s; the scale and convexity parameters for the informative effect of advertising, (ψ_s, ψ_c); the exogenous learning probability, $\hat{\rho}$; the scale parameter regulating the creation of new products, ϕ; and the aggregate productivity of the media sector, A. Apart from A, which can be derived directly from 2.23 using the target values for aggregate advertising expenditure and labor shares and the fraction of time in media, the rest of the 7 parameters are estimated jointly through a Simulated Method of Moments estimation procedure. Apart from the three moments described above concerning the average firm growth by age and the relationship between advertising and sales, at the aggregate level, I target the sales-weighted average markup and standard deviation, the aggregate advertising expenditure as a percentage of GDP, the fraction of time spent in media, and the labor share. Given that there is no physical capital in the model, for comparability, I take the labor share as the share of labor income among labor income and profits, following Cavenaile et al. (2024a). Table 2.1 summarises the results of the calibration. Panel A reports the parameter values, while Panel B reports both the model-implied moments and the empirical ones. The model does well in matching the moments. In addition to the targeted moments, the calibrated model also features an inverted-U relationship between advertising expenditure and relative sales as documented in Cavenaile et al. (2024a).

Note that Compustat is not the ideal dataset to discipline the growth pattern of firms in the model for the following reasons. First, firms do not automatically enter Compustat when they are born, and they may enter at different stages of the life cycle. Second, contrary to the model, firms may grow by expanding to new geographical markets or new product lines. Figure 1.1 plots the average firm relative sales growth rate both in the model and in the data. Note that in the model, if a firm had a constant $\rho_{j,i}$ (this is the case of a firm that has always been the single differentiated firm of the industry), then growth would be monotonically decreasing, pushed by a mechanical force: given that the population is constant, as the firm's customer base expands, the growth rate slows down because (i) a given increase in customers has a smaller relative impact, and (ii) there are fewer non-customers remaining. Things get noisier when there are other competitors and there is entry and exit.

2.3.2 Advertising motives and firm age and size

In this section, I quantify the share of the incentives to advertise attributable to each of the three motives and examine their relationship with firm age and size. The intuition is clear:

Table 1.1: Parameter values and targeted moments

A. Parameters

	Parameter	Description	Calibration	Value
Preferences	β	Discount rate	External	0.98
	σ	CES consumption	Internal	5.0625
	υ	weight of leisure	Uncalibrated	-
Persuasive	ν_s	Scale parameter	Internal	0.1250
	ν_c	Convexity parameter	External	0.2972
Learning	ψ_s	Scale parameter	Internal	0.2194
	ψ_c	Convexity parameter	Internal	0.4500
	$\hat{\rho}$	Exogenous learning	Internal	0.1000
	δ	Mortality rate	External	0.01
Media sector	A	productivity media firms	Internal	3.2087
	$\bar{N}_m$	public sector media	External	$9.5705 \cdot 10^{-5}$
Generic good	A_0	Productivity	Internal	0.5047
Entry/Exit	κ	Exit rate	External	0.1151
	ϕ	Entry scale	Internal	0.6422

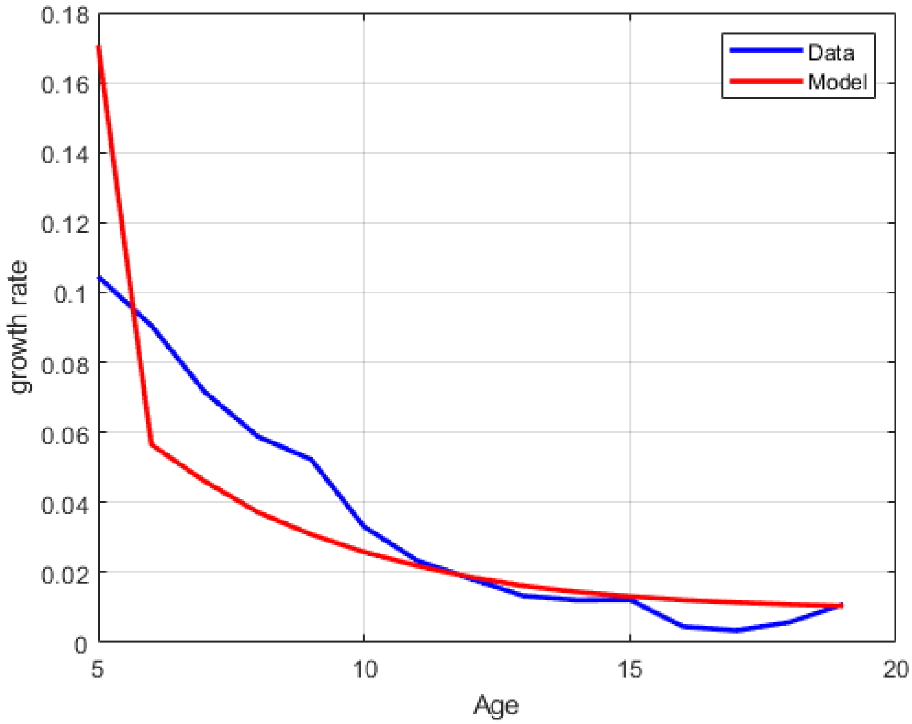
B. Moments

Moment	Data	Model
Sales-weighted average markup	1.3498	1.3281
St.Dev. Markup	0.3460	0.3479
Labor share (capital-adjusted)	0.8359	0.8392
Advertising/GDP	2.2%	2.1535%
Fraction of time in media	0.552	0.552
Intercept (firm growth, age)	0.0784	0.0831
Linear(firm growth, age)	-0.0061	-0.0063
Linear(adv. exp, market share)	0.6710	0.6501

Notes. Panel A reports the parameter values. Panel B reports the simulated and empirical moments. Details on data sources and how these moments are computed can be found in the Appendix 2.7.1

smaller or younger firms—those that are unknown to most consumers—have more potential customers to acquire. In the extreme case, a firm known by all consumers would have no incentive to advertise for informational purposes. Conversely, the anticompetitive motive, which is about retaining market power over the current customers by reducing the probability that they learn about competitors, becomes stronger as the customer base grows. A firm that is unknown to all consumers also has some incentive to prevent them from learning about other firms (since it internalizes that these consumers may eventually become customers, and thus wants them

Figure 1.1: Average firm growth by age

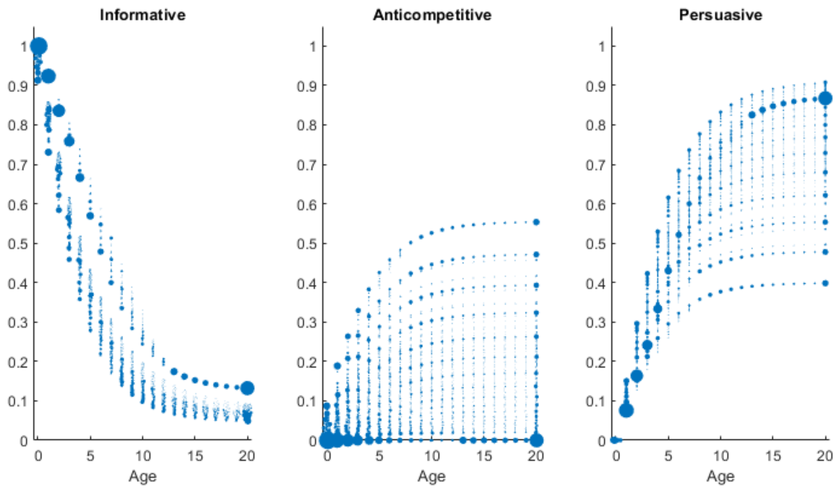


Notes. This figure displays the average firm relative sales growth by age both in the data (blue) and in the model (red). Given that firms typically enter Compustat a few years after their foundation, for comparability I assume age 0 in Compustat corresponds to age 5 in the model.

to know as few goods as possible), but intuitively, the incentive to prevent a consumer from learning about a competitor is higher when the consumer is already a current customer rather than just a potential one. Finally, the persuasive motive also tends to be bigger in older or larger firms. Intuitively, if advertising persuades current customers to spend more, the increase in revenue will be larger if there are more customers.

The decomposition of the FOC of advertising expenditure in (2.20) allows us to see the share of the firm's marginal value of advertising coming from each of the three motives. Using this observation, Figure 1.2 displays the share of the marginal value of advertising attributable to each of the three motives for all the firms in the stationary equilibrium. Three observations can be drawn. First, that there is significant heterogeneity, which indicates that age is far from being a sufficient statistic. This shows that industry dynamics play a key role (i.e. competition matters). Second, despite the variability, it can be observed that the informative motive is negatively associated with age, while the anticompetitive and the persuasive motives are pos-

Figure 1.2: Marginal intensity of the advertising motives by age



Notes. This figure displays the values from the terms of the FOC corresponding to the informative motive (left panel), the anticompetitive motive (middle panel) and the persuasive motive (right panel), for all the firms in the stationary equilibrium, where the relative size of each dot indicates the share of this firm type in the stationary distribution.

itively associated with it. Finally, firm age appears to play a particularly important role in distinguishing the motives during the first 5–10 years of a firm's life.

By aggregating the previous shares, weighted by total industry advertising expenditure, we obtain an indicative breakdown of the advertising expenditure attributable to each motive. This exercise suggests that 57.81% of the incentives correspond to the informative motive, while 33.74% correspond to the persuasive motive and 8.45% to the anticompetitive motive. Figure 1.A2 further repeats this decomposition exercise conditional on the number of firms in the industry. This reveals that the persuasive motive is more important as the number of competitors increases.

Because of the positive link between firm age and size (either sales or customer base, see Figure 1.A3), we obtain qualitatively similar plots when firm size is used instead of age. This numerical result is further supported by the following analytical result:

Proposition 1 *If distribution $\vec{M}_2$ is obtained from $\vec{M}_1$ by adding $\{j\}$ to the awareness sets of some consumers (that is, formally, if $\vec{M}_1$, $\vec{M}_2$ satisfy $M_2(\mathcal{I}\cup\{j\}) - M_1(\mathcal{I}\cup\{j\}) = M_1(\mathcal{I}) - M_2(\mathcal{I}) \geq 0$ for every $\mathcal{I} \in \mathcal{P}_{-j} = \{\mathcal{I} \in \mathcal{P} | j \notin \mathcal{I}\}$), then (for now, the proof is keeping the advertising choices fixed):*

1. *Firm j's informative motive is smaller in $\vec{M}_2$. That is, the informative motive is stronger*

in smaller firms.

2. *Firm j's anticompetitive and persuasive motives are bigger in $\vec{M}_2$. That is, both the anticompetitive and the persuasive motives are stronger in bigger firms.*

Proof. See the Appendix 2.7.5 ■

2.3.3 Counterfactuals shutting motives

How do each of the three motives affect the aggregates? Is the anticompetitive motive necessarily bad? What are the aggregate effects of shutting down advertising? This section addresses these questions. To do so, I compare the baseline economy with three counterfactual scenarios. The first is an economy where firms neglect the anticompetitive motive; that is, they don't internalise that by increasing their advertising expenditure they are effectively reducing the amount of consumers who learn about competitors. To be precise, this is done by removing the anticompetitive component from the firm's first order condition. In the second counterfactual, in addition, firms also neglect the persuasive motive; that is, they don't internalise that advertising increases current customers' spending. In the third one, the informative motive is also shut down, meaning firms don't advertise at all, and so consumers only learn through the exogenous probability; i.e., $\rho_{j,i,t} = \hat{\rho}$. This exercise illustrates what the economy would look like if firms neglected some of the motives to advertise. Such negligence alters firms' decisions, which in turn also has general equilibrium consequences.

Table 1.2 reports some relevant statistics for the counterfactuals and the benchmark. The second row shows the level of the consumption good assuming that the persuasive advertising is deceptive (i.e., consumers make their purchasing decisions based on $\omega_{j,i,t}$, but then they derive utility as if $\omega_{j,i,t} = 1$).

First, as intuition suggests, without the anticompetitive motive, smaller firms face less competition for advertising space, allowing them to grow faster, which increases competition and consequently lowers markups. Additionally, improved growth prospects increase entrepreneurs' incentives to create new products, driving up the entry rate. However, these positive effects on C are mitigated by a negative general equilibrium effect. Removing one incentive to advertise decreases the demand for advertising, which in turn reduces aggregate advertising expenditure,

Table 1.2: Comparison of counterfactuals with firms neglecting the anticompetitive and/or the persuasive motives

	Benchmark	No Anticompetitive	Only Informative	No motive
C	0.7260	0.7268	0.7226	0.6049
C no taste shifter	0.6898	0.6916	0.6935	0.6049
Q	0.5150	0.4860	0.3536	0.0314
Adv/GDP	2.1535	1.9137	1.0054	0
w	0.8392	0.8376	0.8345	0.9010
Avg Number of Firms	1.4148	1.4233	1.4406	1.1087
Sales-Weighted Average Markup	1.3281	1.3260	1.3135	1.1698
Coefficient advertising vs market share	0.6501	-2.5961	-15.0744	.

Notes. In the 'No Anticompetitive' counterfactual, firms make their decisions neglecting the anticompetitive motive; in the 'Only Informative' counterfactual, firms neglect both the anticompetitive and the persuasive motives; and in the 'No motive' firms neglect all the incentives to advertise, so they don't advertise and there is only the exogenous learning, $\rho_{j,i,t} = \hat{\rho}$.

leading to a lower supply of media goods. As a result, consumers spend less time on media, meaning they are less exposed to advertising, which renders advertising less effective. This explains the negligible overall effect on C.[15] Therefore, although the anticompetitive motive has an overall negative effect on consumption, it is not necessarily the case that welfare would be higher in a counterfactual economy without it, due to its contribution on the provision of media goods.

The counterfactual where, in addition, the persuasive motive is shut down suggests that the persuasive motive has an overall positive effect on consumption. The second line shows that this positive effect is due to consumers enjoying the advertised goods more. Similarly to the anticompetitive motive, shutting down the persuasive motive allows smaller firms to capture a larger share of the advertising space, which increases entry and lowers markups. Moreover, as in the first counterfactual, although shutting down the motive has a significant effect on firms' advertising decisions (in this case, advertising expenditure falls by 47.46%), its effect on C is very modest. Thus, on the aggregate, both motives matter mostly through media goods.

Finally, the last counterfactual shows that shutting down advertising would decrease consumption by 16.68% relative to the benchmark. In this counterfactual, the media sector only receives revenue from the public sector. In this counterfactual, firms' customer base only grows via the exogenous learning. This implies that, in the equilibrium, the differentiated firms will tend to

[15] Relatedly, in the Conclusion I discuss that the current specification of $\rho_{j,i,t}$ may exhibit excessive diminishing returns.

be smaller, which decreases the incentives to enter. Given that there are fewer differentiated firms and that they are, on average, smaller in size, the generic goods capture a larger market share, which drives the average markup down.

2.4 Welfare: Planner Problem and Taxation

In this section, I first identify the sources of inefficiency in the model and then solve the informationally-constraint planner problem and compare it to the decentralized equilibrium. Finally, I examine taxation on advertising.

2.4.1 Social planner problem

The model is inefficient for several reasons. First, the dispersion of markups typical in oligopolistic setups leads to labor misallocation in the production sector. Second, when choosing their advertising expenditure, firms do not internalize the entertainment value of media goods, which are financed through advertising. This points to an underprovision of media goods, as in Greenwood et al. (2024). Additionally, there are three sources of inefficiency coming from the advertising choices that are characteristic of the current paper: the anticompetitive motive, the lack of full appropriability, and business-stealing. Note that we can distinguish between inefficiencies in the level of advertising expenditure (or, equivalently, in the prices of advertising space or in the provision of media goods) and inefficiencies in the allocation of advertising space. In this sense, the anticompetitive motive points to too much advertising and shifts the allocation of advertising space toward older firms. The lack of full appropriability, meaning that firms cannot extract the full surplus, pushes towards having too little advertising. Finally, business-stealing here refers to firms not internalizing the losses from the reduction in the consumption of other goods when consumers learn about their product, which pushes toward excessive advertising, especially in industries with more competitors. Moreover, there is also inefficient entry, again due to lack of appropriability and business-stealing.

To assess the importance of these inefficiencies, I solve the following planner problem and compare the resulting equilibrium with the decentralized one.[16] The planner has full control over

[16]This is work in progress. Given that there are numerous sources of inefficiency, a more insightful exercise would be to compare allocations where the planner takes control of one additional decision.

production, media, and entrepreneurial decisions but cannot affect consumers' behavior; that is, the learning process and the choices regarding consumption and media time remain as they are in the decentralized equilibrium. Its goal is to maximize aggregate utility, with all individuals weighted equally. Formally, the planner solves:

$$\max_{\{N_{M,t},N_{j,i,t},N_{e,i,t},p_{j,i,t},\alpha_{j,i,t}\}} U = \sum_{t=0}^{\infty} \beta^t \int_0^1 [\ln C_{\ell t} + L_{\ell t}]\, d\ell$$

$$\text{s.t. } C_{\ell,t} \text{ from (2.5), } C_{\ell,i,t} \text{ from (2.6), } c_{\ell,j,i,t} \text{ from (2.11), and } L_{\ell,t} \text{ from 2.4, with } T_t = Q_t \quad \text{(Consumer choices)}$$

$$y_{j,i,t} = N_{j,i,t}, \quad y_{0,i,t} = A_0 N_{0,i,t}, \quad Q_t = A N_{m,t}^{\frac{1}{2}} \quad \text{(Production functions)}$$

$$1 = N_{m,t} + \int_0^1 \left(\sum_{j\in\{0\}\cup\mathcal{J}_{i,t}} N_{j,i,t} + N_{e,i,t} \right) di \quad , \quad w_t = E_t \quad \text{(Resource constraints)}$$

$$\sum_{j\in\mathcal{J}_{i,t}} \alpha_{j,i,t} = 1, (2.9), (2.10), (2.14), (2.35), (2.36) \quad \text{(Learning process)}$$

$$z_{e,i,t} = \phi N_{e,i,t}^{\frac{1}{2}}, (2.35), (2.36) \quad \text{(Entry and exit)}\ .$$

I leave the details of the solution in the Appendix 2.7.6. The planner sets prices equal to marginal cost times a markup (or a tax) that allows the planner to pay for the labor to produce the media goods and for entry. That is, $p_{j,i,t} = \tau w_t / A_j$, with $\tau = E_t/(w_t N_t^P)$, where N_t^P is the labor used in the production sector.

For the dynamic problem of advertising and media, as in the model introduced above, I focus on the stationary Markov perfect equilibrium. The social planner has to decide on (i) how to allocate the ad space among the differentiated firms of each industry, $\alpha_{j,i,t}$, (ii) how much labor to allocate to the media sector, $N_{m,t}$, and (iii) how much labor to allocate to creating new products in each sector, $N_{e,i,t}$.

First, let's see the social planner choice of $\alpha_{j,i,t}$. The allocation of the ad space has to be such that the marginal social gain of increasing the ad space given to each firm is the same, since otherwise we could improve the allocation. Formally, letting $\ln C_{i,t} = \int_0^1 \ln C_{\ell,i,t} d\ell$ be the total consumption good of industry i, and $U_X = \sum_{t=0}^{\infty} \beta^t \mathbb{E} \ln C_{i,t}$ be the expected life-time utility derived from an industry whose current state is X; it must be

$$\frac{\partial \ln C_{X,t}}{\partial \alpha_{j,X}} + \beta \frac{\partial \mathbb{E} U_{X'}}{\partial \rho_{j,X}} \frac{\partial \rho_{j,X}}{\partial \alpha_{j,X}} = \hat{h}_X \quad \text{for some } \hat{h}_X \text{ and all } j \in \mathcal{J}_X, \text{ together with } \sum_{j\in\mathcal{J}_X} \alpha_{j,X} = 1. \quad (2.26)$$

Note that the anticompetitive motive plays no role in the social planner's allocation of $\alpha_{j,X}$, as the planner directly chooses the ad space occupied by each firm. So, in deciding whether to give more ad space to one firm over another, the planner only considers the utility gains from informing more consumers and from enhancing customers' taste for that good.

Second, let's see the social planner choice of N_m. The planner takes into account that by employing more labor in media it will increase the aggregate quality Q of media, which has two effects: (i) it increases the level of entertainment L; and, by increasing the time spent in media, (ii) it increases the consumption good by increasing the probability of learning goods. The optimal N_m is given by

$$N_m = \frac{N^P}{2}\left(\upsilon Q^2 + \sum_{X\in\Omega}\mu(X)\hat{h}_X\alpha\right). \tag{2.27}$$

Note that, unlike existing literature, here the planner values the provision of media goods even if their entertainment value was negligible (i.e. even if $\upsilon = 0$), due to their role as a vehicle for spreading product awareness. Finally, the labor employed in entry in each industry satisfies:

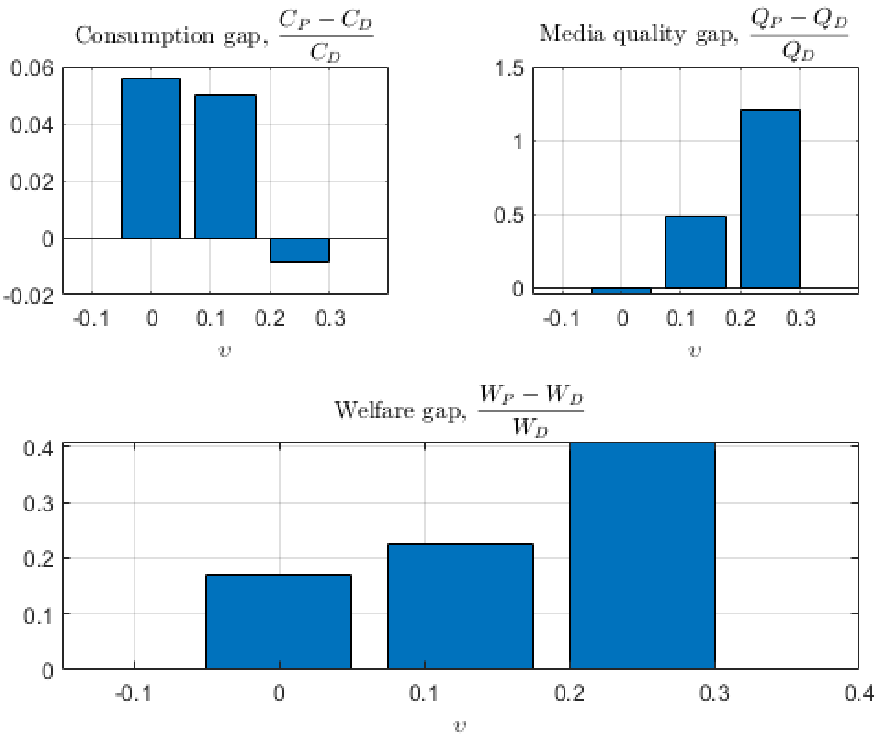
$$N_{e,X} = \left(\frac{\phi N^P}{2}\beta\left(\mathbb{E}_e U_{X'} - \mathbb{E}_{-e} U_{X'}\right)\right)^2, \tag{2.28}$$

where $\mathbb{E}_e U_{X'}$ (resp. $\mathbb{E}_{-e} U_{X'}$) is the expected industry-utility conditional on successfully creating (resp. not creating) a new differentiated good (so the undertainty comes from the probabilities the incumbents exit). The relative wage is $\hat{w} = 1$ as consumers spend all the income they receive, which is w. The labor market clearing, using 2.50 and 2.27 pins down N^P:

$$1 = N^P + N_e + N_m. \tag{2.29}$$

Figure 1.3 compares the planner economy with the decentralized one for different values of the relative utility weight of the entertainment good, υ. As expected, as υ increases, the planner puts more weight on producing media goods, at the expense of consumption, which eventually is lower than in the decentralized equilibrium. More interestingly, when $\upsilon \to 0$ (that is, when spending time on media doesn't provide any direct utility gain to consumers), the supply of media qoods is larger in the decentralized equilibrium, which seems to indicate that, when $\upsilon \to 0$, there is too much advertising. However, it is important to remind that there are inefficiencies both in the level of advertising expenditure as well as in the allocation of the advertising space. Therefore, this doesn't mean that welfare in the decentralized equilibrium would improve if all firms reduced their advertising expenditure proportionally to emulate the same Q as in the planner's equilibrium. In other words, the inefficiencies from the misallocation in the advertising space may be mitigated with the 'overprovision' of media, through its effect on learning.

Figure 1.3: Welfare comparison of the planner and decentralised allocations

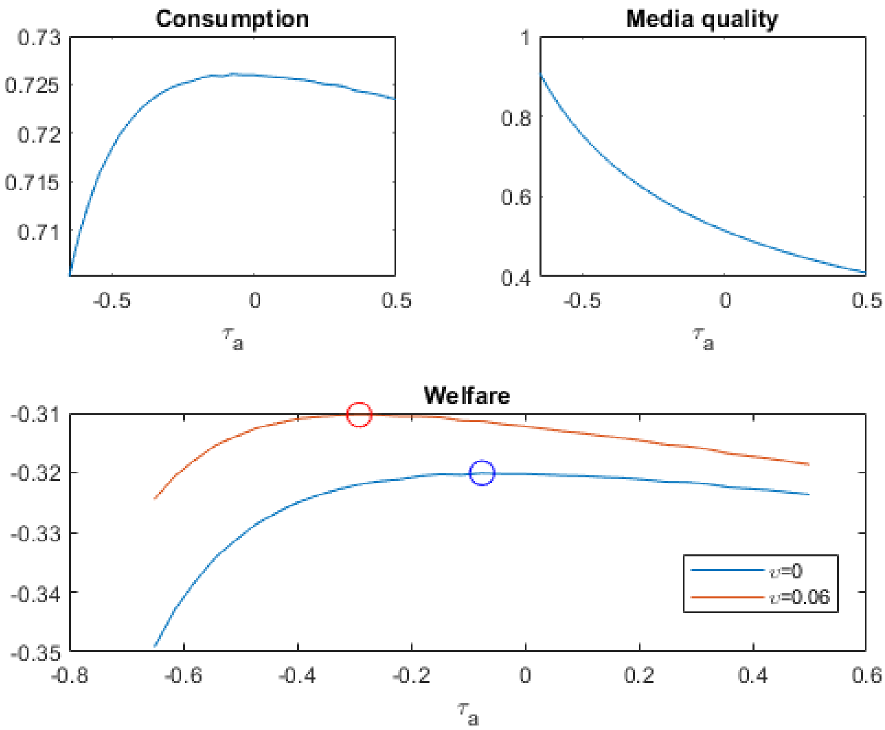


Notes. This figure displays the difference in final output (upper-left panel), in media time (upper-right panel), and welfare (bottom panel) between the planner's equilibrium and the decentralized one, relative to the decentralized one.

2.4.2 Taxing advertising: Uniform tax

Given that the decentralized equilibrium is inefficient, this section explores the welfare gains from the uniform tax on advertising that maximizes welfare. Here, in addition to $e_{j,i,t}$, firms pay $\tau_a e_{j,i,t}$ as taxes to the government, which are distributed as transfers to consumers. Figure 1.4 depicts the effect of this tax on final output, the quality of media, and welfare for different values of υ. As expected, the higher the entertainment value of media, the more valuable a subsidy on advertising becomes, as it magnifies the inefficiency arising from the fact that firms don't internalize the entertainment value of media goods. More interestingly, recall that we have seen that the decentralised equilibrium supplies more media than the planner's equilibrium in the case of $\upsilon \to 0$, which seems to point to an overprovision of media. Actually, it turns out that even when $\upsilon \to 0$, the optimal tax is a subsidy. This suggests that the 'overprovision' of media mitigates the inefficiencies from the misallocation of the advertising space, via increasing the time spent on media and thus the effectiveness of advertising.

Figure 1.4: Welfare under uniform tax on advertising



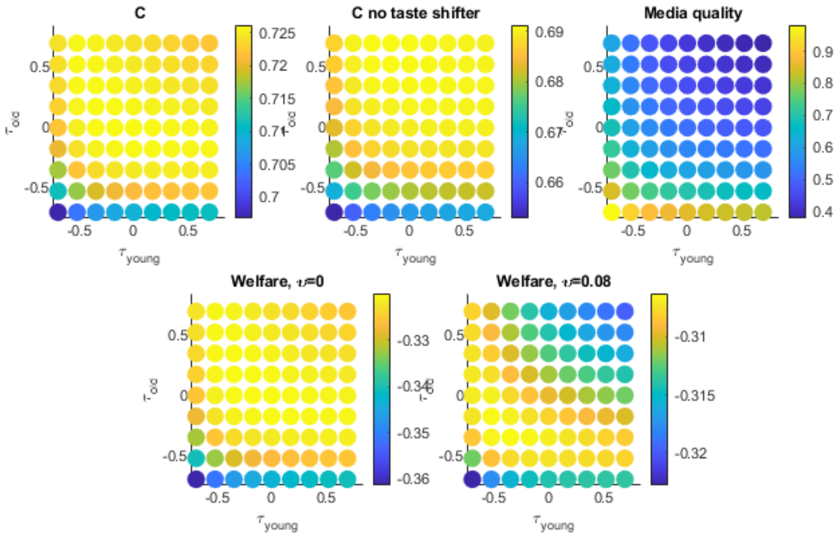
2.4.3 Age-dependent tax

The observation in section 2.3.2 that the informative motive decreases with age, while the persuasive and anticompetitive motives increase with age—and that these motives have different welfare implications—suggests that significant welfare improvements may be achieved by implementing an age-dependent tax rather than applying a uniform tax across all firms. Assume now that firms pay τ_Y if their age is less than the age cutoff $\bar{a}$, and τ_O if their age is greater than or equal to $\bar{a}$. In particular, for this exercise, I have set $\bar{a} = 3$; this means that firms receive different tax treatment during their first three periods of life compared to afterwards. This differential policy treatment makes the vector of ages of the firms an additional state. Note that firms that are at least $\bar{a}$ years old are indistinguishable by age; if all firms are older, then the firm problem is identical to the baseline with a uniform tax. However, for $a_j < \bar{a}$, we need to keep track of the particular age a_j; i.e., how close a firm is to $\bar{a}$ makes a difference. So, if $(a_1, \ldots, a_J)$ is the vector of ages (from older to younger), then the relevant age state is

$\vec{a} = (\hat{a}_1, \ldots, \hat{a}_J)$, where $\hat{a}_j = \min(a_j, \bar{a})$.

Figure 1.5 illustrates the effect of this age-dependent advertising tax on (i) consumption, (ii) consumption under the assumption that the persuasive effect of advertising is deceptive (as described in 2.3.3), (iii) media quality, and (iv) welfare for two values of υ. As with the uniform tax, the gains in consumption from the optimal age-dependent tax are small.

Figure 1.5: Welfare under age-dependent tax on advertising



2.5 Concluding remarks

The informative and persuasive aspects of advertising are widely accepted: firms may advertise to mitigate some information frictions by informing consumers, but also to shift market shares from one firm to another (Bagwell, 2007). On top of this, I highlight that the fact that consumers' attention is limited introduces a novel motive to advertise: firms may want to divert consumers' attention away from competitors.

This paper develops a novel model that accommodates the three motives and allows us to consider how firms build their customer capital and interact with their competitors' customer capital. Additionally, the paper introduces new methods—specifically, piecewise multivariate Newton interpolation and the method used to find the stationary distribution—which we believe will be valuable for future research.

I first use the model to examine how the motives evolve along the firm's life-cycle, their contribution to total advertising, and their aggregate effects. The informative motive, which accounts for around half of the incentives to advertise, is stronger for younger firms, as these are less known. The persuasive and anticompetitive motives, which are stronger in older firms, while relevant from the firms' perspective, have an almost negligible effect on aggregate consumption. Instead, they mostly matter through the provision of entertaining media goods, as they significantly contribute to total advertising expenditure. However, completely shutting down advertising leads to a considerable 16.68% reduction in consumption Finally, given that there are several sources of inefficiency in the model, I compare the planner's allocation with the decentralized one, and study the welfare gains from taxing advertising. I find that advertising should be subsidized, although the gains are small.

Two considerations are relevant in understanding the modest effect on aggregate consumption both when only some motives are shut down and when advertising is taxed. First, as explained in section 2.3.3, there is a general equilibrium effect that mitigates their impact. Second, the model features inherent diminishing returns to advertising: as more consumers become informed, fewer consumers remain to be informed. Therefore, the assumption that the probability of learning also exhibits diminishing returns may imply that the diminishing returns to advertising are too strong, so changes in the exposure to advertising have little impact on learning. Future work should explore the robustness of the results to different specifications of the learning process.

2.6 References

Argente, David, Doireann Fitzgerald, Sara Moreira, and Anthony Priolo, "How do entrants build market share? The role of demand frictions," *American Economic Review: Insights*, 2023, *forthcoming.*

Bagwell, Kyle, *The Economic Analysis of Advertising*, Elsevier, 2007.

Cavenaile, Laurent, Murat Celik, Jesse Perla, and Pau Roldan-Blanco, "A Theory of Dynamic Product Awareness and Targeted Advertising," *Working paper.*

_ , _ , Pau Roldan-Blanco, and Xu Tian, "Style Over Substance? Advertising, Innovation and Endogenous Market Structure," *Journal of Monetary Economics*, 2024a.

Dinlersoz, Emin M. and Mehmet Yorukoglu, "Information and industry dynamics," *American Economic Review*, 2012, *102 (2)*, 884–913.

Einav, Liran, Peter J. Klenow, Jonathan D. Levin, and Raviv Murciano-Goroff, "Customers and Retail Growth," 2022. Working paper.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "The Slow Growth of New Plants: Learning about Demand?," *Economica*, 2016, *83* (329), 91–129.

Gourio, François and Leena Rudanko, "Customer capital," *The Review of Economic Studies*, 2014, *81 (3)*, 1102–1136.

Greenwood, Jeremy, Yueyuan Ma, and Mehmet Yorukoglu, ""You Will:" A Macroeconomic Analysis of Digital Advertising," *The Review of Economic Studies*, 2024, *forthcoming.*

Ignaszak, Marek and Petr Sedlacek, "Customer Acquisition, Business Dynamism and Aggregate and Growth," *Working paper*, 2023.

Manzini, Paola and Marco Mariotti, "Stochastic choice and consideration sets," *Econometrica*, 2014, *84(3)*, 1153–1176.

Molinari, Benedetto and Francesco Turino, "Advertising and aggregate consumption: a Bayesian DSGE assessment," *The Economic Journal*, 2018, *128(613)*, 2106–2130.

Rachel, Lukasz, "Leisure-Enhancing Technological Change," *mimeo*, 2024.

2.7 Appendix

2.7.1 Calibration Appendix: Data sources and Computation of moments

1. **Sales-weighted average markup, sales-weighted standard deviation of markups, labor share, entry rate, and aggregate advertising expenditure as a percentage of GDP.** Taken from Cavenaile et al. (2024a). Following Cavenaile et al. (2024a), given that there is no physical capital in the model, I target the labor share among labor income and profits. Given that $\frac{wL}{wL+\pi+rK} = \frac{wL}{wL+\pi}\frac{wL+\pi}{wL+\pi+rK} = \frac{wL}{wL+\pi}\left(1 - \frac{rK}{wL+\pi+rK}\right)$; then, the target used is obtained from dividing the labor share by one minus the capital share. In the model, given that labor supply is normalised to 1, then labor share equals w.
   In the model, in the stationary distribution the entry and exit rates are equalized, and the exit rate is given by the exogenous probability of exiting κ.

2. **Fraction of time in media.** According to Statista, people in the US spend on average 751 minutes per day in media, which corresponds to the 0.521528 of time. Note that in this measure of media time multitasking is counted separately; that is: it counts the time spend in media while also doing other activities (e.g. commuting to work, breaks at work, listening a podcast while cooking or running), and duplicated media time when using multiple forms of media simultaneously (e.g. watching the TV while using a phone will count double).

3. **Coefficient of a regression of advertising expenditure on relative sales.** This and the growth by age moments are computed using Compustat data for the time period 1976-2018. Both in the model and in the data, I take the logarithm of advertising expenditure and then I standardise it by subtracting their means and dividing by their standard deviation for comparability. In the data, I regress the standardised logarithm of advertising expenses on relative sales of the firm in its SIC4 industry, controlling for the same set of controls used in Cavenaile et al., namely: profitability, leverage, market-to-book ratio, log R&D stock, firm age, the coefficient of variation of the firm's stock price, the number of firms in the industry, and a full set of year and SIC4 industry fixed effects. In the

model, I regress the standardised logarithm of advertising expenses, $p_{a,i,t}e_{j,i,t}$, on market shares, $s_{j,i,t}$, with industry fixed effects. Table 1.A1 shows the results of the empirical regression:

Table 1.A1: Advertising and relative sales in the data

	log advertising expenses
Relative sales	0.671 (0.0448)***
R^2	0.6056
N	40,007

Notes. Robust asymptotic standard errors (in parenthesis) are clustered at the firm level. The sample period is from 1976 to 2018. The regression controls for profitability, leverage, market-to-book ratio, log R&D stock, firm age, the coefficient of variation of the firm's stock price, the number of firms in the industry, and a full set of year and SIC4 industry fixed effects.

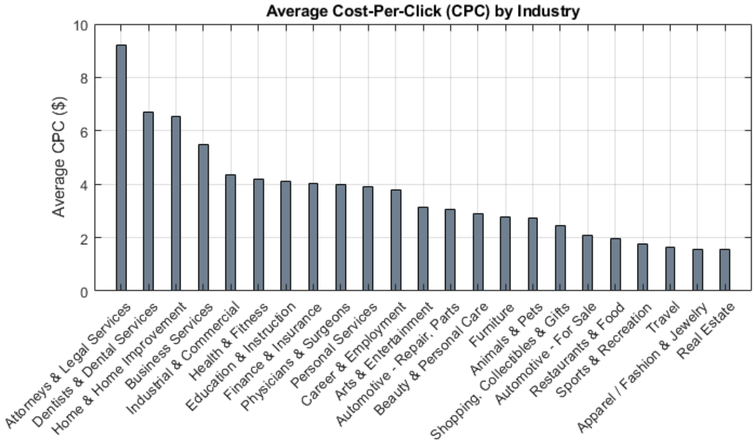
4. **Constant and slope of the fitted line of average firm relative sales growth by age.** In Compustat, I define age as the number of years since the first appearance of the firm in Compustat. First, for comparability with the model, where there is no aggregate growth, I compute growth rates of relative sales of the firm in its SIC4 industry. Second, firms in the data may experience big jumps on sales through expansion to new markets or via mergers and acquisitions, and I am interested in the average evolution of firm growth in the absence of such disruptive events; therefore, I drop all the observations of a firm posterior to a big change in their relative sales. In particular, if a firm's relative sales increase by more than 100% or decrease by more than 50%, this observation and the posterior ones of this firm are dropped. Then, I take the average firm relative sales growth grouping all the observations with the same age. In the model, I redefine age by subtracting 5 (as I am assuming that age 5 in the model corresponds to age 0 in Compustat). Given the average firm relative sales growth by age, $\bar{g}_a$, I define the fitted line $\hat{g}_a = \beta_0 + \beta_1 a$, where a is age. The coefficients β_0 and β_1 are the targeted moments.

5. **Calibration of the public sector financed media $\bar{N}_m$.** According to the US Government Accountability Office, the federal government spent $14.9 billion over the last 10 fiscal years (2014-2023). Then, I use that federal governments spent roughly $1.49

billion per year. In addition, federal appropriations for CPB (Corporation for Public Broadcasting) amounted to \$477 million in fiscal year 2023. So, the estimate I use for public sector spending on media is (\$1.49 + \$0.477) billion, which I divide for the US GDP in 2023, \$27360 billion. This gives 0.008% of GDP, which divided by $w = 0.8359$ gives the $\bar{N}_m = 9.5705 \cdot 10^5$.

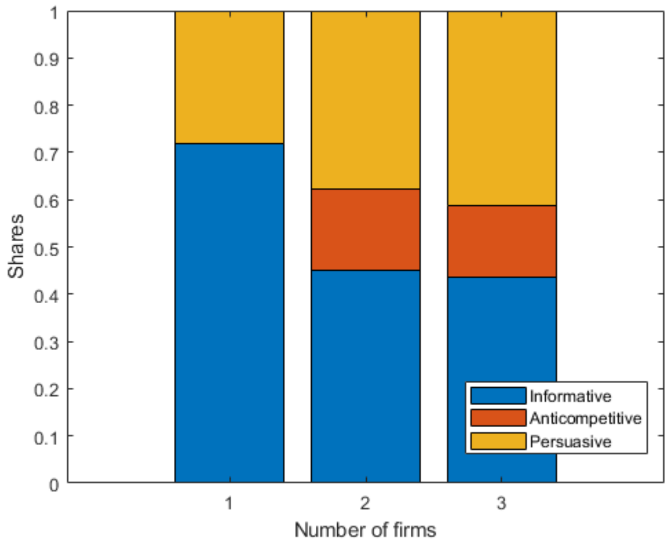
2.7.2 Additional Figures

Figure 1.A1: Average Cost-Per-Click (CPC) in Google search ads by industry



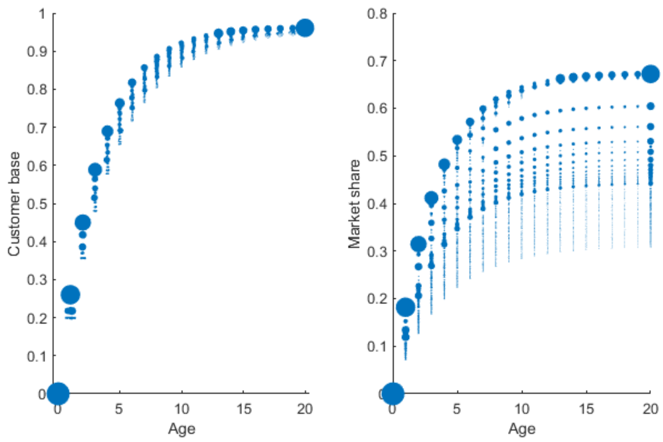
Notes. Adapted from Wordstream (2023). This figure displays the average CPC in Google ads by industry, calculated by dividing the overall cost of a campaign by the number of clicks it received. Each individual click has a different cost as it's determined by the Google Ads auction algorithm.

Figure 1.A2: Decomposition of advertising incentives by motive conditional on age



Notes. This figure displays the average shares of the incentives from the FOC attributable to each motive, weighted by industry advertising expenditure and conditioned on the number of differentiated firms in the industry.

Figure 1.A3: Customer base and market share and age



Notes. This figure displays the relationship of age with customer base (left panel) and market share (right), with the size of the dots indicates the share of this firm type in the stationary distribution.

2.7.3 Preferences

$$\max_{\{\{c_{\ell,j,i,t}\},a_{\ell,t+1},T_{\ell,t}\}} U_\ell = \sum_{t=0}^{\infty} \beta^t \left[\mathbb{E}\frac{C_{\ell,t}^{1-\theta}-1}{1-\theta} + L_{\ell,t} \right],$$

$$\text{s.t. } C_{\ell,t} = \left(\int_0^1 C_{\ell,i,t}^{\frac{\chi-1}{\chi}} di \right)^{\frac{\chi}{\chi-1}}, \qquad L_{\ell,t} = v\left(Q_t T_{\ell,t} - \frac{T_{\ell,t}^2}{2} \right),$$

$$C_{\ell,i,t} = \left(c_{\ell,0,i,t}^{\frac{\sigma-1}{\sigma}} + \sum_{j\in\mathcal{I}_{\ell,i,t}} \omega_{\ell,j,i,t} c_{\ell,j,i,t}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

$$w_t N + r_t a_{\ell t} = \int_0^1 \sum_{j\in\mathcal{I}_{\ell,i,t}} c_{\ell,j,i,t} p_{j,i,t} di + a_{\ell,t+1} - a_{\ell,t} + \tau_t.$$

(for the case $\theta = 1$, $\lim_{\theta\to 1} \frac{c^{1-\theta}}{1-\theta} = \lim_{\theta\to 1} \frac{c^{1-\theta}-1}{1-\theta} + \lim_{\theta\to 1}\frac{1}{1-\theta} = \ln c + \lim_{\theta\to 1}\frac{1}{1-\theta}$)

We can already plug $C_{\ell,t}$ and $C_{\ell,i,t}$ into the objective function.

