Optimal Firm Creation with Speculation †

Valentin Haddad
UCLA and NBER

Paul Ho
Princeton University

Erik Loualiche
University of Minnesota

December 19, 2017

Abstract

Episodes of booming firm creation often coincide with intense speculation on financial markets. Disagreement among investors transforms the economics of optimal firm creation. We characterize the interaction between speculation and classic entry externalities from growth theory through a general entry tax formula for a non-paternalistic planner. The business-stealing effect is mitigated when investors believe they can identify the best firms. Speculation thus increases firm entry but reduces the optimal tax, potentially resulting in under-entry. The appropriability effect also vanishes, leaving only general equilibrium effects on input prices, aggregate demand, or knowledge. As a result, speculation reverses the role of many industry characteristics for efficiency. For instance, as the labor share increases, the optimal tax decreases under agreement but increases under disagreement. Further, economies with identical aggregate properties but a different market structure have the same efficiency with agreement, but call for different policies once financial market speculation is taken into account.

†Haddad: vhaddad@ad.ucla.edu. Ho: pho@princeton.edu. Loualiche: eloualic@umn.edu. We thank seminar participants at UCLA and Minnesota for useful comments. Download the latest version of the paper *here*.

Introduction

Do markets naturally generate the optimal level of firm creation? The discussion around firm creation crystallizes in episodes of abnormally high firm entry. For example, many observers wonder if the startup boom of the last few years is efficient. However, a salient feature of these episodes that has been neglected by the large literature on entry is the intense speculation among investors about which firms will ultimately be successful. In this paper, we study optimal firm creation with speculation. Acknowledging the presence of speculation drastically changes, and sometimes even reverses, the conclusions of standard analysis of entry.

Using a general framework to study optimal firm creation with speculative financial markets, we show that even though high speculation increases firm entry, it can imply less over-entry, often transforming an economy with over-entry into one with under-entry. The role of industry characteristics for efficiency is also reversed. For instance, a lower labor share leads to under-entry rather than over-entry, which is the case under the standard intuition.

We represent speculation using a new specification in which heterogeneous beliefs about which firms will succeed are symmetrically distributed among investors. Market participants agree on the aggregate distribution of productivity but disagree on the relative positions of firms in this distribution. The beliefs about firm productivities are symmetric across households, but disagreement does not wash out in aggregate. In equilibrium, each household invests in their favorite firms, increasing their valuations and incentivizing more firms to enter, thereby starting firm creation booms.[1]

We study the optimality of these outcomes by deriving the optimal tax on firm entry under a non-paternalistic criterion. In particular, we Pareto rank allocations while respecting each household's belief. This is straightforward in our model because the symmetry across households implies that all agents share the same expected utility despite the difference in beliefs. We favor our approach over more paternalistic criteria sometimes used in the literature as our criterion respects the intrinsic lack of information when evaluating new projects. In such an environment investors must rely on their prior and we consider optimal policy which allows them to do so.

The optimal entry tax reveals the classic externalities of growth theory. We show the relative importance of these forces depends crucially on the level of speculation. When firms enter, they do not internalize that they can displace other participants — the business-stealing effect —, affect the surplus of other agents in the economy — the appropriability effect —, and alter the general equilibrium of input markets and aggregate quantities. The first two externalities do not matter with intense speculation, but the last one is left unchanged.

[1] The mechanism of disagreement leading to optimism has been studied in Van den Steen (2004).

This differential effect on externalities leads to the reversal of many implications obtained in settings without speculation.

First, we study an economy where the only reason a firm's entry decision affects competitors is through the ability to participate in the market. This business-stealing externality classically leads to over-entry: firms do not internalize the fact that by entering, they might displace another firm. This externality is weaker in presence of speculation than with agreement. Because each investor views the firms she invests in as more productive than the average firm, all investors are less worried about being displaced by a new entrant. In the high speculation limit, we find that the business-stealing externality disappears altogether if the supply of firms is relatively inelastic. However, when the supply of new firms is elastic, the quantity of firms increases so much that only extremely productive firms survive. Disagreement does not affect the tail of the distribution, and therefore the economy is akin to one with agreement. In more general settings, we focus on settings with inelastic entry, reflecting the high asset prices often observed when many firms enter.