The FOC reads:

$$[c_{\ell jt}]: \qquad \frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} \frac{\partial C_{\ell,t}}{\partial C_{\ell,i,t}} \frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = \mu_{\ell,t} p_{j,i,t},$$

where $\frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} = \beta^t C_{\ell,t}^{-\theta}$, $\frac{\partial C_{\ell,t}}{\partial C_{\ell,i,t}} = C_{\ell,t}^{\frac{1}{\chi}} C_{\ell,i,t}^{-\frac{1}{\chi}}$, and $\frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = C_{\ell,i,t}^{\frac{1}{\sigma}} c_{\ell,j,i,t}^{-\frac{1}{\sigma}} \omega_{\ell,j,i,t}$.

We can break down the FOC into three conditions, by defining $P_{\ell,i,t}$ as $P_{\ell,i,t}C_{\ell,i,t} = \sum_{j\in\mathcal{I}_{\ell,i,t}} c_{\ell,j,i,t}p_{j,i,t}$, and $P_{\ell,t}$ as $P_{\ell,t}C_{\ell,t} = \int_0^1 C_{\ell,i,t}P_{i,t}di$:

1. $\frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} \frac{\partial C_{\ell,t}}{\partial c_{\ell,j,i,t}} = \mu_{\ell,t} \frac{\partial E_{\ell,t}}{\partial C_{\ell,t}} \frac{\partial C_{\ell,t}}{\partial c_{\ell,j,i,t}} \implies \left[\frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} - \mu_{\ell,t}\frac{\partial E_{\ell,t}}{\partial C_{\ell,t}} \right] \frac{\partial C_{\ell,t}}{\partial c_{\ell,j,i,t}} = 0 \implies \beta^t C_{\ell,t}^{-\theta} = \mu_{\ell,t} P_{\ell,t}$.

2. $\frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} \frac{\partial C_{\ell,t}}{\partial C_{\ell,i,t}} \frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = \mu_{\ell,t} \frac{\partial E_{\ell,t}}{\partial C_{\ell,i,t}} \frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} \implies \left[\frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} \frac{\partial C_{\ell,t}}{\partial C_{\ell,i,t}} - \mu_{\ell,t}\frac{\partial E_{\ell,t}}{\partial C_{\ell,i,t}} \right] \frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = 0$, where using from the previous condition that $\mu_{\ell,t} = \beta^t C_{\ell,t}^{-\theta} P_{\ell,t}^{-1}$, we get: $C_{\ell,t}^{\frac{1}{\chi}} C_{\ell,i,t}^{-\frac{1}{\chi}} = \frac{P_{\ell,i,t}}{P_{\ell,t}} \implies C_{\ell,i,t} = C_{\ell,t}\left(\frac{P_{\ell,t}}{P_{\ell,i,t}}\right)^{\chi}$, and plugging it into the definition of $C_{\ell,t}$, we get $P_{\ell,t} = \left(\int_0^1 P_{\ell,i,t}^{1-\chi} di \right)^{\frac{1}{1-\chi}}$. Note that if Cobb-Douglas (i.e. $\chi = 1$), then $E_{\ell,t} = P_{\ell,t}C_{\ell,t} = P_{\ell,i,t}C_{\ell,i,t}$.

3. $\frac{\partial U_{\ell,t}}{\partial C_{\ell,t}} \frac{\partial C_{\ell,t}}{\partial C_{\ell,i,t}} \frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = \mu_{\ell,t} \frac{\partial E_{\ell,t}}{\partial c_{\ell,j,i,t}} = \mu_{\ell,t} p_{j,i,t}$, where using from the previous conditions that $\mu_{\ell,t} = \beta^t C_{\ell,t}^{-\theta} P_{\ell,t}^{-1}$ and $C_{\ell,t}^{\frac{1}{\chi}} C_{\ell,i,t}^{-\frac{1}{\chi}} = \frac{P_{\ell,i,t}}{P_{\ell,t}}$, we get: $\frac{P_{\ell,i,t}}{P_{\ell,t}} C_{\ell,i,t}^{\frac{1}{\sigma}} c_{\ell,j,i,t}^{-\frac{1}{\sigma}} \omega_{\ell,j,i,t} = \frac{p_{j,i,t}}{P_{\ell,t}} \implies c_{\ell,j,i,t} = C_{\ell,i,t}\left(\frac{\omega_{\ell,j,i,t}P_{\ell,i,t}}{p_{j,i,t}}\right)^{\sigma}$, and plugging it into the definition of $C_{\ell,i,t}$, we get: $P_{\ell,i,t} = \left(p_{0,i,t}^{1-\sigma} + \sum_{j\in\mathcal{I}_{\ell,i,t}} \omega_{\ell,j,i,t}^{\sigma} p_{j,i,t}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$.

The FOC for assets is:

$$[a_{\ell,t+1}]: \qquad \mu_{\ell t} = (1 + r_{t+1})\mu_{\ell,t+1}.$$

From the first one, using that $\mu_{\ell,t} = \beta^t C_{\ell,t}^{-\theta} P_{\ell,t}^{-1}$, we get the Euler equation: $\beta^t C_{\ell,t}^{-\theta} P_{\ell,t}^{-1} = (1 + r_{t+1})\beta^{t+1}\mathbb{E}C_{\ell,t+1}^{-\theta} P_{\ell,t+1}^{-1}$, which assuming $\theta = 1$ (i.e. logarithmic preferences on $C_{\ell,t}$), then the expenditure choice is independent of the price indices (so, the awareness set just affects the intratemporal allocation of expenditure).

So, assuming $\chi = \theta = 1$, we have:[17]

$$\begin{aligned} c_{\ell,j,i,t} &= E_{\ell,t} P_{\ell,i,t}^{\sigma-1} p_{j,i,t}^{-\sigma} \omega_{\ell,j,i,t}^{\sigma}, \\ \frac{\mathbb{E} E_{\ell,t+1}}{E_{\ell,t}} &= \beta(1+r_{t+1}). \end{aligned}$$

where E_ℓ is the expenditure of individual ℓ. Next, I show that we can get rid of the expectation, since the expenditure path just depends on the initial level of assets (and, therefore, since all individuals start with the same level of assets, this implies all consumers choose the same expenditure path, $E_{\ell,t} = E_t$). To show this, first consolidate the budget constraint:
To do this, first express the budget constraint as $a_t = (1+r)^{-1}[E_t - w_t + a_{t+1}]$. Next, substitute a_{t+1} using next period's budget constraint (here we will have an expectation coming from the possibility the individual dies:

$$a_t = (1+r)^{-1}\left[E_t - w_t + (1+r)^{-1}\left[\mathbb{E}_t E_{t+1} - w_{t+1} + \mathbb{E}_t a_{t+2}\right]\right].$$

Iterating:

$$a_t = \lim_{\bar{t}\to\infty} \sum_{s=0}^{\bar{t}} (1+r)^{-(s+1)} \left[\mathbb{E}_t E_{t+s} - w_{t+s}\right] + \lim_{\bar{t}\to\infty} (1+r)^{-(\bar{t}+1)} \mathbb{E}_t a_{\bar{t}+1}, \tag{2.30}$$

where the second limit is zero due to the No-Ponzi condition. On the other hand, the Euler condition tells us that $\mathbb{E}_t E_{t+s} = \beta(1+r)\mathbb{E}_t E_{t+s-1}$, and iterating: $\mathbb{E}_t E_{t+s} = \beta^s (1+r)^s E_t$. Using this in 2.30, and that in the stationary equilibrium $w_{t+s} = w_t g^s$, we have:

$$a_t = \sum_{s=0}^{\infty} (1+r)^{-(s+1)+s} \beta^s E_t - \sum_{s=0}^{\infty} (1+r)^{-(s+1)} g^s w_t.$$

And using that the geometric sums are $\sum_{s=0}^{\infty} \beta^s = (1-\beta)^{-1}$ and $\sum_{s=0}^{\infty} (1+r)^{-(s+1)} = (1+r-g)^{-1}$:

$$a_t = \frac{E_t}{(1+r)(1-\beta)} - \frac{w_t}{1+r-g} \implies E_t = \frac{(1+r)(1-\beta)}{1+r-g}\left[(1+r-g)a_t + w_t\right], \tag{2.31}$$

In our stationary equilibrium, the economy is not growing in the aggregate, so $g = 0$ and E_t is constant and equal for all individuals as the initial assets is equal for all individuals. From the Euler equation, this implies that the return to assets r that clears the asset market is such

[17] The result that the consumer choice of expenditure just depends on income and not on the price index is true also out of a stationary equilibrium.

that $1 = \beta(1+r) \implies r = \frac{1-\beta}{\beta}$. So, consumers spend all their income $E_t = ra_t + w_t$, and a_t is constant.

Since, within an industry, the individual is characterised by the awareness set, from now on I use the subindex $\mathcal{I}$, instead of ℓ. The share of expenditure of each consumer on each good they know is $s_{\mathcal{I},j} = \frac{p_j c_{\mathcal{I},j}}{E} = P_{\mathcal{I}}^{\sigma-1} p_j^{1-\sigma} \omega_j^{\sigma} = p_j^{1-\sigma} \omega_j^{\sigma} \left[p_{0,t}^{1-\sigma} + \sum_{k\in\mathcal{I}_\ell} \omega_{\ell,k}^{\sigma} p_k^{1-\sigma}\right]^{-1}$; so, using the definition of markup $\mathcal{M}_j = \frac{p_j A_j}{w}$:

$$s_{\mathcal{I},j} = \left[\left(\frac{1}{\mathcal{M}_j A_0}\right)^{1-\sigma} \left(\frac{1}{\omega_j}\right)^{\sigma} + \sum_{k\in\mathcal{I}_\ell} \left(\frac{\mathcal{M}_k A_j}{\mathcal{M}_j A_k}\right)^{1-\sigma} \left(\frac{\omega_k}{\omega_j}\right)^{\sigma}\right]^{-1}.$$

Next, the choice of media time is straightforward: $\frac{\partial L_{\ell,t}}{\partial T_{\ell,t}} = \upsilon\left(Q_t - T_{\ell,t}\right)$, so $T_t = Q_t$. And so, optimal leisure as a function of Q is: $L_t^* = \upsilon \frac{Q_t^2}{2}$.

2.7.3.1 The effect of learning about another good on $s_{\mathcal{I},j,i}$ and the demand elasticity

Proposition 2 *If $j \in \mathcal{I} \subset \mathcal{I}'$, then:*

1. $s_{\mathcal{I},j,i} > s_{\mathcal{I}',j,i}$.

2. $|\epsilon_{\mathcal{I},j,i}| < |\epsilon_{\mathcal{I}',j,i}|$, *where* $\epsilon_{\mathcal{I},j,i} = \frac{p_{j,i}}{c_{\mathcal{I},j,i}} \frac{\partial c_{\mathcal{I},j,i}}{\partial p_{j,i}}$.

Proof. If $j \in \mathcal{I} \subset \mathcal{I}'$, then, since $\omega_{k,i}, \mathcal{M}_{k,i} > 0$ for all firm k and $\sigma > 1$:

$$\begin{aligned} s_{\mathcal{I},j,i} &= \left[(A_0\mathcal{M}_{j,i})^{\sigma-1}\omega_{j,i}^{-\sigma} + \sum_{k\in\mathcal{I}} \left(\frac{\omega_{k,i}}{\omega_{j,i}}\right)^{\sigma} \left(\frac{\mathcal{M}_{j,i}}{\mathcal{M}_{k,i}}\right)^{\sigma-1}\right]^{-1} \\ &> \left[(A_0\mathcal{M}_{j,i})^{\sigma-1}\omega_{j,i}^{-\sigma} + \sum_{k\in\mathcal{I}} \left(\frac{\omega_{k,i}}{\omega_{j,i}}\right)^{\sigma} \left(\frac{\mathcal{M}_{j,i}}{\mathcal{M}_{k,i}}\right)^{\sigma-1} + \sum_{k\in\mathcal{I}'\setminus\mathcal{I}} \left(\frac{\omega_{k,i}}{\omega_{j,i}}\right)^{\sigma} \left(\frac{\mathcal{M}_{j,i}}{\mathcal{M}_{k,i}}\right)^{\sigma-1}\right]^{-1} = s_{\mathcal{I}',j,i}. \end{aligned}$$

For 2, define $\epsilon_{\mathcal{I},j,i} = \frac{p_{j,i}}{c_{\mathcal{I},j,i}} \frac{\partial c_{\mathcal{I},j,i}}{\partial p_{j,i}}$, and note that $\frac{\partial s_{\mathcal{I},j,i}}{\partial \mathcal{M}_{j,i}} = \frac{\partial s_{\mathcal{I},j,i}}{\partial p_{j,i}} \frac{\partial p_{j,i}}{\partial \mathcal{M}_{j,i}} = \frac{1}{E}\left[c_{\mathcal{I},j,i} + p_{j,i}\frac{\partial c_{\mathcal{I},j,i}}{\partial p_{j,i}}\right] w = \frac{s_{\mathcal{I},j,i}}{\mathcal{M}_{j,i}}(1 + \epsilon_{\mathcal{I},j,i})$. And substituting $\frac{\partial s_{\mathcal{I},j,i}}{\partial \mathcal{M}_{j,i}}$, we have:

$$-\frac{s_{\mathcal{I},j,i}}{\mathcal{M}_{j,i}}(\sigma-1)(1-s_{\mathcal{I},j,i}) = \frac{s_{\mathcal{I},j,i}}{\mathcal{M}_{j,i}}(1+\epsilon_{\mathcal{I},j,i}) \implies \epsilon_{\mathcal{I},j,i} = -\sigma + s_{\mathcal{I},j,i}(\sigma-1).$$

So, using 1, we have $j \in \mathcal{I} \subset \mathcal{I}'$ $0 > \epsilon_{\mathcal{I},j,i} > \epsilon_{\mathcal{I}',j,i}$. ■

2.7.4 Production Firms

2.7.4.1 Derivation of the optimal markup:

The first-order condition is:

$$0 = \frac{\partial \pi_j}{\partial \mathcal{M}_j} = \frac{s_j}{\mathcal{M}_j^2} + (1 - \mathcal{M}_j^{-1}) \frac{\partial s_j}{\partial \mathcal{M}_j}, \tag{2.32}$$

which, using $\frac{\partial s_j}{\partial \mathcal{M}_j} = \sum_{\mathcal{I} \in \mathcal{P}_j} M(\mathcal{I}) s_{\mathcal{I},j}(1 - s_{\mathcal{I},j}) \frac{\sigma - 1}{\mathcal{M}_j}$, we can rewrite as:

$$\frac{s_j}{\mathcal{M}_j} = (\sigma - 1)(1 - \mathcal{M}_j^{-1}) \sum_{\mathcal{I} \in \mathcal{P}_j} M(\mathcal{I}) s_{\mathcal{I},j}(1 - s_{\mathcal{I},j}). \tag{2.33}$$

And dividing by s_j and defining $\bar{s}_j := \sum_{\mathcal{I} \in \mathcal{P}_j} \frac{M(\mathcal{I}) s_{\mathcal{I},j}}{s_j} s_{\mathcal{I},j}$:

$$\frac{1}{\mathcal{M}_j} \left[\sigma - (\sigma - 1)\bar{s}_j\right] = (\sigma - 1)\left(1 - \bar{s}_j\right) \implies \mathcal{M}_j = \frac{\frac{\sigma}{\sigma - 1} - \bar{s}_j}{1 - \bar{s}_j}. \tag{2.34}$$

2.7.5 Derivative with respect to advertising and Proof of Proposition 1

As a preliminary to proof Proposition 1, we need two things. On the one hand, the derivative of $\rho_{j,i}$ with respect to $e_{k,i}$. $\frac{\partial \rho_k}{\partial e_j} = \psi_c \psi_s T^{\psi_c} \alpha_k^{\psi_c - 1} \frac{\partial \alpha_k}{\partial e_j}$; and we distinguish two cases:

- If $J = 1$ (ad space not binding), then $\frac{\partial \alpha_k}{\partial e_j} = 0$, so $\frac{\partial \rho_k}{\partial e_j} = 0$; and $\frac{\partial \alpha_j}{\partial e_j} = \frac{1}{p_a}$, so: $\frac{\partial \rho_j}{\partial e_j} = \psi_c \psi_s T^{\psi_c} \alpha_j^{\psi_c - 1} \frac{1}{p_a}$
- If $J > 1$ (ad space binding), then $\frac{\partial \alpha_k}{\partial e_j} = -\frac{a_k}{\sum_s e_s}$, so $\frac{\partial \rho_k}{\partial e_j} = -\psi_c \psi_s T^{\psi_c} \alpha_k^{\psi_c - 1} \frac{a_k}{\sum_s e_s}$; and $\frac{\partial \alpha_j}{\partial e_j} = \frac{\alpha}{\sum_s e_s} - \frac{\alpha_j}{\sum_s e_s} = \frac{\sum_{k \neq j} \alpha_k}{\sum_s e_s}$, so: $\frac{\partial \rho_j}{\partial e_j} = \psi_c \psi_s T^{\psi_c} \alpha_j^{\psi_c - 1} \frac{\sum_{k \neq j} \alpha_k}{\sum_s e_s}$.
  Note that since $\alpha_j = \alpha \frac{e_j}{\sum_k e_k}$, we can rewrite them as: $\frac{\partial \rho_k}{\partial e_j} = -\psi_c \psi_s T^{\psi_c} \frac{\alpha^{\psi_c}}{(\sum_s e_s)^{\psi_c + 1}} e_k^{\psi_c - 1} e_k$ and $\frac{\partial \rho_j}{\partial e_j} = \psi_c \psi_s T^{\psi_c} \frac{\alpha^{\psi_c}}{(\sum_s e_s)^{\psi_c + 1}} e_j^{\psi_c - 1} \sum_{k \neq j} e_k$

On the other hand, we need the derivative of a function of the next period vector of masses $\vec{M}'$ (like the firms' continuation value). As a first step:

Lemma 1 *If $f : \hat{\vec{M}}' \to \mathbb{R}$, then we have:*

1. $\frac{\partial f}{\partial \rho_j} = \sum_{\mathcal{I} \in \mathcal{P}_{-j}} M(\mathcal{I}) \sum_{\mathcal{I}' \in \mathcal{P}_{-j}, \mathcal{I}' \supseteq \mathcal{I}} \frac{\Theta(\mathcal{I} \to \mathcal{I}')}{1 - \rho_j} \left[\frac{\partial f}{\partial M'(\mathcal{I}' \cup \{j\})} - \frac{\partial f}{\partial M'(\mathcal{I}')}\right]$

2. *For the anticompetitive motive, it will be useful:*

$$\frac{\partial f}{\partial \rho_j} = \sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}} M(\mathcal{I}) \sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}} \frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j} \left[\left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})} - \frac{\partial f}{\partial M'(\mathcal{I}')}\right] - \left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k,j\})} - \frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k\})}\right]\right]$$
$$+ \sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}} (M(\mathcal{I}) + M(\mathcal{I}\cup\{k\})) \sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}} \frac{\Theta(\mathcal{I}\cup\{k\}\to\mathcal{I}'\cup\{k\})}{1-\rho_j} \left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k,j\})} - \frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k\})}\right]$$

Proof. First, recall that $\hat{M}'(\mathcal{I}') = \sum_{\mathcal{I}\in\mathcal{P}(\mathcal{J})} M(\mathcal{I})\Theta(\mathcal{I}\to\mathcal{I}')$.

Next, the derivatives of $\Theta(\mathcal{I}\to\mathcal{I}')$ wrt ρ_j are:

1. If $\mathcal{I}\not\subseteq\mathcal{I}'$: $\frac{\partial\Theta(\mathcal{I}\to\mathcal{I}')}{\partial\rho_j} = 0$

2. If $\mathcal{I}\subseteq\mathcal{I}'$:

   (a) If $j\in\mathcal{I}$: $\frac{\partial\Theta(\mathcal{I}\to\mathcal{I}')}{\partial\rho_j} = 0$

   (b) If $j\in\mathcal{I}'\setminus\mathcal{I}$: $\frac{\partial\Theta(\mathcal{I}\to\mathcal{I}')}{\partial\rho_j} = (1-\delta)\prod_{k\in\mathcal{I}'\setminus(\mathcal{I}\cup\{j\})}\rho_k \prod_{k\notin\mathcal{I}'}(1-\rho_k) = \frac{\Theta(\mathcal{I}\to\mathcal{I}'\setminus\{j\})}{1-\rho_j}$

   (c) If $j\notin\mathcal{I}'$: $\frac{\partial\Theta(\mathcal{I}\to\mathcal{I}')}{\partial\rho_j} = -(1-\delta)\prod_{k\in\mathcal{I}'\setminus\mathcal{I}}\rho_k \prod_{k\notin(\mathcal{I}'\cup\{j\})}(1-\rho_k) = -\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}$

Using this, the derivative of $\hat{M}'(\mathcal{I}')$ wrt ρ_j is:

1. If $j\in\mathcal{I}'$: $\frac{\partial\hat{M}'(\mathcal{I}')}{\partial\rho_j} = \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subset\mathcal{I}'} M(\mathcal{I})\frac{\partial\Theta(\mathcal{I}\to\mathcal{I}')}{\partial\rho_j} = \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subset\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}'\setminus\{j\})}{1-\rho_j}$

2. If $j\notin\mathcal{I}'$: $\frac{\partial\hat{M}'(\mathcal{I}')}{\partial\rho_j} = \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subseteq\mathcal{I}'} M(\mathcal{I})\frac{\partial\Theta(\mathcal{I}\to\mathcal{I}')}{\partial\rho_j} = -\sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subseteq\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}$

And the derivative of a generic function $f:\vec{\hat{M}}'\to\mathbb{R}$ wrt ρ_j is:

$$\frac{\partial f}{\partial\rho_j} = \sum_{\mathcal{I}'\in\mathcal{P}} \frac{\partial f}{\partial M'(\mathcal{I}')}\frac{\partial M'(\mathcal{I}')}{\partial\rho_j} = \sum_{\mathcal{I}'\in\mathcal{P}_j} \frac{\partial f}{\partial M'(\mathcal{I}')}\frac{\partial M'(\mathcal{I}')}{\partial\rho_j} + \sum_{\mathcal{I}'\in\mathcal{P}_{-j}} \frac{\partial f}{\partial M'(\mathcal{I}')}\frac{\partial M'(\mathcal{I}')}{\partial\rho_j}$$
$$= \sum_{\mathcal{I}'\in\mathcal{P}_j} \frac{\partial f}{\partial M'(\mathcal{I}')} \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subset\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}'\setminus\{j\})}{1-\rho_j} - \sum_{\mathcal{I}'\in\mathcal{P}_{-j}} \frac{\partial f}{\partial M'(\mathcal{I}')} \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subseteq\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}$$

Now we are going to merge the two summations using that $\mathcal{P}_j = \{\mathcal{I}\cup\{j\}|\mathcal{I}\in\mathcal{P}_{-j}\}$ [**Proof:** from any set $\mathcal{I}$ that doesn't contain j we can build one by adding j to $\mathcal{I}$ (that is, $\{\mathcal{I}\cup\{j\}|\mathcal{I}\in\mathcal{P}_{-j}\}\subseteq\mathcal{P}_j$), and that from any $\mathcal{I}'$ that contains j we can build another one that doesn't contain j by removing j from $\mathcal{I}'$ (that is, $\mathcal{P}_j = \{(\mathcal{I}'\setminus\{j\})\cup\{j\}|\mathcal{I}'\in\mathcal{P}_j\}\subseteq\{\mathcal{I}\cup\{j\}|\mathcal{I}\in\mathcal{P}_{-j}\}$].

Using this in the previous expression, we get:

$$\frac{\partial f}{\partial\rho_j} = \sum_{\mathcal{I}'\in\mathcal{P}_{-j}} \frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})} \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subseteq\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j} - \sum_{\mathcal{I}'\in\mathcal{P}_{-j}} \frac{\partial f}{\partial M'(\mathcal{I}')} \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subseteq\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}$$
$$= \sum_{\mathcal{I}'\in\mathcal{P}_{-j}} \sum_{\mathcal{I}\in\mathcal{P}_{-j},\mathcal{I}\subseteq\mathcal{I}'} M(\mathcal{I})\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}\left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})} - \frac{\partial f}{\partial M'(\mathcal{I}')}\right]$$
$$= \sum_{\mathcal{I}\in\mathcal{P}_{-j}} M(\mathcal{I}) \sum_{\mathcal{I}'\in\mathcal{P}_{-j},\mathcal{I}'\supseteq\mathcal{I}} \frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}\left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})} - \frac{\partial f}{\partial M'(\mathcal{I}')}\right]$$

where for the last equality, I have used that $\{(\mathcal{I},\mathcal{I}')|\mathcal{I}\in\mathcal{P},\mathcal{I}'\subseteq\mathcal{I}\}=\{(\mathcal{I},\mathcal{I}')|\mathcal{I}'\in\mathcal{P},\mathcal{I}\supseteq\mathcal{I}'\}$.

This proves the first expression of the lemma. For the second:
First, note that above we have shown that $\{\mathcal{I}\cup\{j\}|\mathcal{I}\in\mathcal{P}_{-j}\}=\mathcal{P}_j$, which implies that $\mathcal{P}=\mathcal{P}_{-j}\cup\{\mathcal{I}\cup\{j\}|\mathcal{I}\in\mathcal{P}_{-j}\}$. Analogously, defining $\mathcal{P}_{-k,-j}=\{\mathcal{I}\in\mathcal{P}|j,k\notin\mathcal{I}\}$, we have: $\mathcal{P}_{-j}=\mathcal{P}_{-k,-j}\cup\{\mathcal{I}\cup\{k\}|\mathcal{I}\in\mathcal{P}_{-k,-j}\}$, so the previous expression becomes

$$\begin{aligned}\frac{\partial f}{\partial\rho_j}=\sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}}&M(\mathcal{I})\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}\left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})}-\frac{\partial f}{\partial M'(\mathcal{I}')}\right]\\&+M(\mathcal{I}\cup\{k\})\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-j}\\ \mathcal{I}'\supseteq\mathcal{I}\cup\{k\}}}\frac{\Theta(\mathcal{I}\cup\{k\}\to\mathcal{I}')}{1-\rho_j}\left[\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})}-\frac{\partial f}{\partial M'(\mathcal{I}')}\right]\end{aligned}$$

Now, for the first line, I use the following equivalence: for each $\mathcal{I}\in\mathcal{P}_{-k,-j}$ we have $\{\mathcal{I}'\in\mathcal{P}_{-j}|\mathcal{I}'\supseteq\mathcal{I}\}=\{\mathcal{I}'\in\mathcal{P}_{-k,-j}|\mathcal{I}'\supseteq\mathcal{I}\}\cup\{\mathcal{I}'\cup\{k\}|\mathcal{I}'\in\mathcal{P}_{-k,-j},\mathcal{I}'\supseteq\mathcal{I}\}$. And for the second line, I use the equivalence: for $\mathcal{I}\cup\{k\}$ we have $\{\mathcal{I}'\in\mathcal{P}_{-j}|\mathcal{I}'\supseteq\mathcal{I}\cup\{k\}\}=\{\mathcal{I}'\cup\{k\}|\mathcal{I}'\in\mathcal{P}_{-k,-j},\mathcal{I}'\supseteq\mathcal{I}\}$.

$$\begin{aligned}\frac{\partial f}{\partial\rho_j}=\sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}}&M(\mathcal{I})\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\left[\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}\left(\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})}-\frac{\partial f}{\partial M'(\mathcal{I}')}\right)\right.\\&\left.\qquad+\frac{\Theta(\mathcal{I}\to\mathcal{I}'\cup\{k\})}{1-\rho_j}\left(\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k,j\})}-\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k\})}\right)\right]\\&+M(\mathcal{I}\cup\{k\})\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\frac{\Theta(\mathcal{I}\cup\{k\}\to\mathcal{I}'\cup\{k\})}{1-\rho_j}\left(\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k,j\})}-\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k\})}\right).\end{aligned}$$

Finally, I use that for $\mathcal{I},\mathcal{I}'\in\mathcal{P}_{-k,-j}$ with $\mathcal{I}'\supseteq\mathcal{I}$, we have

$$\Theta(\mathcal{I}\cup\{k\}\to\mathcal{I}'\cup\{k\})=\prod_{h\in\mathcal{I}'\setminus\mathcal{I}}\rho_h\prod_{h\notin\mathcal{I}'\cup\{k\}}(1-\rho_h)\cdot(\rho_k+1-\rho_k)=\Theta(\mathcal{I}\to\mathcal{I}'\cup\{k\})+\Theta(\mathcal{I}\to\mathcal{I}').$$

So, I substitute in the second line $\Theta(\mathcal{I}\to\mathcal{I}'\cup\{k\})=\Theta(\mathcal{I}\cup\{k\}\to\mathcal{I}'\cup\{k\})-\Theta(\mathcal{I}\to\mathcal{I}')$:

$$\begin{aligned}\frac{\partial f}{\partial\rho_j}=\sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}}&M(\mathcal{I})\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}\left[\left(\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{j\})}-\frac{\partial f}{\partial M'(\mathcal{I}')}\right)\right.\\&\left.\qquad-\left(\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k,j\})}-\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k\})}\right)\right]\\&+(M(\mathcal{I})+M(\mathcal{I}\cup\{k\}))\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\frac{\Theta(\mathcal{I}\cup\{k\}\to\mathcal{I}'\cup\{k\})}{1-\rho_j}\left(\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k,j\})}-\frac{\partial f}{\partial M'(\mathcal{I}'\cup\{k\})}\right)\end{aligned}$$

■

With uncertainty:

$$\text{For } \mathcal{I}' \in \mathcal{P}(\mathcal{J}'), \qquad M'(\mathcal{I}') = \begin{cases} \sum_{\{\mathcal{I} \in \mathcal{P}(\mathcal{J}): \mathcal{I} \cap \mathcal{J}' = \mathcal{I}'\}} \hat{M}(\mathcal{I}) & \text{, if } \mathcal{I}' \subseteq \mathcal{J} \\ 0 & \text{, if } \mathcal{I}' \nsubseteq \mathcal{J}, \end{cases} \tag{2.35}$$

where the first case says that two consumers become identical in industry i if all the firms in which they differed exit, whereas the second case says that there are no consumers who are aware of a newborn firm. The last piece of infornation needed to compute expected values is the probabilities that the set of differentiated goods moves from $\mathcal{J}$ to $\mathcal{J}' \subseteq \mathcal{J} \cup \{e\}$, where e denotes an entrant. These probabilities are given by:

$$\text{For } \mathcal{J}' \in \mathcal{P}(\mathcal{J} \cup e)), \qquad Prob\{\mathcal{J} \to \mathcal{J}'\} = \begin{cases} (1 - z_{e,i,t}) \prod_{j \in \mathcal{J} \cap \mathcal{J}'} (1-\kappa) \prod_{j \in \mathcal{J} \setminus \mathcal{J}'} \kappa & \text{, if } e \notin \mathcal{J}' \\ z_{e,i,t} \prod_{j \in \mathcal{J} \cap \mathcal{J}'} (1-\kappa) \prod_{j \in \mathcal{J} \setminus \mathcal{J}'} \kappa & \text{, if } e \in \mathcal{J}'. \end{cases} \tag{2.36}$$

In the model, there is uncertainty on $\mathcal{J}'$, so I am more interested in finding the derivative of the expected value of a function $g : \vec{M}' \to \mathbb{R}$ rather than the derivative of a function $f : \vec{\hat{M}}' \to \mathbb{R}$ (note that f is defined on $\vec{\hat{M}}'$, that is, the next period distribution if there weren't entry and exit, whereas g is defined on the actual next period distribution after the uncertainty has been resolved). Recall that for each $\mathcal{J}' \subseteq \mathcal{J} \cup \{e\}$, the probability of this transition is given by $Prob\{\mathcal{J} \to \mathcal{J}'\}$ defined in 2.36 and the mapping between $\vec{\hat{M}}'$ and $\vec{M}'$ is given by $F_{\mathcal{J},\mathcal{J}'} : \vec{\hat{M}}' \to \vec{M}'$ defined in 2.35:

$$F_{\mathcal{J},\mathcal{J}'}(\vec{\hat{M}}') = \begin{cases} M'(\mathcal{I}) = \sum_{\{\mathcal{I} \in \mathcal{P}(\mathcal{J}): \mathcal{I} \cap \mathcal{J}' = \mathcal{I}'\}} \hat{M}(\mathcal{I}) & \text{, for } \mathcal{I}' \subseteq \mathcal{J} \\ M'(\mathcal{I}) = 0 & \text{, for } \mathcal{I}' \nsubseteq \mathcal{J}. \end{cases}$$

Going in the reverse order, each $\mathcal{I} \in \mathcal{P}(\mathcal{J})$ is associated to $\mathcal{I}' = \mathcal{I} \cap \mathcal{J}' \in \mathcal{P}(\mathcal{J}')$; therefore $\frac{\partial g\left(F_{\mathcal{J},\mathcal{J}'}(\vec{\hat{M}}')\right)}{\partial \hat{M}'(\mathcal{I})} = \frac{\partial g(\vec{M}')}{\partial \hat{M}'(\mathcal{I})} = \frac{\partial g(\vec{M}')}{\partial M'(\mathcal{I} \cap \mathcal{J}')} \frac{\partial M'(\mathcal{I} \cap \mathcal{J}')}{\partial \hat{M}'(\mathcal{I})} = \frac{\partial g(\vec{M}')}{\partial M'(\mathcal{I} \cap \mathcal{J}')}$.

Then, we can apply this to the result of the case without entry and exit and we have:

$$\frac{\partial g\left(F_{\mathcal{J},\mathcal{J}'}(\vec{\hat{M}}')\right)}{\partial \rho_j} = \sum_{\mathcal{I} \in \mathcal{P}_{-j}} M(\mathcal{I}) \sum_{\mathcal{I}' \in \mathcal{P}_{-j}, \mathcal{I}' \supseteq \mathcal{I}} \frac{\Theta(\mathcal{I} \to \mathcal{I}')}{1 - \rho_j} \left[\frac{\partial g}{\partial M'((\mathcal{I}' \cup \{j\}) \cap \mathcal{J}')} - \frac{\partial g}{\partial M'(\mathcal{I}' \cap \mathcal{J}')} \right].$$

Note that if $j \notin \mathcal{J}'$, then this derivative is 0 since in this case $\frac{\partial g}{\partial M'((\mathcal{I}' \cup \{j\}) \cap \mathcal{J}')} = \frac{\partial g}{\partial M'(\mathcal{I}' \cap \mathcal{J}')}$.

With this, the expected value is defined as:

$$\mathbb{E} g\left(\vec{M}\right) = \sum_{\mathcal{J}' \subseteq \mathcal{J} \cup \{e\}} Prob\{\mathcal{J} \to \mathcal{J}'\} g\left(F_{\mathcal{J},\mathcal{J}'}(\vec{\hat{M}}')\right).$$

Proof of the Proposition 1: For the informative motive it is straightforward from applying 1 of Lemma 1, together with the note on the Uncertainty case:

$$\frac{\partial \mathbb{E}V_j(\mathcal{J}',\vec{M}')}{\partial \rho_j}\frac{\partial \rho_j}{\partial e_j} = \Bigg[\sum_{\mathcal{J}'\subseteq \mathcal{J}\cup\{e\}} \text{Prob}\{\mathcal{J}\to\mathcal{J}'\}\sum_{\mathcal{I}\in\mathcal{P}_{-j}} M(\mathcal{I}) \sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-j}\\ \mathcal{I}'\supseteq\mathcal{I}}} \frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_j}$$
$$\cdot\left(\frac{\partial V_j(\mathcal{J}',\vec{M}')}{\partial M'((\mathcal{I}'\cup\{j\})\cap\mathcal{J}')} - \frac{\partial V_j(\mathcal{J}',\vec{M}')}{\partial M'(\mathcal{I}'\cap\mathcal{J}')}\right)\Bigg]\cdot\frac{\partial \rho_j}{\partial e_j} > 0,$$

where the positive comes from the fact that $\frac{\partial V_j(\mathcal{J}',\vec{M}')}{\partial M'((\mathcal{I}'\cup\{j\})\cap\mathcal{J}')} \geq \frac{\partial V_j(\mathcal{J}',\vec{M}')}{\partial M'(\mathcal{I}'\cap\mathcal{J}')}$, since firm j's value increases more if we add a consumer that besides $\mathcal{I}'\cap\mathcal{J}'$ she is also aware of j (there is equality if j has exited in the scenario with $\mathcal{J}'$). For the result that the informative motive decreases if we add $\{j\}$ to some consumers that weren't aware, note that the above expression is a summation over the awareness sets that don't contain j, and the change described implies a reduction of the masses in these sets.

For the anticompetitive motive, I use 2 of 1, which together with the note on the Uncertainty case, implies:

$$\sum_{k\neq j}\left(-\frac{\partial \mathbb{E}V_j(\mathcal{J}',\vec{M}')}{\partial \rho_k}\right)\left(-\frac{\partial \rho_k}{\partial e_j}\right) = \sum_{k\neq j}\sum_{\mathcal{J}'\subseteq\mathcal{J}\cup\{e\}} \text{Prob}\{\mathcal{J}\to\mathcal{J}'\}$$
$$\cdot\Bigg\{\sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}} M(\mathcal{I})\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\frac{\Theta(\mathcal{I}\to\mathcal{I}')}{1-\rho_k}\underbrace{\left[\left(\frac{\partial V_j}{\partial M'(\mathcal{I}')}-\frac{\partial V_j}{\partial M'(\mathcal{I}'\cup\{k\})}\right)-\left(\frac{\partial V_j}{\partial M'(\mathcal{I}'\cup\{j\})}-\frac{\partial V_j}{\partial M'(\mathcal{I}'\cup\{j,k\})}\right)\right]}_{<0}$$
$$+\sum_{\mathcal{I}\in\mathcal{P}_{-k,-j}}(M(\mathcal{I})+M(\mathcal{I}\cup\{j\}))\sum_{\substack{\mathcal{I}'\in\mathcal{P}_{-k,-j}\\ \mathcal{I}'\supseteq\mathcal{I}}}\frac{\Theta(\mathcal{I}\cup\{j\}\to\mathcal{I}'\cup\{j\})}{1-\rho_k}\underbrace{\left[\frac{\partial V_j}{\partial M'(\mathcal{I}'\cup\{j\})}-\frac{\partial V_j}{\partial M'(\mathcal{I}'\cup\{j,k\})}\right]}_{>0}\Bigg\}$$
$$\cdot\left(-\frac{\partial \rho_k}{\partial e_j}\right) > 0,$$

where the intuition for the negative sign of the first underbrace is that the firm value is more affected if it is a customer who learns about another good (since she will reduce the spending in j) rather than if it is a non-customer who learns about another good. And the positive sign in the second underbrace is because the value of a firm decreases if a customer learns about another good. So, we see that the anticompetitive motive increases more if $M(\mathcal{I}\cup\{j\})$ increases rather than $M(\mathcal{I})$ (direct from taking the derivatives with respect to $M(\mathcal{I}\cup\{j\})$ and $M(\mathcal{I})$ in the previous expression. Finally, that the anticompetitive motive is positive follows from observing that $\Theta(\mathcal{I}\to\mathcal{I}') = (1-\rho_j)\Theta(\mathcal{I}\cup\{j\}\to\mathcal{I}'\cup\{j\}) < \Theta(\mathcal{I}\cup\{j\}\to\mathcal{I}'\cup\{j\})$, and then the negative part of the second line is offset by the third line.

The result for the persuasive motive is straightforward from observing that **??** is a summation over the awareness sets that contain j.

2.7.6 Social planner problem

$$\max_{\{N_{M,t},N_{j,i,t},h_{e,i,t},p_{j,i,t},\alpha_{j,i,t}\}} U = \sum_{t=0}^{\infty} \beta^t \int_0^1 [\ln C_{\ell t} + L_{\ell t}]\, d\ell$$

s.t. $C_{\ell,t}$ from (2.5), $C_{\ell,i,t}$ from (2.6), $c_{\ell,j,i,t}$ from (2.11), and $L_{\ell,t}$ from 2.4, with $T_t = Q_t$

$$y_{j,i,t} = N_{j,i,t}, \quad y_{0,i,t} = A_0 N_{0,i,t}, \quad Q_t = AN_{m,t}^{\varphi} \quad \text{(Production functions)}$$

$$1 = N_{m,t} + \int_0^1 \left(\sum_{j\in\{0\}\cup\mathcal{J}_{i,t}} N_{j,i,t} + N_{e,i,t} \right) di \quad , \quad w_t = E_t \quad \text{(Resource constraints)}$$

$$\sum_{j\in\mathcal{J}_{i,t}} \alpha_{j,i,t} = \alpha_i, (2.9), (2.10), (2.14), (2.35), (2.36) \quad \text{(Learning process)}$$

$$z_{e,i,t} = \phi N_{e,i,t}^{\frac{1}{2}}, (2.35), (2.36) \quad \text{(Entry and exit)}$$

Plugging $C_{\ell,t}$ and $L_{\ell,t}$ with $T_t = Q_t$ into the objective function and interchanging the integrals over ℓ and i:

$$\max_{\{N_{M,t},N_{e,i,t},N_{j,i,t},p_{j,i,t},\alpha_{j,i,t}\}} U = \int_0^1 \int_0^1 \sum_{t=0}^{\infty} \beta^t \ln C_{\ell,i,t} d\ell di + \sum_{t=0}^{\infty} \beta^t \upsilon \frac{Q_t^2}{2}$$

s.t. $C_{\ell,i,t}$ from (2.6), $c_{\ell,j,i,t}$ from (2.11)

$$y_{j,i,t} = N_{j,i,t}, \quad y_{0,i,t} = A_0 N_{0,i,t}, \quad Q_t = AN_{m,t}^{\varphi} \quad \text{(Production functions)}$$

$$1 = N_{m,t} + \int_0^1 \left(\sum_{j\in\{0\}\cup\mathcal{J}_{i,t}} N_{j,i,t} + N_{e,i,t} \right) di \quad , \quad w_t = E_t \quad \text{(Resource constraints)}$$

$$\sum_{j\in\mathcal{J}_{i,t}} \alpha_{j,i,t} = \alpha_i, (2.9), (2.10), (2.14), (2.35), (2.36) \quad \text{(Learning process)}$$

$$z_{e,i,t} = \phi N_{e,i,t}^{\frac{1}{2}}, (2.35), (2.36) \quad \text{(Entry and exit)}$$

The planner decides how much to produce for each individual and, accordingly, sets the prices that induce the consumers to consume these quantities. Let N_t^P be the labor used to produce all the goods in the production sector. Then, The FOC for $c_{\mathcal{I},j,i,t}$ reads:

$$[c_{\mathcal{I},j,i,t}]: \quad \beta^t \frac{1}{C_t} \frac{\partial C_t}{\partial C_{i,t}} \frac{\partial C_{i,t}}{\partial C_{\mathcal{I},i,t}} \frac{\partial C_{\mathcal{I},i,t}}{\partial c_{\mathcal{I},j,i,t}} = \beta^t \lambda \frac{\partial N_t^P}{\partial C_t} \frac{\partial C_t}{\partial C_{i,t}} \frac{\partial C_{i,t}}{\partial C_{\mathcal{I},i,t}} \frac{\partial C_{\mathcal{I},i,t}}{\partial c_{\mathcal{I},j,i,t}} \tag{2.37}$$

1. Dividing both sides $\frac{\partial C_t}{\partial C_{i,t}} \frac{\partial C_{i,t}}{\partial C_{\mathcal{I},i,t}} \frac{\partial C_{\mathcal{I},i,t}}{\partial c_{\mathcal{I},j,i,t}} > 0$, and defining $\hat{P}_t = \frac{w_t N_t^P}{C_t}$, we get:

$$\frac{1}{C_t} = \lambda \frac{\hat{P}_t}{w_t} \implies \lambda = \frac{w_t}{\hat{P}_t C_t} \tag{2.38}$$

2. $\ln C_t = \int_0^1 \ln C_{i,t} di$. Dividing both sides of 2.37 by $\frac{\partial C_{i,t}}{\partial C_{\mathcal{I},i,t}} \frac{\partial C_{\mathcal{I},i,t}}{\partial c_{\mathcal{I},j,i,t}} > 0$, letting $N_{i,t}$ be the

the labor used in sector i, $\frac{\partial N_t^P}{\partial C_t}\frac{\partial C_t}{\partial C_{i,t}} = \frac{\partial N_t^P}{\partial C_{i,t}}$ and defining $\hat{P}_{i,t} = \frac{w_t N_{i,t}}{C_{i,t}}$, we get:

$$\frac{1}{C_{i,t}} = \lambda \frac{\hat{P}_{i,t}}{w_t} \implies \lambda = \frac{w_t}{\hat{P}_{i,t} C_{i,t}} \implies C_{i,t} = C_t \frac{\hat{P}_t}{\hat{P}_{i,t}} \tag{2.39}$$

where for the last expression I have used 2.38. Plugging $C_{i,t}$ into the definition of C_t, we get:

$$\ln \hat{P}_t = \int_0^1 \ln \hat{P}_{i,t} di \tag{2.40}$$

3. $\ln C_{i,t} = \int_0^1 \ln C_{\ell,i,t} d\ell$. Dividing both sides of 2.37 by $\frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} > 0$, letting $N_{\ell,i,t}$ be the the labor used in sector i by ℓ, $\frac{\partial N_{i,t}}{\partial C_{i,t}}\frac{\partial C_{i,t}}{\partial C_{\ell,i,t}} = \frac{\partial N_{i,t}}{\partial C_{\ell,i,t}}$ and defining $\hat{P}_{\ell,i,t} = \frac{w_t N_{\ell,i,t}}{C_{\ell,i,t}}$, we get:

$$\frac{1}{C_{\ell,i,t}} = \lambda \frac{\hat{P}_{\ell,i,t}}{w_t} \implies \lambda = \frac{w_t}{\hat{P}_{\ell,i,t} C_{\ell,i,t}} \implies C_{\ell,i,t} = C_{i,t} \frac{\hat{P}_{i,t}}{\hat{P}_{\ell,i,t}} \tag{2.41}$$

where for the last expression I have used 2.39. Plugging $C_{\ell,i,t}$ into the definition of $C_{i,t}$, we get:

$$\ln \hat{P}_{i,t} = \int_0^1 \ln \hat{P}_{\ell,i,t} di \tag{2.42}$$

4. $C_{\ell,i,t}$ given by 2.6. Letting $N_{\ell,j,i,t}$ be the the labor used in good j in sector i by ℓ, $\frac{\partial N_{\ell,i,t}}{\partial C_{\ell,i,t}}\frac{\partial C_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = \frac{\partial N_{\ell,i,t}}{\partial c_{\ell,j,i,t}} = \frac{1}{A_j}$, we get:

$$\frac{1}{C_{\ell,i,t}} \left(\frac{C_{\ell,i,t}}{c_{\ell,j,i,t}}\right)^{\frac{1}{\sigma}} \omega_{j,i,t} = \lambda \frac{1}{A_j} \implies \left(\frac{C_{\ell,i,t}}{c_{\ell,j,i,t}}\right)^{\frac{1}{\sigma}} \omega_{j,i,t} = \frac{w_t}{\hat{P}_{\ell,i,t} A_j} \implies c_{\ell,j,i,t} = C_{\ell,i,t} \hat{P}^{\sigma}_{\mathcal{I},i,t} \left(\omega_{j,i,t} \frac{A_j}{w_t}\right)^{\sigma} \tag{2.43}$$

where I have used λ from 2.41. Plugging $c_{\ell,j,i,t}$ into the definition of $C_{\ell,i,t}$, we get:

$$\hat{P}_{\mathcal{I},i,t} = \left(\left(\frac{A_0}{w_t}\right)^{\sigma-1} + \sum_{j\in\mathcal{I}} \omega^{\sigma}_{j,i,t} \left(\frac{1}{w_t}\right)^{\sigma-1}\right)^{\frac{1}{1-\sigma}} \tag{2.44}$$

Since we have $\hat{P}_{\ell,i,t} C_{\ell,i,t} = \hat{P}_{i,t} C_{i,t} = \hat{P}_t C_t = w_t N_t^P$ and λ from 2.38, then we have:

$$c_{\mathcal{I},j,i,t} = w_t N_t^P \hat{P}^{\sigma-1}_{\mathcal{I},i,t} \left(\omega_{j,i,t} \frac{A_j}{w_t}\right)^{\sigma}, \quad N_t^P = \frac{1}{\lambda} \tag{2.45}$$

Comparing this with the consumer choices:

$$c_{\mathcal{I},j,i,t} = E_t P^{\sigma-1}_{\mathcal{I},i,t} p^{-\sigma}_{j,i,t} \omega^{\sigma}_{j,i,t}, \quad P_{\mathcal{I},i,t} = \left(p^{1-\sigma}_{0,i,t} + \sum_{j\in\mathcal{I}} \omega^{\sigma}_{j,i,t} p^{1-\sigma}_{j,i,t}\right)^{\frac{1}{1-\sigma}}$$

It is straightforward to check that the planner can induce the consumer to consume the quantities in 2.45 by setting prices equal to the marginal cost times a markup (or a tax) equal to the ratio of expenditure to the production costs; i.e. $p_{j,i,t} = \frac{w_t}{A_j}\tau_t$, where $\tau_t = \frac{E_t}{w_t N_t^P}$.

The particular level of τ affects the level of consumption, but not the share of expenditure allocated to each good, since, as seen in the following expression, $s_{\mathcal{I},j,i,t}$ is independent of τ (that is, τ doesn't distort how N_t^P is allocated among the production goods):

$$s_{\mathcal{I},j,i,t} = \frac{p_{j,i,t}c_{\mathcal{I},j,i,t}}{E_{\mathcal{I},i,t}} = \frac{\tau\frac{w_t}{A_j}c_{\mathcal{I},j,i,t}}{\tau w_t N_{\mathcal{I},i,t}} = \omega_{j,i,t}^{\sigma}\left(\frac{w_t}{A_j\hat{P}_{\mathcal{I},i,t}}\right)^{1-\sigma} \implies s_{\mathcal{I},j,i} = \frac{\omega_{j,i}^{\sigma}}{A_0^{\sigma-1} + \sum_{k\in\mathcal{I}}\omega_{k,i}^{\sigma}} \tag{2.46}$$

Using that $c_{\mathcal{I},j,i} = A_j N_{\mathcal{I},j,i} = A_j\frac{N_{\mathcal{I},j,i}}{N_{\mathcal{I},i}}N_{\mathcal{I},i} = A_j s_{\mathcal{I},j,i}\frac{E_t}{w_t}\frac{1}{\tau}$; then, we can write $C_{\mathcal{I},i}$ as:

$$C_{\mathcal{I},i,t} = \frac{1}{\tau}\frac{E_t}{w_t}\left((A_0 s_{\mathcal{I},0,i,t})^{\frac{\sigma-1}{\sigma}} + \sum_{k\in\mathcal{I}}\omega_{k,i,t}(s_{\mathcal{I},k,i,t})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{2.47}$$

and combining this with $\frac{E_t}{\tau} = w_t N_t^P = C_{\mathcal{I},i,t}\hat{P}_{\mathcal{I},i,t}$, we get (where for the second equality I use 2.46):

$$\left((A_0 s_{\mathcal{I},0,i,t})^{\frac{\sigma-1}{\sigma}} + \sum_{k\in\mathcal{I}}\omega_{k,i,t}(s_{\mathcal{I},k,i,t})^{\frac{\sigma-1}{\sigma}}\right) = \left(\frac{w_t}{\hat{P}_{\mathcal{I},i,t}}\right)^{\frac{\sigma-1}{\sigma}} = \omega_{j,i,t}s_{\mathcal{I},j,i,t}^{-\frac{1}{\sigma}} \tag{2.48}$$

Next, I move to the advertising part of the planner problem. We will use the following derivatives:

$$\frac{\partial C_{\mathcal{I},i,t}}{\partial\omega_{j,i}} = \frac{\sigma}{\sigma-1}C_{\mathcal{I},i,t}^{\frac{1}{\sigma}}\left[s_{\mathcal{I},j,i,t}^{\frac{\sigma-1}{\sigma}} + \frac{\sigma-1}{\sigma}\left(\sum_{k\in\mathcal{I}}\omega_{k,i,t}s_{\mathcal{I},k,i,t}^{\frac{-1}{\sigma}}\frac{\partial s_{\mathcal{I},k,i,t}}{\partial\omega_{j,i}} + A_0^{\frac{\sigma-1}{\sigma}}s_{\mathcal{I},0,i,t}^{\frac{-1}{\sigma}}\frac{\partial s_{\mathcal{I},0,i,t}}{\partial\omega_{j,i}}\right)\right]$$

$$\frac{\partial s_{\mathcal{I},j,i,t}}{\partial\omega_{j,i}} = s_{\mathcal{I},j,i,t}(1-s_{\mathcal{I},j,i,t})\frac{\sigma}{\omega_{j,i,t}}, \qquad \frac{\partial s_{\mathcal{I},k,i,t}}{\partial\omega_{j,i}} = -s_{\mathcal{I},j,i,t}s_{\mathcal{I},k,i,t}\frac{\sigma}{\omega_{j,i,t}}$$

$$\frac{\partial\omega_{j,i}}{\partial\alpha_{j,i,t}} = \nu_c\nu_s T^{\nu_c}\alpha_{j,i,t}^{\nu_c-1} = \frac{T}{\alpha_{j,i,t}}\frac{\partial\omega_{j,i}}{\partial T}$$

The term in parenthesis of the first line can be rewritten as (in the second expression, I use 2.48):

$$\left(\omega_{j,i,t} - \sum_{k\in\mathcal{I}}\omega_{k,i,t}s_{\mathcal{I},k,i,t}^{\frac{\sigma-1}{\sigma}} - A_0^{\frac{\sigma-1}{\sigma}}s_{\mathcal{I},0,i,t}^{\frac{\sigma-1}{\sigma}}\right)s_{\mathcal{I},j,i,t}\frac{\sigma}{\omega_{j,i,t}} = \left(1 - s_{\mathcal{I},j,i,t}^{-\frac{1}{\sigma}}\right)s_{\mathcal{I},j,i,t}\sigma < 0$$

so, the term in the parenthesis is negative. And we have:

$$\frac{\partial\ln C_{\mathcal{I},i,t}}{\partial\omega_{j,i}} = \frac{\sigma}{\sigma-1}C_{\mathcal{I},i,t}^{\frac{1-\sigma}{\sigma}}\left[s_{\mathcal{I},j,i,t}^{\frac{\sigma-1}{\sigma}} + (\sigma-1)\left(1 - s_{\mathcal{I},j,i,t}^{-\frac{1}{\sigma}}\right)s_{\mathcal{I},j,i,t}\right] = \frac{\sigma}{\sigma-1}\left(\frac{s_{\mathcal{I},j,i,t}}{C_{\mathcal{I},i,t}}\right)^{\frac{\sigma-1}{\sigma}}\left[1 + (\sigma-1)\left(s_{\mathcal{I},j,i,t}^{\frac{1}{\sigma}} - 1\right)\right]$$

So:

$$\frac{\partial\ln C_{\mathcal{I},i,t}}{\partial\alpha_{j,i}} = \left(\frac{s_{\mathcal{I},j,i,t}}{C_{\mathcal{I},i,t}}\right)^{\frac{\sigma-1}{\sigma}}\left[\frac{\sigma}{\sigma-1} + \sigma\left(s_{\mathcal{I},j,i,t}^{\frac{1}{\sigma}} - 1\right)\right]\nu_c\nu_s T^{\nu_c}\alpha_{j,i,t}^{\nu_c-1}$$

For the dynamic problem of advertising/media, it is useful to define $U_X = \int_0^1 \sum_{t=0}^{\infty} \beta^t \ln C_{\ell,i,t} d\ell$ as the expected life-time industry-consumption utility of an industry with the current industry state being X.

The social planner has to decide on (i) how much labor to allocate to the media sector, $N_{m,t}$, and (ii) how to allocate the ad space among the differentiated firms of each industry, $\alpha_{j,i,t}$.

First, let's see the social planner choice of $\alpha_{j,i,t}$. The allocation of the ad space has to be such that the marginal social gain of increasing the ad space given to each firm is the same, since otherwise we could improve the allocation. Formally, it must be $\beta \frac{\partial \mathbb{E} U_{X'}}{\partial \rho_{j,i}} \frac{\partial \rho_{j,i}}{\partial \alpha_{j,i}} + \frac{\partial \ln C_{X,t}}{\partial \alpha_{j,X}} = \hat{h}_X$ for some $\hat{h}_X$ and all $j \in \mathcal{J}_X$, together with $\sum_{j \in \mathcal{J}_X} \alpha_{j,X} = \alpha_X$.

Second, let's see the social planner choice of N_m.

$$\left[\frac{\partial L}{\partial Q} + \sum_{X \in \Omega} \mu_t(X) \sum_{j \in \mathcal{J}_X} \left[\beta \frac{\partial \mathbb{E} U_{X'}}{\partial \rho_{j,X}} \frac{\partial \rho_{j,X}}{\partial T} + \frac{\partial \ln C_{X,t}}{\partial \omega_{j,X}} \frac{\partial \omega_{j,X}}{\partial T}\right] \frac{\partial T}{\partial Q}\right] \frac{\partial Q}{\partial N_m} = \lambda$$

where $\frac{\partial L}{\partial \tilde{Q}} = \upsilon Q$, $\frac{\partial T}{\partial Q} = 1$ (if $Q < 1$, otherwise it is 0). Also, using that $\frac{\partial \rho_{j,X}}{\partial T} = \frac{\alpha_{j,X}}{T} \frac{\partial \rho_{j,X}}{\partial \alpha_{j,X}}$, $\frac{\partial \omega_{j,X}}{\partial T} = \frac{\alpha_{j,X}}{T} \frac{\partial \omega_{j,X}}{\partial \alpha_{j,X}}$, and $\frac{\partial Q}{\partial N_m} = \varphi \frac{Q}{N_m}$

$$\left[\upsilon Q + \sum_{X \in \Omega} \mu_t(X) \sum_{j \in \mathcal{J}_X} \left[\beta \frac{\partial \mathbb{E} U_{X'}}{\partial \rho_{j,X}} \frac{\partial \rho_{j,X}}{\partial \alpha_{j,X}} + \frac{\partial \ln C_{X,t}}{\partial \alpha_{j,X}}\right] \frac{\alpha_{j,X}}{T}\right] \varphi \frac{Q}{N_m} = \lambda$$

and using that $\frac{\partial \mathbb{E} U_{X'}}{\partial \rho_{j,X}} \frac{\partial \rho_{j,X}}{\partial \alpha_{j,X}} + \frac{\partial \ln C_{X,t}}{\partial \alpha_{j,X}} = \hat{h}_X$ for some value $\hat{h}_X$ and all j, that $\sum_j \alpha_{j,X} = \alpha_X$, and $Q = T$, then the condition for N_m reads:

$$\upsilon Q^2 + \sum_{X \in \Omega} \mu(X) \hat{h}_X \alpha_X = \frac{\lambda}{\varphi} N_m \tag{2.49}$$

Finally, the labor employed in entry in each industry satisfies:

$$\lambda = \frac{\phi}{2} N_{e,X}^{-\frac{1}{2}} \beta \left(\mathbb{E}_e U_{X'} - \mathbb{E}_{-e} U_{X'}\right) \tag{2.50}$$

where $\mathbb{E}_e U_{X'}$ (resp. $\mathbb{E}_{-e} U_{X'}$) is the expected industry-utility conditional on successfully creating (resp. not creating) a new differentiated good (so the expectation comes from the probabilities the incumbents exit).

Using 2.45, 2.50 and 2.49, the labor market clearing condition reads:

$$1 = N^P + N_e + N_m \implies \lambda = 1 + \upsilon Q^2 \varphi + \sum_{X \in \Omega} \mu(X) \left(\varphi \hat{h}_X \alpha_X + \left(\frac{\phi}{2} \beta \left(\mathbb{E}_e U_{X'} - \mathbb{E}_{-e} U_{X'}\right)\right)^2 \lambda^{-1}\right) \tag{2.51}$$

Note that this clearly implies $\lambda > 1$. Finally, the budget constraint implies the relative wage is 1, $\hat{w} = \frac{w}{E} = 1$. Therefore the planner's markup (or tax) is $\tau = \frac{1}{N^P \hat{w}} = \frac{1}{N^P} = \lambda > 1$.

2.7.7 Proof of convergence to an ergodic distribution and uniqueness

Uniqueness:

Let τ be the first period that we arrive at state $\mathcal{J} = \emptyset$, and $P_{t,0}(X)$ be he probability that we are at X after t periods starting from $\mathcal{J} = \emptyset$; then the probability we are at state X starting from a given state is:

$P_t\{X\} = \sum\limits_{k=1}^{t} P\{\tau = k\}P_{t-k,0}\{X\} + P\{\tau > t\}P_t\{X|\tau > K\}$

As $t \to \infty$, $P\{\tau > t\} \to 0$ since every period there is a positive probability that all differentiated firms die and we arrive at $\mathcal{J} = \emptyset$. Therefore, this tells us that if $P_{t,0}\{X\}$ converges (which later I prove that this is the case), then, the only stationary distribution we can have is $P_0(X) = \lim_{t\to\infty} P_{t,0}(X)$.

The set of possible states is at most countably infinite

This is a consequence of two things: (i) from a given state you can directly move to a finite number of states; (ii) with probability 1 any industry will pass through the state $\mathcal{J} = \emptyset$ at some point in time. Just as in the proof of Uniqueness, (ii) is telling us that the only stationary distribution we can have (if any, since I haven't proved this yet) is the one we would converge to starting from the state $\mathcal{J} = \emptyset$, which (i) tells us that at most will have a countably infinite number of different states.

Convergence (Existence)

Suppose there are $n \in \mathbb{N} \cup \{\infty\}$ possible states and the probability of moving from state j to state i is $a_{i,j}$, then the transition matrix is

$$Q = \begin{pmatrix} 1 - \sum_{j=2}^{n} a_{j,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & 1 - \sum_{j\neq 2} a_{j,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & 1 - \sum_{j=1}^{n} a_{n,n} \end{pmatrix}$$

Let $m_t = (m_{1,t}, \ldots, m_{n,t})$ be the vector of masses in each state, and call $M_t := m_{t+1} - m_t = (Q - \mathbb{I}_n)m_t$; so $M_{i,t} = \sum_{k\neq i} m_{k,t}a_{i,k} - m_{i,t}\sum_{k\neq i} a_{k,i}$.

Lemma 2 $\sum_{k=1}^{n} M_{k,t} = 0$

Proof. Given that $M_{i,t} = \sum_{k\neq i} m_{k,t}a_{i,k} - m_{i,t}\sum_{k\neq i} a_{k,i}$; then

$\sum_{i=1}^{n} M_{i,t} = \sum_{i=1}^{n}\left[\sum_{k\neq i} m_{k,t}a_{i,k} - m_{i,t}\sum_{k\neq i} a_{k,i}\right] = \sum_{i=1}^{n}\sum_{k\neq i} m_{k,t}a_{i,k} - \sum_{i=1}^{n}\sum_{k\neq i} m_{i,t}a_{k,i}$

$= \sum_{i=1}^{n}\sum_{k\neq i} m_{k,t}a_{i,k} - \sum_{k=1}^{n}\sum_{i\neq k} m_{k,t}a_{i,k}$.

So, we just need to see that $\{k \neq i | i, k \in \{1, \ldots, n\}\} = \{i \neq k | i, k \in \{1, \ldots, n\}\}$, which is clearly satisfied by symmetry of the $\neq$-relationship.

■

And the following lemma expresses M_{t+q} for $q \in \mathbb{N}$ in terms of M_t:

Lemma 3 *For any* $q \in \mathbb{N}$, $M_{t+q} = Q^q M_t$, *with* $M_{i,t+q} = \left(1 - \sum_{k\neq i} a_{k,i}^{(q)}\right) M_{i,t} + \sum_{k\neq i} a_{i,k}^{(q)} M_{k,t}$, *where* $a_{i,k}^{(q)}$ *is the probability of moving from* k *to* i *in* q *periods.*

Proof. By definition, $M_{t+q} = (Q - \mathbb{I}_n)m_{t+q} = (Q - \mathbb{I}_n)Q^q m_t = (Q^{q+1} - Q^q)m_t = Q^q(Q - \mathbb{I}_n)m_t = Q^q M_t$.

■

My main goal here is to study the convergence of M_t towards the null vector; so, we want to establish some result that compares M_t to M_{t+q} for some $q \in \mathbb{N}$. Since M_t is an n-dimensional object, it is important to specify under which metric. To see the importance of this, let's see a counterexample that shows that not necessarily each component of M_t has to monotonically decrease in absolute value:

Lemma 4 *It is not necessarily true that* $|m_{t+1}(k) - m_t(k)| \geq |m_{t+2}(k) - m_{t+1}(k)|$ *for all* k.

Proof. Suppose that m_t is only non-zero in position i, where $m_{i,t} = 1$. Then

$$M_t = (Q - \mathbb{I}_n)m_t = \begin{pmatrix} a_{1,i} & \cdots & -\sum\limits_{k\neq i} a_{k,i} & \cdots & a_{n,i} \end{pmatrix}^t, \quad m_{t+2} - m_{t+1} = \begin{pmatrix} B_1 & \cdots & B_i & \cdots & B_n \end{pmatrix}^t,$$

where

$$B_j = a_{j,1}\left(1 - \sum_{k\neq i} a_{k,i} - \sum_{k\neq j} a_{k,j}\right) + \sum_{k\notin\{i,j\}} a_{j,k}a_{k,1} \text{ for } j \neq i,$$

$$B_i = -\left(\sum_{k\neq i} a_{k,i}\right)\left(1 - \sum_{k\neq i} a_{k,i}\right) + \sum_{k\neq i} a_{i,k}a_{k,i}.$$

Then, we can find a counterexample by just supposing $a_{1,i} = 0$ and that there exists $k \notin \{1, i\}$ such that $a_{1,k}a_{k,i} > 0$. Then, we have: $|m_{1,t+1} - m_{1,t}| = a_{1,i} = 0 < a_{1,k}a_{k,i} \leq \sum_{k\notin\{1,i\}} a_{1,k}a_{k,i} =$

$|B_1| = |m_{1,t+2} - m_{1,t+1}|$

■

The norm that will prove useful is $||M_t|| := \max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} M_{k,t}|\}$. Define $\mathcal{B}^+ := \{i \in \{1,\dots,n\} | M_{i,t} > 0\}$ and $\mathcal{B}^- := \{i \in \{1,\dots,n\} | M_{i,t} < 0\}$

Proposition 3 *It is satisfied that* $\max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} M_{k,t+1}|\} \leq \max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} M_{k,t}|\} = \sum_{k\in\mathcal{B}^+} M_{k,t}$ *Further, if* $i \in \mathcal{B}^+$, $j \in \mathcal{B}^-$ *and there exists* $\ell \in \{1,\dots,n\}$ *such that* $a^{(q)}_{\ell,i}$, $a^{(q)}_{\ell,j} > 0$ *(*ℓ *can be equal to* i *or* j*, which means that this state has period smaller or equal than* q*), then* $\max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} M_{k,t+q}|\} < \max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} M_{k,t}|\} = \sum_{k\in\mathcal{B}^+} M_{k,t}$

Proof. First, $\max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} M_{k,t}|\} = \sum_{k\in\mathcal{B}^+} M_{k,t}$ since

$$\max_{\mathcal{A}\subset\{1,\dots,n\}} \{|\sum_{k\in\mathcal{A}} [m_{t+1}(k) - m_t(k)]\,|\} = \max\{\sum_{k\in\mathcal{B}^+} M_{k,t}, -\sum_{k\notin\mathcal{B}^+} M_{k,t}\},$$

and, by Lemma 2, both terms have the same value . Next, from Lemma 3, for any given $q \in \mathbb{N}$ and $\mathcal{A} \subset \{1,\dots,n\}$, we have:

$$\sum_{k\in\mathcal{A}} M_{k,t+q} = \sum_{k\in\mathcal{A}} \left[(1 - \sum_{j\neq k} a^{(q)}_{j,k}) M_{k,t} + \sum_{j\neq k} a^{(q)}_{k,j} M_{j,t}\right].$$

And I group the terms with same $M_{i,t}$. Let's focus first on the positive terms (i.e. $M_{i,t}$ for $i \in \mathcal{B}^+$):

- If $i \in \mathcal{B}^+ \cap \mathcal{A}$, then: (i) for $k = i$ we have the term $(1 - \sum_{j\neq i} a^{(q)}_{j,i}) M_{i,t}$; (ii) for each $k \in \mathcal{A}\backslash\{i\}$, we have the term $a^{(q)}_{k,i} M_{i,t}$.

- If $i \in \mathcal{B}^+ \setminus \mathcal{A}$: for each $k \in \mathcal{A}$, we have the term $a^{(q)}_{k,i} M_{i,t}$.

Then, the positive terms can be written as:

$$\begin{aligned}
&\sum_{i\in\mathcal{B}^+\cap\mathcal{A}} M_{i,t} \left(1 - \sum_{j\neq i} a^{(q)}_{j,i} + \sum_{k\in\mathcal{A}\backslash\{i\}} a^{(q)}_{k,i}\right) + \sum_{i\in\mathcal{B}^+\setminus\mathcal{A}} M_{i,t} \left(\sum_{k\in\mathcal{A}} a^{(q)}_{k,i}\right) \\
&\qquad = \sum_{i\in\mathcal{B}^+\cap\mathcal{A}} M_{i,t} \left(1 - \sum_{k\notin\mathcal{A}} a^{(q)}_{k,i}\right) + \sum_{i\in\mathcal{B}^+\setminus\mathcal{A}} M_{i,t} \left(\sum_{k\in\mathcal{A}} a^{(q)}_{k,i}\right)
\end{aligned}$$

And using that for any i $\sum_{k\neq i} a_{k,i}^{(q)} \in [0,1]$, we have

$$\sum_{i\in\mathcal{B}^+\cap\mathcal{A}} M_{i,t}\left(1-\sum_{k\notin\mathcal{A}} a_{k,i}^{(q)}\right) + \sum_{i\in\mathcal{B}^+\setminus\mathcal{A}} M_{i,t}\left(\sum_{k\in\mathcal{A}} a_{k,i}^{(q)}\right) \leq \sum_{i\in\mathcal{B}^+} M_{i,t} \tag{2.52}$$

Analogously, for the negative terms (i.e $M_{i,t}$ for $i \in \mathcal{B}^-$):

- If $i \in \mathcal{A}\cap\mathcal{B}^-$, then: (i) for $k = i$ we have the term $(1-\sum\limits_{j\neq i} a_{j,i}^{(q)})M_{i,t}$; (ii) for each $k \in \mathcal{A}\setminus\{i\}$, we have the term $a_{k,i}^{(q)} M_{i,t}$
- If $i \in \mathcal{B}^- \setminus \mathcal{A}$: for each $k \in \mathcal{A}$, we have the term $a_{k,i}^{(q)} M_{i,t}$

Then, the negative terms can be written as:

$$\sum_{i\in\mathcal{B}^-\cap\mathcal{A}} M_{i,t}\left(1-\sum_{j\neq i} a_{j,i}^{(q)} + \sum_{k\in\mathcal{A}\setminus\{i\}} a_{k,i}^{(q)}\right) + \sum_{i\in\mathcal{B}^-\setminus\mathcal{A}} M_{i,t}\left(\sum_{k\in\mathcal{A}} a_{k,i}^{(q)}\right)$$
$$= \sum_{i\in\mathcal{B}^-\cap\mathcal{A}} M_{i,t}\left(1-\sum_{k\notin\mathcal{A}} a_{k,i}^{(q)}\right) + \sum_{i\in\mathcal{B}^-\setminus\mathcal{A}} M_{i,t}\left(\sum_{k\in\mathcal{A}} a_{k,i}^{(q)}\right).$$

So, again using that for any i $\sum_{k\neq i} a_{k,i}^{(q)} \in [0,1]$, we have

$$\sum_{i\in\mathcal{B}^-\cap\mathcal{A}} (-M_{i,t})\left(1-\sum_{k\notin\mathcal{A}} a_{k,i}^{(q)}\right) + \sum_{i\in\mathcal{B}^-\setminus\mathcal{A}} (-M_{i,t})\left(\sum_{k\in\mathcal{A}} a_{k,i}^{(q)}\right) \leq -\sum_{i\in\mathcal{B}^-} M_{i,t} \tag{2.53}$$

The first part of the proposition follows directly from the fact that the previous two inequalities for $q = 1$ imply that for any $\mathcal{A} \subset \{1,\ldots,n\}$, we have $|\sum_{k\in\mathcal{A}} M_{k,t+1}| \leq \sum_{k\in\mathcal{B}^+} M_{k,t}$, and so the inequality is also true for the maximum.

For the second part of the proposition, suppose the condition holds and I will show by contradiction that we cannot find any $\mathcal{A}$ such that $|\sum_{k\in\mathcal{A}} M_{k,t+q}| \leq \sum_{k\in\mathcal{B}^+} M_{k,t}$ holds with equality, and so the inequality has to be strict.

In order for the equality to hold, it must be one of the following two cases:

- Case A: The positive terms are equal to the upper bound, and the negative terms are zero. For this to be the case, we need: (i) for $i \in \mathcal{B}^+ \cap \mathcal{A}$, $\sum_{k\notin\mathcal{A}} a_{k,i}^{(q)} = 0$; (ii) for $i \in \mathcal{B}^+ \setminus \mathcal{A}$, $\sum_{k\in\mathcal{A}} a_{k,i}^{(q)} = 1$; (iii) for $i \in \mathcal{B}^- \cap \mathcal{A}$, $\sum_{k\notin\mathcal{A}} a_{k,i}^{(q)} = 1$; and (iv) for $\mathcal{B}^- \setminus \mathcal{A}$, it must be $\sum_{k\in\mathcal{A}} a_{k,i}^{(q)} = 0$

  If $i \in \mathcal{A}$, then condition (i) implies that also $\ell \in \mathcal{A}$, since otherwise we would have the contradiction $0 = \sum_{k \notin \mathcal{A}} a_{k,i}^{(q)} \geq a_{\ell,i}^{(q)} > 0$. If $i \notin \mathcal{A}$, then condition (ii) again implies that $\ell \in \mathcal{A}$, since otherwise we would have the contradiction $1 = \sum_{k \in \mathcal{A}} a_{k,i}^{(q)} \leq 1 - a_{\ell,i}^{(q)} < 0$. Therefore, ℓ must be in $\mathcal{A}$ in order for the positive terms to reach the upper bound.

Next, if $j \in \mathcal{A}$, then condition (iii) implies that $\ell \notin \mathcal{A}$, since otherwise we would have the contradiction $1 = \sum_{k \notin \mathcal{A}} a_{k,j}^{(q)} \leq 1 - a_{\ell,j}^{(q)} < 1$. So, the only possibility is that $j \notin \mathcal{A}$, but then condition (iv) contradicts that $\ell \in \mathcal{A}$, since then we would have the contradiction $0 = \sum_{k \in \mathcal{A}} a_{k,j}^{(q)} \geq a_{\ell,j}^{(q)} > 0$. Therefore, Case A is not possible.

- Case B: The positive terms are equal to zero, and the negative terms are equal to the lower bound. For this to be the case, we need: (i) for $i \in \mathcal{B}^+ \cap \mathcal{A}$, $\sum_{k \notin \mathcal{A}} a_{k,i}^{(q)} = 1$; (ii) for $i \in \mathcal{B}^+ \setminus \mathcal{A}$, $\sum_{k \in \mathcal{A}} a_{k,i}^{(q)} = 0$; (iii) for $i \in \mathcal{B}^- \cap \mathcal{A}$, $\sum_{k \notin \mathcal{A}} a_{k,i}^{(q)} = 0$; and (iv) for $\mathcal{B}^- \setminus \mathcal{A}$, it must be $\sum_{k \in \mathcal{A}} a_{k,i}^{(q)} = 1$

  Analogously as in the previous case, we get to the conclusion that this case is not possible. If $i \in \mathcal{A}$, then (i) implies $\ell \notin \mathcal{A}$, since otherwise $1 = \sum_{k \notin \mathcal{A}} a_{k,i}^{(q)} \leq 1 - a_{\ell,i}^{(q)} < 1$. If $i \notin \mathcal{A}$, then (ii) also implies that $\ell \notin \mathcal{A}$, since otherwise $0 = \sum_{k \in \mathcal{A}} a_{k,i}^{(q)} \geq a_{\ell,i}^{(q)} > 0$. So, it must be $\ell \notin \mathcal{A}$.

  If $j \in \mathcal{A}$, (iii) implies the contradiction $0 = \sum_{k \notin \mathcal{A}} a_{k,j}^{(q)} \geq a_{\ell,j} > 0$. But if $j \notin \mathcal{A}$, (iv) also implies the contradiction $1 = \sum_{k \in \mathcal{A}} a_{k,j}^{(q)} \leq 1 - a_{\ell,j} < 1$. So, we conclude that this case is not possible.

■

This proposition tells us that a sufficient condition to guarantee convergence to an ergodic distribution is that whenever we are not in a stationary distribution, we can find states that have changed in opposite directions in the previous iteration (period) such that there exists some state which can be reached from each of the two states with positive probability in the same number of periods (in other words, if two points start from state i and j respectively, there is positive probability they will meet at some future period).

The next definitions and proposition show a sufficient condition for this condition to hold:

Definition 1 *A Markov chain is* ***irreducible*** *if for any pair of states i, j, there exists $q \in \mathbb{N}$ such that $a_{j,i}^{(q)} > 0$. (that is, it is possible to get to any state from any other state)*

Definition 2 *Let the* ***longitude of the shortest path between two states*** *i, j be $d_{i,j} =$*

$\min\{q \in \mathbb{N} | a_{i,j}^{(q)} > 0\}$ *(that is, the smallest number of periods required to go from one state to the other).*

Proposition 4 *In an irreducible Markov chain that contains at least one state i with $d_{i,i} = 1$, as long as we are not in the stationary distribution, it is always possible to find states $j \in \mathcal{B}^+$ and $k \in \mathcal{B}^-$, and a state ℓ such that $a_{\ell,j}^{(q)}, a_{\ell,k}^{(q)} > 0$ for some $q \in \mathbb{N}$ (and so, in such Markov chain we can guarantee convergence to an ergodic distribution).*

Proof. If we are not in a stationary distribution then there are j with $M_{j,t} \neq 0$, and by Lemma 2 there must be $j \in \mathcal{B}^+$ and $k \in \mathcal{B}^-$. .Let i be the state such that $d_{i,i} = 1$. Then, it is sufficient to see that we can find $q \in \mathbb{N}$ such that $a_{i,j}^{(q)}, a_{i,k}^{(q)} > 0$, which is straightforward. We can check that $q := \max(d_{i,j}, d_{i,k})$ satisfies this (intuitively, the first to arrive from one of the two states then stays with positive probability in i, and at some point the one that started from the other state will also arrive to i). Without loss of generality, assume $\max(d_{i,j}, d_{i,k}) = d_{i,k}$. $a_{i,k}^{d_{i,k}} > 0$ by definition of $d_{i,k}$. But also $a_{i,j}^{d_{i,k}} \geq a_{i,j}^{d_{i,j}} a_{i,i}^{(d_{i,k}-d_{i,j})} \geq a_{i,j}^{d_{i,j}} \left(a_{i,i}^{(1)}\right)^{d_{i,k}-d_{i,j}} > 0$

■

So, in the Uniqueness section I proved that the only possible stationary distribution is the one we would obtain if the initial sate is $\mathcal{J} = \emptyset$ (if this converges). Now, the previous proposition tells us that $P_{t,0}(X)$ converges, since the Markov chain obtained is irreducible (if a state is possible, it means that there was positive probability of arriving to it starting from $\mathcal{J} = \emptyset$; and, from any state, there is probability 1 of eventually going back to $\mathcal{J} = \emptyset$) and the state $\mathcal{J} = \emptyset$ satisfies that the longitude of its shortest path connecting it to itself is 1 (with positive probability there will be no entrant and we stay at $\mathcal{J} = \emptyset$).

2.7.8 Summary of the method to solve the model

1. First, for each possible number of firms J:

   - Define the different awareness sets $\mathcal{P}_J$. There are 2^J awareness sets (think on how many different ways we can assign $\{0, 1\}$ to J variables).
   - Define the N_J grid nodes we will use, $\vec{M}_n$, $n = 1, \ldots, N_J$. Each node is a vector of the mass of consumers in each awareness set. As we have seen, the solutions

of the model are functions of the form $f(\mathcal{J}, \vec{M})$ on a continuous m-dimensional space, with $m = 2^J - 1$ ($\vec{M}$ is $(m+1)$-dimensional, but since the masses have to add up to 1, one is redundant). To deal with this exponentially increasing state space and alleviate the curse of dimensionality, I introduce a piecewise multivariate Newton interpolation method described in detail in section 2.7.9. Using this method, increasing the number of grid points leads to a better approximation, as in standard univariate methods using a grid and linear interpolation, with the advantage that the higher degree of the interpolating polynomial allows to reduce the number of necessary grid points for a given fit.[18]

Also, note that $\mathcal{J}$ has information of the identity of the firm. Therefore, some nodes are just a reordering of firms, so I use this to avoid solving again nodes that are just a reordering of a node that has already been solved.

2. Define initial guesses for the aggregate states w and T, as well as initialize the policy functions for advertising expenditure and entry; that is, assign a value for the grid nodes $\{\{\{e_{j,n}\}_{j=1}^{J}\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$ and entry $\{\{N_{e,n}\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$.

3. Given the aggregate states:

   (a) Solve the firm problem:

      i. Given the guess of the policy functions for advertising expenditures and entry:

         - Solve the static price-setting problem. Note that this has to be updated in every iteration of the firm problem because the advertising choices affect the demand shifters $\omega_{j,i,t}$. This gives us profits and markups at each node: $\{\{\mathcal{M}_j(J, \vec{M}_n), \pi_j(J, \vec{M}_n)\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$.
         - Solve for the value function implied by the policy functions of advertising and entry and the profit function. Note that this implies solving a linear system on $\{\{V(J, \vec{M}_n)\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$.

      ii. Given the functions for markups and firm value found in the previous points, $\{\{\mathcal{M}_j(J, \vec{M}_n), V_j(J, \vec{M}_n)\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$, compute the best responses:

[18] Using piecewise interpolation is important because increasing the degree of an interpolating polynomial doesn't necessary lead to a better approximation (Runge's phenomenon).

$\{\{\{e'_{j,n}, N'_{e,n}\}_{j=1}^{J}\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$ (i.e. the optimal choice keeping the competitors' choices fixed). If the difference between these best responses and the previous guess is small enough, we are done (in this case we have found a Nash equilibrium); otherwise, update the guesses and go back to (i).

(b) Solve for the unique stationary distribution given the solution of the firm problem (in particular, we need the policy functions for the advertising space $\{\{\{\alpha_{j,n}\}_{j=1}^{J}\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$ and entry $\{\{N_{e,n}\}_{n=1}^{N_J}\}_{J=1}^{\bar{J}}$ and entry in an industry with $J = 0$: $N_{e,0}$). For the details of the method, see section 2.7.8.1

4. Given the firm policy functions and the stationary distribution, compute the implied aggregates w and T, using 2.24, 2.23, together with $T = Q$. If the difference between the guesses and the implied values of w and T are close enough, we are done; otherwise, update the new guesses for w and T and go back to 3.

2.7.8.1 Method used to find the stationary distribution

The method has two parts.

1. In section 2.7.7, I show that the set of industry states observed in the stationary distribution is at most countably infinite; and so the stationary distribution is a discrete probability function defined on a potentially infinite set of points, and so, computationally, the set of states needs to be bounded some way. In the following I describe the approach used in the baseline to bound the set of states. As a robustness, I compare the stationary distribution obtained from this approach to the one obtained by bounding the space by a grid (that is, restricting $\vec{M}$ to take only values from a grid). The baseline approach tends to be much faster.

   (a) Given that in section 2.7.7 I show that the unique stationary distribution is the one we would obtain if the initial sate is $\mathcal{J} = \emptyset$, then:

   - I initialize the *List* of states with this state. For each state in the *List*, I store (1) the number of firms, (2) the vector of masses corresponding to this state, (3) the vector of ages, and (4) the probability *Prob*, which I now describe. *Prob* is the probability of going from state $\mathcal{J} = \emptyset$ to the particular state X in the shortest path from $\mathcal{J} = \emptyset$ to X. That is, for this initial state $\mathcal{J} = \emptyset$, we have $Prob = 1$.

- To facilitate the process of looking up whether we have already encountered a state before (i.e. whether a state is already in *List*, I order the states in a library *LibraryStates* lexicographically based on (i) the number of active firms, (ii) the vector of ages, and (iii) the vector of masses. Initially, *LibraryStates* = 1.
- I also initialize $iter = 0$ and the list of states I will take as starting point in the following iteration, $NewStates_{iter}$. Initially, $NewStates_1 = 1$.
- Finally, we need the transition matrix with the probabilities of going from each state to the others. However, since this matrix is very sparse (the states are just directly connected to few others) and storing the whole matrix with all the zeros would be highly costly for memory storage (and solving the system would also be very slow), I only store the non-zero elements of the transition matrix in a *Library*, where each book contains three pieces of information (the books are ordered lexicographically based on the same order of these three pieces of information): the row in the transition matrix (i.e., state of origin), the column in the transition matrix (i.e., state of destination), and the value in this position of the matrix resulting from subtracting the transition matrix from the identity matrix. I initialize it as $Library = [1, 1, 1]$ (the third 1 is the 1 from the identity matrix).

(b) Then, as long as $NewStates_{iter+1}$ is not empty, increase iter by 1 and do the following for each state $s \in NewStates_{iter+1}$:

   i. Calculate the next period vector of masses if there weren't entry/exit, and the probability of an entrant. Then, for each of the possible cases of entry/exit, letting q be the probability of the particular event of entry/exit, I do the following:

   ii. Look up whether this state is already in *List*, using the order in *LibraryStates*. Here is where I bound the problem.

      - If the probability *Prob* is above a threshold, then I check for an exact match (that is, they match in the three elements: J, the vector of ages, and the vector of masses (note that, although the vector of ages is not a state in the baseline firm problem, it is useful to distinguish it for the quantitative exercises).

- If the probability *Prob* is below the threshold (that is, it is a rare state), then I just check for J and the vector of ages. If there is no state in *List* matching J and the vector of ages, then we will treat this state as a new state; otherwise, I will treat it as if it were identical to the first state in *List* with the same J and ages. The intuition is that, although the vector of ages is not a sufficient statistic (because history matters), it serves as a good first approximation. The other boundary I set is on the firm age; in particular, I don't distinguish ages above a threshold (which I set to 20 years old). The intuition is that for firms older than 20 years old very few consumers remain unaware of the firm, so the error from not distinguishing older firms is negligible.

iii. If the outcome from the previous point is that it is not a new state, then we index it by s' equal to the index of the state we have matched it to and go to (iv); else, if it is a new state, then we index it by s' equal to the current size of *List* plus one and do the following:

- Add the one of the identity matrix to *Library*; that is: add $[s', s', 1]$.
- Add the four pieces of information relative to this state in *List*. *Prob* will be equal to the *Prob* of state s time q.
- Add s' to $NewStates_{iter+1}$.

iv. Add $[s, s', -q]$ to the *Library*. If there is already an element at position $[s, s']$, then just add $-q$.