Market participation is not the only margin through which firms affect welfare. They can also have an impact on input markets, or other aggregate dimensions such as aggregate demand or aggregate knowledge. For example, a fixed supply of labor input introduces two more discrepancies between the social value of a firm and its private value. First, households supplying labor collect some of the surplus created by firms, which pushes towards under-entry. Second, more entry leads to a more productive set of active firms and a higher demand for labor. This yields higher equilibrium wages, which hurts the profits of other firms, pushing towards over-entry. In a high speculation episode, the appropriability effect becomes negligible as speculators value the firms they invest in much more than the average firm in the market, which is the one driving labor income. In contrast, the proportional impact of wages does not depend on firm productivity in a model with a fixed labor share, leaving this externality unchanged. Therefore, with high speculation, only the latter externality survives.

The particular externalities that survive with high speculation lead to a reversal in the role of industry characteristics. With agreement, a smaller labor share implies a smaller share of surplus for workers and increases over-entry. With disagreement, the planner instead focuses on the general equilibrium effect of firm entry the wedge. This effect becomes weaker with a smaller labor share, leading to less over-entry. In the limit where the labor share is zero, the economy is efficient with high speculation. Similarly, the Pareto tail exponent of the productivity distribution has a positive effect on the entry tax with agreement, but a negative effect with high speculation. Extensions to the framework reveal that these reversals occur more generally.

The observation that only general equilibrium externalities survive is true beyond the effect on input markets. With differentiated goods, positive aggregate demand externalities are unchanged by speculation. With knowledge spillovers, positive knowledge externalities survive. Both of these positive forces can lead to situations where there is actually under-entry with speculation. If these externalities are strong enough, the conditions for over-entry (under-entry) under agreement imply under-entry (over-entry) with disagreement.

Our analysis of these general equilibrium mechanisms highlights that the precise structure of the economy is crucial for optimal entry decisions in presence of speculation. In our setting, with agreement, only the aggregate production function matters for the optimal tax. With disagreement, the specific forces and externalities underlying the aggregate production function impact the optimal tax. For example, introducing aggregate demand effects (Aghion and Howitt, 1992; Grossman and Helpman, 1992) or knowledge externalities (Romer, 1986, 1990) does not alter the macroeconomic features and optimal entry tax of our baseline model under agreement. However, both features change the optimal tax under disagreement. Our results emphasize the importance of understanding the microeconomic structure of firms' interaction in situations with disagreement.

Our setup is flexible and maintains its tractability under a number of other features present in the literature on firm creation. In particular, we extend our tax formulas to setups where a variable number of firms participate in the ex-post market, either exogenously or endogenously as in Hopenhayn (1992) and Melitz (2003). In these richer models, again we find that speculation affects the economics of optimal entry. For instance, the role of the Frisch elasticity of labor supply and the elasticity of firm participation on output markets are reversed with disagreement.

Our results suggest that it is important, during an episode of high innovation, to determine if the high entry is coming from high overall productivity prospects, actual or perceived, or from speculation. We show that different combinations of expected productivity and disagreement can lead to the same price and quantity of entry, but imply very different optimal entry taxes. Information on financial markets will help differentiate those situations. Starting from primitives, one could directly measure heterogeneity in opinions about firms. Other equilibrium objects are also informative: speculation lead to more specialized portfolios.

We introduce our model of speculation with business stealing and the welfare analysis in Section 1. Section 2 considers optimal entry in richer general equilibrium models. We consider the practical implications of our model in Section 3.

Literature Review

TO BE COMPLETED.

1 Business Stealing with Speculation

Further, economies with identical aggregate properties but a different market structure have the same efficiency with agreement, but call for different policies once financial market speculation is taken into account.

In this section, we lay out our framework for speculation and introduce our welfare approach in the context of a simple model. This first model focuses on one of the classic forces behind overentry, business stealing, and ignores all other sources of firms' interaction. Firms compete for a fixed number of slots to produce in equilibrium. When a firm enters it does not internalize that it might take up the slot of another firm. We show that the business stealing externality is weaker with speculation, and completely disappears in high speculation environments with high asset prices. The results of this section constitute a building block for the more complete study of entry externalities of Section 2.