2. In the second part, we need to solve for the stationary distribution. The matrix found in the previous step is singular (note that the sum of all the elements in row s is $1 - \sum_{s'} p_{s,s'} = 1 - 1 = 0$, where $p_{s,s'}$ is the probability of moving from state s to s'). So, we need to add a new condition to have a compatible and determinate system: it is the condition that the solution must add up to 1; so, I add to *Library* $[s, 0, 1]$, for all the states s. Then, we

also need the vector of independent coefficients, which again is very sparse (there is only one non-zero value), so again I store it in a library called $LibraryB = [0, 1]$.

2.7.9 Multivariate Newton Interpolation

Suppose we want to interpolate a function $f : \mathbb{R}^m \to \mathbb{R}$ by a polynomial of m variables and degree n.

Definition 3 ***(Generating points):*** *For each dimension $i = 1, \ldots, m$, we define $n + 1$ points $x_{i,k}$, $k = 0, \ldots, n$. $\{\{x_{i,k}\}_{k=0}^{n}\}_{i=1}^{m}$ are called the generating points.*

Definition 4 ***(Multiindices):*** *Let $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \Lambda_{m,n} := \{\vec{\alpha} \in \{0, \ldots, n\}^m | \sum\limits_{i=1}^{m} \alpha_i \leq n\}$, and $\vec{x}_{\vec{\alpha}} = (x_{1,\alpha_1}, \ldots, x_{m,\alpha_m})$.*

The cardinal of $\Lambda_{m,n}$ (i.e. the number of different multiindices) is given by $N(m, n) = \binom{n+m}{n}$ (to see this, you can think of $1^{\alpha_0} x_1^{\alpha_1} \cdots x_m^{\alpha_m}$ with $\sum_{i=0}^{m} \alpha_i = n$, which we can transcribe as $\underbrace{1 \ldots 1}_{\alpha_0} \# \underbrace{x_1 \ldots x_1}_{\alpha_1} \# \ldots \# \underbrace{x_m \ldots x_m}_{\alpha_m}$; so the problem of finding the number of different multi-indices is equivalent to finding the number of different ways we can choose m boxes from $n + m$ boxes (i.e. the position of the m hashtags), which is $\binom{n+m}{m}$).

Definition 5 ***(Newton polynomial):*** $w_{\vec{\alpha}}(\vec{x}) = \prod\limits_{i=1}^{m} \prod\limits_{k=0}^{\alpha_i - 1} (x_i - x_{i,k})$.

Definition 6 *The m-dimensional* ***Newton interpolating polynomial*** *of degree n of the function f is $p_{m,n}(\vec{x}) = \sum\limits_{\vec{\alpha} \in \Lambda_{m,n}} a_{\vec{\alpha}} w_{\vec{\alpha}}(\vec{x})$, satisfying $f(\vec{x}_{\vec{\alpha}}) = p_{m,n}(\vec{x}_{\vec{\alpha}})$, for all $\vec{\alpha} \in \Lambda_{m,n}$.*

Lemma 5 *Note that given $\vec{\beta}, \vec{\alpha} \in \Lambda_{m,n}$, if $\beta_i - 1 \geq \alpha_i$, then $w_{\vec{\beta}}(\vec{x}_{\vec{\alpha}})$ contains the term $(x_{i,\alpha_i} - x_{i\alpha_i}) = 0$.*

Corollary 1 *Then, $f(\vec{x}_{\vec{\alpha}}) = p_{m,n}(\vec{x}_{\vec{\alpha}}) = \sum\limits_{k_m=-1}^{\alpha_m - 1} \cdots \sum\limits_{k_1=-1}^{\alpha_1 - 1} \prod\limits_{s_m=0}^{k_m} (x_{m,\alpha_m} - x_{m,s_m}) \cdots \prod\limits_{s_m=0}^{k_m} (x_{1,\alpha_1} - x_{1,s_1}) a_{(k_1+1,\ldots,k_m+1)}$*

To allow generality, I define:

Definition 7 *Given $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$, define:*
(i) $\vec{\alpha}^{(i,k)} = (\alpha_1, \ldots, \alpha_{i-1}, k, \alpha_{i+1}, \ldots, \alpha_m)$ (i.e. $\vec{\alpha}^{(i,k)}$ equals $\vec{\alpha}$ except k in position i)

(ii) $\vec{\alpha}^{(k)} = (\alpha_{k+1}, \ldots, \alpha_m)$.

(iii) $\vec{\beta}^{(i,k)} = (\vec{\alpha}^{(i,k-1)}, \ldots, \vec{\alpha}^{(i,0)}, \ldots, \vec{\alpha}^{(m,\alpha_m-1)}, \ldots, \vec{\alpha}^{(m,0)})$, *and let* $\vec{\beta}^{(i,0)} = \vec{\beta}^{(i+1,\alpha_{i+1}-1)}$ *and* $\vec{\beta}^{(m,0)} = \emptyset$.

Definition 8 ***(Divided differences):***
$f[\vec{\alpha}, \vec{\beta}^{(i,b)}] = \sum_{k_i=b-1}^{\alpha_i-1} \cdots \sum_{k_1=-1}^{\alpha_1-1} \prod_{s_i=b}^{k_i} (x_{m,\alpha_m} - x_{m,s_m}) \cdots \prod_{s_1=0}^{k_1} (x_{1,\alpha_1} - x_{1,s_1}) a_{(k_1+1,\ldots,k_i+1,\vec{\alpha}^{(i)})}$ *if* $\alpha_i > b$*; and* $f[\vec{\alpha}, \vec{\beta}^{(i,b)}] = f[\vec{\alpha}, \vec{\beta}^{(i-1,0)}]$ *otherwise.*

Note that by Corollary 1, and since $\vec{\alpha}^{(m)} = \emptyset$, then $f[\vec{\alpha}, \vec{\beta}^{(m,0)}] = f(\vec{x}_{\vec{\alpha}})$. The algorithm to find the coefficients $a_{\vec{\alpha}}$ is defined as follows:

1. Start setting $i = m$ and $b = 0$.

2. If there is some $\vec{\alpha} \in \Lambda_{m,n}$ such that $\alpha_i > b$, then:

   (a) For all the $\vec{\alpha} \in \Lambda_{m,n}$ such that $\alpha_i > b$: Noting that $f[\vec{\alpha}^{(i,b)}, \vec{\beta}^{(i,b)}]$ contains all the terms of $f[\vec{\alpha}, \vec{\beta}^{(i,b)}]$ with $k_i = b - 1$, and so the remaining terms will all contain $(x_{i,\alpha_i} - x_{i,b})$; then:

$$\begin{aligned} f[\vec{\alpha}, \vec{\beta}^{(i,b+1)}] &= \frac{f[\vec{\alpha}, \vec{\beta}^{(i,b)}] - f[\vec{\alpha}^{(i,b)}, \vec{\beta}^{(i,b)}]}{x_{i,\alpha_i} - x_{i,b}} \\ &= \sum_{k_i=b}^{\alpha_i-1} \cdots \sum_{k_1=-1}^{\alpha_1-1} \prod_{s_i=b+1}^{k_i} (x_{m,\alpha_m} - x_{m,s_m}) \cdots \prod_{s_1=0}^{k_1} (x_{1,\alpha_1} - x_{1,s_1}) a_{(k_1+1,\ldots,k_i+1,\vec{\alpha}^{(i)})} \end{aligned}$$

   (b) For all the $\vec{\alpha} \in \Lambda_{m,n}$ such that $\alpha_i \le b$, then $f[\vec{\alpha}, \vec{\beta}^{(i,b+1)}] = f[\vec{\alpha}, \vec{\beta}^{(i,b)}]$ (satisfies the definition since $\alpha_i \le b < b+1$, so $f[\vec{\alpha}, \vec{\beta}^{(i,b+1)}] = f[\vec{\alpha}, \vec{\beta}^{(i,b)}] = f[\vec{\alpha}, \vec{\beta}^{(i-1,0)}]$

   Set $b = b + 1$, and go back to step 2.

3. If $\alpha_i \le b$ for all $\vec{\alpha} \in \Lambda_{m,n}$ (which is satisfied if and only if $b \le n$), then make $f[\vec{\alpha}, \vec{\beta}^{(i-1,0)}] = f[\vec{\alpha}, \vec{\beta}^{(i,b)}]$, and set $i = i - 1$ and $b = 0$. If $i = 0$, we are done; otherwise, go back to step 2.

All is left to do is to show that the $f[\vec{\alpha}, \vec{\beta}^{(0,0)}] = a_{\vec{\alpha}}$ for all $\vec{\alpha} \in \Lambda_{m,n}$. Given that the divided difference of $\vec{\alpha}$ just changes when we apply (2a) to it, then it is sufficient to see that in the last time that we select $\vec{\alpha}$ for (2a) it is $f[\vec{\alpha}, \vec{\beta}^{(i,b+1)}] = a_{\vec{\alpha}}$; since then it will be $f[\vec{\alpha}, \vec{\beta}^{(0,0)}] = f[\vec{\alpha}, \vec{\beta}^{(i,b+1)}] = a_{\vec{\alpha}}$.

Proof. If we have used $a_{\vec{\alpha}}$ in (2a), it means that $\alpha_i > b$, which implies that exactly one of the following is true:

1. $\alpha_i > b+1$, in which case $a_{\vec{\alpha}}$ would also be selected in the next iteration, contradicting it was the last time it was selected;

2. $\alpha_i = b+1$, in which case $a_{\vec{\alpha}}$ it is the last iteration for variable i that $a_{\vec{\alpha}}$ is selected. In this case there are two possibilities:

   - $\alpha_k > 0$ for some $k < i$, in which case in iteration $(k, 0)$ $\vec{\alpha}$ would be selected, contradicting the hypothesis.

   - $\alpha_k = 0$ for all $k < i$, in which case we have:

$$
\begin{aligned}
f[\vec{\alpha}, \vec{\beta}^{(i,b+1)}] &= \sum_{k_i=\alpha_i-1}^{\alpha_i-1} \cdots \sum_{k_1=-1}^{\alpha_1-1} \prod_{s_i=\alpha_i}^{k_i} (x_{m,\alpha_m} - x_{m,s_m}) \cdots \prod_{s_1=0}^{k_1} (x_{1,\alpha_1} - x_{1,s_1}) a_{(k_1+1,\ldots,k_i+1,\vec{\alpha}^{(i)})} \\
&= \prod_{s_i=\alpha_i}^{\alpha_i-1} (x_{m,\alpha_m} - x_{m,s_m}) \cdots \prod_{s_1=0}^{-1} (x_{1,\alpha_1} - x_{1,s_1}) a_{(0,\ldots,0,\alpha_i,\vec{\alpha}^{(i)})} \\
&= a_{(0,\ldots,0,\alpha_i,\vec{\alpha}^{(i)})} = a_{\vec{\alpha}}
\end{aligned}
$$

■

Chapter 3

Fiscal Policy, Competition and Growth

3.1 Introduction

Fiscal consolidation, a term used to describe government efforts to reduce fiscal deficits and lower public debt, has become a priority for many countries. Significant increases in spending following the 2008-09 financial crisis, the Covid-19 pandemic, and recent shocks such as fluctuating commodity prices and the Ukraine conflict have placed considerable strain on public finances. A significant number of developing and emerging economies are grappling with debt distress, while many advanced economies—even if not yet in the same situation—face a pressing need for fiscal consolidation (IMF, 2020). Strengthening policy buffers will be essential to address infrastructure gaps, respond to climate change, support aging populations, and prepare for future economic shocks.

There is broad agreement that fiscal consolidation tends to have negative short-term effects on growth, primarily through fiscal multipliers, while supporting higher growth in the long run by crowding in private investment and reducing policy uncertainty (Balasundharam et al., 2023). Interestingly, the transmission channels typically cited—such as lower interest rates and reduced uncertainty—are mechanisms more commonly associated with short-run dynamics, as captured in standard macroeconomic models designed to study business cycle fluctuations rather than long-term growth. Indeed, much of the existing analysis of the long-run effects of fiscal consolidation relies on this class of models. Yet, it is well established that the fundamental driver of

[1]This chapter is joint work with Giammario Impullitti and Antonin Bergeaud.

long-run growth is innovation and technological progress. This suggests that a comprehensive assessment of the long-term implications of debt and deficit reduction should explicitly consider their impact on innovation and the sustained productivity growth it enables.

Structural reforms—such as product and labour market liberalisation—have the potential to mitigate the contractionary effects typically associated with fiscal consolidation. In particular, product market competition is widely recognised as a key driver of innovation and long-run economic growth.[2] This is particularly relevant in light of the observed rise in market power across many countries in recent decades,[3] which has coincided with a prolonged slowdown in productivity growth across a broad set of countries, including many advanced economies.

Does market power matter for the long-run effect of fiscal policy? What are the long-run effects of fiscal consolidation in economies characterized by high market power and low productivity growth. Is there complementarity between fiscal and competition policies, and how does it shape the long-run impact of fiscal consolidation?

To address these questions, we develop an endogenous growth model with variable markups to study the effects of fiscal policy on innovation and productivity growth, and to analyse the role of product market competition in shaping the transmission of these policies. Our framework builds on the class of step-by-step Schumpeterian growth models, in which a small number of oligopolistic firms compete strategically for market leadership within each product line (e.g. Aghion et al., 2005). The economy features a 'mixed market structure', where large, highly productive firms coexist with a competitive fringe of smaller, less productive firms. Within each industry, a small number of heterogeneous superstar firms engage in Cournot competition for market shares and invest in innovation to improve their productivity and maintain their leadership position. Firm-level productivity differences arise both from initial heterogeneity at entry and from subsequent success in innovation. In parallel, each industry is populated by a continuum of smaller firms that operate at a lower productivity level and capture a limited share of the market. These fringe firms are separated from the superstar firms by an exogenously given productivity gap, but retain the ability to invest in innovation and potentially ascend to

[2]See the seminal contributions of Aghion et al. (2001, 2005), and Griffith and Reenen (2021) for a recent review of the theoretical and empirical literature on competition and growth.

[3]A range of indicators of market power—including markups and market concentration measures—have risen over the past three decades in many advanced economies (e.g. Autor et al., 2020; De Loecker et al., 2020; Bajgar et al., 2019; Diez et al., 2021).

the superstar class. While the inflow of new superstar firms is driven by successful innovation from the competitive fringe, exit from the superstar group occurs either when firms fail to cover a fixed operating cost or when their relative productivity falls sufficiently behind that of the industry leader. Finally, a mass of entrepreneurs can enter the market by paying an entry cost to join the competitive fringe. In this setting, long-run economic growth is ultimately driven by the innovation efforts of superstar firms within each industry, which push the technological frontier forward.

The model generates non-degenerate distributions of sales, employment, and markups within each industry, which are shaped by the number of superstar firms, the distribution of their productivities, and the relative productivity and mass of the competitive fringe. This rich industry structure enables us to discipline the model using firm-level data, allowing it to replicate key cross-sectional patterns observed in the data—including the dispersion of productivity and markups across firms.

Moreover, the model inherits the nuanced relationship between competition and growth that characterises step-by-step innovation frameworks. On the one hand, firms have strong incentives to innovate in order to escape intense competition and secure a more dominant market position—a mechanism known as the *escape competition* effect, which generates a positive relationship between competition and innovation. On the other hand, greater competition can erode the expected rents from innovation by reducing post-innovation markups, thereby discouraging investment in R&D—a channel often referred to as the discouragement or *Schumpeterian* effect. The interplay between these two opposing forces provides the model with the necessary flexibility to capture the often non-monotonic or non-linear relationship between competition and innovation documented in the empirical literature.

In our model economy, the government levies taxes on consumption and labour income. In the baseline specification, we assume that the government operates under a balanced budget, using the proceeds from these taxes to finance lump-sum transfers to households. We later extend the framework to allow for fiscal deficits and public debt accumulation. Labour supply is endogenous, as households derive utility from both consumption and leisure. Importantly, the size of the market—a key determinant of innovation incentives—depends on labour supply, which responds to changes in wages through the standard income and substitution effects.

Consequently, fiscal policy—through its effects on consumption and labour taxation—influences the size of the market by altering households' labour supply decisions.

The effects of fiscal policy on long-run innovation and growth operate through the two central mechanisms linking competition and innovation in our framework: the escape competition effect and the Schumpeterian effect. For instance, a fiscal expansion that stimulates labour supply increases the size of the market, boosting demand and firm profits, which in turn fosters innovation. However, in the absence of firm entry, higher profits reduce the incentive to innovate through the escape competition channel, as firms become more profitable without needing to outpace competitors. Crucially, in our model, firm entry responds endogenously to changes in market size: a larger market encourages entry, intensifies competition, and strengthens incentives for innovation in order to escape rivals. This *pro-competitive* effect of fiscal policy amplifies the positive impact of fiscal expansions on innovation and growth, highlighting the important role of product market competition in mediating the transmission and effectiveness of fiscal policy.

We use administrative French firm-level data, FICUS and FARE from Insee-DGFiP to calibrate the model and explore the impact of fiscal policy numerically. We begin by examining whether economies characterised by higher market power—where firms enjoy greater pricing power—exhibit stronger or weaker responses to changes in fiscal policy compared to more competitive environments. Since product market competition is endogenous in our framework, we then explore how fiscal expansions affect both competition and innovation. Comparing these outcomes with those generated by a version of the model in which competition is exogenously fixed—and thus unresponsive to policy changes—allows us to isolate and quantify the role of competition in mediating the effects of fiscal policy. We find that market structure endogeneity amplifies the effect of fiscal policy on growth by 10%. Finally, we investigate the complementarity between fiscal and competition policies. Specifically, we contrast the growth effects of policy regimes in which fiscal and competition instruments are used in isolation with those in which they are deployed jointly. This analysis sheds light on the potential for competition policy to mitigate the adverse effects of fiscal consolidation on long-run growth.

Literature review. We use the frontier version of the step-by-step Schumpeterian growth model recently developed by (Cavenaile et al., 2021). This framework builds on the standard step-by-step model—widely used to study the relationship between competition and innovation—by introducing a key improvement: a fully-fledged entry process. In the standard model, each industry is characterized by a duopoly, and even when entry is allowed, the entrant simply replaces one of the incumbent firms. Markups depend on the productivity gap between the leader and the follower. In contrast, the new model features non-degenerate distributions of sales, employment, markups, and profits within each industry. It allows for an arbitrarily large and endogenous number of oligopolistically competing firms, providing a richer and more realistic characterization of industry dynamics. The presence of a competitive fringe captures the realistic coexistence of both small and large firms within an industry, allowing for more empirically plausible market share distributions and firm life cycles. In particular, entrants do not immediately displace incumbents as industry leaders but instead typically start small and grow over time.

We extend the framework along three dimensions. First, we introduce fixed operating costs for superstar firms, thereby incorporating a selection channel in the model. Second, we assume decreasing returns to labor to allow the size of the fringe to matter for the degree of product market competition in each industry. Third, we add endogenous labor supply and fiscal policy.

There is a recent literature analyzing the link between competition and the effects of monetary policy. Variable markups imply that firms do no fully pass cost reductions onto prices, as part of them go into increasing profit margins. Via this *incomplete pass-through* channel, reduction in borrowing costs brought about by expansionary monetary policy have a weaker impact on output (e.g Ferrando et al., 2021, forthcoming). Moreover, *strategic complementarity* in firm-level pricing, typical of oligopolistic models can increase price-stickiness thereby generating stronger money non-neutrality, stronger output response to monetary policy changes (e.g. Mongey, 2021; Wang and Werning, 2022). While this line of research focuses on short-run effects of monetary policy, others have analysed long-run effects and highlighted a third channel shaping the interaction between competition and monetary policy which operates via *credit constraints*. Higher product market competition means lower profits and therefore more need for external funds. Thus as firms are more dependable on credit, monetary expansion aimed at reducing the cost of credit have a stronger impact on firms' investment decisions than in an economy with higher profits (Aghion et al., 2019).

To the best of our knowledge, there is no work on the interaction between competition and fiscal policy. While monetary policy acts directly on a price, the cost of borrowing, allowing firms to increase investment and hiring, fiscal policy operates directly on demand, stimulating firms growth via an expansion of the size of the market.[4] The size of the market plays an important role for competition, as larger markets typically promote entry. Thus, the linkage between competition and fiscal policy operates via a different channel compared to monetary policy, the *market size* channel produced by demand management.

3.2 The model

3.2.1 Environment

Preferences. Time is continuous. There is a unit mass of identical consumers who choose their consumption and labor supply to maximize lifetime utility:

$$U = \int_0^{\infty} e^{-\rho t} \; [\ln C_t + \gamma \ln(1 - L_t)] \, dt, \tag{3.1}$$

where L_t is the labor supplied by the individual per unit of time, and C_t is the final consumption good, which is a Cobb-Douglas aggregator of the continuum of industry goods:

$$\ln C_t = \int_0^1 C_{i,t} di. \tag{3.2}$$

In each industry, there is a fringe producing a homogeneous good and an endogenous discrete number of superstar firms each one producing a single differentiated good. The industry consumption good is a CES aggregator of these goods:

$$C_{i,t} = \left(C_{f,i,t}^{\frac{\sigma-1}{\sigma}} + \sum_{j=1}^{N_{i,t}} c_{j,i,t}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad , \quad \sigma > 1, \tag{3.3}$$

[4]Monetary policy operates via a direct effect on the cost of borrowing in all its standard transmission channels. In the 'traditional money-view' policy operates through the expectation channel and affects the user cost of capital, while in the "credit-channel view" it operates through its effects on external finance and credit constraints (Bernanke and Gertler, 1995).

where $C_{f,i,t}$ is the quantity consumed from the good of the fringe in industry i, and $c_{j,i,t}$ is the quantity consumed from superstar j in industry i.

Households own all the assets in the economy, and face the budget constraint:

$$\dot{a}_t = r_t a_t + (1 - \tau_L) w_t L_t - (1 + \tau_C) P_t C_t + TR_t, \tag{3.4}$$

where a_t is the level of assets, w_t is the wage rate, P_t is the price index of the final consumption good, which is the numeraire, τ_L and τ_C are taxes on labor and consumption, respectively, and TR_t is a lump-sum transfer.

Production technology and market structure. The fringe is populated by a mass $m_{i,t}$ of small firms, all producing the same homogeneous good. Let lowercase $y_{f,i,t}$ and $\ell_{f,i,t}$ refer to the output and labor of a single firm in the fringe, and $Y_{f,i,t}$ and $L_{f,i,t}$ be the the output and labor of the fringe as a whole.

Both fringe and superstar firms face a production function with decreasing returns to scale in labor

$$y_{f,i,t} = A_{f,i,t} \ell^{\alpha}_{f,i,t} \ , \quad y_{j,i,t} = A_{j,i,t} \ell^{\alpha}_{j,i,t}, \ \ j = 1, \ldots, N_{i,t}, \quad \alpha \in (0, 1], \tag{3.5}$$

where $A_{j,i,t}$ and $\ell_{j,i,t}$ are the productivity and labor, respectively, of superstar j in industry i and time t. Analogously, $A_{f,i,t}$ and $\ell_{f,i,t}$ are the productivity and labor of the fringe. All firms in the fringe of industry i have the same productivity given by $A_{f,i,t} = A_{1,i,t}\zeta$, where $A_{1,i,t}$ is the productivity of the leader in industry i.

In addition to the variable labor costs of production, superstar firms need to pay a fixed cost f (in labor units) at every t in order to operate; otherwise, they exit.

Innovation by superstars. Superstar firms employ labor, $h_{j,i,t}$, to improve their productivity. If successful, the productivity of that superstar, $A_{j,i,t}$, is multiplied by a step $\lambda > 1$. Success in innovation arrives at the Poisson rate $\rho_{j,i,t}$ given by

$$\rho_{j,i,t} = Z h^{\psi}_{j,i,t}, \ \ \psi \in (0, 1). \tag{3.6}$$

Entry an exit to superstars. Each small firm in the fringe can also employ labor $h_{f,i,t}$ with the aim to become a superstar in their industry. If successful, and as long as the current number of superstars is below an exogenous number $\bar{N}$, there will be a new superstar in the industry that starts with productivity $A_{j,i,t} = A_{1,i,t}\lambda^{-a_e}$, where $a_e \in \mathbb{N}$. Success in innovation by a small firm arrives at the Poisson rate $\rho_{f,i,t}$ (so, the industry Poisson rate of a new superstar is $\bar{\rho}_{f,i,t} = m_{i,t}\rho_{f,i,t}$) given by

$$\rho_{f,i,t} = Z_f h_{f,i,t}^{\psi_f}, \quad \psi_f \in (0,1). \tag{3.7}$$

There are two possible ways a superstar can exit. On the one hand, since the firm has to incur a fixed cost f to operate as a superstar, a firm will endogenously exit when its value is negative. On the other hand, as in Cavenaile et al. (2021), a superstar exits (or loses its superstar status) when it falls more than $\bar{a}$ gaps below the industry leader (that is, if a superstar is $\bar{a}$ gaps behind the leader and the leader innovates, then the firm exits from the pool of superstars). As in Cavenaile et al. (2021), our interpretation of the exogenous $\bar{a}$ gap is that firms below this gap are not large enough to strategically interact with the other superstars.

Entry and exit to the fringe. As in Cavenaile et al. (2021), we assume an exogenous death rate of small firms, δ, and a measure one of entrepreneurs who employ $h_{e,t}$ researchers to get a Poisson rate $\rho_{e,t} = Z_e h_{e,t}^{\frac{1}{2}}$ of starting a new small firm. New small firms are randomly allocated to an industry, so the mass of small firms is the same in all industries, $m_{i,t} = m_t$. Successful entrepreneurs sell their firm on a competitive market at its full value and remain in the set of entrepreneurs, so that the mass of entrepreneurs is unchanged.

3.2.2 Equilibrium

For the derivations, see the Appendix.

Consumer's Problem. The optimal consumption and labor choices are given by:

$$\frac{\dot{E}_t}{E_t} = r_t - \rho, \tag{3.8}$$

$$E_t = P_t C_t = C_{i,t} P_{i,t}, \tag{3.9}$$

$$c_{j,i,t} = C_{i,t}\left(\frac{P_{i,t}}{p_{j,i,t}}\right)^{\sigma}, \quad C_{f,i,t} = C_{i,t}\left(\frac{P_{i,t}}{P_{f,i,t}}\right)^{\sigma}, \tag{3.10}$$

$$\frac{\gamma}{1-L_t} = \frac{(1-\tau_L)w_t}{(1+\tau_c)E_t}. \tag{3.11}$$

where $E_t = P_t C_t$ is expenditure, $P_{i,t}$ is the price index of the industry, given by $P_{i,t} = \left(\sum_{j=1}^{n_{i,t}} p_{j,i,t}^{1-\sigma} + P_{f,i,t}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$, and, given that all small firms in the fringe produce the same homogeneous good, the price index of the fringe, $P_{f,i,t}$, equals the price set by each small firm (i.e. $P_{f,i,t} = p_{f,i,t}$). Equation 3.11 governs the consumption-leisure choice, and states that as taxes, either on consumption or labor, increase, the labor supplied decreases.

Quantity setting. Market clearing implies that the quantities consumed equal the quantities produced (i.e. $y_{j,i,t} = c_{j,i,t}$, $Y_{f,i,t} = C_{f,i,t}$, $Y_{i,t} = C_{i,t}$, and $Y_t = C_t$). Superstar firms within the same industry compete a la Cournot. They choose quantity $y_{j,i,t}$ to maximize their profits $\pi_{j,i,t} = y_{j,i,t}p_{j,i,t} - w_t \ell_{j,i,t}$, given the quantities of the other superstars and the fringe, and internalizing that its choice has an effect on the industry quantity $Y_{i,t}$. This delivers an optimal markup that is increasing in the market share of the firm, $s_{j,i,t} = \frac{p_{j,i,t}y_{j,i,t}}{P_{i,t}Y_{i,t}}$:

$$\mathcal{M}_{j,i,t} = \frac{\sigma}{\sigma-1}\left[1 - s_{j,i,t}\right]^{-1}. \tag{3.12}$$

Superstar value function and innovation One way to express the industry state is by the vector of productivity gaps from the leader $\theta = \{a_1, \ldots, a_N\}$, where $a_R = \ln\left(\frac{A_1}{A_R}\right)\frac{1}{\ln\lambda}$ (that is, the productivity of the industry leader is $A_1 = A_R\lambda^{a_R}$).[5] This vector not only gives us information about the relative productivities, but also about the number of superstars in the industry. A key feature of the firm problem is the presence of a fixed cost. Due to the fixed cost, some superstars may find that their value of continuing operations is negative. Here, two considerations need to be made. First, as it is well known, there may not be a unique Nash equilibrium if firms decide simultaneously whether to stay or exit; therefore, we impose a sequentiality on the exit decision, in particular assuming that firms with higher value choose

[5] This is just one way to write the sufficient information needed to specify the industry state; another possibility (which is the one we use for the code) is to state it in consecutive gaps, in which case we would define $a_R := \ln\left(\frac{A_R}{A_{R+1}}\right)\frac{1}{\ln\lambda}$, $R = 1, \ldots, n-1$.

earlier whether to stay or exit (solving by backward induction, this is equivalent to say that firms with lower productivity exit first).[6] Second, to solve the model in general, we need to allow for mix strategies in the exit decision; that is, a firm may randomise with some probability whether to stay or exit the moment it becomes indifferent (i.e. when its value becomes 0). If the result of the randomisation is to stay, the firm stays at least until the industry state changes again.

To make their exit decisions, firms need to compute their value of continuing operations, $V^S_{j,\theta,t}$; that is, the value of firm j in an industry of type θ if all firms in the industry decide to stay. In contrast, let $V_{j,\theta,t}$ be the expected value of firm j when the industry state becomes θ (before exit movements). That is, letting $\theta = (a_1, \ldots, a_{N_\theta})$, $\theta^{(k)} = (a_1, \ldots, a_k)$ (i.e. θ truncated at k), $\hat{N}_\theta \in [0, N_\theta]$ be the expected number of firms that stay when the state θ is reached, and $\hat{j}_\theta \in \mathbb{N}$ be the greatest integer smaller or equal than $\hat{N}_\theta$; then: $V_{j,\theta,t} = (\hat{N}_\theta - \hat{j}_\theta)V^S_{j,\theta^{(\hat{j}_\theta+1)},t} + (\hat{j}_\theta + 1 - \hat{N}_\theta)V^S_{j,\theta^{(\hat{j}_\theta)},t}$ for $j \leq \hat{j}_\theta$, and $V_{j,\theta,t} = 0$ otherwise.

The firm problem to obtain the values of continuing operations, $V^S_{j,\theta,t}$, can be expressed as the following Hamilton-Jacobi-Bellman equation (the detailed derivations are in the Appendix 3.6.2.4):

$$\rho V^S_{j,\theta,t} = \max_{h_{j,\theta,t}} \Bigg\{ E_t s_{j,\theta}(1 - \alpha \mathcal{M}^{-1}_{j,\theta}) - w_t(f + h_{j,\theta,t}) + \sum_{k=1}^{N_\theta} \rho_{k,\theta,t}\Big(V_{j'(\theta,k),\theta'(\theta,k),t} - V^S_{j,\theta,t}\Big) + \bar{\rho}_{f,\theta,t}\Big(V_{j'(\theta,f),\theta'(\theta,f),t} - V^S_{j,\theta,t}\Big)\Bigg\}$$

$$\text{s.t.} \quad \rho_{j,\theta,t} = Z h^{\psi}_{j,\theta,t}.$$

The first two terms inside the brackets are production profits and research and fixed costs. The third term captures that the firm internalises that with arrival rate $\rho_{k,\theta,t}$ the incumbent firm k will innovate, in which case the industry state will become $\theta'(\theta, k)$, which doesn't take into account exit (as described above, exit is taken into account by $V_{j'(\theta,k),\theta'(\theta,k),t}$), and $j'(\theta, k)$ is the new position of firm j in the ranking of productivities.[7] More explicitly, $\theta'(\theta, k)$ is obtained in the following steps. First, if the leader innovates, then the productivity gap of all the followers increases by 1 (i.e. a_R becomes $a_R + 1$ for all $R > 1$); whereas if the innovator is the incumbent $R > 1$, then only the productivity gap of firm R is reduced by 1 (i.e. a_R becomes $a_R - 1$).

[6]Note that given that what happens in one industry state affects the value in other states, this sequentiality is defined not only within an industry state but across industry states. That is: the firm with the lowest value across all industry states exit first.

[7]For instance, if $\theta = \{0, 2, 2\}$ and firm 3 innovates, then it moves up to position 2 in the ranking.

Second, and this is just to avoid redundant industry states, we reorder firms from lower to higher productivity gap, and $j'(\theta, k)$ maps the previous position on the ranking to the new one. Finally, the fourth term in the HJB equation captures that with arrival rate $\bar{\rho}_{f,\theta,t} = m\rho_{f,\theta,t}$ some small firm will become a superstar, in which case the new industry state (before exit movements) is $\theta'(\theta, f)$. The only thing that changes from the previous explanation is the first step. In this case, the productivity gaps of the incumbents are unchanged, but we need to add to θ a new productivity gap for the entrant, a_e.

Given that, as we will see $\frac{w_t}{E_t}$ is constant in the stationary equilibrium and that production profits are linear in E_t

$$\pi_{j,\theta,t} = E_t s_{j,\theta}(1 - \alpha \mathcal{M}_{j,\theta}^{-1}).$$

Then, using guess and verify, it is straightforward to check that the value function is linear in E_t, $V_{j,\theta,t} = \hat{V}_{j,\theta} E_t$.

The optimal amount of researchers employed by firm j if all firms in industry θ stay is given by

$$h_{j,\theta}^{1-\psi} = Z\psi \frac{E}{w}\Big(\hat{V}_{j'(\theta,j),\theta'(\theta,j)} - \hat{V}_{j,\theta}^{S}\Big). \tag{3.13}$$

See Figure 2.1 for the model-implied relationship between industry market share and research, as well as between total industry research and the industry Herfindahl index.

When computing the values of continuing operations, $\hat{V}_{j,\theta}^{S}$, using the system of HJB equations, firms need to take into account the exit decisions of firms at each industry state to compute $\hat{V}_{j,\theta}$ (in particular, as stated above, they need the expected number of firms that stay, $\hat{N}_\theta$). In the equilibrium of the firms problem, the expected number of firms that stay at each industry state, $\hat{N}_\theta$, must be consistent; that is, if $\hat{N}_\theta < N_\theta$, then $\hat{V}_{N_\theta,\theta}^{S} \leq 0$, and if $\hat{N}_\theta = N_\theta$, then $\hat{V}_{N_\theta,\theta}^{S} \geq 0$. Otherwise, the lowest productivity superstar firm would have an incentive to deviate from its exit decision.

Small firm value function and entry into superstars. Given that there is a continuum of small firms producing the same homogeneous good, small firms are price takers and $\mathcal{M}_{f,i} = 1$. As with superstars, small firms' production profits are linear in E_t

$$\pi_{f,\theta,t} = E_t(1 - \alpha)\frac{s_{f,\theta}}{m},$$

where I have dropped all subindices from the mass of small firms m because, as we will see, in the stationary equilibrium it is constant and the same across industries.

So, the value of small firms is also linear in E_t, and they solve:

$$
\begin{aligned}
(\rho+\delta)v^f(\theta) = \max_{\{h_{f,\theta}\}} \Big\{ & (1-\alpha)\frac{s_{f,\theta}}{m} - \frac{w}{E}h_{f,\theta} + \rho_{f,\theta}\hat{V}_{f'(\theta,f),\theta'(\theta,f)} \\
& + \sum_{j=1}^{N_\theta} \rho_{j,\theta,t}\Big(v^f(\theta'(\theta,j)) - v^f(\theta)\Big) + \bar{\rho}_{f,\theta}\Big(v^f(\theta'(\theta,f)) - v^f(\theta)\Big) \Big\} \\
\text{s.t. } \rho_{f,\theta} = Z_f h_{f,\theta}^{\psi_f}, &
\end{aligned}
$$

where $v^f(\theta)$ is the value of a small firm in industry state θ, $\bar{\rho}_{f,\theta} = \int \rho_{f,\theta} df = m\rho_{f,\theta}$, and $\hat{V}_{f'(\theta,f),\theta'(\theta,f)}$ is the value of becoming a new superstar.[8] The first term are the profits of a small firm in that industry. The next two terms are the costs and gains of innovating to become a new superstar, whereas the other two terms capture the change in the value of remaining in the fringe when some superstar innovates and when another small firm becomes a superstar, respectively. The optimal research choice is given by:

$$
h_{f,\theta}^{1-\psi_f} = Z_f \psi_f \frac{E}{w}\hat{V}_{f'(\theta,f),\theta'(\theta,f)}. \tag{3.14}
$$

Entry into small firms. The entrepreneurs solve the value problem

$$
\rho v^e = \max_{h_e} \left\{ \rho_e \sum_\theta v^f(\theta)\mu(\theta) - \frac{w}{E}h_e \right\}, \quad \text{s.t. } \rho_e = Z_e h_e^{\frac{1}{2}},
$$

where $\mu(\theta)$ is the mass of industries in state θ. That is, successful entrepreneurs become a small firm in a randomly allocated industry. So, the optimal research by entrepreneurs is given by:

$$
h_e^{\frac{1}{2}} = \frac{Z_e}{2}\frac{E}{w}\sum_\theta v^f(\theta)\mu(\theta). \tag{3.15}
$$

In the stationary equilibrium, the mass of small firms must be constant, that is, the mass of entrants into small firms must equal the mass of small firms who exit; that is: $\rho_e = m\delta$.

3.2.2.1 Law of Motion of the Distribution of Industry States

With the optimal research and entry decisions, we can obtain the infinitesimal transition matrix. Let $\mu_t(\theta)$ be the mass of industries with state θ at time t (post exit movements), $p(\theta, \theta')$ be the time derivative of the probability of transiting from industry state θ to θ' (pre exit movements),

[8] The value of a small firm is infinitesimal, the value of small firms in θ is $\int_f v^f(\theta)df = mv^f(\theta)$.

and $G(\theta, \theta')$ be the probability that exit movements turn θ into θ'. That is, using the same notation as in section 3.2.2, then: (i) $G(\theta, \theta') = \hat{N}_\theta - \hat{j}_\theta$ if $\theta' = \theta^{(\hat{j}_\theta+1)}$ with $\hat{j}_\theta + 1 \leq N_\theta$; (ii) $G(\theta, \theta') = 1 + \hat{j}_\theta - \hat{N}_\theta$ if $\theta' = \theta^{(\hat{j}_\theta)}$; and (iii) $G(\theta, \theta') = 0$ otherwise.
Since in the stationary distribution it must be $\mu_t(\theta) = \mu(\theta)$, then we have (see Appendix 3.6.3 for the detailed derivation):

$$0 = \frac{d\mu_t(\theta)}{dt} = \sum_{\theta' \in \Theta} \mu_t(\theta') \sum_{\theta'' \neq \theta'} p(\theta', \theta'') G(\theta'', \theta) + \mu_t(\theta) p(\theta, \theta) \quad , \quad \theta \in \Theta. \tag{3.16}$$

And $p(\theta, \theta')$ is non-zero only when (i) $\theta' = \theta$, $p(\theta, \theta) = -\sum_j^{N_\theta} \rho_{j,\theta} - \rho_{e,\theta}$, or when (ii) θ' is the state we obtain from θ when only one new firm becomes a superstar and no incumbent innovates (and before any potential exit), or when (iii) only one incumbent j innovates and there is no new superstar (and before any potential exit).
Calling Q the infinitesimal matrix and ordering the industry states $\Theta = \{\theta_1, \ldots, \theta_T\}$, then the element in row i and column j of Q is $(Q)_{i,j} = \sum_{k \neq i} p(\theta_i, \theta_k) G(\theta_k, \theta_j)$ if $j \neq i$ and $(Q)_{i,i} = p(\theta_i, \theta_i) + \sum_{k \neq i} p(\theta_i, \theta_k) G(\theta_k, \theta_i)$, and (3.16) can be written as $Q^t \vec{\mu} = \vec{0}$. From the above description, it is clear that $\sum_{j=1}^T Q_{i,j} = 0$;[9] so, Q has not full rank. The extra condition we need to solve for the stationary distribution is $\sum_{i=1}^T \mu(\theta_i) = 1$.

3.2.2.2 Labour Market Clearing

Labour is used in production by superstars and small firms, and in research by superstars, small firms, and entrepreneurs. So, the labour market clearing condition writes:

$$L_S = \sum_{\theta \in \Theta} \mu(\theta) \left(\sum_{j=1}^{N_\theta} (L_{j,\theta} + h_{j,\theta} + f) + L_{f,\theta} + m h_{f,\theta} \right) + h_e, \tag{3.17}$$

where the labour supply L_S is given by $L_S = 1 - \gamma \frac{(1+\tau_c)}{(1-\tau_L)} \frac{E}{w}$. Note that in the stationary equilibrium $\frac{E}{w}$ is constant.

3.2.2.3 Aggregate Production Function and Growth

As derived in the Appendix 3.6.5, the aggregate production function can be written as:

[9] $\sum_{j=1}^T Q_{i,j} = p(\theta_i, \theta_i) + \sum_{j=1}^T \sum_{k \neq i} p(\theta_i, \theta_k) G(\theta_k, \theta_j) = p(\theta_i, \theta_i) + \sum_{k \neq i} p(\theta_i, \theta_k) \sum_{j=1}^T G(\theta_k, \theta_j) = \sum_{k=1}^T p(\theta_i, \theta_k) = 0.$

$$\ln Y_t = \alpha \ln\left(\alpha\frac{E_t}{w_t}\right) + \int_0^1 \ln A_{1,i,t}di + \frac{\sigma}{\sigma-1}\sum_{\theta\in\Theta}\mu_t(\theta)\ln\left(\left(\zeta m^{1-\alpha}s_{f,\theta}^{\alpha}\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j=1}^{N_\theta}\left(\hat{A}_{j,\theta}\left(\frac{s_{j,\theta}}{\mathcal{M}_{j,\theta}}\right)^{\alpha}\right)^{\frac{\sigma-1}{\sigma}}\right),$$

which decomposes the aggregate output in three terms. The first term is linked to the relative wage; the second is the geometric mean of the industry leaders' productivity level; and the third refers to the composition of industry states. In a stationary equilibrium, $\mu_t(\theta) = \mu(\theta)$, as well as $\frac{E_t}{w_t}$ is constant; and so in the stationary equilibrium, all the growth comes from the second term. In particular, the growth rate in the stationary distribution is:

$$g = \frac{\dot{Y}_t}{Y_t} = \int_0^1 \frac{\dot{A}_{1,i,t}}{A_{j,i,t}}di = \ln\lambda\sum_{\theta}\mu(\theta)\rho_{1,\theta}. \tag{3.18}$$

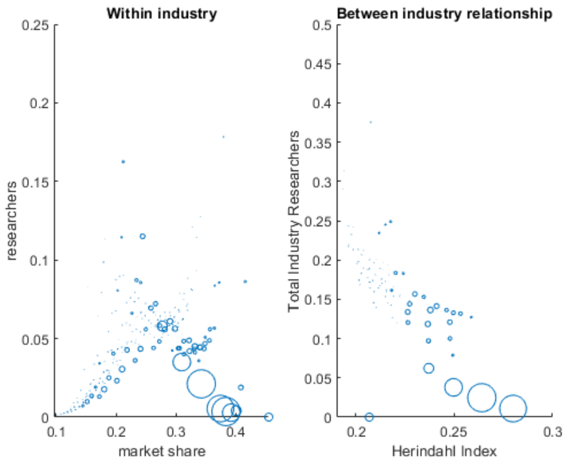
That is, the economy grows at a rate equal to the average probability a leader innovates times the log of the innovation step size.

3.3 Quantitative Analysis

3.3.1 Calibration

In this section, we describe the calibration of the model. We estimate the model for France at an annual frequency and set the discount rate to $\rho = 0.04$. For the parameters that computationally bound the problem, we set the maximum number of superstars to $\bar{N} = 4$ and the maximum number of steps behind the leader to $\bar{a} = 5$. Additionally, we set the number of gaps for a new superstar to $a_e = 3$. We also set the degree of decreasing returns to scale to $\alpha = 0.9$, following the value used in the Appendix of Cavenaile et al. (2021), and the death rate of small firms in the fringe to $\delta = 0.0531$, corresponding to the entry rate in the data. This leaves us 10 parameters to estimate: the elasticity of substitution parameter, σ; the relative productivity of the fringe, ζ; the consumption-leisure parameter, γ; the innovation step-size, λ; the innovation scale parameters of superstars, small firms, and entrepreneurs, Z, Z_f, Z_e, respectively; the innovation convexity parameters of superstars and small firms, ψ, ψ_f, respectively; and the operating fixed cost, f. These parameters are estimated jointly through a Simulated Method of Moments estimation procedure. In particular, given a point in the space $\{\sigma, \zeta, \lambda, Z, Z_f, \psi, \psi_f, f\}$, γ is set such that employment in the model perfectly matches the employment rate in the data, and Z_e is set such that the mass of the fringe is normalized to 1. The other 8 parameters are set to target: the sales-weighted average and variance of the

Figure 2.1: Research within and between industries



Notes. On the left, we plot the researchers employed by superstars, $h_{j,\theta}$ for each j and θ, against the firm's industry market share. On the right, total researchers demanded by superstars, $\sum_j h_{j,\theta}$ for each θ, against the Herfindahl index. The size of each circle is proportional to the frequency we observe it in the stationary distribution.

log of markups using FARE, the labor share, the growth rate, the average share of $R\&D$ to sales for manufacturing firms from EU KLEMS, the average share of fixed costs to total firm costs, taken from De Ridder (2024), and the linear coefficient and the implied top point from a regression of innovation expenditure and relative sales and its square. For more details on how these moments are calculated in the model, see Appendix 3.6.4.

Table 2.1 summarizes the results of the calibration. Panel A reports the parameter values, while Panel B reports both the model-implied moments and the empirical ones. Overall, the model does well in matching the moments.

Figure 2.1 shows, on the left, the (targeted) inverted-U relationship between $R\&D$ spending and market share in the stationary equilibrium of the calibrated model; on the right, it shows that competition is positively related to overall industry $R\&D$ spending. In both panels, the size of each circle is proportional to its frequency in the stationary equilibrium.

3.3.2 Fiscal Policy Effects: the Role of Market Structure Endogeneity

A sufficient statistic to study the effect of a change in either taxes on consumption or labor is given by $\tau = (1 - \tau_L)/(1 + \tau_C) - 1$; therefore, we will refer as fiscal policy as changing τ. To

Table 2.1: Parameter values and targeted moments

A. Parameters

Parameter	Description	Value
ρ	Discount rate	0.0400
σ	Elasticity of substitution between superstars and fringe	9.7139
$\ln\left(\frac{A_{1,i,t}}{A_{f,i,t}}\right)$	Gap between leader and fringe	0.6045
γ	Consumption-Leisure parameter	0.2306
$\ln\lambda$	Innovation step-size	0.0826
Z	Incumbents' innovation scale parameter	0.8033
ψ	Incumbents' innovation convexity parameter	0.4574
Z_f	Small firms' innovation scale parameter	0.1651
ψ_f	Small firms' innovation convexity parameter	0.5393
Z_e	Entrants' innovation scale parameter	0.3638
δ	Small firms' exit rate	0.0531
f	fixed cost	0.0150
α	Decreasing returns production	0.9000

B. Moments

Moment	Data	Model
Average of log(Markup)	0.39	0.4023
Variance of log(Markup)	0.25	0.2503
Employment rate	0.75	0.7500
Labor share	0.65	0.6922
Growth rate	0.93%	0.9283%
R&D share	6%	5.9724%
Entry rate	5.31%	5.31%
β_1 (regression innovation, relative sales)	1.66	1.6612
Top point (regression innovation, relative sales)	0.3673	0.2810
Fixed cost as share of costs	0.1030	0.1030

Notes. Panel A reports the parameter values. Panel B reports the simulated and empirical moments. Details on how these moments are computed can be found in the Appendix 3.6.4

examine the role of market structure endogeneity in the response of the economy to fiscal policy, we compare the effects of increasing τ from 0 to $\tau_0 = 0.0301$, where τ_0 is the tax that achieves a tax revenue equivalent to 5% of GDP,[10] in the baseline model, as well as in a counterfactual economy where market structure is exogenous. In particular, to obtain this counterfactual, we

[10]That is, assuming $\tau_C = \tau_L = \tau_0$, then $\tau_0 wL + \tau_0 E = 0.05Y$.

fix: (i) the stationary distribution, $\{\mu(\theta)\}_{\theta\in\Theta}$; (ii) the endogenous exit decisions of superstars, $\{\hat{N}_\theta\}_{\theta\in\Theta}$; and (iii) the mass of firms in the fringe, m.

Table 2.2 presents the results of this experiment. We see that the effect on growth would be 10% more muted if the role of endogenous market structure were neglect. The increase in τ leads to a reduction in labor supply, which in turn reduces market size. This lowers profitability, discouraging innovation by superstars as well as innovation to become a superstar or entering the fringe. As a result, both the number of superstars and the mass of the fringe decline. As shown in Figure 2.1, in the calibrated model, more competitive industries are positively associated with innovation intensity. Thus, these changes in market structure—toward less competitive industries—amplify the negative effect of an increase in τ on growth. Finally, note that the effect on both research by the fringe and by entrepreneurs is more muted in the baseline than in the counterfactual. As said above, the increase in τ decreases the incentives to innovate both for small firms and entrepreneurs (this is true both in the baseline and the counterfactual); however, in the baseline, this leads to a shift towards industries with fewer superstars and smaller fringe, and in industries with fewer firms the incentives to enter are larger, which mitigates the effect on h_f and h_e in the baseline.

Table 2.2: Effects of Fiscal Policy with Endogenous vs. Exogenous Market Structure

	Δ **g**	$\Delta\sum_{\theta,j}\mu(\theta)s_{j,\theta}\mathcal{M}_{j,\theta}$	Δm	$\Delta\sum_{\theta}\mu(\theta)N_\theta$	$\Delta\sum_{\theta}\mu(\theta)h_{f,\theta}$	Δh_e
Baseline	-0.0045	0.0002	-0.0039	-0.0011	-0.0000	-0.0002
Exogenous Market structure	-0.0040	0	0	0	-0.0001	-0.0003

Notes. We compare the effects of increasing τ (where $\tau = (1-\tau_L)/(1+\tau_C)-1$) from 0 to 0.0301 under the baseline economy, where market structure adjusts endogenously, with a counterfactual scenario in which market structure remains fixed.

3.3.3 Policy complementarity

Here, we examine whether there is complementarity between fiscal and competition policies. For competition policy, we consider, on the one hand, subsidies to innovation by small firms aiming to become superstars, $\tau_f w h_{f,\theta}$, and, on the other hand, subsidies to entrepreneurs for creating new small firms, $\tau_e w h_e$. To assess the complementarity between policies, we analyze how the effect of an increase in taxes, τ, from 0 to τ_0, changes when we introduce one of the following competition policies: (i) a subsidy to innovation by small firms, $\tau_f = \tau_{f,0}$; (ii) a subsidy to entrepreneurs, $\tau_e = \tau_{e,0}$; or (iii) both a subsidy to small firms and to entrepreneurs, $\tau_f = \tau_e = \tau_{f,1}$. In particular, as in the previous section, τ_0, $\tau_{f,0}$, $\tau_{e,0}$, and $\tau_{f,1}$ are calibrated so that, when implemented individually, they imply tax revenue (or public spending) equivalent to 5% of GDP.

Table 2.3 presents the effects of fiscal and competition policies on key aggregates when implemented separately and in combination. Focusing on the growth rate, we observe that both types of subsidies—to innovation by small firms and to entrepreneurs—amplify the effect of fiscal policy. For instance, the increase in τ causes the growth rate to decline by 0.0064 (= 0.2514 – 0.2578) when both subsidies are in place, compared to a decline of 0.0045 when no other policy is implemented. Similarly, we find that the effect of fiscal policy on the mass of the fringe is amplified particularly in the presence of a subsidy to entrepreneurs, while the effect on the number of superstars is amplified especially when there is a subsidy to small firms aiming to become superstars. Finally, note that a subsidy to innovation by small firms actually discourages entrepreneurs from creating new firms, thereby reducing the mass of the fringe. Although entrepreneurs get some gains from the subsidy, as they internalize that becoming a superstar is easier, these gains are dominated by the negative effect on their value resulting from the increase in the number of superstars, which reduces the profits of firms in the fringe as well as those of a prospective new superstar.

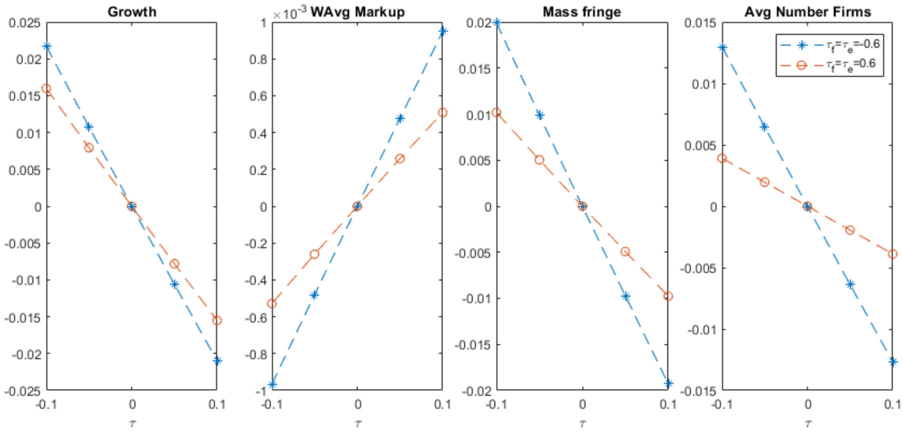
Figure 2.2 shows visually the same result that the effect of fiscal policy is amplified (i.e. steeper slope) in the presence of competition policies that promote entry.

Table 2.3: Policy Complementarity

	Δ g	$\Delta \sum_{\theta,j} \mu(\theta) s_{j,\theta} \mathcal{M}_{j,\theta}$	Δm	$\Delta \sum_{\theta} \mu(\theta) N_\theta$	$\Delta \sum_{\theta} \mu(\theta) h_{f,\theta}$	Δh_e
$\tau = \tau_0$	-0.0045	0.0002	-0.0039	-0.0011	-0.0000	-0.0002
$\tau_f = \tau_{f,0}$	0.3221	-0.0282	-0.0775	0.8576	0.0562	-0.0032
$+ \tau = \tau_0$	0.3159	-0.0279	-0.0808	0.8532	0.0555	-0.0033
$\tau_e = \tau_{e,0}$	0.0374	-0.0430	0.9227	0.1788	0.0013	0.0574
$+ \tau = \tau_0$	0.0324	-0.0428	0.9152	0.1770	0.0012	0.0568
$\tau_f = \tau_e = \tau_{f,1}$	0.2578	-0.0493	0.6228	0.7172	0.0222	0.0348
$+ \tau = \tau_0$	0.2514	-0.0490	0.6167	0.7130	0.0219	0.0344

Notes. We compare the effects of different tax structures individually and in combination with $\tau = \tau_0$, where $\tau = (1-\tau_L)/(1+\tau_C) - 1$, and $\tau_0 = 0.0301$, $\tau_{f,0} = -0.8165$, $\tau_{e,0} = -0.7131$, and $\tau_{f,1} = -0.6393$.

Figure 2.2: Comparing Fiscal Policy Effects Under Entry Subsidies and Taxes



Notes. This figure shows the changes (relative to the case $\tau = 0$) in growth (measured in percentage points), sales-weighted average markup, mass of fringe firms, and the number of superstar firms, for five equally spaced values of the effective tax rate, from $\tau = -0.1$ to $\tau = 0.1$. The blue curve represents a subsidy to entry ($\tau_f = \tau_e = -0.6$), while the orange curve represents a tax on entry ($\tau_f = \tau_e = 0.6$).

3.4 Conclusion

This paper develops a step-by-step growth model with endogenous market structure to study the role of competition in shaping the effects of fiscal policy on long-run growth, and the complementarities between fiscal and competition policies. We find that the market structure response amplifies the effect of fiscal policy.

In future work, we plan to incorporate government expenditure and debt into the model to explore fiscal consolidation, as outlined in the introduction. Additionally, we will explore transitional dynamics, which will help to capture the adjustment process toward the stationary equilibrium.

3.5 References

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

_ , **Emmanuel Farhi, and Enisse Kharroubi**, "Monetary Policy, Product Market Competition and Growth," *Economica*, 2019, *86*, 431–470.

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

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

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

Balasundharam, Vybhavi, Olivier Basdevant, Dalmacio Benicio, Andrew Ceber, Yujin Kim, Luca Mazzone, Hoda Selim, and Yongzheng Yang, "Fiscal Consolidation: Taking Stock of Success Factors, Impact, and Design," IMF Working Paper WP/23/63, International Monetary Fund March 2023.

Bernanke, Ben and Mark Gertler, "Inside the Black Box: the Credit Channel of Monetary Policy Transmission," *Journal of Economic Perspectives*, 1995.

Cavenaile, Laurent, Murat Celik, and Xu Tian, "Are Markups Too High? Competition, Strategic Innovation, and Industry Dynamics," *mimeo*, 2021.

Diez, Federico, Jiayue Fan, and Carolina Villegas-Sanchez, "Global Declining Competition?," *Journal of International Economics*, September 2021, *132* (1), 100–112.

Ferrando, Annalisa, Peter McAdam, Filippos Petroulakis, and Xavier Vives, "Product Market Structure and Monetary Policy Evidence from the Euro Area," *ECB Working Paper Series*, 2021, *2632.*

— , — , — , and — , "Product Market Structure and Monetary Policy Evidence from the Euro Area," *American Economic Review PP*, forthcoming.

Griffith, Rachel and John Van Reenen, "Product Market Competition, Creative Destruction and innovation," POID working paper 2021.

IMF, "G20 Surveillance Note," Technical Report, International Monetary Fund 2020.

Loecker, Jan De, Jan Eckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 2020, *135* (1), 581–644.

Mongey, Simon, "Market Structure and Monetary Policy Non-Neutrality," *NBER Working*

Papers, 2021, *N. 29233*, 257–283.

Ridder, Maarten De, "Market Power and Innovation in the Intangible Economy," *American Economic Review*, 2024, (forthcoming).

Wang, Olivier and Ivan Werning, "Dynamic Oligopoly and Price Stickiness," *American Economic Review*, 2022, *112*, 2815–49.

3.6 Appendix

3.6.1 Preferences with taxes (on consumption and leisure)

$$U = \max_{\{c_{j,it}, a_{t+1}, L_t\}} \int_0^{\infty} e^{-\rho t} \; [\ln C_t + \gamma \ln(1 - L_t)] \, dt$$

$$\text{s.t. } \dot{a}_t = r_t a_t + (1 - \tau_L) w_t L_t - (1 + \tau_c) P_t C_t + TR_t$$

$$\ln C_t = \int_0^1 \ln C_{i,t} di \ , \quad C_{i,t} = \left(C_{f,i,t}^{\frac{\sigma-1}{\sigma}} + \sum_{j=1}^{n_{i,t}} c_{j,i,t}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad , \quad \sigma > 1$$

Writing the present-value Hamiltonian, the optimal solution of this problem is given by the optimality conditions:

$$[c_{j,i,t}] : \ \left[e^{-\rho t} \frac{1}{C_t} - \mu_t (1 + \tau_c) P_t \right] \frac{\partial C_t}{\partial c_{j,i,t}} = 0 \implies \mu_t = e^{-\rho t} \frac{1}{(1 + \tau_c) E_t}$$

$$[C_{f,i,t}] : \ e^{-\rho t} \, C_{i,t}^{-1} C_{i,t}^{\frac{1}{\sigma}} C_{f,i,t}^{-\frac{1}{\sigma}} - \mu_t (1 + \tau_c) P_{f,i,t}$$

$$[a_t] : \ -\dot{\mu}_t = \mu_t r_t \implies \frac{\dot{E}_t}{E_t} = r_t - \rho$$

$$[L_t] : e^{-\rho t} \frac{\gamma}{1 - L_t} = \mu_t (1 - \tau_L) w_t \implies \frac{\gamma}{1 - L_t} = \frac{(1 - \tau_L)}{(1 + \tau_c)} \frac{w_t}{E_t}$$

The condition of $c_{j,i,t}$ can also be written as:

$$[c_{j,i,t}] : \ \left[e^{-\rho t} \, C_{i,t}^{-1} - \mu_t (1 + \tau_c) P_{i,t} \right] \frac{\partial C_{i,t}}{\partial c_{j,i,t}} = 0$$

which substituting for $\mu_t = e^{-\rho t} \frac{1}{(1+\tau_c) C_t P_t}$ implies:

$$E_t = P_t C_t = C_{i,t} P_{i,t}$$

Plugging into C_t: $\ln C_t = \ln C_t + \ln P_t - \int_0^1 \ln P_{i,t} di$, so: $\ln P_t = \int_0^1 \ln P_{i,t} di$.

Finally, the condition of $c_{j,i,t}$ can be written as:

$$[c_{j,i,t}] : \ e^{-\rho t} \, C_{i,t}^{-1} C_{i,t}^{\frac{1}{\sigma}} c_{j,i,t}^{-\frac{1}{\sigma}} - \mu_t (1 + \tau_c) p_{j,i,t} = 0$$

which substituting for $\mu_t = e^{-\rho t} \frac{1}{(1+\tau_c) C_t P_t}$, and $P_t C_t = C_{i,t} P_{i,t}$, we have:

$$C_{i,t}^{\frac{1}{\sigma}} c_{j,i,t}^{-\frac{1}{\sigma}} = \frac{p_{j,i,t}}{P_{i,t}} \implies c_{j,i,t} = C_{i,t} \left(\frac{P_{i,t}}{p_{j,i,t}} \right)^{\sigma}$$

Analogously, we get $C_{f,i,t} = C_{i,t} \left(\frac{P_{i,t}}{P_{f,i,t}} \right)^{\sigma}$. And plugging it into $C_{i,t}$:

$$C_{i,t} = C_{i,t} P_{i,t}^{\sigma} \left(P_{f,i,t}^{1-\sigma} + \sum_{j=1}^{n_{i,t}} p_{j,i,t}^{1-\sigma} \right)^{\frac{\sigma}{\sigma-1}} \implies P_{i,t} = \left(P_{f,i,t}^{1-\sigma} + \sum_{j=1}^{n_{i,t}} p_{j,i,t}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

with $P_{f,i,t} = p_{f,i,t}$ because each small firm in the same industry produces the same good.[11]

[11] With monopolistic competition, we have: $P_{f,i,t} = \left(\int_0^{m_t} p_{f,i,t}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = m_t^{-\frac{1}{\sigma-1}} \frac{\sigma}{\sigma-1} \frac{w_t}{A_{1,i,t} \lambda^{-a_f}}$.

Market shares: Letting $s_{j,i,t} = \frac{p_{j,i,t}y_{j,i,t}}{P_{i,t}Y_{i,t}}$, using $y_{j,i,t} = Y_{i,t}\left(\frac{P_{i,t}}{p_{j,i,t}}\right)^{\sigma}$, and the definition of $Y_{i,t}$, we have:

$$s_{j,i,t} = \left(\frac{y_{j,i,t}}{Y_{i,t}}\right)^{1-\frac{1}{\sigma}} = \left[\left(\frac{Y_{f,i,t}}{y_{j,i,t}}\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j'=1}^{n_{i,t}}\left(\frac{y_{j',i,t}}{y_{j,i,t}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{-1}$$

3.6.2 Production Firms

3.6.2.1 Quantities and profits as function of shares and markups

Given that $y_{j,i} = A_{j,i}\ell_{j,i}^{\alpha}$, then $\ell_{j,i} = \left(\frac{y_{j,i}}{A_{j,i}}\right)^{\frac{1}{\alpha}}$; so, marginal cost equals $MgC_{j,i} = \frac{w}{\alpha}\frac{\ell_{j,i}}{y_{j,i}} = \frac{w}{\alpha}\frac{\ell_{j,i}^{1-\alpha}}{A_{j,i}}$. On the other hand, $\mathcal{M}_{j,i} = \frac{p_{j,i}}{MgC_{j,i}} = \frac{p_{j,i}y_{j,i}\alpha}{\ell_{j,i}w} \implies \ell_{j,i} = \alpha\frac{s_{j,i}}{\mathcal{M}_{j,i}}\frac{E}{w}$ So:

$$y_{j,i} = A_{j,i}\left(\frac{s_{j,i}}{\mathcal{M}_{j,i}}\frac{E}{w}\alpha\right)^{\alpha}$$

Analogously, for the fringe, since each small firm sets price equal to marginal cost (i.e. $p_{f,i} = \frac{w}{\alpha}\frac{\ell_{f,i}}{y_{f,i}} = \frac{w}{\alpha}\frac{\ell_{f,i}}{Y_{f,i}}m_i$, since by symmetry we have $Y_{f,i} = m_i y_{f,i}$); so, isolating $\ell_{f,i}$ from the price condition, we have: $\ell_{f,i} = \alpha\frac{s_{f,i}}{m_i}\frac{E}{w}$, where $s_{f,i} = \frac{p_{f,i}Y_{f,i}}{E}$; so:

$$Y_{f,i} = m_i y_{f,i} = m_i A_{f,i}\ell_{f,i}^{\alpha} = A_{f,i}m_i^{1-\alpha}\left(\alpha s_{f,i}\frac{E}{w}\right)^{\alpha}$$

For profits:

$$\pi_{j,i} = Es_{j,i} - w\ell_{j,i} = Es_{j,i} - E\alpha\frac{s_{j,i}}{\mathcal{M}_{j,i}} = E(1-\alpha\mathcal{M}_{j,i}^{-1})s_{j,i}$$

And for a small firm in the fringe:

$$\pi_{f,i} = p_{f,i}y_{f,i} - w\ell_{f,i} = E\frac{s_{f,i}}{m_i} - w\alpha\mathcal{M}_{f,i}^{-1}\frac{s_{f,i}}{m_i} = E(1-\alpha)\frac{s_{f,i}}{m_i}$$

3.6.2.2 Cournot Competition

$$\max_{y_{j,i,t}} \; \pi_{j,i,t} = Es_{j,i} - \frac{w}{A_{j,i}^{\frac{1}{\alpha}}}y_{j,i}^{\frac{1}{\alpha}}$$

$$\text{s.t. } p_{j,i,t}y_{j,i,t} = E_t s_{j,i,t}, \quad s_{j,i,t} = \left[\left(\frac{Y_{f,i,t}}{y_{j,i,t}}\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j'=1}^{n_{i,t}}\left(\frac{y_{j',i,t}}{y_{j,i,t}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{-1}$$

The derivative of the market share $s_{j,i,t}$ with respect to its own quantity $y_{j,i,t}$ is:

$$\frac{\partial s_{j,i,t}}{\partial y_{j,i,t}} = \frac{\sigma-1}{\sigma} s_{j,i,t}^2 \left[\left(\frac{Y_{f,i,t}}{y_{j,i,t}} \right)^{\frac{\sigma-1}{\sigma}} \frac{1}{y_{j,i,t}} + \sum_{k \neq j} \left(\frac{y_{k,i,t}}{y_{j,i,t}} \right)^{\frac{\sigma-1}{\sigma}} \frac{1}{y_{j,i,t}} \right]$$

$$= \frac{\sigma-1}{\sigma} \frac{s_{j,i,t}^2}{y_{j,i,t}} \left[\left(\frac{Y_{f,i,t}}{y_{j,i,t}} \right)^{\frac{\sigma-1}{\sigma}} + \sum_{k=1}^{n_{i,t}} \left(\frac{y_{k,i,t}}{y_{j,i,t}} \right)^{\frac{\sigma-1}{\sigma}} - 1 \right] = \frac{\sigma-1}{\sigma} \frac{s_{j,i,t}}{y_{j,i,t}} \left[1 - s_{j,i,t} \right]$$

So, the FOC is

$$\frac{\partial \pi_{j,i,t}}{\partial y_{j,i,t}} = E \frac{\partial s_{j,i,t}}{\partial y_{j,i,t}} - \frac{1}{\alpha} \frac{w}{A_{j,i,t}^{\frac{1}{\alpha}}} y_{j,i,t}^{\frac{1-\alpha}{\alpha}} = E \frac{\sigma-1}{\sigma} \frac{s_{j,i,t}}{y_{j,i,t}} \left[1 - s_{j,i,t} \right] - \frac{p_{j,i,t}}{\mathcal{M}_{j,i,t}} = 0$$

So, given that $s_{j,i,t} = \frac{p_{j,i,t} y_{j,i,t}}{E}$, the optimal markup is

$$\mathcal{M}_{j,i,t} = \frac{\sigma}{\sigma-1} \left[1 - s_{j,i,t} \right]^{-1}$$

3.6.2.3 Algorithm static solver Matlab:

We start with a guess of $\{s_{j,i}\}$, $s_{f,i}$. With these, compute the optimal markups $\{\mathcal{M}_{j,i}\}$ ($\mathcal{M}_{f,i} = 1$) using 3.12. With markups and shares, we can compute labor relative to one firm in the fringe $\ell_{j,i}^f = \frac{\ell_{j,i}}{\ell_{f,i}} = m_i \frac{s_{j,i}}{s_{f,i}} \frac{1}{\mathcal{M}_{j,i}}$. Then, the quantities relative to the fringe are: $y_{j,i}^f = \frac{y_{j,i}}{y_{f,i}} = \frac{A_{j,i}}{A_{f,i}} \frac{(\ell_{j,i}^f)^\alpha}{m_i}$. Finally, we can compute the shares (and iterate):

$$s_{j,i} = \left[\left(\frac{1}{y_{j,i}^f} \right)^{\frac{\sigma-1}{\sigma}} + \sum_{j'=1}^{N_i} \left(\frac{y_{j',i}^f}{y_{j,i}^f} \right)^{\frac{\sigma-1}{\sigma}} \right]^{-1}$$

3.6.2.4 Derivation of the HJB equation

Let $V_{j,\theta}^S$ be the value of firm j conditional on all firms of industry state θ staying, and $V_{j,\theta'}$ be the expected value of firm j in industry state θ' (it takes into account the exit strategy of firms in θ'). That is, the $V_{j,\theta}$'s can be computed knowing $V_{j,\theta}^S$'s and the exit strategies (following the expression in the main text). The $V_{j,\theta}^S$'s satisfy:

$$V_{j,\theta,t}^S = (\pi_{j\theta,t'} - w_{t'} h_{j,\theta,t'}) \Delta t + e^{-\int_t^{t+\Delta t} r_u du} \left(\sum_{\theta' \neq \theta} P_{\Delta t}(\theta \to \theta') V_{j,\theta',t+\Delta t} + P_{\Delta t}(\theta \to \theta) V_{j,\theta,t+\Delta t}^S \right),$$

where $t' \in (t, t+\Delta t)$ and $P_{\Delta t}(\theta \to \theta')$ the probability of moving from state θ to θ' in the time interval Δt.

$$V_{j,\theta,t}^S \left(1 - e^{-\int_t^{t+\Delta t} r_u du} \right) = (\pi_{j,\theta,t'} - w_{t'} h_{j,\theta,t'}) \Delta t$$
$$+ e^{-\int_t^{t+\Delta t} r_u du} \left(\sum_{\theta' \neq \theta} P_{\Delta t}(\theta \to \theta') V_{j,\theta',t+\Delta t} + P_{\Delta t}(\theta \to \theta) V_{j,\theta,t+\Delta t}^S - V_{j,\theta,t}^S \right).$$

For the left-hand side, note that the first order Taylor expansion gives us $1 - e^{-\int_t^{t+\Delta t} r_u du} = r_{t+\Delta t} \Delta t$. Then, dividing by Δt and making $\Delta \to 0$:

$$r_t V^S_{j,\theta,t} = \pi_{j,\theta,t'} - w_{t'} h_{j,\theta,t'} + \lim_{\Delta t \to 0} \frac{\sum_{\theta' \neq \theta} P_{\Delta t}(\theta \to \theta') V_{j,\theta',t+\Delta t} + P_{\Delta t}(\theta \to \theta) V^S_{j,\theta,t+\Delta t} - V^S_{j,\theta,t}}{\Delta t}.$$

Let's focus on the limit.

$$\begin{aligned}
&\lim_{\Delta t \to 0} \frac{\sum_{\theta' \neq \theta} P_{\Delta t}(\theta \to \theta') V_{j,\theta',t+\Delta t} + P_{\Delta t}(\theta \to \theta) V^S_{j,\theta,t+\Delta t} - V^S_{j,\theta,t}}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\sum_{\theta' \neq \theta} V_{j,\theta',t+\Delta t} \left(P_{\Delta t}(\theta \to \theta') - P_0(\theta \to \theta')\right) + V^S_{j,\theta,t+\Delta t}(P_{\Delta t}(\theta \to \theta) - P_0(\theta \to \theta)) + P_0(\theta \to \theta) \left(V_{j,\theta,t+\Delta t} - V^S_{j,\theta,t}\right)}{\Delta t} \\
&= \sum_{\theta' \neq \theta} V^E_{j,\theta',t} \frac{dP_0(\theta \to \theta')}{dt} + V^S_{j,\theta,t} \frac{dP_0(\theta \to \theta)}{dt} + \dot{V}^S_{j,\theta',t},
\end{aligned}$$

where for the first equality I have used that $P_0(\theta \to \theta') = 0$ and $P_0(\theta \to \theta) = 1$, and for the second equality I have used that if all firms in θ decided to stay at t, then they will also stay at $t + \Delta t$ if nothing has changed, i.e. $\lim_{\Delta t \to 0} V_{j,\theta,t+\Delta t} = V^S_{j,\theta,t}$.

Now, focus on the terms $\frac{dP_0(\theta \to \theta')}{dt}$. Suppose the J firms of the industry innovate $k_1, \ldots, k_J \in \mathbb{N}^J$ steps respectively. Then, the probability this happens in the interval Δt of time is

$$P_{\Delta t}(k_1, \ldots, k_J) = \prod_{j=1}^{J} e^{-\int_t^{t+\Delta t} \rho_{j,i,u} du} \frac{\left(\int_t^{t+\Delta t} \rho_{j,i,u} du\right)^{k_j}}{k_j!}.$$

And the derivative with respect to Δt is, making $\Delta t \to 0$:

$$\frac{dP_{\Delta t}(k_1, \ldots, k_J)}{d\Delta t} = 0 + \sum_{\{j | k_j = 1\}} \rho_{j,i,t} \prod_{j' \neq j} \frac{0^{k_{j'}}}{k_{j'}!} + \sum_{\{j | k_j = 0\}} (-\rho_{j,i,t}) \prod_{j' \neq j} \frac{0^{k_{j'}}}{k_{j'}!}.$$

So, for this derivative to be different than 0, it must be either no firm innovates, $k_j = 0$ for all j:

$$\frac{dP_{\Delta t}(0, \ldots, 0)}{d\Delta t} = -\sum_{j=1}^{J} \rho_{j,i,t},$$

or only one firm j innovates (note that here I am not distinguishing between incumbents and entrants):

$$\frac{dP_{\Delta t}(0, \ldots, 1, \ldots, 0)}{d\Delta t} = \rho_{j,i,t}.$$

I denote by $\theta'(\theta, j)$ the state we obtain when incumbent j achieves an innovation, and by $\theta'(\theta, f)$ the state we obtain when a small firm successfully becomes a superstar. So, the previous derivatives become:

$$\sum_{\theta' \neq \theta} V_{j,\theta',t} \frac{dP_0(\theta \to \theta')}{dt} + V^S_{j,\theta,t} \frac{dP_0(\theta \to \theta)}{dt} + \dot{V}^S_{j,\theta',t} = \sum_{k=1}^{N_\theta} \rho_{k,\theta,t} \left(V_{j,\theta'(\theta,k),t} - V^S_{j,\theta,t}\right) + \rho_{e,\theta,t} \left(V_{j,\theta'(\theta,e),t} - V^S_{j,\theta,t}\right) + \dot{V}^S_{j,\theta,t}$$

Finally, using the Euler equation $r_t = \rho + \frac{\dot{E}_t}{E_t}$, we get the HJB equation:

$$\left(\rho + \frac{\dot{E}_t}{E_t} - \frac{\dot{V}^S_{j,\theta,t}}{V^S_{j,\theta,t}}\right) V^S_{j,\theta,t} = \pi_{j,\theta,t} - w_t h_{j,\theta,t} + \sum_{k=1}^{N_\theta} \rho_{k,\theta,t} \left(V_{j,\theta'(\theta,k),t} - V^S_{j,\theta,t}\right) + \rho_{e,\theta,t} \left(V_{j,\theta'(\theta,e),t} - V^S_{j,\theta,t}\right)$$

and in a stationary equilibrium $\frac{\dot{E}_t}{E_t} = \frac{\dot{V}^S_{j,\theta,t}}{V^S_{j,\theta,t}}$

3.6.3 Stationary distribution

Let $P_{\Delta t}(\theta,\theta')$ be the probability of transiting from θ to θ' in the lapse of time Δt (before exit movements), and $G(\theta,\theta')$ the probability that exit movements turn θ into θ'. Then, from the condition that in the stationary distribution it must be $\mu_{t+\Delta t}(\theta)=\mu_t(\theta)=\mu(\theta)$, we have:

$$0=\frac{\mu_{t+\Delta t}(\theta)-\mu_t(\theta)}{\Delta t}=\frac{\sum\limits_{\theta'\in\Theta}\mu_t(\theta')\sum\limits_{\theta''\neq\theta'}P_{\Delta t}(\theta',\theta'')G(\theta'',\theta)+\mu_t(\theta)P_{\Delta t}(\theta,\theta)-\mu_t(\theta)}{\Delta t}$$
$$=\sum_{\theta'\in\Theta}\mu_t(\theta')\sum_{\theta''\neq\theta'}\frac{P_{\Delta t}(\theta',\theta'')-0}{\Delta t}G(\theta'',\theta)+\mu_t(\theta)\frac{P_{\Delta t}(\theta,\theta)-1}{\Delta t}$$

And so, since $P_0(\theta',\theta)=0$ when $\theta'\neq\theta$, and $P_0(\theta,\theta)=1$, when $\Delta t\to 0$ we have the expression of the main text.

Here we can see why we needed the assumption that firms only decide to exit when a change of state occurs: if there were some state with $\mu_t(\theta)>0$ and $G(\theta,\theta)<1$, then $\lim\limits_{dt\to 0^+}\mu_{t+dt}(\theta)=G(\theta,\theta)\mu_t(\theta)<\mu_t(\theta)$, and so $\frac{\mu_{t+\Delta t}(\theta)-\mu_t(\theta)}{\Delta t}$ would not be well-defined. That is, in such case, $\mu_t(\theta)$ would never be strictly positive for a positive interval of time, and so it would be just as if we just allowed firms to stay or exit with probability 1, and so in cases where in the baseline model we have $\mu(\theta)>0$ with $G(\theta,\theta)<1$, in this alternative setting we wouldn't have a stationary equilibrium.

3.6.4 Estimation

First, we present the details on how we compute the different moments in the model:

1. $m(1)$: Sales-weighted Average of log markups: $m(1)=\sum\limits_{X\in\Omega}\mu(X)\sum\limits_{j=0}^{J_X}s_{j,X}\ln\mathcal{M}_{j,X}$.
2. $m(2)$: Sales-weighted variance of log markups:
   $m(2)=\sum\limits_{X\in\Omega}\mu(X)\sum\limits_{j=0}^{J_X}s_{j,X}\left(\ln\mathcal{M}_{j,X}\right)^2-m(1)^2$
3. $m(3)$: Sales-weighted Average share of fixed costs:
   $m(3)=\sum\limits_{X\in\Omega}\mu(X)\left(s_{f,X}\frac{h_{f,X}}{h_{f,X}+L_{f,X}}+\sum\limits_{j=1}^{J_X}s_{j,X}\frac{f+h_{j,X}}{f+h_{j,X}+L_{j,X}}\right)$
4. $m(4)$: Labor share: $m(4)=\frac{wL}{Y}=\frac{wL}{E+wh_e}$
5. $m(5)$: Growth rate: $m(5)=100\sum\limits_{X\in\Omega}\mu(X)\sum\limits_{j\in\text{Leader}_X}\rho_{j,X}\ln\lambda$, with $\rho_{j,X}=Zh_{j,X}^{\psi}$
6. $m(6)$: R&D share: $m(6)=100\sum\limits_{X\in\Omega}\mu(X)\sum\limits_{j=0}^{J_X}\frac{wh_{j,X}}{s_{j,X}}\frac{1}{J_{h,X}}$, where $J_{h,X}=\sum_{j=0}^{J_X}\mathbb{1}_{h_{j,X}>0}$ is the number of firms with positive research.

7. $m(7), m(8)$: Linear coefficient from the following regression, and top point:
   Define (i) $y_{j,X} = wh_{j,X} - \frac{1}{J_X}\sum_{j'=1}^{J_X} wh_{j',X}$, (ii) $x_{1,j,X} = s_{j,X} - \frac{1}{J_X}\sum_{j'=1}^{J_X} s_{j',X}$, and (iii) $x_{2,j,X} = s^2_{j,X} - \frac{1}{J_X}\sum_{j'=1}^{J_X} s^2_{j',X}$, which capture already the industry fixed effects; and regress $y_{j,X} = \beta_1 x_{1,j,X} + \beta_2 x_{2,j,X} + u_{j,X}$
   Define $m(7) = \beta_1$ and $m(8) = -\frac{\beta_1}{2\beta_2}$

Second, we discuss which moments are mot directly related to each parameter. First, the elasticity of substitution, σ, and the relative productivity of small firms, A_f, help match the average markup and the labour share. The link of these moments with the elasticity of substitution is straightforward, as σ enters directly into the expression of the markup, and markups are inversely related to labour demand. On the other hand, a higher A_f implies a reallocation of market share towards the fringe, which affects both moments directly, as the fringe has no profits and so a higher market share for the fringe decreases the average markup and increases labour demand, and indirectly, via reducing the market power of superstars.
Second, we associate the scale and convexity parameters for both superstars and small firms, as well as the step size innovation λ, to the growth rate, the R&D expenditure share, and, as in Cavenaile et al. (2021), we also use the empirically observed inverted U-shape relationship between market share and innovation. In particular, as they do, we require the linear and quadratic coefficients of the following regression using the model data to be as close as possible to the coefficients of the same regression using the empirical data.

$$R\&Dexp_{j,i} = \alpha_{0,i} + \alpha_1 s_{j,i} + \alpha_2 s^2_{j,i} + u_{j,i}$$

where $a_{0,i}$ denotes the industry fixed effects, $R\&Dexp_{j,i}$ is the R&D expenditure of firm j in industry i, and $s_{j,i}$ is the industry market share of firm j in industry i.
An increase in the innovation step-size, λ, directly increases growth and innovation incentives (because it increases the magnitude of a successful innovation). In addition, keeping everything else fixed, an increase in the step-size increases the dispersion of productivities among superstars, and so, the dispersion of markups.
Finally, given all other parameters, we set the innovation scale parameter of entrants, Z_e, to normalise the mass of firms in the fringe to 1 (that is, from the definition of ρ_e and the stationary

condition $\rho_e = m\delta$, we set $Z_e = \delta h_e^{-\frac{1}{2}}$).

3.6.5 Aggregate Production Function and Growth

Remember that $y_{j,i,t} = A_{j,i,t}\left(\frac{s_{j,\theta}}{\mathcal{M}_{j,\theta}}\frac{E}{w}\alpha\right)^{\alpha}$, where $A_{1,i,t}$ is the leader's productivity and $\hat{A}_{j,i,t} = \frac{A_{j,i,t}}{A_{1,i,t}}$. Similarly, for the fringe, since the markup in small firms equals 1 and their productivity is $A_{f,i,t} = A_{1,i,t}\zeta$:

$$Y_{f,i,t} = A_{1,i,t}\zeta m^{1-\alpha}\left(s_{f,\theta}\frac{E}{w}\alpha\right)^{\alpha}$$

Then, the aggregate production function writes

$$\ln Y_t = \frac{\sigma}{\sigma-1}\int_0^1 \ln\left(\left(A_{1,i,t}\zeta m^{1-\alpha}\left(s_{f,i,t}\frac{E_t}{w_t}\alpha\right)^{\alpha}\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j=1}^{N_{i,t}}\left(\hat{A}_{j,i,t}A_{1,i,t}\left(\frac{s_{j,i,t}}{\mathcal{M}_{j,i,t}}\frac{E_t}{w_t}\alpha\right)^{\alpha}\right)^{\frac{\sigma-1}{\sigma}}\right)di$$

which can be rewritten as:

$$\ln Y_t = \alpha\ln\left(\alpha\frac{E_t}{w_t}\right) + \int_0^1 \ln A_{1,i,t}di + \frac{\sigma}{\sigma-1}\int_0^1 \ln\left(\left(\zeta m^{1-\alpha}s_{f,i,t}^{\alpha}\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j=1}^{N_{i,t}}\left(\hat{A}_{j,i,t}\left(\frac{s_{j,i,t}}{\mathcal{M}_{j,i,t}}\right)^{\alpha}\right)^{\frac{\sigma-1}{\sigma}}\right)di$$

Finally, note that $s_{j,i,t}$, $\mathcal{M}_{j,i,t}$ and $\hat{A}_{j,i,t}$ are fully determined by knowing θ; so:

$$\ln Y_t = \alpha\ln\left(\alpha\frac{E_t}{w_t}\right) + \int_0^1 \ln A_{1,i,t}di + \frac{\sigma}{\sigma-1}\sum_{\theta\in\Theta}\mu_t(\theta)\ln\left(\left(\zeta m^{1-\alpha}s_{f,\theta}^{\alpha}\right)^{\frac{\sigma-1}{\sigma}} + \sum_{j=1}^{N_{\theta}}\left(\hat{A}_{j,\theta}\left(\frac{s_{j,\theta}}{\mathcal{M}_{j,\theta}}\right)^{\alpha}\right)^{\frac{\sigma-1}{\sigma}}\right)$$

In a stationary equilibrium, $\mu_t(\theta) = \mu(\theta)$, as well as $\frac{E_t}{w_t}$ is constant; and so in the stationary equilibrium, all the growth comes from the second term, namely, the geometric mean of the industry leaders' productivity level. In particular, the growth rate in the stationary distribution is:

$$g = \frac{\dot{Y}_t}{Y_t} = \int_0^1 \frac{\dot{A}_{1,i,t}}{A_{j,i,t}}di = \sum_{\theta}\mu(\theta)\rho_{1,\theta}\ln\lambda$$

3.6.6 Summary of the code to solve the model

In this section, we summarise the code used to solve for the stationary equilibrium of the model. In this stationary equilibrium, the relative wage is constant, and the state space of the firm problem consists on all the possible industry states. The industry state consists on a vector telling us the number of firms and their productivity gaps. More explicitly, letting $\bar{N}$ be the maximum number of superstars, $\bar{n}$ be the maximum gap relative to the industry leader allowed, ordering firms from higher to lower productivity and letting $\theta = (\theta_1, \ldots, \theta_{N_\theta - 1})$ be the vector of consecutive productivity gaps of an industry with N firms (that is, $\theta_i = \ln\left(\frac{A_R}{A_{R+1}}\right)\frac{1}{\ln\lambda}$); then the industry state space is

$\{\theta = (\theta_1, \ldots, \theta_{N-1}) \mid N \in \{1, \ldots, \bar{N}\},\ \theta_1, \ldots, \theta_{N-1} \in \mathbb{N},\ \sum_{i=1}^{n-1} \theta_i \leq \bar{n}\}$

Static price-setting problem: The price-setting problem is solved as follows for each industry state and obtain $\pi_{j,\theta}$.