1.1 Model

We consider a two-period model where firms are created at date $t = 0$ and consumption and production are realized at date $t = 1$.

1.1.1 Setup

The economy is populated by households, firms, and firm creators. The commodities in the economy are a date-0 consumption good in fixed supply, blueprints produced by households and used to create firms, and a date-1 consumption good produced by firms.

Firm Creators. There is an continuum of short-lived firm creators operating a technology to create at date 0 the firms that will produce at date 1. Each firm creator can use a unit blueprint to create a new firm. If they create a firm, they sell it on competitive financial markets at date 0. They do not have any information about the firms they create. We assume firm creators participate in competitive markets for blueprints and firms, taking their respective prices p_b and p_i as given. The firm creator problem is therefore:

$$\max_{c \in \{0,1\}} c \cdot (p_i - p_b). \tag{1.1}$$

Firm creators are owned by households and have aggregate profits Π.

Firms. A continuum of firms, indexed by i and with total mass M_e, is created in equilibrium. At date 1, each firm's productivity a is revealed and production occurs. We assume that only the most productive mass M of firms is able to produce. This allocation of production slots is the key assumption to capture the business stealing externality; we discuss it below. This implies that given the cumulative distribution function (cdf) of productivities in the population F, only firms above a cutoff $\underline{a}$ are able to produce, with:

$$\underline{a} := F^{-1}\left(1 - \frac{M}{M_e}\right). \tag{1.2}$$

The profit function of a firm with productivity a is given by:

$$\pi(a) = a^{\eta} \cdot \mathbf{1}\left\{a \geq \underline{a}\right\}, \tag{1.3}$$

where η determines how differences in productivities translate into differences in generated profits. In this setting, the marginal active firm collects positive profits. While this feature enhances the tractability of our framework, we show it is not what drives our results in Section 1.4.

Households. There is a unit mass of households indexed by j. At date 0, household j is endowed with a fixed unit of consumption good c_0 as well as her share of firm creators. At date 0, each household decides how many blueprints to supply, b_j, and the share she invests in each firm on financial markets, $\{s_i^j\}_i$. Blueprints are produced at a convex cost $W(b_j) = 1/(\theta+1) f_e \left(M_e/M\right)^{\theta+1}$, where the parameter θ controls the elasticity of supply of blueprints and f_e the level of production costs. On financial markets, we assume that agents can only take long positions in claims to firms. At date 0, households have a view on the distribution of future productivities, which we describe in detail hereafter. They also behave competitively and take prices as given. Hence, they solve the problem:

$$\max_{c_0, s_i^j \geq 0, b_j} \; c_0 + \mathbf{E}^j\left\{\int s_i^j \pi_i di\right\} - W(b_j) \tag{1.4}$$

$$\text{s.t. } c_0 + \int s_i^j p_i di \leq 1 + p_b b_j + \Pi, \tag{1.5}$$

where $\mathbf{E}^j$ is the household j expectation.

Markets. There are two markets in this economy: the market for blueprints and the market for claims to firms profits (the stock market). Hence we can write market clearing conditions for each market across all households:

$$\int b_j dj = M_e,$$

$$\int s_i^j dj = 1.$$

Finally, to concentrate on situations where fewer firms produce than are created, we assume:[2]

$$W'(M) \leq \int_{-\infty}^{+\infty} \pi(a) dF(a).$$

Business Stealing. This economy is not purely competitive: we assume the production slots are allocated to the most productive firms rather than traded on a market.[3] This missing market gives rise to the business stealing externality because firms do not internalize that they might take up the slot of another firm.

The assumption of a missing market to allocate production slots is plausible, in particular in the context of innovation. For instance intellectual property law often provides exclusive use of a technology to its inventor. Formally, we could assume there are M processes to produce the homogenous good. The first firm to discover a process gets its exclusive use. The speed of discovery for a process is perfectly correlated with the productivity type a.