Let $\hat{y}_{j,\theta} = \frac{y_{j,\theta}}{Y_{f,\theta}}$ be the quantity produced by superstar j relative to the quantity of the fringe, and $\theta = (\theta_1, \ldots, \theta_{N_\theta - 1})$ be the vector of consecutive productivity gaps of the superstars in the industry and $\theta_f \in \mathbb{R}$ the exogenous gap of the fringe to the industry leader. Then, solving the problem involves solving the following system of N_θ equations in the N_θ unknowns $\hat{y}_{j,\theta}$, $j = 1, \ldots, N$:

$$\hat{y}_{j,\theta} = \frac{s_j}{s_f} \mathcal{M}_{j,\theta}^{-1} \lambda^{\theta_f - \sum_{i=1}^{j-1} \theta_{i,\theta}},\ j = 1, \ldots, N_\theta, \text{ with}$$

$$\mathcal{M}_{j,\theta} = \left[\frac{\eta-1}{\eta} - \frac{\sigma-1}{\sigma} s_{j,\theta} - \tilde{s}_{j,\theta}\left(\frac{\eta-\sigma}{\eta\sigma}\right)\right]^{-1},\ \tilde{s}_{j,\theta} = \left[\sum_{j'=1}^{N_\theta}\left(\frac{\hat{y}_{j',\theta}}{\hat{y}_{j,\theta}}\right)^{\frac{\eta-1}{\eta}}\right]^{-1} \text{ and}$$

$$s_{j,\theta} = \tilde{s}_{j,\theta}\left[\tilde{s}_{j,\theta}^{\frac{\eta(\sigma-1)}{(\eta-1)\sigma}} \hat{y}_{j,\theta}^{-\frac{\sigma-1}{\sigma}} + 1\right]^{-1}.$$

Dynamic problem: For the dynamic problem, let N_θ be the number of firms of industry state θ, and $S_\theta \in [0, 1]$ be the probability that firm $j = N_\theta$ decides to stay. If $S_\theta = 0$, then the least productive superstar exits with probability 1, so we never observe state θ (and then we need to check the truncated state at $N_\theta - 1$).

To solve, I follow these steps:

1. Start with some guesses for the relative wage $\frac{w}{E}$ and for the mass of small firms $\hat{m}_f$, the amount of researchers $\hat{h}_{j,\theta}$, $\hat{h}_{f,\theta}$, and the exit decisions $\hat{N}_\theta$, for all $\theta \in \Theta$.

2. Given the relative wage and the exit decisions, solve the firm problem:

   (a) Given the guesses for $h_{j,\theta}$, $h_{f,\theta}$ and m_f, calculate the arrival rates $\rho_{j,\theta} = Z h_{j,\theta}^{\psi}$ and $\hat{\rho}_{f,\theta} = m_f Z_f h_{f,\theta}^{\psi_f}$.

   (b) Given these arrival rates, solve the system of linear equations on $\hat{V}_{j,\theta}^S$, $\theta \in \Theta$, implied by the HJB equation of the superstars problem. Here, for the different possible innovators, we need to find the state we would obtain if that innovation happens, taking into account the expectation of firms exiting in the new state.

   (c) Given the guess of the value function, compute $h_{j,\theta}$ and $h_{f,\theta}$ according to (3.13) and (3.14), respectively.

(d) Update the innovation choices and repeat from (a) until convergence.

3. Update the exit decision. For each industry state $\theta = (\theta_1, \ldots, \theta_{N_\theta - 1})$, check whether $\hat{V}^S_{N_\theta,\theta} < 0$. If so, then reduce S_θ (at most up to 0). Otherwise, if $\hat{V}^S_{N_\theta,\theta} > \epsilon$, then increase S_θ (at most up to 1), where $\epsilon > 0$ is some tolerance allowed because, given numerical error, the value will very rarely be exactly 0.

   If there have not been any updates, go to the next step; otherwise, repeat step 2.

4. Given the optimal research choices, compute the infinitesimal transition matrix and solve for the associated stationary distribution, as described in section 3.2.2.1.

5. Given the value function for superstars and the arrival rates implied by the research choices of superstars and small firms obtained before, solve the system of linear equations on v^f_θ, $\theta \in \Theta$, implied by the HJB equation of the small firms problem.

6. Given the solutions μ_θ and v^f_θ of the previous steps, compute the researchers employed by entrants h_e using (3.15). And update the guess of the mass of small firms (in the stationary equilibrium, it must be $\frac{Z_e h_e^{\frac{1}{2}}}{\delta} = m_f$). In the baseline, the mass of small firms is normalised to 1, so when calibrating, instead of updating m_f, at this point I update the parameter Z_e so that entry equals the exit rate. Repeat this step until convergence of m_f (or Z_e).

7. Given the guess of the stationary distribution and research choices, compute aggregate labour used in production and research (by superstars, small firms, and entrants), and solve for the relative wage from the labour market clearing condition 3.17. If the difference between this relative wage and the previous guess is small enough, then we are done; otherwise, go back to step 3.

Chapter 4

New sufficient conditions for Comparative statics using the Farkas' Lemma

4.1 Introduction

When working with models, it is natural to ask how a variable reacts to a change in a parameter or a state variable, and, in particular, in which direction. In other words, one would like to carry out comparative statics. However, the conditions required to apply the available comparative statics results are rarely satisfied—except in very stylized models. So one must settle for numerical analysis (that is, studying the change of the variable in particular points of the parameter and/or state spaces).

This paper develops a novel approach to do comparative statics that builds on a classical result from linear algebra, which, as we argue, has been largely underutilized in economics: the Farkas' Lemma.[1] Farkas' Lemma—originally proved by Farkas (1902)—essentially states that exactly one of the following holds: either (1) a linear system (the primal system) has a non-negative solution, or (2) there exists a vector that satisfies a related dual system of linear inequalities. The power of Farkas' Lemma is that, while it may be hard to prove that (2) does not hold, it may be very easy to prove that (2) holds; in which case we know that (1) does not. Then,

[1] Some examples where Farkas' Lemma has been applied in economics include Vohra (2006), Belhaj and Deroïan (2012), or Gollier and Kimball (2018).

with some modifications, one can utilize this to establish sufficient conditions for the sign of a specific element of the solution vector of (1); that is, the sign of the comparative static. In addition, we show that this is not only useful for sign analysis, but also for deriving bounds on the value of the comparative static itself. We show that the bounds in Norris (2025) can also be obtained—and even improved—with the approach presented in this paper.

The appeal of this approach using Farkas' Lemma lies in the intuitiveness of the strategy it follows and its generality. That is, on the one hand, the logic behind the method is easy to grasp; and on the other hand, although the sufficient conditions provided rely on specific conditions, the method itself does not. It is therefore a tool that is always available and worth considering when facing any comparative statics problem.

The paper is organized as follows: Section 4.2 presents the theoretical framework and provides sufficient conditions for comparative statics under various assumptions. It also discusses strategies for applying the method when none of these conditions are met. Section 4.3 illustrates the approach with an application to an oligopoly model with differentiated goods and CES preferences. Finally, Section 4.4 concludes.

Literature. This paper contributes to a long literature that has sought to determine comparative statics. Norris (2025) uses a result in the linear algebra literature (Ostrowski (1952)) that bounds the value of the elements in the inverse of a strictly diagonally dominant matrix, without inverting the matrix. Their focus is on the bounds of the comparative static, but these are also informative of the sign when both bounds have the same sign. Relative to Norris (2025), the approach presented here is more general, in the sense that it can reproduce their results and is not limited to systems that satisfy (or some transformation does) the diagonal dominance condition.

A second strand of the literature, utilizing lattice theory (Topkis (1976), Milgrom and Roberts (1990)), is referred to as monotone comparative statics (Vives (1990), Milgrom and Shannon (1994), Villas-Boas (1997), Quah (2007), Barthel and Sabarwal (2018)), the foundations of which are laid in Milgrom and Roberts (1990) Another group of papers have explored comparative statics in aggregative games, that is, games in which each player's payoff depends on the own player action and an aggregate of the actions of the rest (Corchón (1994), Acemoglu and

Jensen (2013)).

Regarding the application of comparative statics in oligopoly models with differentiated products under CES preferences, this paper is, to the best of my knowledge, the first to establish the sign of the comparative statics on prices, quantities, and profits for all firms in response to a shock. Dixit (1986) identifies the sign of comparative statics in an oligopoly model with homogeneous products. Milgrom and Roberts (1990) and Norris (2025) establish uniqueness of the equilibrium, and are able to establish the sign of comparative statics of all players only in cases where the shock solely has a direct effect on a single firm.

4.2 Theory

4.2.1 Setting of interest and notation

Consider n variables u_j, $j \in \mathcal{J} := \{1, \ldots, n\}$, and n conditions that determine jointly these variables, $\{f_i(u_1, \ldots, u_n; t) = 0\}_{i \in \mathcal{J}}$, where t is a parameter. Differentiating the conditions with respect to t, we get the following linear system

$$\sum_{j \in \mathcal{J}} \frac{\partial f_i}{\partial u_j} \frac{du_j}{dt} + \frac{\partial f_i}{\partial t} = 0, \quad \forall i \in \mathcal{J}.$$

We are interested in how the variables react to a change in t; that is, we are interested in $x_j := \frac{du_j}{dt}$. In the following of this section, acknowledging that we face a linear system $Ax = c$, we are going to define the $n \times n$ matrix A, with the element on row i and column j being $(A)_{i,j} = a_{i,j} := \frac{\partial f_i}{\partial u_j}$, $\forall i, j \in \mathcal{J}$, and the $n \times 1$ matrix c, with $c_i := -\frac{\partial f_i}{\partial t}$, $\forall i \in \mathcal{J}$.

4.2.2 Farkas' Lemma and some corollaries

Theorem 1 *(Farkas' Lemma): Let A be a $m \times n$ matrix and $c \in \mathbb{R}^m$. Exactly one of the following is true:*

1. *$Ax = c$ and $x \geq 0$ for some $x \in \mathbb{R}^n$.*
2. *$y^t A \leq 0$ and $y^t c > 0$ for some $y \in \mathbb{R}^m$.*

Proof. See Appendix **??**.

■

Although the Farkas' Lemma by itself only tells us whether all the elements of the vector of unknowns x are non-negative, to allow for different combinations of signs one can consider the following transformations of the original linear system:

Let $\mathcal{K} := \{(k_1, \ldots, k_n) : k_j \in \{0,1\}, j \in \mathcal{J}\}$ be the collection of multiindices associated to all the possible combinations of signs of the n unknowns, and let K denote an element of $\mathcal{K}$. Then, the system $Ax = c$ is satisfied if and only if the system $A_K x_K = c$ is satisfied, with:

$$A_K = \begin{pmatrix} (-1)^{k_1} a_{1,1} & \cdots & (-1)^{k_n} a_{1,n} \\ (-1)^{k_1} a_{2,1} & \cdots & (-1)^{k_n} a_{2,n} \\ \vdots & \ddots & \vdots \\ (-1)^{k_1} a_{n,1} & \cdots & (-1)^{k_n} a_{n,n} \end{pmatrix} \quad x_K = \begin{pmatrix} (-1)^{k_1} x_1 \\ (-1)^{k_2} x_2 \\ \vdots \\ (-1)^{k_n} x_n \end{pmatrix}.$$

The power of Farkas' lemma is in the fact that, while it may be very difficult to prove that there does not exist $y \in \mathbb{R}^n$ such that condition 2 of Farkas' Lemma is satisfied, it may be relatively easy to find one $y \in \mathbb{R}^n$ that does satisfy such condition 2. This will be the main strategy of this paper: if x_{j_0} is hypothesized to be positive (resp. negative), we are going to assume by contradiction that $k_{j_0} = 1$ (resp. $k_{j_0} = 0$) and search for vectors $y \in \mathbb{R}^n$ that satisfy condition 2 of Farkas' Lemma for $K \in \mathcal{K}_1$ (resp. $\mathcal{K}_0$), with $\mathcal{K}_s := \{K \in \mathcal{K} : k_{j_0} = s\}$, $s = 0, 1$.

In the following, we are going to assume the following normalization:

Condition 1 *The system $Ax = c$ satisfies:*

$$A = \begin{pmatrix} 1 & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & 1 & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & 1 \end{pmatrix}.$$

Note that assuming $a_{j,j} = 1$ is without any loss of generality: (i) if $a_{j,j} < 0$, then redefine $a'_{j,i} = -a_{j,i}$, $\forall i \in \mathcal{J}$, and $c'_j = -c_j$; and (ii) if $a_{j,j} > 0$, but $a_{j,j} \neq 1$, then redefine $a'_{i,j} = a_{j,j}^{-1} a_{i,j}$, $\forall i \in \mathcal{J}$ and $x'_j = a_{j,j} x_j$, which preserves the sign.[2]

Definition 9 *We are going to say that two systems $Ax = c$ and $A'x' = c'$ are:*

1. ***Equivalent in signs****, when $sign(x_i) = sign(x'_i)$ for all $i = 1, \ldots, n$.*

[2] And if the interest is in the bounds on x_j rather than the sign: if the bounds are $x'_j \in [l'_j, u'_j]$, then $x_j \in [a_{j,j}^{-1} l'_j, a_{j,j}^{-1} u'_j]$.

2. ***Equivalent in*** x_i*, when they preserve the value of the i-th variable (i.e.* $x_i = x_i'$*).*

Lemma 6 *Any system* $Ax = c$ *can be transformed (non-uniquely) into an equivalent in signs one satisfying Condition 1.*

Proof. See Appendix 4.6.2.

■

4.2.3 Comparative Statics: Sign

In this section, we use the previous strategy to find sufficient conditions for the sign of x_{j_0} under different assumptions on A: in section 4.2.3.1 we assume A satisfies Condition 2 of column diagonal dominance; in section 4.2.3.2, in addition to Condition 2, we assume the off-diagonal elements of A are non-positive; in section 4.2.3.3, in contrast, we assume the off-diagonal elements are non-negative; in section 4.2.3.4, we assume A satisfies row (instead of column) diagonal dominance. Finally, we discuss some complementary strategies we can follow: (i) a change of variable to get a system where we can use the main results, (ii) consider a reduced system neglecting *dispensable* variables, (iii) if the sign of some variable is already known.

4.2.3.1 Proposition 1

One condition that proves useful to get some first set of sufficient conditions is the following one of column diagonal dominance. The intuition is that the effect of a change in u_j on condition f_j is always greater in absolute value than the sum of effects the change in u_j causes to conditions f_i for $i \in \mathcal{J} \setminus \{j\}$.

Condition 2 *(Column diagonal dominance):* $\sum_{i \neq j} |a_{i,j}| \leq |a_{j,j}| = a_{j,j} = 1$, $\forall j \in \mathcal{J}$.

which is a bit less restrictive than the strict column diagonal dominance, which is the one used in Norris (2025):[3]

Condition 3 *(Strict column diagonal dominance):* $\sum_{i \neq j} |a_{i,j}| < |a_{j,j}| = a_{j,j} = 1$, $\forall j \in \mathcal{J}$.

[3] This assumption has long been identified as useful in economics (McKenzie, 1960).

We leave as a next step to check conditions that guarantee that A can be transformed into A' satisfying column diagonal dominance using the operations defined in the Proof of Lemma 6 (i.e. row multiplication by a scalar, and column multiplication by a positive scalar). A first step in this direction is given in Section 4.2.3.4 for a matrix that satisfies row diagonal dominance.

Proposition 5 *If Condition 2 is satisfied, then a sufficient condition for sign*$(x_{j_0}) =$ *sign*(c_{j_0}) *is:*

$$|c_{j_0}| > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|,$$

where $|y_j^*|$ *is the unique limit of the decreasing sequence* $\{|y_j^{(m)}|\}_{m=0}^{\infty}$*, defined by* $|y_j^{(0)}| = 1$ *and* $|y_j^{(m)}| = |a_{j_0,j}| + \sum_{k \in \mathcal{J} \setminus \{j_0, j\}} |y_k^{(m-1)}||a_{k,j}|$.

Proof. See Appendix 4.6.2. ■

Corollary 2 *If Condition 2 is satisfied, then a sufficient condition for sign*$(x_{j_0}) =$ *sign*(c_{j_0}) *is:*

$$|c_{j_0}| > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |\hat{y}_j||c_j|,$$

where $|\hat{y}_j| \in [y_j^*, 1]$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$.

If Condition 3 is satisfied, then a sufficient condition for sign$(x_{j_0}) =$ *sign*(c_{j_0}) *is:*

$$|c_{j_0}| \geq \sum_{j \in \mathcal{J} \setminus \{j_0\}} |\hat{y}_j||c_j|,$$

where $|\hat{y}_j| \in [y_j^*, 1]$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$ *and* $|\hat{y}_{j_1}||c_{j_1}| > |y_{j_1}^*||c_{j_1}|$ *for at least one* $j_1 \in \mathcal{J} \setminus \{j_0\}$.

Some examples of $\hat{y}$ *include:*

1. $|\hat{y}_j| = 1$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$.

2. *(Theorem 1, Norris)* $|\hat{y}_j| = \max_{j_1 \in \mathcal{J} \setminus \{j_0\}} A_{c,j_1}$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$*, where* $A_{c,j} := \sum_{k \in \mathcal{J} \setminus \{j\}} |a_{k,j}|$.

3. $|\hat{y}_j| = H$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$ *with* $H := \frac{\max_{j \in \mathcal{J} \setminus \{j_0\}} |a_{j_0,j}|}{1 - \max_{j \in \mathcal{J} \setminus \{j_0\}} \sum_{k \in \mathcal{J} \setminus \{j_0, j\}} |a_{k,j}|}$.[4]

4. *(Proposition 1, Norris)* $|\hat{y}_j| = A_{c,j}$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$.

Proof. See Appendix 4.6.3. ■

[4]The least demanding condition of the form $|c_{j_0}| > H \sum_{j \in \mathcal{J} \setminus \{j_0\}} |c_j|$ is achieved with $H = H^* := \max_{j \in \mathcal{J} \setminus \{j_0\}} \frac{|a_{j_0,j}|}{1 - \sum_{k \in \mathcal{J} \setminus \{j_0, j\}} |a_{k,j}|}$, as shown in Proposition 11.

4.2.3.2 Proposition 2: Sufficient condition when $a_{i,j} \leq 0$ for all $i \neq j$

Condition 4 *The matrix A satisfies $a_{i,j} \leq 0$ for all $i, j \in \mathcal{J}$ with $i \neq j$.*

Proposition 6 *If A satisfies Conditions 2 and 4, then:*

1. *If* $c_{j_0} > \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j < 0\}} |y_j^*||c_j| - \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j \geq 0\}} |y_{j,0}^*||c_j|$, *then* $x_{j_0} \geq 0$.
2. *If* $-c_{j_0} > \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j > 0\}} |y_j^*||c_j| - \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j \leq 0\}} |y_{j,0}^*||c_j|$, *then* $x_{j_0} \leq 0$.

where $|y_j^*|$ is the limit of the sequence defined in Proposition 5, and $|y_{j,0}^*|$ is the unique limit of the increasing sequence $\{|y_j^{(m)}|\}_{m=0}^{\infty}$, defined by $|y_{j,0}^{(0)}| = 0$ and $|y_{j,0}^{(m)}| = |a_{j_0,j}| + \sum_{j_1 \in \mathcal{J} \setminus \{j_0, j\}} |y_{j_1,0}^{(m-1)}||a_{k,j}|$. If A satisfies Condition 3, then $|y_{j,0}^*| = |y_j^*|$.

Proof. See Appendix 4.6.7.

■

The conditions implied by Milgrom and Roberts are direct from Proposition 6; in other words, whenever the conditions to apply Milgrom and Roberts are satisfied, we can also apply Proposition 6:

Corollary 3 *(Milgrom and Roberts): If A satisfies Condition 2 and $a_{i,j} \leq 0$ for all $i \neq j$ and $c_j \geq 0$ (resp. $c_j \leq 0$) for all j; then $x_j \geq 0$ (resp. $x_j \leq 0$) for all j.*

Proof. Straightforward, since if $c \geq 0$ (resp. $c \leq 0$), then $\{j \neq j_0 : c_j < 0\} = \emptyset$ (resp. $\{j \neq j_0 : c_j > 0\} = \emptyset$).

■

Corollary 4 *If A satisfies Conditions 2 and 4, then:*

1. *If* $c_{j_0} > \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j < 0\}} |\hat{y}_j||c_j| - \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j \geq 0\}} |\hat{y}_{j,0}||c_j|$, *then* $x_{j_0} \geq 0$.
2. *If* $-c_{j_0} > \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j > 0\}} |\hat{y}_j||c_j| - \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j \leq 0\}} |\hat{y}_{j,0}||c_j|$, *then* $x_{j_0} \leq 0$.

where $|\hat{y}_j| \in [y_j^*, 1]$, $|\hat{y}_{j,0}| \in [0, y_{j,0}^*]$ for all $j \in \mathcal{J} \setminus \{j_0\}$.
If, in addition, Condition 3 is satisfied and exist either $(|\hat{y}_{j_1}| - |y_{j_1}^*|)|c_{j_1}| > 0$ or $(|\hat{y}_{j_1,0}| - |y_{j_1}^*|)|c_{j_1}| < 0$ for some $j_1, j_2 \in \mathcal{J} \setminus \{j_0\}$, then we can replace $>$ for $\geq$ in the previous conditions. Some examples of $|\hat{y}_j|, |\hat{y}_{j,0}|$, $\forall j \in \mathcal{J} \setminus \{j_0\}$, include (any combination of these $|\hat{y}_j|$ and $|\hat{y}_{j,0}|$ work):

- $|\hat{y}_j| = 1$
- $|\hat{y}_j| = \max_{j_1 \in \mathcal{J}\setminus\{j_0\}} A_{c,j_1}$
- $|\hat{y}_j| = A_{c,j}$
- $|\hat{y}_j| = H$ *as defined in Corollary 2.*

- $|\hat{y}_{j,0}| = 0$
- $|\hat{y}_{j,0}| = \min_{j_1 \in \mathcal{J}\setminus\{j_0\}} |a_{j_0,j_1}|$
- $|\hat{y}_{j,0}| = |a_{j_0,j}|$
- $|\hat{y}_{j,0}| = H_1 := \frac{\min_{j \in \mathcal{J}\setminus\{j_0\}} |a_{j_0,j}|}{1 - \min_{j \in \mathcal{J}\setminus\{j_0\}} \sum_{k \in \mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|}$.

Proof. See Appendix 4.6.8. ■

4.2.3.3 Proposition 3: Sufficient condition when $a_{i,j} \geq 0$ for all $i \neq j$

Here we are going to use the following condition, which intuitively tells us that the direct effect of x_j to condition f_{j_0} dominates the indirect effects of x_j to condition f_{j_0} through its effect on the other variables:

Condition 5 *A satisfes that* $|a_{j_0,j}| \geq \sum_{j_1 \notin \{j_0,j\}} |a_{j_0,j_1}||a_{j_1,j}|$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$.

Proposition 7 *If A satisfies Conditions 2 and 5, then:*

1. *If* $c_{j_0} > \sum_{\{j \in \mathcal{J}\setminus\{j_0\}: |a_{j_0,j}|c_j > 0\}} |\bar{y}_{j,1}||c_j| - \sum_{\{j \in \mathcal{J}\setminus\{j_0\}: |a_{j_0,j}|c_j \leq 0\}} |\bar{y}_{j,0}||c_j|$, *then* $x_{j_0} \geq 0$.
2. *If* $-c_{j_0} > \sum_{\{j \in \mathcal{J}\setminus\{j_0\}: |a_{j_0,j}|c_j < 0\}} |\bar{y}_{j,1}||c_j| - \sum_{\{j \in \mathcal{J}\setminus\{j_0\}: |a_{j_0,j}|c_j \geq 0\}} |\bar{y}_{j,0}||c_j|$, *then* $x_{j_0} \leq 0$.

where $|\bar{y}_{j,0}|, |\bar{y}_{j,1}|$ *are the unique limits of the increasing and decreasing, respectively, sequences* $\{|y_{j,0}^{(m)}|\}_{m=0}^{\infty}$, $\{|y_{j,1}^{(m)}|\}_{m=0}^{\infty}$, *defined by*

$$|y_{j,0}^{(0)}| = 0 \ , \qquad |y_{j,0}^{(m)}| = |a_{j_0,j}| - \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} |y_{j_1,1}^{(m-1)}||a_{j_1,j}|,$$
$$|y_{j,1}^{(0)}| = |a_{j_0,j}| \ , \quad |y_{j,1}^{(m)}| = |a_{j_0,j}| - \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} |y_{j_1,0}^{(m-1)}||a_{j_1,j}|.$$

If A satisfies Condition 3, then $|\bar{y}_{j,0}| = |\bar{y}_{j,1}| = |\bar{y}_j|$.

Proof. See Appendix 4.6.9.

■

Corollary 5 *If A satisfies Conditions 2 and 5, then:*

1. *If* $c_{j_0} > \sum_{\{j \in \mathcal{J}\setminus\{j_0\}: |a_{j_0,j}|c_j > 0\}} |\hat{y}_{j,1}||c_j| - \sum_{\{j \in \mathcal{J}\setminus\{j_0\}: |a_{j_0,j}|c_j \leq 0\}} |\hat{y}_{j,0}||c_j|$, *then* $x_{j_0} \geq 0$.

2. If $-c_{j_0} > \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:|a_{j_0,j}|c_j<0\}} |\hat{y}_{j,1}||c_j| - \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:|a_{j_0,j}|c_j\geq 0\}} |\hat{y}_{j,0}||c_j|$*, then* $x_{j_0} \leq 0$.

where $|\hat{y}_{j,1}| \in [\bar{y}_{j,1}, 1]$, $|\hat{y}_{j,0}| \in [0, \bar{y}_{j,0}]$ *for all* $j \in \mathcal{J} \setminus \{j_0\}$.

If, in addition, Condition 3 is satisfied and exist either $(|\hat{y}_{j_1,1}| - |\bar{y}_{j_1,1}|)|c_{j_1}| > 0$ *or* $(|\hat{y}_{j_1,0}| - |\bar{y}_{j_1,0}|)|c_{j_1}| < 0$ *for some* $j_1, j_2 \in \mathcal{J}\setminus\{j_0\}$*, then we can replace* $>$ *for* $\geq$ *in the previous conditions. An example is given by* $|\hat{y}_{j,1}| = |a_{j_0,j}|$, $|\hat{y}_{j,0}| = 0$, $\forall j \in \mathcal{J} \setminus \{j_0\}$.

4.2.3.4 Row diagonal dominance

Condition 6 *(Row diagonal dominance):* $\sum_{i\neq j} |a_{j,i}| \leq |a_{j,j}| = a_{j,j} = 1$ *for all* $j \in \mathcal{J}$.

Lemma 7 *If A satisfies Condition 6 but not Condition 2 (i.e.* $\exists j \in \mathcal{J}$ *such that* $\sum_{i\in\mathcal{J}\setminus\{j\}} |a_{i,j}| > 1$*); then* $\exists j_0 \in \mathcal{J}$ *such that* $\sum_{i\in\mathcal{J}\setminus\{j_0\}} |a_{i,j_0}| < 1$.

Proof. See Appendix 4.6.10. ■

An almost direct consequence of this lemma is the following result:

Proposition 8 *Any matrix satisfying Condition 6 and* $\min_j A_{c,j} > 0$ *can be transformed (non-uniquely) into an equivalent in signs one satisfying Condition 2. In addition, this can be done multiplying the rows and columns by strictly positive scalars.*

Proof. See Appendix 4.6.10. ■

Note that x_j for $j \in \mathcal{J}_0 := \{j \in \mathcal{J} : A_{c,j} = 0\}$ is irrelevant for the other variables; so the approach in this case would be to do the analysis on the reduced system with the variables x_j, $\forall j \in \mathcal{J} \setminus \mathcal{J}_0$, and then study the implications for x_j, $\forall j \in \mathcal{J}_0$. Direct from the proof of this Proposition, we have:

Corollary 6 *If the system* $Ax = c$ *satisfies Condition 6 with* $\min_j A_{c,j} > 0$ *but not Condition 2, a system equivalent in signs that do satisfy Condition 2 (and so we can potentially apply the previous sufficient conditions) is given by* $A'x' = c'$*, with:*

For $i \in \mathcal{J}_0$*:* $a'_{i,j} = a_{i,j}\frac{A_{c,i}}{A_{c,j}}, \forall j \in \mathcal{J}_0, \quad a'_{i,j} = a_{i,j}A_{c,i}, \forall j \notin \mathcal{J}_0, \quad c'_i = c_i A_{c,i}, \quad x'_i = x_i A_{c,i},$

For $i \notin \mathcal{J}_0$*:* $a'_{i,j} = a_{i,j}\frac{1}{A_{c,j}}, \forall j \in \mathcal{J}_0, \quad a'_{i,j} = a_{i,j}, \forall j \notin \mathcal{J}_0, \quad c'_i = c_i, \quad x'_i = x_i.$

Proof. Direct from the transformation used in the proof of Proposition 8. ■

This substantially expands the set of systems where we can do comparative statics relative to the result with row diagonal dominance in Norris (2025), which is restricted to x_{j_0} such that $c_j = 0$ for all $j \in \mathcal{J} \setminus \{j_0\}$.

4.2.3.5 Transformations, System Reduction, and Known Signs

Transformation into a system equivalent in x_{j_0} If we do the change of variable $x = x' + z$, where z is such that $z_{j_0} = 0$; then the system $Ax = c$ is equivalent in variable x_{j_0} to $Ax' = c' = c - Az$. Therefore, if we can apply any of the previous propositions to the system $Ax' = c'$ for the sign of x'_{j_0}; then we know the sign of x_{j_0}.

Corollary 7 *If A satisfies Condition 4 (i.e. off-diagonal elements are non-positive), and exists some $k \in \mathcal{J}$ such that (i) $sign(c_k) = -sign(c_j)$ for all $j \in \mathcal{J}\backslash\{k\}$ and (ii) $|c_k| \leq \min_{j\in\mathcal{J}\backslash\{k\}} \left\{\frac{|c_j|}{|a_{j,k}|}\right\}$; then $sign(x_j) = sign(c_j)$ for all $j \in \mathcal{J} \setminus \{k\}$.*

Proof. See Appendix 4.6.11.3.

■

The following result states that Condition 3 guarantees the existence of a change of variable that achieves a system equivalent in x_{j_0} to which we can apply Proposition 5.

Proposition 9 *If A satisfies Condition 3, there exists $z \in \mathbb{R}^n$ such that making the change of variable $x = x^* + z$, we get a system $Ax^* = c^*$ equivalent in x_{j_0} to $Ax = c$ such that $c^* = (0, \ldots, c^*_{j_0}, \ldots, 0)^t$, and Proposition 5 tells us that $sign(x_{j_0}) = sign(c^*_{j_0})$. An explicit expression of $c^*_{j_0}$ is given by:*

$$c^*_{j_0} = c_{j_0} + \sum_{k=0}^{K}(-1)^{k+1} \sum_{j_1\notin\{j_0\}} \sum_{j_2\notin\{j_0,j_1\}} \cdots \sum_{j_{k+1}\notin\{j_0,j_k\}} \prod_{s=0}^{k} a_{j_s,j_{s+1}} c_{j_{k+1}}.$$

Proof. See Appendix 4.6.11.4.

■

Some examples where the expression of $c^*_{j_0}$ simplifies considerably (the derivations are in Appendix 4.6.11.4):

Example 1 *(Case $n = 2$):*

$$c^*_{j_0} = c_{j_0} - a_{j_0,j_1} c_1.$$

□

Example 2 *(Case $n = 3$):*

$$c^*_{j_0} = c_{j_0} + \frac{a_{j_0,j_2}a_{j_2,j_1} - a_{j_0,j_1}}{1 - a_{j_1,j_2}a_{j_2,j_1}} c_{j_1} + \frac{a_{j_0,j_1}a_{j_1,j_2} - a_{j_0,j_2}}{1 - a_{j_2,j_1}a_{j_1,j_2}} c_{j_2}.$$

□

Example 3 *Suppose $a_{i,j} = a$ for all $i \neq j$ and $a_{j,j} = 1$. To satisfy Condition 2, it must be $a \leq (n-2)^{-1}$. Then:*

$$c^*_{j_0} = c_{j_0} - \sum_{j_1 \notin \{j_0\}} \frac{ac_{j_1}}{1+(n-2)a}.$$

□

System Reduction: dispensable variables

Definition 10 *We are going to say that x_j for $j \in \mathcal{I}$ are dispensable for the sign of x_{j_0} if $a_{k,j} = 0$ for all $j \in \mathcal{I}$ and $k \notin \mathcal{I}$.*

We say they are dispensable because we can set $y_j = 0$ for all $j \in \mathcal{I}$ and then the conditions $(y^t A)_j \leq 0$ for all $j \in \mathcal{I}$ are satisfied; and so it is sufficient to consider the system obtained after deleting the rows and columns $\mathcal{I}$.

Known signs Suppose we know that $x_j \geq 0$ for all $j \in \mathcal{I}_+$ and $x_j \leq 0$ for all $j \in \mathcal{I}_-$, but we are uncertain about the sign of x_j for all $\mathcal{I}$. Then, it will be useful to define:

Definition 11 *We are going to say that x_j for $j \in \mathcal{I}$ are conditionally dispensable for the sign of x_{j_0} if, given sign(x_j) for $j \in \mathcal{I}$, we can find y, with $y_j = 0$ for all $j \in \mathcal{I}$, such that $y^t A \leq 0$ and $y^t c > 0$.*

4.2.4 Comparative Statics: Bounds

Although the purpose of this paper is to establish some sufficient conditions for the sign of the comparative statics, the previous theory can also be used to establish bounds on the value of the comparative statics. The way to proceed to find the upper (resp. lower) bound on x_{j_0} implied by a given proposition of Section 4.2.3 is as follows: Consider the change of variable $x'_{j_0} = x_{j_0} - z_{j_0}$, which transforms the system $Ax = c$ into $Ax' = c'$ with $c'_j = c_j - a_{j,j_0} z_{j_0}$, $\forall j \in \mathcal{J}$, and find the minimal (resp. maximal) z_{j_0} such that the given proposition of Section 4.2.3 allows us to state that $x'_{j_0} \leq 0$ (resp. $x'_{j_0} \geq 0$).

For example, using Proposition 5, we obtain the following result. The subsequent corollary shows that the interval implied by these bounds is included in the interval implied by the bounds in Norris (2025). To ease notation, we define:

$$B_{j_0}(\hat{y}, i_1, i_2) := \frac{c_{j_0} + (-1)^{i_1} \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|}{1 + (-1)^{i_2} \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||a_{j,j_0}|}.$$

Proposition 10 *Using the sufficient condition of Proposition 5, we have the following bounds on x_{j_0}, where y_j^* is the limit defined in Proposition 5.*

For an upper bound:

- *If $-c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: $x_{j_0} \leq U^*_{-,j_0} := B_{j_0}(y^*, 0, 0) < 0$.*
- *If $-c_{j_0} \leq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: $x_{j_0} \leq U^*_{+,j_0} := B_{j_0}(y^*, 0, 1)$.*

For a lower bound:

- *If $c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: $x_{j_0} \geq L^*_{+,j_0} := B_{j_0}(y^*, 1, 0) > 0$.*
- *If $c_{j_0} \leq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: $x_{j_0} \geq L^*_{-,j_0} := B_{j_0}(y^*, 1, 1)$.*

Proof. See Appendix 4.6.5. ■

Corollary 8 *If Condition 2 is satisfied and $|\hat{y}_j| \in [y_j^*, 1]$ for all $j \in \mathcal{J} \setminus \{j_0\}$, then:*

For an upper bound:

- *If $-c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|$, then also $-c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$; and: $x_{j_0} \leq U^*_{-,j_0} \leq B_{j_0}(\hat{y}, 0, 0)$.*
- *If $-c_{j_0} \leq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|$: $U^*_{-,j_0} \leq U^*_{+,j_0} \leq B_{j_0}(\hat{y}, 0, 1)$.*

For a lower bound:

- *If $c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|$, then also $c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$; and: $x_{j_0} \geq L^*_{+,j_0} \geq B_{j_0}(\hat{y}, 1, 0)$.*
- *If $c_{j_0} \leq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|$: $L^*_{+,j_0} \geq L^*_{-,j_0} \geq B_{j_0}(\hat{y}, 1, 1)$.*

Some examples of $\hat{y}$ include the ones defined in Corollary 2.

Proof. See Appendix 4.6.6. ■

4.3 Applications

4.3.1 Oligopoly with differentiated goods

4.3.1.1 Setup

Assume an oligopoly of n firms, each one producing one differentiated good and competing a la Bertrand. Appendix 4.6.12 includes the derivations, as well with Cournot competition instead of Bertrand.

Firm j's production function is given by $y_j = A_j \ell_j$, and so profits write $\pi_j = s_j(1 - \mathcal{M}_j^{-1})$, where $\mathcal{M}_j = \frac{p_j}{w} A_j$ is the markup.

Consumers have Constant Elasticity of Substitution (CES) preferences. In particular, the representative consumer solves

$$\max_{\{y_j\}_{j\in\mathcal{J}}} Y = \left(\sum_{j\in\mathcal{J}} y_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \text{ s.t. } E = PY = \sum_{j\in\mathcal{J}} p_j y_j.$$

With this, the market share is given by:

$$s_j = \left[\sum_{j'=1}^{J} \left(\frac{A_j}{A_{j'}} \frac{\mathcal{M}_{j'}}{\mathcal{M}_j}\right)^{1-\sigma}\right]^{-1}. \tag{4.1}$$

Firm j solves

$$\max_{\mathcal{M}_j} \ \pi_j = s_j(1 - \mathcal{M}_j^{-1}) \text{ , s.t. } s_j \text{ given by (4.1).}$$

We are going to show that: (i) $\frac{d\mathcal{M}_{j_0}}{dA_{j_0}} \geq 0$, $\frac{d\mathcal{M}_j}{dA_{j_0}} \leq 0$, $\forall j \in \mathcal{J} \setminus \{j_0\}$; and (ii) $\frac{d\pi_{j_0}}{dA_{j_0}} \geq 0$, $\frac{d\pi_j}{dA_{j_0}} \leq 0$, $\forall j \in \mathcal{J} \setminus \{j_0\}$.

4.3.1.2 Effect on Markups

Lemma 8 *The first-order condition reads:*

$$F_j(\mathcal{M}_1, \ldots, \mathcal{M}_n) := \mathcal{M}_j^{-1} - (\sigma - 1)(1 - \mathcal{M}_j^{-1})(1 - s_j) = 0.$$

We will need the following derivatives:

$$\frac{\partial F_j}{\partial \mathcal{M}_j} = -\frac{\sigma-1}{\mathcal{M}_j}\left[1 - s_j + \frac{s_j}{\mathcal{M}_j}\right], \qquad \frac{\partial F_i}{\partial \mathcal{M}_j} = \frac{\sigma-1}{\mathcal{M}_j}\frac{s_i}{\mathcal{M}_i}\frac{s_j}{1-s_i},$$

$$\frac{\partial F_{j_0}}{\partial A_{j_0}} = \frac{\sigma-1}{A_{j_0}}\frac{s_{j_0}}{\mathcal{M}_{j_0}}, \qquad \frac{\partial F_j}{\partial A_{j_0}} = -\frac{\sigma-1}{A_{j_0}}\frac{s_j}{\mathcal{M}_j}\frac{s_{j_0}}{1-s_j}.$$

And we have the linear system $Ax = c$, where $(A)_{i,j} = a_{i,j} := -\frac{\partial F_i}{\partial \mathcal{M}_j}$, $x_j := \frac{d\mathcal{M}_j}{dA_{j_0}}$, and $c_j := \frac{\partial F_j}{\partial A_{j_0}}$. We can obtain a system $A'x' = c'$ equivalent in signs by applying operations of the type described in Lemma 6. In particular, multiply row i of both A and c by $b_i = \frac{A_{j_0}}{s_{j_0}} \frac{\mathcal{M}_i(1-s_i)}{(\sigma-1)}$, and column j of A by $e_j = \frac{s_{j_0}}{A_{j_0}} \frac{\mathcal{M}_j}{s_j} d_j$ (i.e. change of variable $x'_j = \frac{x_j}{e_j}$), where $d_j = \left(\frac{\mathcal{M}_j}{s_j}(1-s_j)^2 + 1 - s_j \right)^{-1}$. So, we have:

$$(A')_{j,j} = (A)_{j,j} b_j e_j = d_j \left(\frac{\mathcal{M}_j}{s_j}(1-s_j)^2 + 1 - s_j \right) = 1 \ , \quad (A')_{i,j} = (A)_{i,j} b_i e_j = -s_i d_j,$$
$$c'_{j_0} = c_{j_0} b_{j_0} = 1 - s_{j_0} \ , \qquad c'_i = c_i b_i = -s_i.$$

It is straightforward to check that A' satisfies Condition 3, since $\sum_{i \in \mathcal{J} \setminus \{j\}} |(A')_{i,j}| = \sum_{i \in \mathcal{J} \setminus \{j\}} s_i d_j \leq (1-s_j) d_j < 1 = |(A')_{j,j}|$, $\forall j \in \mathcal{J}$.

First, Corollary 2, using $|\hat{y}_j| = 1$ is sufficient to show $x'_{j_0} \geq 0$, since $c'_{j_0} = 1 - s_{j_0} \geq \sum_{j \in \mathcal{J} \setminus \{j_0\}} s_j \geq \sum_{j \in \mathcal{J} \setminus \{j_0\}} |c'_j|$.

Second, we can show $x'_j \leq 0$ using Corollary 7, making $k = j_0$, since it satisfies $c'_{j_0} = 1 - s_{j_0} > 0 > -s_i = c'_i$, and (ii) $\frac{|c_{j_0}|}{|(A')_{j_0,j_0}|} \leq d_{j_0}^{-1} = \frac{s_i}{d_{j_0} s_i} = \frac{|c'_i|}{|a_{i,j_0}|}$, $\forall i \in \mathcal{J} \setminus \{j_0\}$.

4.3.1.3 Effect on Profits

Now, we add $\{G_j := s_j(1 - \mathcal{M}_j^{-1}) - \pi_j = 0\}_{j \in \mathcal{J}}$ to the conditions $\{F_j = 0\}_{j \in \mathcal{J}}$. We have:

Lemma 9 *We need the following derivatives:*

$$\frac{\partial G_j}{\partial \pi_j} = -1 \ , \qquad \frac{\partial G_i}{\partial \pi_j} = 0,$$
$$\frac{\partial G_j}{\partial \mathcal{M}_j} = 0 \text{ due to the FOC,} \qquad \frac{\partial G_i}{\partial \mathcal{M}_j} = \frac{s_i s_j}{\mathcal{M}_j \mathcal{M}_i (1-s_i)},$$
$$\frac{\partial G_{j_0}}{\partial A_{j_0}} = \frac{s_{j_0}}{\mathcal{M}_{j_0} A_{j_0}} \ , \qquad \frac{\partial G_i}{\partial A_{j_0}} = -\frac{s_{j_0} s_i}{\mathcal{M}_i A_{j_0} (1-s_i)}.$$

Now, the linear system $Ax = c$ is given by

$$A := \left\{ \begin{array}{cc} (A)_{i,j} = -\frac{\partial G_i}{\partial \pi_j} & (A)_{i,n+j} = -\frac{\partial G_i}{\partial \mathcal{M}_j} \\ (A)_{n+i,j} = -\frac{\partial F_i}{\partial \pi_j} = 0 & (A)_{n+i,n+j} = -\frac{\partial F_i}{\partial \mathcal{M}_j} \end{array} \right\}_{i,j \in \mathcal{J}} , \quad c := \left\{ \begin{array}{c} c_i = \frac{\partial G_i}{\partial A_{j_0}} \\ c_{n+i} = \frac{\partial F_i}{\partial A_{j_0}} \end{array} \right\}_{i \in \mathcal{J}} .$$

Again, we can obtain a system equivalent in signs by applying operations of the type described in Lemma 6. In particular, multiply row i of both A and c by $b_i = \mathcal{M}_i(1-s_i)\frac{A_{j_0}}{s_{j_0}}$, and row $n+i$ by $b_{n+i} = \frac{\mathcal{M}_i(1-s_i)}{\sigma-1} \frac{A_{j_0}}{s_{j_0}}$, for $i \in \mathcal{J}$, and column j and $n+j$ of A by $e_j = b_j^{-1}$ and $e_{n+j} = \frac{s_{j_0} \mathcal{M}_j}{A_{j_0} s_j} d_j$, respectively, for $j \in \mathcal{J}$, where d_j is defined as before. The transformed system writes:

$$A' = \left\{ \begin{array}{cccc} (A')_{i,i} = 1 \ , & (A')_{i,j} = 0, \forall j \in \mathcal{J} \setminus \{i\} \ , & (A')_{i,n+i} = 0 & (A')_{i,n+j} = -s_i d_j, \forall j \in \mathcal{J} \setminus \{i\} \\ (A')_{n+i,i} = 0 \ , & (A')_{n+i,j} = 0, \forall j \in \mathcal{J} \setminus \{i\} \ , & (A')_{n+i,n+i} = 1 & (A')_{n+i,n+j} = -s_i d_j, \forall j \in \mathcal{J} \setminus \{i\} \end{array} \right\}_{i \in \mathcal{J}}$$

$$c' = \left\{ \ c'_{j_0} = 1 - s_{j_0} \ , \quad c'_i = -s_i, \forall i \in \mathcal{J} \setminus \{j_0\} \ , \quad c'_{n+j_0} = 1 - s_{j_0} \ , \quad c'_{n+i} = -s_i, \forall i \in \mathcal{J} \setminus \{j_0\} \ \right\}_{i \in \mathcal{J}} .$$

This A' does not directly satisfy the column diagonal dominance of Condition 2, but we will simplify the problem using the notions of *dispensable* variables. First, note that if we are interested in the sign of x_j for some $j \in \mathcal{J}$, then x_i for all $i \in \mathcal{J} \setminus \{j\}$ are dispensable in the sense of Definition 10. Second, we use that we know the signs $x_{n+j_0} \geq 0$ and $x_{n+j} \leq 0$, $\forall j \in \mathcal{J} \setminus \{j_0\}$ (that is, $k_{n+j} = 1$, $\forall j \in \mathcal{J} \setminus \{j_0\}$). Next, consider two cases.

1. If $j = j_0$, we are going to see that x_{n+j_0} is conditionally dispensable in the sense of Definition 11. Indeed, if $k_{j_0} = 1$, we can find $y \in \mathbb{R}^{2n}$ with $y_i = 0$ for $i \in \{n+j_0\} \cup \mathcal{J} \setminus \{j_0\}$ (e.g. setting $y_{j_0} = 1 + \epsilon$, for $\epsilon > 0$, and $y_j = 1$ for all $j \neq j_0$) such that $y^t A'_K \leq 0$ and $y^t c' > 0$. Indeed, noting that the zeros in y imply $(y^t A'_K)_i = (-1)^{k_i} (y^t A')_i = (-1)^{k_i} \left((1+\epsilon)(A')_{j_0,i} + \sum_{j \in \mathcal{J} \setminus \{j_0\}} (A')_{n+j,i} \right)$; so:

   (a) $(-1)^{k_{j_0}} (y^t A')_{j_0} = -(1+\epsilon) \leq 0$, and $(y^t A')_{n+j_0} = \sum_{i \in \mathcal{J} \setminus \{j_0\}} (-s_i d_{j_0}) \leq 0$.

   (b) For $i \in \mathcal{J} \setminus \{j_0\}$: $(y^t A')_i = 0 \leq 0$, and
   $(-1)^{k_{n+i}} (y^t A')_{n+i} = -\left(-(1+\epsilon) s_{j_0} d_i + 1 + \sum_{j \in \mathcal{J} \setminus \{j_0, i\}} (-s_j d_i) \right) \leq 0$, using $\sum_{j \in \mathcal{J} \setminus \{i\}} s_j d_i = (1 - s_i) d_i < 1$.

   (c) $y^t c' = (1+\epsilon)(1 - s_{j_0}) + \sum_{j \in \mathcal{J} \setminus \{j_0\}} s_j > 0$.

2. If $j \in \mathcal{J} \setminus \{j_0\}$, we are going to see that x_{n+j} for all $j \in \mathcal{J} \setminus \{j_0\}$ are conditionally dispensable in the sense of Definition 11. Indeed, if $k_j = 0$, we can find $y \in \mathbb{R}^{2n}$ with $y_i = 0$ for $i \in \{n+j_0\} \cup \mathcal{J} \setminus \{j_0\}$ (e.g. setting $y_j = -1$, and $y_{n+j_0} = -s_j d_{j_0}$) such that $y^t A'_K \leq 0$ and $y^t c' > 0$. Indeed, noting that the zeros in y imply $(y^t A'_K)_i = (-1)^{k_i} (y^t A')_i = (-1)^{k_i} \left((-1)(A')_{j,i} - s_j d_{j_0} (A')_{n+j_0,i} \right)$; so:

   (a) $(-1)^{k_j} (y^t A')_j = -1 \leq 0$, and $(-1)^{k_i} (y^t A')_i = 0 \leq 0$, $\forall i \in \mathcal{J} \setminus \{j_0\}$.

   (b) For $i \in \mathcal{J} \setminus \{j_0, j\}$: $(-1)^{k_i} (y^t A')_{n+i} = -\left(-(-s_j d_i) - s_j d_{j_0} (-s_{j_0} d_i) \right) \leq 0$,
   $(-1)^{k_j} (y^t A')_{n+j} = -\left(-(0) - s_j d_{j_0} (-s_{j_0} d_j) \right) \leq 0$, and
   $(-1)^{k_{j_0}} (y^t A')_{n+j_0} = -\left(-(-s_j d_{j_0}) - s_j d_{j_0} \right) \leq 0$.

   (c) $y^t c' = -(-s_j) + -s_j d_{j_0} (1 - s_{j_0}) > 0$, where I have used that $d_{j_0}(1 - s_{j_0}) \in (0, 1)$.

4.4 Conclusion

This paper introduces a novel approach to comparative statics, building on Farkas' Lemma, which enables both sign analysis and the derivation of bounds on comparative statics. By offering an intuitive and general method, this framework overcomes the limitations of traditional approaches that rely on strict conditions. We have applied this method to derive sufficient conditions under various assumptions, replicating and improving earlier results. Moreover, we believe there is further potential to extend it and derive sufficient conditions for a broader range of problems.

The next step is to derive better sufficient conditions conditional on knowing the sign of some variables. This could derive in an iterative method, where in the first iteration we determine the signs of those variables that satisfy some of the conditions in Corollaries 2, 4, 5; and in the next iterations we would use the *relaxed* conditions on the remaining variables given the known signs.

In a follow-up paper, I plan to use this theory to do comparative statics in dynamic games. This is a bit more challenging, as the value function is defined recursively.

4.5 References

Acemoglu, Daron and Martin Kaae Jensen, "Aggregate comparative statics," *Games and Economic Behavior*, 2013, *81*, 27–49.

Barthel, Anne-Christine and Tarun Sabarwal, "Directional Monotone Comparative Statics," *Economic Theory*, 2018, *66 (3)*, 557–591.

Belhaj, Mohamed and Frédéric Deroïan, "Risk taking under heterogenous revenue sharing," *Journal of Development Economics*, 2012, *98 (2)*, 192–202.

Corchón, Luis C., "Comparative statics for aggregative games the strong concavity case," *Mathematical Social Sciences*, 1994, *28 (3)*, 151–165.

Dixit, Avinash, "Comparative Statics for Oligopoly," *International Economic Review*, 1986, pp. 107–122.

Farkas, Julius, "Theorie der einfachen Ungleichungen," *Journal für die reine und angewandte Mathematik*, 1902, *124*, 1–27.

Gollier, Christian and Miles S. Kimball, "New methods in the classical economics of uncertainty: Comparing risks," *The Geneva Risk and Insurance Review*, 2018, *43*, 5–23.

McKenzie, Lionel, "Matrices with Dominant Diagonals and Economic Theory," 1960.

Milgrom, Paul and Chris Shannon, "Monotone Comparative Statics," *Econometrica*, 1994, *62 (1)*, 157–180.

_ **and John Roberts**, "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica: Journal of the Econometric Society*, 1990, pp. 1255–1277.

Norris, Jordan, "Bounds High Dimensional Comparative Statics," *Working Paper*, 2025.

Ostrowski, A.M., "Note on Bounds for Determinants with Dominant Principal Diagonal," *Proceedings of the American Mathematical Society*, 1952, *3 (1)*, 26–30.

Quah, John K.H., "The Comparative Statics of Constrained Optimization Problems," *Econometrica*, 2007, *75 (2)*, 401–431.

Topkis, Donald M., "The Structure of Sublattices of the Product of n Lattices," *Pacific Journal of Mathematics*, 1976, *65*, 525–532.

Villas-Boas, J. Miguel, "Comparative Statics of Fixed Points," *Journal of Economic Theory*, 1997, *73 (1)*, 183–198.

Vives, Xavier, "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, 1990, *19*, 305–321.

Vohra, Rakesh V., "The Ubiquitous Farkas Lemma," in Frank B. Alt, Michael C. Fu, and Bruce L. Golden, eds., *Perspectives in Operations Research: Papers in Honor of Saul Gass' 80th Birthday*, Springer, 2006, pp. 199–210.

4.6 Appendix

4.6.1 Proof Lemma 6

Proof. Let $a_{i,j} = (A)_{i,j}$ be the element in row i and column j of matrix A. We are going to use two types of transformations that preserve the sign of the unknowns:

1. *(Row multiplication by scalar)* For any $b_i \in \mathbb{R}$, the system $Ax = c$ is equivalent to $A'x = c'$, with $(A')_{i,j} = a_{i,j}b_i$ and $(c')_i = (c)_i b_i$, and $(A')_{i',j} = a_{i',j}$ and $c'_{i'} = c_{i'}$ for all $i' \neq i$.

2. *(Column multiplication by positive scalar)* For any $d_j \in \mathbb{R}_+$, the system $Ax = c$ is equivalent to $A'x' = c$, with $(A')_{i,j} = a_{i,j}d_j$ and $x'_j = x_j/d_j$, and $(A')_{i,j'} = a_{i,j'}$ and $x'_{j'} = x_{j'}$ for all $j' \neq j$, given that if x satisfies $Ax = c$, then the x' defined above clearly satisfies $A'x' = c$ and x' has the same signs as x.

To transform the system $Ax = c$ into one that satisfies Condition 1: first, we apply 1 to row $j \in \mathcal{J}$ using $b_j = \frac{a_{j,j}}{|a_{j,j}|}e_j$, for some $e_j > 0$. Second, we apply 2 to column $j \in \mathcal{J}$ of A using $d_j = \frac{1}{|a_{j,j}|}\frac{1}{e_j}$. This transformations achieve a diagonal of ones, and the non-unicity result is shown by the freedom on $e_j > 0$ for $j \in \mathcal{J}$.

■

4.6.2 Proof Proposition 5

Proof. We want a sufficient condition for $sign(x_{j_0})$ (let $k^*_{j_0} = \mathbb{1}_{x_{j_0} \geq 0}$) in a system that satisfies Condition 2. The way we will proceed is as follows: First, suppose by contradiction that $k_{j_0} = 1 - k^*_{j_0}$ and find a vector y that satisfies $y^t A' \leq 0$, with the normalisation $|y_{j_0}| = 1$, and, if we succeed, then $y^t c > 0$ gives us a condition on c.

In particular, we are going to find a vector y of the form: $y_{j_0} = (-1)^{k_{j_0}+1}$, and $y_j = (-1)^{k_j+1}|y_j|$, for some $|y_j| \in [0, 1]$ that satisfies $y^t A \leq 0$; that is:

1. Given this structure of y, a sufficient condition for $y^t c > 0$ is:

$$y^t c = (-1)^{k_{j_0}+1} c_{j_0} + \sum_{j \in \mathcal{J} \setminus \{j_0\}} y_j c_j \geq (-1)^{k_{j_0}+1} c_{j_0} - \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j||c_j| > 0$$

   So, for $x_{j_0} \geq 0$ (i.e. $k^*_{j_0} = 0 = 1 - k_{j_0}$) a sufficient condition would be $c_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j||c_j|$; and for $x_{j_0} \leq 0$ (i.e. $k^*_{j_0} = 1 = 1 - k_{j_0}$), a sufficient condition would be $-c_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j||c_j|$.

2. $(y^t A)_{j_0} \leq 0$ is guaranteed by Condition 2:

$$(y^tA)_{j_0} = (-1)^{k_{j_0}} \left((-1)^{k_{j_0}+1} + \sum_{j\in\mathcal{J}\setminus\{j_0\}} y_j a_{j,j_0} \right) \leq -1 + \sum_{j\in\mathcal{J}\setminus\{j_0\}} |a_{j,j_0}| \leq 0$$

3. For $j \in \mathcal{J} \setminus \{j_0\}$: $(y^tA)_j \leq 0$ requires:

$$(y^tA)_j = -|y_j| + (-1)^{k_j} \left((-1)^{k_{j_0}+1} a_{j_0,j} + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} y_k a_{k,j} \right)$$

   We are going to construct a decreasing sequence $\{\{|y_j^{(m)}|\}_{j\in\mathcal{J}\setminus\{j_0\}}\}_{m\in\mathbb{N}}$ such that for each $m \in \mathbb{N}$ and each $j \in \mathcal{J} \setminus \{j_0\}$ it is satisfied $-|y_j^{(m)}| + |a_{j_0,j}| + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |y_k^{(m)}||a_{k,j}| \leq 0$. Initialize $|y_j^{(0)}| = 1$, which clearly satisfies the conditions due to Condition 2. Now, suppose that $\{|y_j^{(m-1)}|\}_{j\in\mathcal{J}\setminus\{j_0\}}$ satisfies $\{|a_{j_0,j}| + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |y_k^{(m-1)}||a_{k,j}| \leq |y_j^{(m-1)}|\}_{j\in\mathcal{J}\setminus\{j_0\}}$, and define $\{|y_j^{(m)}| = |a_{j_0,j}| + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |y_k^{(m-1)}||a_{k,j}| \leq |y_j^{(m-1)}|\}_{j\in\mathcal{J}\setminus\{j_0\}}$. Note that, since by construction $|y_j^{(m)}| \leq |y_j^{(m-1)}|$ for all $j \in \mathcal{J} \setminus \{j_0\}$, then also $\{|a_{j_0,j}| + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |y_k^{(m)}||a_{k,j}| \leq |y_j^{(m)}|\}_{j\in\mathcal{J}\setminus\{j_0\}}$. Given that $\{y_j^{(m)}\}_{m\in\mathbb{N}}$ is a decreasing sequence and bounded by $[0, 1]$, then it converges to a unique $|y^*|$.

■

4.6.3 Proof Corollary 2

Proof. $|c_{j_0}| > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j| \geq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$, and so by Proposition 5, we have $sign(x_{j_0}) = sign(c_{j_0})$.

For the second part, note that Condition 3 implies $\max_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*| < 1$, since:

$$\begin{aligned}
||y^{(m)}||_\infty &= \max_{j\in\mathcal{J}\setminus\{j_0\}} \left\{ |a_{j_0,j}| + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |y_k^{(m-1)}||a_{k,j}| \right\} \\
&\leq \max_{j\in\mathcal{J}\setminus\{j_0\}} \left\{ |a_{j_0,j}| + ||y^{(m-1)}||_\infty \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}| \right\} \\
&< \max_{j\in\mathcal{J}\setminus\{j_0\}} \left\{ |a_{j_0,j}| + ||y^{(m-1)}||_\infty (1 - |a_{j_0,j}|) \right\} \leq ||y^{(m-1)}||_\infty
\end{aligned}$$

And, so, we can find $\hat{y}$ satisfying the conditions of the statement (as long as $\max_{j\in\mathcal{J}\setminus\{j_0\}} |c_j| > 0$) and we have $|c_{j_0}| \geq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j| \geq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$, and so by Proposition 5, we have $sign(x_{j_0}) = sign(c_{j_0})$.

For the examples, note on the one hand that $y_j^* \leq y_j^{(1)} = A_{c,j} \leq \max_{j\in\mathcal{J}\setminus\{j_0\}} A_{c,j} \leq 1$. On the other hand, given that we have seen that the sequence converges, then it has to be $y_j^* = \sum_{i=0}^{\infty} B^i B_0 \leq ||\sum_{i=0}^{\infty} B^i B_0||_\infty \leq \sum_{i=0}^{\infty} ||B||_\infty^i ||B_0||_\infty = \frac{||B_0||_\infty}{1-||B||_\infty} = \frac{\max_{j\in\mathcal{J}\setminus\{j_0\}} |a_{j_0,j}|}{1-\max_{j\in\mathcal{J}\setminus\{j_0\}} \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|}$ ■

4.6.4 Proof of the least demanding condition with $|y_j| = H$ for all $j \in \mathcal{J} \setminus \{j_0\}$

Proposition 11 *If Condition 2 is satisfied, then a sufficient condition for* $sign(x_{j_0}) = sign(c_{j_0})$ *is:*

$$|c_{j_0}| > H \sum_{j\in\mathcal{J}\setminus\{j_0\}} |c_j| \quad , \textit{with } H := \max_{j\in\mathcal{J}\setminus\{j_0\}} \frac{|a_{j_0,j}|}{1 - \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|}$$

Proof. Here, we follow the same strategy of the proof of Proposition 5, but restricting the vector y to be of the form $y_{j_0} = (-1)^{1-k_{j_0}}$, and $y_j = (-1)^{1-k_j} H$, for some $H \in [0,1]$ that satisfies $y^t A \leq 0$; that is:

1. Given this structure of y, a sufficient condition for $y^t c > 0$ is $|c_{j_0}| > H \sum_{j\in\mathcal{J}\setminus\{j_0\}} |c_j|$.
2. $(y^t A)_{j_0} \leq 0$ is guaranteed by Condition 2 and $|y_j| \leq 1$:

$$(y^t A)_{j_0} = (-1)^{k_{j_0}} \left((-1)^{1-k_{j_0}} + \sum_{j\in\mathcal{J}\setminus\{j_0\}} y_j a_{j,j_0} \right) \leq -1 + \sum_{j\in\mathcal{J}\setminus\{j_0\}} |a_{j,j_0}| \leq 0$$

3. For $j \in \mathcal{J} \setminus \{j_0\}$: $(y^t A)_j \leq 0$ requires:

$$\begin{aligned} (y^t A)_j &= -|y_j| + (-1)^{k_j} \left((-1)^{1-k_{j_0}} a_{j_0,j} + \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} y_k a_{k,j} \right) \\ &\leq -|y_j| + (-1)^{k_j+1-k_{j_0}} a_{j_0,j} + H \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}| \leq 0 \end{aligned}$$

   And so, given that k_j is arbitrary, we want $|y_j| \geq |a_{j_0,j}| + H \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|$. That is, we want H that satisfies $H \geq |a_{j_0,j}| + H \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|$ for all $j \in \mathcal{J} \setminus \{j_0\}$; in other words: $H \geq H_j := \frac{|a_{j_0,j}|}{1-\sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|}$.[5] In particular, we are interested in the minimal H that satisfies these inequalities, that is $H = \max_{j\in\mathcal{J}\setminus\{j_0\}} H_j$; let $h \in \mathcal{J} \setminus \{j_0\}$ be such that $H_h = H$.

   To see that this condition is more general than the one in Theorem 1 of Norris (2025) (i.e. we want to show that $H \leq \max_{j\in\mathcal{J}\setminus\{j_0\}} A_{c,j}$), first note that $H \leq 1$ since by Condition 2, we have $H = \frac{|a_{j_0,h}|}{1-\sum_{k\in\mathcal{J}\setminus\{j_0,h\}} |a_{k,h}|} \leq \frac{1-\sum_{k\in\mathcal{J}\setminus\{j_0,h\}} |a_{k,h}|}{1-\sum_{k\in\mathcal{J}\setminus\{j_0,h\}} |a_{k,h}|} = 1$. Then, we have $H = |a_{j_0,h}| + H \sum_{k\in\mathcal{J}\setminus\{j_0,h\}} |a_{k,h}| \leq \sum_{k\in\mathcal{J}\setminus\{j_0\}} |a_{k,h}| = A_{c,h} \leq \max_{j\in\mathcal{J}\setminus\{j_0\}} A_{c,j}$

■

[5] Note that Condition 2 guarantees that the denominator is non-negative.

4.6.5 Proof Proposition 10

Proof. Consider x' with $x'_{j_0} = x_{j_0} = x_{j_0} - z_{j_0}$ and $x'_j = x_j$. So, $Ax = c$ can be rewritten as $Ax' = c' = c - Az$. Note that $c'_j = c_j - a_{j,j_0} z_{j_0}$ for $j \in \mathcal{J}$.

For a lower bound on x_{j_0}, we want to find the maximal z_{j_0} such that the condition in Proposition 5 allows us to state that $x'_{j_0} = x_{j_0} - z_{j_0} \geq 0$. That is, we want to find the maximal z_{j_0} for which we have $c'_{j_0} = c_{j_0} - z_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c'_j| = \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j - a_{j,j_0} z_{j_0}|$. A sufficient condition for this is:

$$c_{j_0} - z_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*| \left(|c_j| + |a_{j,j_0}||z_{j_0}|\right) \implies z_{j_0} + |z_{j_0}| \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||a_{j,j_0}| < c_{j_0} - \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|$$

And there are two cases:

1. If we know that $z_{j_0} \geq 0$ (recall that $c_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|$ implies $x_{j_0} \geq 0$, and so the lower bound is also non-negative): $z_{j_0} < L_{+,j_0} := \frac{c_{j_0} - \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|}{1 + \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||a_{j,j_0}|}$
   So: $x_{j_0} \geq L_{+,j_0} - \epsilon$ for $\epsilon > 0$ arbitrarily small.

2. If $z_{j_0} < 0$: $z_{j_0} < L_{-,j_0} := \frac{c_{j_0} - \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|}{1 - \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||a_{j,j_0}|}$
   So: $x_{j_0} \geq L_{-,j_0} - \epsilon$ for $\epsilon > 0$ arbitrarily small.

For an upper bound on x_{j_0}, we want to find the minimal z_{j_0} such that the condition in Proposition 5 allows us to state that $x'_{j_0} = x_{j_0} - z_{j_0} \leq 0$. That is, we want to find the minimal z_{j_0} for which we have $-c'_{j_0} = -c_{j_0} + z_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c'_j| = \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j - a_{j,j_0} z_{j_0}|$. A sufficient condition for this is:

$$z_{j_0} - c_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*| \left(|c_j| + |a_{j,j_0}||z_{j_0}|\right) \implies z_{j_0} - |z_{j_0}| \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||a_{j,j_0}| > c_{j_0} + \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|$$

And there are two cases:

1. If we know that $z_{j_0} \leq 0$ (recall that $-c_{j_0} > \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|$ implies $x_{j_0} \leq 0$, and so the upper bound is also non-positive): $z_{j_0} > U_{-,j_0} := \frac{c_{j_0} + \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|}{1 + \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||a_{j,j_0}|}$
   So: $x_{j_0} \leq U_{-,j_0} + \epsilon$ for $\epsilon > 0$ arbitrarily small.

2. If $z_{j_0} > 0$: $z_{j_0} > U_{+,j_0} := \frac{c_{j_0} + \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||c_j|}{1 - \sum_{j \in \mathcal{J} \setminus \{j_0\}} |y_j^*||a_{j,j_0}|}$
   So: $x_{j_0} \leq U_{+,j_0} + \epsilon$ for $\epsilon > 0$ arbitrarily small.

■

4.6.6 Proof Corollary 8

Proof. Next, we want to see that using $|\hat{y}_j| \in [|y_j^*|, 1]$ for all $j \in \mathcal{J} \setminus \{j_0\}$ can only worsen the bounds (i.e. widens the interval). That is, we want to see that $\frac{\partial U_{-,j_0}}{\partial|\hat{y}_j|}, \frac{\partial U_{+,j_0}}{\partial|\hat{y}_j|} \geq 0$ and $\frac{\partial L_{-,j_0}}{\partial|\hat{y}_j|}, \frac{\partial L_{+,j_0}}{\partial|\hat{y}_j|} \leq 0$.

For the upper bounds:

- If $-c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j| \geq 0$: then we know that $x_{j_0} \leq 0$, and so we know that the upper bound is non-positive; that is, the relevant bound is U_{-,j_0}. The derivative with respect to $|\hat{y}_k|$ is:

$$\frac{\partial U_{-,j_0}}{\partial|\hat{y}_k|} = \frac{|c_k|\left(1+\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right) - |a_{k,j_0}|\left(c_{j_0}+\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||c_j|\right)}{\left(1+\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right)^2} \geq 0$$

  where I have used that from the condition we have $-c_{j_0} > 0$.

- If $-c_{j_0} \leq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: then Proposition 11 is not sufficient to determine whether $x_{j_0} \leq 0$, and so, the upper bound is positive; that is the relevant bound is U_{+,j_0}. The derivative with respect to $|\hat{y}_k|$ is:

$$\begin{aligned}\frac{\partial U_{+,j_0}}{\partial|\hat{y}_k|} &= \frac{|c_k|\left(1-\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right) + |a_{k,j_0}|\left(c_{j_0}+\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||c_j|\right)}{\left(1-\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right)^2} \\ &\geq \frac{|a_{k,j_0}|\left(\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||c_j| - \sum_{j\in\mathcal{J}\setminus\{j_0\}}|y_j^*||c_j|\right)}{\left(1-\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right)^2} \geq 0\end{aligned}$$

  where in the first inequality I have used the condition $-c_{j_0} \leq \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$, and for the second I have used $|\hat{y}_j| \geq |y_j^*|$.

For the lower bounds:

- If $c_{j_0} > \sum_{j\in\mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: then we know that $x_{j_0} \geq 0$, and so we know that the lower bound is non-negative; that is, the relevant bound is L_{+,j_0}. The derivative with respect to $|\hat{y}_k|$ is:

$$\frac{\partial L_{+,j_0}}{\partial|\hat{y}_k|} = \frac{-|c_k|\left(1+\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right) - |a_{k,j_0}|\left(c_{j_0}-\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||c_j|\right)}{\left(1+\sum_{j\in\mathcal{J}\setminus\{j_0\}}|\hat{y}_j||a_{j,j_0}|\right)^2} \leq 0$$

where I have used the condition $c_{j_0} > 0$.

- If $c_{j_0} \leq \sum_{j \in \mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$: then Proposition 5 is not sufficient to determine whether $x_{j_0} \geq 0$, and so, the lower bound is negative; that is the relevant bound is L_{-,j_0}. The derivative with respect to $|\hat{y}_k|$ is:

$$\frac{\partial L_{-,j_0}}{\partial |\hat{y}_k|} = \frac{-|c_k|\left(1 - \sum_{j \in \mathcal{J}\setminus\{j_0\}} |\hat{y}_j||a_{j,j_0}|\right) + |a_{k,j_0}|\left(c_{j_0} - \sum_{j \in \mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|\right)}{\left(1 - \sum_{j \in \mathcal{J}\setminus\{j_0\}} |\hat{y}_j||a_{j,j_0}|\right)^2}$$

$$\leq \frac{|a_{k,j_0}|\left(\sum_{j \in \mathcal{J}\setminus\{j_0\}} |y_j^*||c_j| - \sum_{j \in \mathcal{J}\setminus\{j_0\}} |\hat{y}_j||c_j|\right)}{\left(1 - \sum_{j \in \mathcal{J}\setminus\{j_0\}} |\hat{y}_j||a_{j,j_0}|\right)^2} \leq 0$$

where in the first inequality I have used the condition $c_{j_0} \leq \sum_{j \in \mathcal{J}\setminus\{j_0\}} |y_j^*||c_j|$, and for the second I have used $|\hat{y}_j| \geq |y_j^*|$.

■

4.6.7 Proof Proposition 6

Notation:

Let $\mathcal{R}_{+,j} := \{j_1 \in \mathcal{J} \setminus \{j_0, j\} : a_{j_0,j}a_{j_0,j_1}a_{j_1,j} > 0\}$ and $\mathcal{R}_{-,j} := \{j_1 \in \mathcal{J} \setminus \{j_0, j\} : a_{j_0,j}a_{j_0,j_1}a_{j_1,j} \leq 0\}$. Also, let $i_j = \mathbb{1}_{\{k_j + k_{j_0} + k_{a_{j_0,j}} \text{ odd}\}}$, and let $\mathcal{I} := \{(i_1, \ldots, i_{j_0-1}, i_{j_0+1}, \ldots, i_n) : i_j \in \{0, 1\}\}$ be the collection of multiindices, and let I denote an element of $\mathcal{I}$.[6] Ans define the restriction $\mathcal{I}_{\{i_j = s\}} = \{I \in \mathcal{I} : I_j = i_j = s\}$ for $s = 0, 1$.

Proof. We want a sufficient condition for $sign(x_{j_0})$ (let $k_{j_0}^* = \mathbb{1}_{x_{j_0} \geq 0}$) in a system that satisfies Condition 2 and 4. The way we will proceed is as follows (as in the proof of Proposition 5): First, suppose by contradiction that $k_{j_0} = 1 - k_{j_0}^*$ and find a vector y that satisfies $y^t A' \leq 0$, with the normalisation $|y_{j_0}| = 1$, and, if we succeed, then $y^t c > 0$ gives us a condition on c.

However, now we want to exploit that we know the sign of the elements $a_{i,j}$. We consider a vector y of the form: $y_{j_0} = (-1)^{1-k_{j_0}}$, and $y_j = (-1)^{k_{j_0} + k_{a_{j_0,j}}} \sum_{I \in \mathcal{I}} |y_{j,I}|$, with $|y_{j,I}| \in [0, 1]$, $\forall j \in \mathcal{J} \setminus \{j_0\}$ and $\forall I \in \mathcal{I}$. Condition 2 of Farkas' lemma requires:

1. $(y^t A)_{j_0} \leq 0$ is guaranteed by Condition 2 and $|y_j| \leq 1$:

[6] Note that $\#\mathcal{I} = 2^{n-1}$.

$$(y^tA)_{j_0} = (-1)^{k_{j_0}} \left((-1)^{1-k_{j_0}} + \sum_{j \in \mathcal{J}\setminus\{j_0\}} y_j a_{j,j_0} \right) \leq -1 + \sum_{j \in \mathcal{J}\setminus\{j_0\}} |a_{j,j_0}| \leq 0$$

2. For $j \in \mathcal{J} \setminus \{j_0\}$: $(y^tA)_j \leq 0$ requires:

$$(y^tA)_j = (-1)^{k_j} \sum_{I \in \mathcal{I}} \mathbb{1}_I \Big((-1)^{1-k_{j_0}+k_{a_{j_0,j}}} |a_{j_0,j}| + (-1)^{k_{j_0}+k_{a_{j_0,j}}} |y_{j,I}| + \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} (-1)^{k_{j_0}+k_{a_{j_0,j_1}}+k_{a_{j_1,j}}} |y_{j_1,I}||a_{j_1,j}| \Big) \leq 0$$

Extracting common factors and using the definition of i_j:

$$(y^tA)_j = (-1)^{i_j} \sum_{I \in \mathcal{I}} \mathbb{1}_I \left(|y_{j,I}| - |a_{j_0,j}| + \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} (-1)^{k_{a_{j_0,j}}+k_{a_{j_0,j_1}}+k_{a_{j_1,j}}} |y_{j_1,I}||a_{j_1,j}| \right) \leq 0.$$

Using the definitions of $\mathcal{R}_{+,j}$ and $\mathcal{R}_{-,j}$, then:

$$(y^tA)_j = (-1)^{i_j} \sum_{I \in \mathcal{I}} \mathbb{1}_I \left(|y_{j,I}| - |a_{j_0,j}| + \sum_{j_1 \in \mathcal{R}_{+,j}} |y_{j_1,I}||a_{j_1,j}| - \sum_{j_1 \in \mathcal{R}_{-,j}} |y_{j_1,I}||a_{j_1,j}| \right) \leq 0.$$

We are going to construct two sequences: (i) a decreasing one $\{|y_{j,1}^{(m)}|\}_{m\in\mathbb{N}}$ associated to $i_j = 1$, and (ii) an increasing one $\{|y_{j,0}^{(m)}|\}_{m\in\mathbb{N}}$ associated to $i_j = 0$, with $|y_{j,1}^{(m)}| \geq |y_{j,0}^{(m)}|$ for all $m \in \mathbb{N}$, such that:

$$\begin{aligned} |y_{j,0}^{(m)}| &\leq |a_{j_0,j}| - \sum_{j_1 \in \mathcal{R}_{+,j}} |y_{j_1,1}^{(m)}||a_{j_1,j}| + \sum_{j_1 \in \mathcal{R}_{-,j}} |y_{j_1,0}^{(m)}||a_{j_1,j}| \\ &\leq |a_{j_0,j}| - \sum_{j_1 \in \mathcal{R}_{+,j}} |y_{j_1,0}^{(m)}||a_{j_1,j}| + \sum_{j_1 \in \mathcal{R}_{-,j}} |y_{j_1,1}^{(m)}||a_{j_1,j}| \leq |y_{j,1}^{(m)}|. \end{aligned} \tag{4.2}$$

Note that for this to be possible it has to be $0 \leq |a_{j_0,j}| - \sum_{j_1 \in \mathcal{R}_{+,j}} |y_{j_1,1}^{(m)}||a_{j_1,j}| + \sum_{j_1 \in \mathcal{R}_{-,j}} |y_{j_1,0}^{(m)}||a_{j_1,j}|$ for $\forall j \in \mathcal{J} \setminus \{j_0\}$. This is clearly satisfied if Condition 4 holds, since $a_{i,j}$ for all $i \neq j$ implies $\mathcal{R}_{+,j} = \emptyset$ for $\forall j \in \mathcal{J} \setminus \{j_0\}$, so the conditions simplify to:

$$|y_{j,0}^{(m)}| \leq |a_{j_0,j}| + \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} |y_{j_1,0}^{(m)}||a_{j_1,j}| \leq |a_{j_0,j}| + \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} |y_{j_1,1}^{(m)}||a_{j_1,j}| \leq |y_{j,1}^{(m)}|. \tag{4.3}$$

We can initialize the sequence by setting $|y_{j,0}^{(0)}| = 0$ and $|y_{j,1}^{(0)}| = 1$, which clearly satisfy these conditions:

$$|y_{j,0}^{(0)}| = 0 \leq |a_{j_0,j}| + \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} 0|a_{j_1,j}| \leq |a_{j_0,j}| + \sum_{j_1 \in \mathcal{J}\setminus\{j_0,j\}} |a_{j_1,j}| \leq 1 = |y_{j,1}^{(0)}|.$$

Now, suppose $|y^{(m-1)}_{j,0}|, |y^{(m-1)}_{j,1}|$ satisfy 4.3, and define $|y^{(m)}_{j,0}|, |y^{(m)}_{j,1}|$ by:

$$\begin{aligned}|y^{(m-1)}_{j_1,0}| \leq |a_{j_0,j}| + \sum_{j_1\in\mathcal{J}\setminus\{j_0,j\}} |y^{(m-1)}_{j_1,0}||a_{j_1,j}| =: |y^{(m)}_{j,0}| \\ \leq |y^{(m)}_{j,1}| := |a_{j_0,j}| + \sum_{j_1\in\mathcal{J}\setminus\{j_0,j\}} |y^{(m-1)}_{j_1,1}||a_{j_1,j}| \leq |y^{(m-1)}_{j_1,1}|.\end{aligned}$$

For the examples: for $m=0$ $|y^{(0)}_{j,0}| = 0$, $|y^{(0)}_{j,1}| = 1$; for $m=1$ $|y^{(1)}_{j,0}| = |a_{j_0,j}|$, $|y^{(1)}_{j,1}| = A_{c,j}$.
Given that $\{y^{(m)}_{j,0}\}_{m\in\mathbb{N}}$ and $\{y^{(m)}_{j,1}\}_{m\in\mathbb{N}}$ are increasing and decreasing sequences respectively and bounded by $[0,1]$, then they are convergent to $y^*_{j,0}$ and $y^*_{j,1}$, respectively. Note that $|y^*_{j,1}| = |y^*_j|$, where $|y^*_j|$ is the limit of the sequence of Proposition 5 (since they have the same initial condition and updating rule).
In matricial form, for $s=0,1$, $|y^{(m)}_s| = B_0 + |y^{(m-1)}_s||A_{(j_0)}|$, where B_0 is the $1\times(n-1)$ matrix with elements $|a_{j_0,j}|$ for $j\in\mathcal{J}\setminus\{j_0\}$, and $|A_{(j_0)}|$ consists on the absolute values of matrix A without row and column j_0. Then, $\lim_{m\to\infty} |y^{(m)}_s| = \lim_{m\to\infty}\left(\sum_{i=0}^{m-1} B_0|A_{(j_0)}|^i + |y^{(0)}||A_{(j_0)}|^m\right)$, and if $\lim_{m\to\infty} |y^{(0)}||A_{(j_0)}|^m = 0$, then the limit doesn't depend on the initial value and so $y^*_{j,0} = y^*_{j,1} = y^*_j$.

Note that if A satisfies Condition 3, then $\lim_{m\to\infty} |||y^{(0)}_s||A_{(j_0)}|^m||_1 \leq \lim_{m\to\infty} |||y^{(0)}_s|||_1|||A_{(j_0)}|||_1^m = 0$.

3. Given this structure of y, a sufficient condition for $y^tc > 0$ is:

$$\begin{aligned}y^tc &= (-1)^{1-k_{j_0}}c_{j_0} + \sum_{I\in\mathcal{I}} \mathbb{1}_I \sum_{j\in\mathcal{J}\setminus\{j_0\}} (-1)^{k_{j_0}+k_{a_{j_0,j}}+k_{c_j}}|y_{j,I}||c_j| \\ &= (-1)^{k_{j_0}} \sum_{I\in\mathcal{I}} \mathbb{1}_I \left[-c_{j_0} + \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:a_{j_0,j}c_j>0\}} |y_{j,I}||c_j| - \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:a_{j_0,j}c_j<0\}} |y_{j,I}||c_j|\right] > 0\end{aligned} \tag{4.4}$$

So, for $x_{j_0}\geq 0$ (i.e. $k^*_{j_0} = 0 = 1-k_{j_0}$) a sufficient condition would be (note that I include the possibility of equality, $a_{j_0,j}c_j = 0$, in the negative terms)

$$c_{j_0} > \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:a_{j_0,j}c_j>0\}} \max_I |y_{j,I}||c_j| - \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:a_{j_0,j}c_j\leq 0\}} \min_I |y_{j,I}||c_j|$$

and for $x_{j_0}\leq 0$ (i.e. $k^*_{j_0} = 1 = 1-k_{j_0}$), a sufficient condition would be

$$-c_{j_0} > \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:a_{j_0,j}c_j<0\}} \max_I |y_{j,I}||c_j| - \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:a_{j_0,j}c_j\geq 0\}} \min_I |y_{j,I}||c_j|$$

And for Condition 4, $a_{i,j}\leq 0$, $\forall i\neq j$:

$$c_{j_0} > \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:|a_{j_0,j}|c_j<0\}} |y_{j,1}^{(m)}||c_j| - \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:c_j\geq 0\}} |y_{j,0}^{(m)}||c_j|$$

$$-c_{j_0} > \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:|a_{j_0,j}|c_j>0\}} |y_{j,1}^{(m)}||c_j| - \sum_{\{j\in\mathcal{J}\setminus\{j_0\}:c_j\leq 0\}} |y_{j,0}^{(m)}||c_j|$$

And, since $|y_{j,1}^{(m)}|$ is decreasing in m and $|y_{j,0}^{(m)}|$ is increasing in m, then the conditions are less demanding as m increases.

■

4.6.8 Proof Corollary 4

Proof. The only thing that is not straightforward is the part of H_1. For this, first note that

$$\sum_{j_1}(B^m)_{i,j_1}B_{0,j_1} = \sum_{j_1}\sum_{j_2}(B^{m-1})_{i,j_2}(B)_{j_2,j_1}B_{0,j_1}$$
$$\geq \sum_{j_2}(B^{m-1})_{i,j_2}\sum_{j_1}(B)_{j_2,j_1}\min_j B_{0,j} \geq \left(\min_i \sum_{j_1}(B)_{i,j_1}\right)^m \min_j B_{0,j}.$$

So: $y_j^* = \sum_{i=0}^{\infty} B^i B_0 \geq H_1 := \frac{\min_{j\in\mathcal{J}\setminus\{j_0\}} |a_{j_0,j}|}{1-\min_{j\in\mathcal{J}\setminus\{j_0\}} \sum_{k\in\mathcal{J}\setminus\{j_0,j\}} |a_{k,j}|}$. ■

4.6.9 Proof Proposition 7

Proof. We follow the same procedure as in the Proof of Proposition 6, but now $a_{i,j} \geq 0$, $\forall i \neq j$. Again, suppose by contradiction that $k_{j_0} = 1 - k_{j_0}^*$ we consider a vector y of the form: $y_{j_0} = (-1)^{1-k_{j_0}}$, and $y_j = (-1)^{k_{j_0}+k_{a_{j_0,j}}} \sum_{I\in\mathcal{I}} |y_{j,I}|$, with $|y_{j,I}| \in [0,1]$, $\forall j \in \mathcal{J}\setminus\{j_0\}$ and $\forall I \in \mathcal{I}$. Condition 2 of Farkas' lemma requires:

1. $(y^tA)_{j_0} \leq 0$ is guaranteed by Condition 2 and $|y_j| \leq 1$:

$$(y^tA)_{j_0} = (-1)^{k_{j_0}}\left((-1)^{1-k_{j_0}} + \sum_{j\in\mathcal{J}\setminus\{j_0\}} y_j a_{j,j_0}\right) \leq -1 + \sum_{j\in\mathcal{J}\setminus\{j_0\}} |a_{j,j_0}| \leq 0$$

2. For $j \in \mathcal{J}\setminus\{j_0\}$: a sufficient condition for $(y^tA)_j \leq 0$ is given by (4.2), which, using that $a_{i,j} \geq 0$ $\forall i \neq j$ implies $\mathcal{R}_{-,j} = \emptyset$, writes:

$$|y_{j,0}^{(m)}| \leq |a_{j_0,j}| - \sum_{j_1\in\mathcal{J}\setminus\{j_0,j\}} |y_{j_1,1}^{(m)}||a_{j_1,j}| \leq |a_{j_0,j}| - \sum_{j_1\in\mathcal{J}\setminus\{j_0,j\}} |y_{j_1,0}^{(m)}||a_{j_1,j}| \leq |y_{j,1}^{(m)}| \qquad (4.5)$$

Note that for this to be possible it has to be $0 \leq |a_{j_0,j}| - \sum_{j_1 \in \mathcal{J} \setminus \{j_0,j\}} |y_{j_1,1}^{(m)}||a_{j_1,j}|$ for $\forall j \in \mathcal{J} \setminus \{j_0\}$. And note that Condition 5 guarantees this if we have $|y_{j_1,1}^{(m)}| \leq |a_{j_0,j_1}|$, $\forall j_1 \in \mathcal{J} \setminus \{j_0\}$. Then, we can initialize the two sequences with $|y_{j,0}^{(0)}| = 0 \leq |a_{j_0,j}| = |y_{j,1}^{(0)}|$.

Now, suppose $|y_{j,0}^{(m-1)}|, |y_{j,1}^{(m-1)}|$ satisfy 4.5, and define $|y_{j,0}^{(m)}|, |y_{j,1}^{(m)}|$ by:

$$|y_{j_1,0}^{(m-1)}| \leq |a_{j_0,j}| - \sum_{j_1 \in \mathcal{J} \setminus \{j_0,j\}} |y_{j_1,1}^{(m-1)}||a_{j_1,j}| =: |y_{j,0}^{(m)}|$$

$$\leq |y_{j,1}^{(m)}| := |a_{j_0,j}| - \sum_{j_1 \in \mathcal{J} \setminus \{j_0,j\}} |y_{j_1,0}^{(m-1)}||a_{j_1,j}| \leq |y_{j_1,1}^{(m-1)}|$$

Given that $\{y_{j,0}^{(m)}\}_{m \in \mathbb{N}}$ and $\{y_{j,1}^{(m)}\}_{m \in \mathbb{N}}$ are increasing and decreasing sequences respectively and bounded by $[0, |a_{j_0,j}|]$, then they are convergent to $\bar{y}_{j,0}$ and $\bar{y}_{j,1}$, respectively. And if $\lim_{m \to \infty} |A_{(j_0)}|^m = 0$, where $|A_{(j_0)}|$ is the same matrix defined in Appendix 4.6.7, then the limit doesn't depend on the initial value and so $\bar{y}_{j,0} = \bar{y}_{j,1} = \bar{y}_j$. As shown in the proof of Proposition 6, Condition 3 is a sufficient for this.

3. For $y^t c > 0$, following from (4.4):

   for $x_{j_0} \geq 0$ (i.e. $k_{j_0}^* = 0 = 1 - k_{j_0}$) a sufficient condition would be (note that I include the possibility of equality, $a_{j_0,j} c_j = 0$, in the negative terms)

$$c_{j_0} > \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j > 0\}} |\bar{y}_{j,1}||c_j| - \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j \leq 0\}} |\bar{y}_{j,0}||c_j|$$

   and for $x_{j_0} \leq 0$ (i.e. $k_{j_0}^* = 1 = 1 - k_{j_0}$), a sufficient condition would be

$$-c_{j_0} > \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j < 0\}} |\bar{y}_{j,1}||c_j| - \sum_{\{j \in \mathcal{J} \setminus \{j_0\} : |a_{j_0,j}| c_j \geq 0\}} |\bar{y}_{j,0}||c_j|$$

■

4.6.10 Proofs Row diagonal dominance

Proof of Lemma 7:

By contradiction, suppose $\sum_{i \neq j} |a_{i,j}| \geq 1$ for all $j \in \mathcal{J}$, then:

$$\sum_{j_1 \in \mathcal{J}} 1 \geq \sum_{j_1 \in \mathcal{J}} \sum_{j_2 \in \mathcal{J} \setminus \{j_1\}} |a_{j_1,j_2}| = \sum_{j_2 \in \mathcal{J}} \sum_{j_1 \in \mathcal{J} \setminus \{j_2\}} |a_{j_1,j_2}| \geq \sum_{j_1 \in \mathcal{J} \setminus \{j_0\}} |a_{j_1,j_0}| + \sum_{j_2 \in \mathcal{J} \setminus \{j_0\}} \sum_{j_1 \in \mathcal{J} \setminus \{j_2\}} 1 > \sum_{j_1 \in \mathcal{J}} 1$$

where I have used that A satisfies row diagonal dominance for the first inequality, the contradiction hypothesis for the second inequality, and that A doesn't satisfy the column diagonal dominance for the last strict inequality. This gives us a contradiction, which proves the result.

□

Proof of Proposition 8:

If A satisfies Condition 2, we are done; otherwise Lemma 7 tells us that $\mathcal{J}_0 := \{j \in \mathcal{J} : A_{c,j} < 1\}$ is not empty. Then, multiply each row $i \in \mathcal{J}_0$ by $A_{c,i} < 1$. On the one hand, it clearly preserves row diagonal dominance: $A_{c,i} \geq \sum_{j \in \mathcal{J}\setminus\{i\}} |a_{j,i}| A_{c,i}$. On the other hand, now column i satisfies diagonal dominance: $A_{c,i} = \sum_{j \in \mathcal{J}\setminus\{i\}} |a_{j,i}| \geq \sum_{j \in \mathcal{J}_0\setminus\{i\}} |a_{j,i}| A_{c,j} + \sum_{j \in \mathcal{J}\setminus\mathcal{J}_0} |a_{j,i}|$. The only thing missing is to have ones in the diagonal, for which we multiply each column $j \in \mathcal{J}_0$ by $A_{c,j}^{-1} < \infty$, since $A_{c,j} > 0$.

□

4.6.11 Other

4.6.11.1 Review Matricial norms induced by vector norms:

Definition 12 *Given the vector norm $||\cdot||_p$, $1 \leq p \leq \infty$, we define the (induced) matrix norm $||A||_p := \sup_{x \neq 0} \left\{ \frac{||Ax||_p}{||x||_p} \right\}$.*

It is straightforward from the definition that $||ABx||_p = ||A(Bx)||_p \leq ||A||_p ||Bx||_p \leq ||A||_p ||B||_p ||x||_p$; therefore: $||AB||_p = \sup_{x\neq 0} \left\{ \frac{||ABx||_p}{||x||_p} \right\} \leq \sup_{x \neq 0} \left\{ \frac{||A||_p ||B||_p ||x||_p}{||x||_p} \right\} = ||A||_p ||B||_p$. As a corollary: $||A^k x||_p \leq ||A^k||_p ||x||_p \leq (||A||_p)^k ||x||_p$; so if $||A||_p < 1$, then $\lim_{k\to\infty} ||A^k x||_p = 0$. In particular, for $p = 1, \infty$:

- The matrix norm induced by the vector norm $||x||_\infty = \max_j \{|x_j|\}$. Note that $\frac{||Ax||_\infty}{||x||_\infty} = \frac{\max_i \{|\sum_j a_{i,j} x_j|\}}{\max_j |x_j|} = \max_i \{|\sum_j a_{i,j} \frac{x_j}{\bar{x}}|\} \leq \max_i \{\sum_j |a_{i,j}| \left|\frac{x_j}{\bar{x}}\right|\} \leq \max_i \{\sum_j |a_{i,j}|\}$, and this upper bound is attained taking $x_j = \frac{a_{i,j}}{|a_{i,j}|} \bar{x}$.

- Now, for the vector norm $||x||_1 = \sum_j |x_j|$. Note that $\frac{||Ax||_1}{||x||_1} = \frac{\sum_i \left|\sum_j a_{i,j} x_j\right|}{\sum_j |x_j|} = \sum_i \left| \sum_j a_{i,j} \frac{x_j}{\sum_k x_k} \right| \leq \sum_i \sum_j |a_{i,j}| \left| \frac{x_j}{\sum_k |x_k|} \right| = \sum_j \left(\sum_i |a_{i,j}| \right) \left| \frac{x_j}{\sum_k |x_k|} \right| \leq \max_j \left\{ \sum_i |a_{i,j}| \right\} \sum_j \left| \frac{x_j}{\sum_k |x_k|} \right| = \max_j \left\{ \sum_i |a_{i,j}| \right\}$, and this upper bound is attained taking $x_{j^*} = 1$ and $x_j = 0$ for $j \neq j^*$, for j^* such that $\sum_i |a_{i,j^*}| = \max_j \left\{ \sum_i |a_{i,j}| \right\}$.

Lemma 10 *Let $x, y \in \mathbb{R}^n$ and $A \in \mathbb{R}^n \times \mathbb{R}^n$, then $||y^t A x||_1 \leq ||y||_\infty ||A||_1 ||x||_1$.*

Proof. $||y^t A x||_1 = |\sum_i y_i \sum_j (A)_{i,j} x_j| \leq \sum_i \sum_j |y_i (A)_{i,j} x_j| \leq \sum_j \max_k \{|\sum_i y_i (A)_{i,k}|\} |x_j|$
$\leq \max_k \{|\sum_i \max_h \{y_h\} (A)_{i,k}|\} ||x||_1 = ||y||_\infty ||A||_1 ||x||_1$

■

Condition 2 guarantees that $||B||_1 < 1$ and so we have the convergence result we wanted:

$$\lim_{k\to\infty} ||c^{(k+1,j_0)}||_1 = \lim_{k\to\infty} ||B^k c^{(0,j_0)}||_1 \le \lim_{k\to\infty} ||B^k||_1 ||c^{(0,j_0)}||_1 \le \lim_{k\to\infty} (||B||_1)^k \, ||c^{(0,j_0)}||_1 = 0$$

So, for $j \neq j_0$: $\lim_{k\to\infty} |c_j^{(k)}| \le \lim_{k\to\infty} \sum_{j' \neq j_0} |c_{j'}^{(k)}| = \lim_{k\to\infty} ||c^{(k,j_0)}||_1 = 0$.

And for j_0: $\lim_{k\to\infty} |c_{j_0}^{(k)} - c_{j_0}^{(k-1)}| = \lim_{k\to\infty} |a' B c^{(k-2,j_0)}| = 0$. So, $\{c_{j_0}^{(k)}\}_k$ is a Cauchy sequence on $\mathbb{R}$, and since $(\mathbb{R}, |\cdot|)$ is a complete metric space, it is convergent: $\lim_{k\to\infty} c_{j_0}^{(k)} = c_{j_0}^*$.

4.6.11.2 Neumann series expansion of the inverse

Note that if $\lim_{k\to\infty} B^k = 0$, then $\lim_{K\to\infty} (I - B) \sum_{k=0}^{K} B^k = \lim_{K\to\infty} \sum_{k=0}^{K} B^k - \sum_{k=1}^{K+1} B^k = \lim_{K\to\infty} I - B^{K+1} = I$. And so $(I - B)^{-1} = \sum_{k=0}^{\infty} B^k$. In our case, Condition 3 guarantees that $B = I - A$ satisfies $||I - A||_1 < 1$ and so $\lim_{k\to\infty} B^k = 0$; therefore, we have $A^{-1} = \sum_{k=0}^{\infty} (I - A)^k$, and so the solution to our system is given by $x = A^{-1}c = \sum_{k=0}^{\infty} B^k c$.

Lemma 11 *We have $b_{j,j} := (B)_{j,j} = 0$, $b_{i,j} = (B)_{i,j} = -a_{i,j}$ for all $i \neq j$, and so:*

$$x_{j_0} = c_{j_0} + \sum_{k=1}^{\infty} (-1)^k \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_1\}} \cdots \sum_{j_k \notin \{j_{k-1}\}} \prod_{s=0}^{k-1} a_{j_s, j_{s+1}} c_{j_k}$$

Proof. For $k = 0$: $B^0 c = c$. For $k = 1$: $(B^1 c)_{j_0} = \sum_{j_1 \notin \{j_0\}} b_{j_0,j_1} c_{j_1}$. Suppose by induction that for $k = K$: $(B^K c)_{j_0} = \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_1\}} \cdots \sum_{j_K \notin \{j_{K-1}\}} \prod_{s=0}^{K-1} b_{j_s,j_{s+1}} c_{j_K}$. Then, for $k = K+1$: $(B^{K+1}c)_{j_0} = (BB^K c)_{j_0} = \sum_{i \notin \{j_0\}} b_{j_0,i} \sum_{j_1 \notin \{i\}} \sum_{j_2 \notin \{j_1\}} \cdots \sum_{j_K \notin \{j_{K-1}\}} \prod_{s=0}^{K-1} b_{j_s,j_{s+1}} c_{j_K}$, and redefining j_s as j_{s+1} for $s = 1, \ldots, K$ and defining $j_1 = i$, we have the result. Then, we just need to change $b_{i,j} = (-1)a_{i,j}$. ■

4.6.11.3 Proof of Corollary 7

Consider the change of variable $x = x' + z$ with z a vector of zeros except in position k, $z_k = \mu$, with $\mu \in \mathbb{R}$. So $Ax' = c'$, with $c'_j = c_j - \mu a_{j,k}$, $\forall j \in \mathcal{J}$, is a system equivalent in x_j for all $j \in \mathcal{J} \setminus \{k\}$. Using that for all $j \in \mathcal{J} \setminus \{k\}$, we have $a_{j,k} \le 0$ and $k_{c_j} = 1 - k_{c_k}$ (where, for a generic x: $k_x = 1$ if $x \ge 0$ and 0 otherwise), then $c'_j = (-1)^{k_{c_j}} (|c_j| - |c_k||a_{j,k}|)$; so, $sign(c_j) = sign(c'_j)$ for $j \in \mathcal{J} \setminus \{k\}$ if and only if $|c_j| - |c_k||a_{j,k}| \ge 0 \iff |c_k| \le \frac{|c_j|}{|a_{j,k}|}$, for all $j \in \mathcal{J} \setminus \{k\}$. Then, if c_k satisfies this condition, we have that $(-1)^{1-k_{c_k}} c \ge 0$, and so Proposition 6 then tells us that for all $j \in \mathcal{J} \setminus \{k\}$: $sign(x_j) = sign(x'_j) = sign(c'_j) = sign(-c_k) = sign(c_j)$.

□

4.6.11.4 Proof of Proposition 9

Consider the change of variables $x = x' + z$ with z defined by $z_{j_0} = 0$ and $z_j = c_j$, $\forall j \in \mathcal{J} \setminus \{j_0\}$, which leads to a system $Ax' = c'$ equivalent in variable x_{j_0}. Now, with the aim to iterate this operation, define $c^{(0)} := c$, $x^{(0)} := x$, and for $m \in \{0\} \cup \mathbb{N}$, define $z^{(m)}$ defined as $z_{j_0}^{(m)} = 0$, and $z_j^{(m)} = c_j^{(m)}$, $\forall j \in \mathcal{J} \setminus \{j_0\}$. And define $c^{(m+1)} := c^{(m)} - Az^{(m)}$, and $x^{(m+1)} = x^{(m)} - z^{(m)}$ (and since by construction of $z^{(m)}$ we have $z_{j_0}^{(m)} = 0$ for all m, then $x_{j_0}^{(m)} = x_{j_0}$ for all m, and so $Ax^{(m)} = c^{(m)}$ is equivalent in x_{j_0} to $Ax = c$). We are going to show that if A satisfies Condition 3, then the sequence $\{c^{(m)}\}_{m \in \mathbb{N}}$ converges to a vector of the form $c^* = (0, \ldots, c_{j_0}^*, \ldots, 0)^t$; and so, that we converge to a system equivalent in x_{j_0} to which we can apply Proposition 5, which then would tell us that $sign(x_{j_0}) = sign(c_{j_0}^*)$.

To this aim, let's first introduce some notation. For a generic vector $x \in \mathbb{R}^n$, we will denote $x_{(j)}$ the vector in $\mathbb{R}^{n-1}$ the vector obtained from eliminating the j-th element in x. Analogously, for a generic matrix $A \in \mathbb{R}^n \times \mathbb{R}^n$, $A_{(j)}$ will denote the matrix in $\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ obtained from eliminating row and column j from A, and $(A)_{(r,j_0)}$ denote the row j_0 of matrix A without the element j_0. With this, note that we have

$$\begin{aligned} c_{j_0}^{(m+1)} &= c_{j_0}^{(m)} - (A)_{(r,j_0)} c_{(j_0)}^{(m)} \\ c_{(j_0)}^{(m+1)} &= c_{(j_0)}^{(m)} - A_{(j_0)} c_{(j_0)}^{(m)} = B c_{(j_0)}^{(m)} \end{aligned}$$

where we have defined $B := \mathbb{1}_{n-1} - A_{(j_0)}$, with $\mathbb{1}_{n-1}$ the $(n-1)$-dimensional identity matrix. Then, we have:

Lemma 12 *$\{c^{(m)}\}_m$ defined above converges to a vector of the form $c^* = (0, \ldots, c_{j_0}^*, \ldots, 0)^t$.*

Proof. Condition 3 guarantees that $||B||_1 < 1$ and so we have the convergence result we wanted, since this guarantees that

$$\lim_{m \to \infty} ||c_{(j_0)}^{(m+1)}||_1 = \lim_{k \to \infty} ||B^m c_{(j_0)}^{(0)}||_1 \leq \lim_{m \to \infty} (||B||_1)^m \, ||c_{(j_0)}^{(0)}||_1 = 0$$

So, for $j \neq j_0$: $\lim_{m \to \infty} |c_j^{(m)}| \leq \lim_{m \to \infty} \sum_{j' \in \mathcal{J} \setminus \{j_0\}} |c_{j'}^{(m)}| = \lim_{m \to \infty} ||c_{(j_0)}^{(0)}||_1 = 0.$

And for j_0: $\lim_{m \to \infty} |c_{j_0}^{(m)} - c_{j_0}^{(m-1)}| = \lim_{m \to \infty} |(A)_{(r,j_0)} c_{(j_0)}^{(m)}| = 0$. So, $\{c_{j_0}^{(m)}\}_m$ is a Cauchy sequence on $\mathbb{R}$, and since $(\mathbb{R}, |\cdot|)$ is a complete metric space, it is convergent: $\lim_{m \to \infty} c_{j_0}^{(m)} = c_{j_0}^*$.

■ Then, the z of the proposition is given by $z = \sum_m z^{(m)}$. And using Neumann expansion series, we get an explicit expression of $c_{j_0}^*$:

Lemma 13 *We have:* $c_{j_0}^* = c_{j_0} + \sum_{k=0}^{K} (-1)^{k+1} \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_0, j_1\}} \cdots \sum_{j_{k+1} \notin \{j_0, j_k\}} \prod_{s=0}^{k} a_{j_s, j_{s+1}} c_{j_{k+1}}$

Proof. Direct from Lemma 11.

■

Example 4 *(Case $n = 2$): If $n = 2$:*

$$c^*_{j_0} = c_{j_0} + \sum_{k=0}^{\infty}(-1)^1 a' c^{(k,j_0)} = c_{j_0} + \sum_{k=0}^{\infty}(-1)^{k+1} \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_0, j_1\}} \cdots \sum_{j_{k+1} \notin \{j_0, j_k\}} \prod_{s=0}^{k} a_{j_s, j_{s+1}} c_{j_{k+1}}$$

$$= c_{j_0} + \sum_{k=0}^{0}(-1)^{k+1} \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_0, j_1\}} \cdots \sum_{j_{k+1} \notin \{j_0, j_k\}} \prod_{s=0}^{k} a_{j_s, j_{s+1}} c_{j_{k+1}}$$

$$= c_{j_0} - a_{j_0, j_1} c_1$$

□

Example 5 *(Case $n = 3$): If $n = 3$:*

$$\sum_{k=0}^{\infty}(-1)^1 a' c^{(k,j_0)} = \sum_{k=0}^{\infty}(-1)^{k+1} \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_0, j_1\}} \cdots \sum_{j_{k+1} \notin \{j_0, j_k\}} \prod_{s=0}^{k} a_{j_s, j_{s+1}} c_{j_{k+1}}$$

$$= \sum_{k=0}^{\infty}(-1)^{k+1} \sum_{s \in \{1,2\}} a_{j_0, j_s} \left(a_{j_s, j_{3-s}} a_{j_{3-s}, j_s}\right)^{\lfloor \frac{k}{2} \rfloor} a_{j_s, j_{3-s}}^{\frac{k}{2} - \lfloor \frac{k}{2} \rfloor} \left(\mathbb{1}_{\{k\ odd\}} c_{j_{3-s}} + \mathbb{1}_{\{k\ even\}} c_{j_s}\right)$$

$$= \sum_{k=0}^{\infty} \sum_{s \in \{1,2\}} a_{j_0, j_s} \left(\left(a_{j_s, j_{3-s}} a_{j_{3-s}, j_s}\right)^k a_{j_s, j_{3-s}} c_{j_{3-s}} - \left(a_{j_s, j_{3-s}} a_{j_{3-s}, j_s}\right)^k c_{j_s}\right)$$

$$= \sum_{s \in \{1,2\}} a_{j_0, j_s} \frac{a_{j_s, j_{3-s}} c_{j_{3-s}} - c_{j_s}}{1 - a_{j_s, j_{3-s}} a_{j_{3-s}, j_s}} = \sum_{s \in \{1,2\}} \frac{a_{j_0, j_{3-s}} a_{j_{3-s}, j_s} - a_{j_0, j_s}}{1 - a_{j_s, j_{3-s}} a_{j_{3-s}, j_s}} c_{j_s}$$

So:

$$c^*_{j_0} = c_{j_0} + \frac{a_{j_0, j_2} a_{j_2, j_1} - a_{j_0, j_1}}{1 - a_{j_1, j_2} a_{j_2, j_1}} c_{j_1} + \frac{a_{j_0, j_1} a_{j_1, j_2} - a_{j_0, j_2}}{1 - a_{j_2, j_1} a_{j_1, j_2}} c_{j_2}$$

□

Example 6 *Suppose $a_{i,j} = a$ for all $i \neq j$ and $a_{j,j} = 1$. To satisfy Condition 2, it must be $a \leq (n-2)^{-1}$. Then, Lemma 13 writes:*

$$c^*_{j_0} = c_{j_0} + \sum_{k=0}^{\infty}(-1)^{k+1} \sum_{j_1 \notin \{j_0\}} \sum_{j_2 \notin \{j_0, j_1\}} \cdots \sum_{j_{k+1} \notin \{j_0, j_k\}} \prod_{s=0}^{k} a c_{j_{k+1}}$$

$$= c_{j_0} - \sum_{j_1 \notin \{j_0\}} a c_{j_1} + \sum_{k=1}^{\infty}(-1)^{k+1} \sum_{j_{k+1} \notin \{j_0\}} \sum_{j_k \notin \{j_0, j_{k+1}\}} \cdots \sum_{j_1 \notin \{j_0, j_2\}} a^{k+1} c_{j_{k+1}}$$

$$= c_{j_0} - \sum_{j_1 \notin \{j_0\}} a c_{j_1} + \sum_{k=1}^{\infty}(-1)^{k+1} \sum_{j_1 \notin \{j_0\}} (n-2)^k a^{k+1} c_{j_1}$$

Note that

$$\sum_{k=1}^{\infty}(-1)^{k+1}(n-2)^k a^{k+1} = \sum_{k=0}^{\infty}(-1)^k(n-2)^{k+1}a^{k+2} = (n-2)a^2\sum_{k=0}^{\infty}(-1)^k[(n-2)a]^k$$
$$= (n-2)a^2\sum_{k=0}^{\infty}[(n-2)a]^{2k}\left[1-(n-2)a\right] = (n-2)a^2\frac{1-(n-2)a}{1-[(n-2)a]^2} = \frac{(n-2)a^2}{1+(n-2)a}$$
$$c^*_{j_0} = c_{j_0} - \sum_{j_1\notin\{j_0\}} ac_{j_1} + \sum_{j_1\notin\{j_0\}}\frac{(n-2)a^2}{1+(n-2)a}c_{j_1} = c_{j_0} - \sum_{j_1\notin\{j_0\}} ac_{j_1}\left[1-\frac{(n-2)a}{1+(n-2)a}\right]$$
$$= c_{j_0} - \sum_{j_1\notin\{j_0\}}\frac{ac_{j_1}}{1+(n-2)a}$$

□

4.6.12 Appendix Applications

4.6.12.1 Setup

Consumers. Preferences are CES, in particular, the representative consumer solves

$$\max_{\{y_j\}_{j\in\mathcal{J}}} Y = \left(\sum_{j\in\mathcal{J}} y_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \text{ s.t. } E = PY = \sum_{j\in\mathcal{J}} p_j y_j$$

$$\textbf{Solution: } y_j = EP^{\sigma-1}p_j^{-\sigma} = Y\left(\frac{P}{p_j}\right)^{\sigma} = Y^{1-\sigma}\left(\frac{E}{p_j}\right)^{\sigma}, \text{ with } P = \left(\sum_{j\in\mathcal{J}} p_j^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

With this, the market share is given by $s_j = \frac{p_j y_j}{E} = \left(\frac{p_j}{P}\right)^{1-\sigma} = \left(\frac{y_j}{Y}\right)^{1-\frac{1}{\sigma}}$, which can be expanded (both in prices and in quantities):

$$s_j = \left[\sum_{j'=1}^{J}\left(\frac{p_{j'}}{p_j}\right)^{1-\sigma}\right]^{-1} = \left[\sum_{j'=1}^{J}\left(\frac{A_j}{A_{j'}}\frac{\mathcal{M}_{j'}}{\mathcal{M}_j}\right)^{1-\sigma}\right]^{-1} = \left[\sum_{j'=1}^{J}\left(\frac{y_{j'}}{y_j}\right)^{\frac{\sigma-1}{\sigma}}\right]^{-1}$$

Lemma 14 *(Derivatives of s_j, P and Y):*

1. $\frac{\partial P}{\partial p_j} = \left(\frac{p_j}{P}\right)^{-\sigma} = s_j^{-\frac{\sigma}{1-\sigma}}$

2. $\frac{\partial Y}{\partial y_j} = Y^{\frac{1}{\sigma}}y_j^{-\frac{1}{\sigma}} = s_j^{-\frac{1}{\sigma-1}}$

3. $\frac{\partial s_j}{\partial y_j} = \frac{\sigma-1}{\sigma}s_j\left[\frac{1}{y_j} - \frac{1}{Y}\frac{\partial Y}{\partial y_j}\right] = \frac{\sigma-1}{\sigma}\frac{s_j}{y_j}\left[1-\left(\frac{y_j}{Y}\right)^{1-\frac{1}{\sigma}}\right] = \frac{\sigma-1}{\sigma}\frac{s_j}{y_j}(1-s_j)$

4. $\frac{\partial s_j}{\partial y_k} = -\frac{\sigma-1}{\sigma}s_j\frac{1}{Y}\frac{\partial Y}{\partial y_k} = -\frac{\sigma-1}{\sigma}\frac{s_j}{y_k}\left(\frac{y_k}{Y}\right)^{1-\frac{1}{\sigma}} = -\frac{\sigma-1}{\sigma}\frac{s_j}{y_k}s_k$

5. $\frac{\partial s_j}{\partial p_j} = (1-\sigma)s_j\left[\frac{1}{p_j} - \frac{1}{P}\frac{\partial P}{\partial p_j}\right] = (1-\sigma)\frac{s_j}{p_j}\left[1-\left(\frac{p_j}{P}\right)^{1-\sigma}\right] = (1-\sigma)\frac{s_j}{p_j}(1-s_j)$. *Note that* $\frac{\partial s_j}{\partial \mathcal{M}_j} = \frac{\partial s_j}{\partial p_j}\frac{\partial p_j}{\partial \mathcal{M}_j} = (1-\sigma)\frac{s_j}{p_j}(1-s_j)\frac{w}{A_j} = (1-\sigma)\frac{s_j}{\mathcal{M}_j}(1-s_j)$

6. $\frac{\partial s_j}{\partial \mathcal{M}_k} = \frac{\partial s_j}{\partial p_k}\frac{w}{A_k} = -(1-\sigma)s_j\frac{1}{P}\frac{\partial P}{\partial p_k}\frac{w}{A_k} = -(1-\sigma)\frac{s_j}{\mathcal{M}_k}\left(\frac{p_k}{P}\right)^{1-\sigma} = (\sigma-1)\frac{s_j}{\mathcal{M}_k}s_k$.

7. *When we are taking quantities as given, A_j has no effect on market shares; but when we take markups as given, they do, and they are given by:* $\frac{\partial s_j}{\partial A_j} = s_j(1-s_j)\frac{\sigma-1}{A_j}$, *and* $\frac{\partial s_j}{\partial A_k} = -s_j s_k\frac{\sigma-1}{A_k}$

Firms. Firm j's production function is given by $y_j = A_j \ell_j$, and so profits write $\pi_j = s_j(1-\mathcal{M}_j^{-1})$, where $\mathcal{M}_j = \frac{p_j}{w}A_j$ is the markup. Firm j solves $\max_{p_j} s_j(1-\mathcal{M}_j^{-1})$, s.t $\mathcal{M}_j = \frac{p_j}{w}A_j$, and $s_j = \left(\frac{p_j}{P}\right)^{1-\sigma}$.

Proof of Lemma 8: For the FOC:

$$\frac{\partial \pi_j}{\partial \mathcal{M}_j} = \frac{\partial s_j}{\partial \mathcal{M}_j}(1-\mathcal{M}_j^{-1}) + s_j\mathcal{M}_j^{-2} = (1-\sigma)\frac{s_j}{\mathcal{M}_j}(1-s_j)(1-\mathcal{M}_j^{-1}) + s_j\mathcal{M}_j^{-2}$$
$$= \frac{s_j}{\mathcal{M}_j}\left[\mathcal{M}_j^{-1} - (\sigma-1)(1-s_j)(1-\mathcal{M}_j^{-1})\right]$$

1. $\frac{\partial F_j}{\partial \mathcal{M}_j} = -\frac{1}{\mathcal{M}_j^2} - \frac{\sigma-1}{\mathcal{M}_j^2}(1-s_j) + (\sigma-1)(1-\mathcal{M}_j^{-1})\frac{\partial s_j}{\partial \mathcal{M}_j} = -\frac{\sigma-1}{\mathcal{M}_j}\left[1 - s_j - \frac{1}{(\sigma-1)(1-s_j)}\frac{\partial s_j}{\partial \mathcal{M}_j}\right]$, where the last equality follows from the FOC.[7] And substitute for $\frac{\partial s_j}{\partial \mathcal{M}_j}$.
   And $\frac{\partial F_j}{\partial \mathcal{M}_k} = (\sigma-1)(1-\mathcal{M}_j^{-1})\frac{\partial s_j}{\partial \mathcal{M}_k} = \frac{1}{\mathcal{M}_j}\frac{1}{1-s_j}\frac{\partial s_j}{\partial \mathcal{M}_k}$, where the last equality comes from the FOC. And substitute for $\frac{\partial s_j}{\partial \mathcal{M}_k}$.

2. $\frac{\partial F_j}{\partial A_j} = \frac{\partial F_j}{\partial s_j}\frac{\partial s_j}{\partial A_j}$ and $\frac{\partial F_j}{\partial A_k} = \frac{\partial F_j}{\partial s_j}\frac{\partial s_j}{\partial A_k}$, were we need the expressions of $\frac{\partial s_j}{\partial A_j}$ and $\frac{\partial s_j}{\partial A_k}$ from Lemma 14, and that $\frac{\partial F_j}{\partial s_j} = (\sigma-1)(1-\mathcal{M}_j^{-1}) = \frac{1}{\mathcal{M}_j}\frac{1}{1-s_j}$, where the last equality follows from the FOC.

□

Proof of Lemma 9: First: (i) $\frac{\partial G_j}{\partial s_j} = 1-\mathcal{M}_j^{-1}$, (ii) $\frac{\partial G_j}{\partial \mathcal{M}_j} = \frac{s_j}{\mathcal{M}_j^2}$, and $\frac{\partial G_j}{\partial \pi_j} = -1$.

Then:

1. $\frac{\partial G_j}{\partial \mathcal{M}_j} = 0$ due to the FOC, and $\frac{\partial G_i}{\partial \mathcal{M}_j} = \frac{\partial G_i}{\partial s_i}\frac{\partial s_i}{\partial \mathcal{M}_j} = (1-\mathcal{M}_i^{-1})(\sigma-1)\frac{s_i s_j}{\mathcal{M}_j} = \frac{s_i s_j}{\mathcal{M}_j\mathcal{M}_i(1-s_i)}$

2. $\frac{\partial G_j}{\partial A_j} = \frac{\partial G_j}{\partial s_j}\frac{\partial s_j}{\partial A_j} = (1-\mathcal{M}_j^{-1})(\sigma-1)\frac{s_j(1-s_j)}{A_j} = \frac{s_j}{\mathcal{M}_j A_j}$, and $\frac{\partial G_i}{\partial A_j} = \frac{\partial G_i}{\partial s_i}\frac{\partial s_i}{\partial A_j} = -(1-\mathcal{M}_i^{-1})(\sigma-1)\frac{s_j s_i}{A_j} = -\frac{s_j s_i}{\mathcal{M}_i A_j(1-s_i)}$

□

4.6.12.2 Cournot Competition

Firm j solves $\max_{y_j} s_j(1-\mathcal{M}_j^{-1})$, s.t $\mathcal{M}_j = \frac{p_j}{w}A_j$, $p_j = EY^{\frac{1-\sigma}{\sigma}}y_j^{-\frac{1}{\sigma}}$, and $s_j = \left(\frac{y_j}{Y}\right)^{1-\frac{1}{\sigma}}$. We are going to show that: (i) $\frac{d\mathcal{M}_j}{dA_j} \geq 0$, $\frac{d\mathcal{M}_k}{dA_j} \leq 0$; and (ii) $\frac{dy_j}{dA_j} \geq 0$, $\frac{dy_k}{dA_j} \geq 0$

Lemma 15 *The FOC writes:*

$$F_j(y_1, \ldots, y_n) := \frac{\sigma-1}{\sigma}(1-s_j) - \mathcal{M}_j^{-1} = 0$$

We will need the following derivatives:

[7] In particular, substitute $-\frac{1}{\mathcal{M}_j^2} = -\frac{1}{\mathcal{M}_j}(\sigma-1)(1-\mathcal{M}_j^{-1})(1-s_j) = -\frac{\sigma-1}{\mathcal{M}_j}(1-s_j) + \frac{\sigma-1}{\mathcal{M}_j^2}(1-s_j)$ in the first term, and $(\sigma-1)(1-\mathcal{M}_j^{-1}) = \frac{1}{\mathcal{M}_j(1-s_j)}$ in the third term.

$$\frac{\partial F_j}{\partial y_j} = -\left(\frac{\sigma-1}{\sigma}\right)^2 \frac{s_j(1-s_j)}{y_j} \ , \qquad \frac{\partial F_i}{\partial y_j} = \left(\frac{\sigma-1}{\sigma}\right)^2 \frac{s_i s_j}{y_j}$$
$$\frac{\partial F_{j_0}}{\partial A_{j_0}} = \frac{1}{A_{j_0}\mathcal{M}_{j_0}} \ , \qquad \frac{\partial F_j}{\partial A_{j_0}} = 0$$

Proof. For the FOC, first, note that $\frac{\partial p_j}{\partial y_j} = p_j \left[\frac{1-\sigma}{\sigma}\frac{1}{Y}\frac{\partial Y}{\partial y_j} - \frac{1}{\sigma}\frac{1}{y_j}\right] = \frac{p_j}{y_j}\left[\frac{1-\sigma}{\sigma}\left(\frac{y_j}{Y}\right)^{1-\frac{1}{\sigma}} - \frac{1}{\sigma}\right] = -\frac{p_j}{y_j}(\frac{\sigma-1}{\sigma}s_j + \frac{1}{\sigma}) = -\frac{p_j}{y_j}(s_j + \frac{1}{\sigma}(1-s_j))$

$$\frac{\partial \pi_j}{\partial y_j} = \frac{\partial s_j}{\partial y_j}(1-\mathcal{M}_j^{-1}) + s_j\mathcal{M}_j^{-2}\frac{\partial \mathcal{M}_j}{\partial p_j}\frac{\partial p_j}{\partial y_j} = \frac{\sigma-1}{\sigma}\frac{s_j}{y_j}(1-s_j)(1-\mathcal{M}_j^{-1}) - s_j\mathcal{M}_j^{-2}\frac{A_j}{w}\frac{p_j}{y_j}(s_j + \frac{1}{\sigma}(1-s_j))$$
$$= \frac{s_j}{y_j}\left[\frac{\sigma-1}{\sigma}(1-s_j)(1-\mathcal{M}_j^{-1}) - \mathcal{M}_j^{-1}(s_j + \frac{1}{\sigma}(1-s_j))\right]$$
$$= \frac{s_j}{y_j}\left[\frac{\sigma-1}{\sigma}(1-s_j) - (1-s_j)\mathcal{M}_j^{-1}\left(\frac{\sigma-1}{\sigma} + \frac{1}{\sigma}\right) - s_j\mathcal{M}_j^{-1}\right] = \frac{s_j}{y_j}\left[\frac{\sigma-1}{\sigma}(1-s_j) - \mathcal{M}_j^{-1}\right]$$

1. $\frac{\partial F_j}{\partial y_j} = \frac{\partial F_j}{\partial s_j}\frac{\partial s_j}{\partial y_j} = -\left(\frac{\sigma-1}{\sigma}\right)^2\frac{s_j(1-s_j)}{y_j}$, and $\frac{\partial F_i}{\partial y_j} = \frac{\partial F_i}{\partial s_i}\frac{\partial s_i}{\partial y_j} = \left(\frac{\sigma-1}{\sigma}\right)^2\frac{s_i s_j}{y_j}$

2. $\frac{\partial F_j}{\partial A_j} = \frac{\partial F_j}{\partial \mathcal{M}_j}\frac{\partial \mathcal{M}_j}{\partial A_j} = \frac{1}{\mathcal{M}_j A_j}$, and $\frac{\partial F_i}{\partial A_j} = 0$

where we have used: (i) $\frac{\partial F_j}{\partial \mathcal{M}_j} = \frac{1}{\mathcal{M}_j^2}$; (ii) $\frac{\partial F_j}{\partial s_j} = -\frac{\sigma-1}{\sigma}$; (iii) $\frac{\partial \mathcal{M}_j}{\partial A_j} = \frac{\mathcal{M}_j}{A_j}$ ■

Then, we have the linear system $Ax = c$, where $(A)_{i,j} = a_{i,j} := -\frac{\partial F_i}{\partial y_j}$, $x_j := \frac{dy_j}{dA_{j_0}}$, and $c_j := \frac{\partial F_j}{\partial A_{j_0}}$.

Next, we obtain a system equivalent in signs by applying operations of the type described in Lemma 6. We suggest two alternatives:

One is to multiply column j of A by $e_j = \frac{1}{A_{j_0}\mathcal{M}_{j_0}}\left(\frac{\sigma-1}{\sigma}\right)^{-2}\frac{y_j}{s_j}\frac{1}{1-s_j}$ (i.e. change of variable $x'_j = e_j^{-1}x_j$), and row i of both A and c by $b_i = A_{j_0}\mathcal{M}_{j_0}$. So, we have:

$$(A')_{j,j} = (A)_{j,j}b_j e_j = 1 \ , \quad (A')_{i,j} = (A)_{i,j}b_i e_j = -\frac{s_i}{1-s_j}$$
$$c'_{j_0} = c_{j_0}b_{j_0} = 1 \ , \qquad c'_i = c_i b_i = 0$$

It is straightforward to check that the system satisfies column diagonal dominance of Condition 2, since $1 - s_j \geq \sum_{k\in\mathcal{J}\setminus\{j\}} s_k$.

In addition, using Corollary 4 $c_j > \sum_{\{i\in\mathcal{J}\setminus\{j\}:|a_{j,i}|c_i<0\}} |\hat{y}_i||c_i| - \sum_{\{i\in\mathcal{J}\setminus\{j\}:|a_{j,i}|c_i\geq 0\}} |\hat{y}_{i,0}||c_i|$, with $|\hat{y}_i| = 1$ and $|\hat{y}_{i,0}| = |(A')_{j,i}| = \frac{s_j}{1-s_i}$, we have:

1. $c_{j_0} = 1 > 0 = -\sum_{\{j\in\mathcal{J}\setminus\{j_0\}:c_j\geq 0\}} |(A')_{j_0,j}||c_j|$, and so $x_{j_0} \geq 0$

2. $c_j = 0 > -\frac{s_j}{1-s_i} = -\sum_{\{i\in\mathcal{J}\setminus\{j\}:c_i\geq 0\}} |(A')_{j,i}||c_i|$, and so $x_j \geq 0$, $\forall j \in \mathcal{J}\setminus\{j_0\}$

The other alternative is to multiply each column j of A by y_j, and then it is straightforward to check that the resulting matrix satisfies row diagonal dominance of Condition 6, and $\min_j A_{c,j} > 0$, so we can apply Corollary 6, and since $c_j = 0, \forall j \in \mathcal{J}\setminus\{j_0\}$, it is straightforward that $sign(x_{j_0}) = sign(c_{j_0})$, it can also be shown that $sign(x_j) = sign(c_{j_0})$, $\forall j \in \mathcal{J}\setminus\{j_0\}$, using, in addition, Corollary 4.

Effect on markups:

Lemma 16 *We will need the following derivatives:*

$$\frac{\partial F_j}{\partial \mathcal{M}_j} = \mathcal{M}_j^{-2} + \frac{(\sigma-1)^2}{\sigma}\frac{s_j}{\mathcal{M}_j}(1-s_j)\ , \quad \frac{\partial F_i}{\partial \mathcal{M}_j} = -\frac{(\sigma-1)^2}{\sigma}\frac{s_j}{\mathcal{M}_j}s_i$$
$$\frac{\partial F_{j0}}{\partial A_{j0}} = -\frac{(\sigma-1)^2}{\sigma}\frac{s_{j0}}{A_{j0}}(1-s_{j0})\ , \quad \frac{\partial F_j}{\partial A_{j0}} = \frac{(\sigma-1)^2}{\sigma}\frac{s_{j0}}{A_{j0}}s_j$$

Proof.

1. $\frac{\partial F_j}{\partial \mathcal{M}_j} = \mathcal{M}_j^{-2} - \frac{\sigma-1}{\sigma}\frac{\partial s_j}{\partial \mathcal{M}_j} = \mathcal{M}_j^{-2} + \frac{(\sigma-1)^2}{\sigma}\frac{s_j}{\mathcal{M}_j}(1-s_j)$, and $\frac{\partial F_k}{\partial \mathcal{M}_j} = -\frac{(\sigma-1)^2}{\sigma}\frac{s_j}{\mathcal{M}_j}s_k$

2. $\frac{\partial F_j}{\partial A_j} = \frac{\partial F_j}{\partial s_j}\frac{\partial s_j}{\partial A_j} = -\frac{\sigma-1}{\sigma}s_j(1-s_j)\frac{\sigma-1}{A_j}$. On the other hand, $\frac{\partial F_k}{\partial A_j} = \frac{\partial F_k}{\partial s_k}\frac{\partial s_k}{\partial A_j} = \frac{\sigma-1}{\sigma}s_j s_k \frac{\sigma-1}{A_j}$.

■

Then, we have the linear system $Ax = c$, where $(A)_{i,j} = \frac{\partial F_i}{\partial \mathcal{M}_j}$, $x_j := \frac{d\mathcal{M}_j}{dA_{j0}}$, and $c_j := -\frac{\partial F_j}{\partial A_{j0}}$.

We can obtain a system $A'x' = c'$ equivalent in signs by applying operations of the type described in Lemma 6. In particular, multiply row i of both A and c by $b_i = \frac{\sigma}{(\sigma-1)^2}\frac{A_{j0}}{s_{j0}}$, and column j of A by $e_j = \frac{s_{j0}}{A_{j0}}\frac{\mathcal{M}_j}{s_j}d_j$ (i.e. change of variable $x'_j = \frac{x_j}{e_j}$), where $d_j = \left[(\mathcal{M}_j s_j)^{-1}\frac{\sigma}{(\sigma-1)^2} + 1 - s_j\right]^{-1}$. So, we have:

$$(A')_{j,j} = (A)_{j,j}b_j e_j = 1\ , \quad (A')_{i,j} = (A)_{i,j}b_i e_j = -s_i d_j$$
$$c'_{j0} = c_{j0}b_{j0} = 1 - s_{j0}\ , \quad c'_i = c_i b_i = -s_i$$

So, this transformed system is exactly the same as the case of Bertrand.