Another motivation for the incompleteness is the difficulty to establish markets for what has not been encountered yet. Indeed it is often the case that nobody owns something before it is discovered. For instance how could we trade nuclear power before Marie Curie discovered radioactivity? Formally, we could assume that to produce, a firm needs one unit of an indivisible good that has not been discovered yet —the unknown ingredient—and that only M of those exists in nature. Again firms with a higher type a, find the ingredient faster than others.[4]

1.1.2 Beliefs

To capture the notion of speculation, we assume that households have heterogenous beliefs about which firms will be successful.

Households agree on the ex-ante population distribution of firm productivity F, which we assume follows a Pareto distribution: $F(a) = 1 - a^{-\gamma}$ for $a \geq 1$. However, once firms are created, agents have different priors on which firms will be successful. Each agent observes all the firms organized in packets of size n, where n is an integer. She believes she knows the exact ranking of productivity draws of the firms within each packet. For instance, if a firm is predicted to be the most productive of its packet, the agent perceives it has productivity drawn from F^n, the distribution of the maximum of n independent draws of F. We assume the composition of packets and the order of firms within packets is drawn in an i.i.d. equiprobable fashion across agents and firms.

[2] The case where $M_e = M$ is straightforward to analyze.

[3] We study perfectly competitive benchmarks in Section 2.

[4] Other frictions in the allocations of productive positions can also lead to a business stealing effect, network goods, or fixed number of slots due to institutional constraints—see Borjas and Doran (2012) for evidence in the context of scientific research.

The parameter n controls the intensity of disagreement and therefore the speculation it generates. When households see packets of size $n = 1$, they consider all firms to be the same, with productivity drawn from F. As n increases, households can compare more firms to each other. Thus, they have stronger views on which firms are likely to have high productivity, the best of each packet. These views differ across households because they see different packets and differing rankings.

This set of assumptions has a couple of notable implications which facilitate our analysis. First, while they disagree on which firms will succeed, all agents agree on the population distribution of firms at date 1. This is because the population of all n firms in a packet is always viewed as having distribution F. Hence, they also agree on all aggregate outcomes: the threshold $\underline{a}$ as well as market conditions which play a role in the richer settings in Section 2. Second, beliefs are completely symmetric across households. This alleviates the typical issue of having to keep track of the entire distribution of beliefs and allocations. In particular, all households will have the same level of utility in equilibrium, rendering welfare comparison straightforward.

1.1.3 Competitive Equilibrium

The competitive equilibrium of this economy is defined as follows. Firm creators maximize profits from selling their firms taking wages as given. Households maximize their perceived expected utility by choosing their optimal blueprint discovery effort and optimal portfolio, taking the price of blueprints and of firms as given. Firms maximize profits given their production status.

Firm creators' profit maximization implies that the price of a firm equals the price of a blueprint,

$$p_i = p_b.$$

Denote the multiplier on household j's budget constraint by λ_j. Household optimization is characterized by the following first-order conditions with respect to date-0 consumption, investment in a firm i purchased by the household, and blueprint provision:

$$\begin{aligned} \lambda_j &= 1 \\ \mathbf{E}^j\{\pi(a_i)\} &= \lambda_j p_i \\ W'(b) &= p_b \end{aligned}$$

Because the equilibrium is symmetric, each investor only buys the firms she prefers within each packet, in quantity n, thereby guaranteeing market clearing on the market for firms.

Combining these conditions and market clearing for blueprints, we obtain the following entry condition characterizing equilibrium firm creation:

$$W'(M_e) = V^{(n)}(M_e) = \int_{\underline{a}}^{+\infty} \pi(a) dF^n(a). \tag{1.6}$$

The marginal cost of creating an additional firm, $W'(M_e)$, is equal to the expected profits to an investor who favors it, $V^{(n)}(M_e)$. It is immediate that higher levels of disagreement n result in higher firm entry M_e and a higher price of firms on financial markets $W'(M_e)$.

With our simple specification of profits, we have $V^{(n)}(M_e) = \mathcal{I}_n(M_e, \eta)$, where we define:

$$\mathcal{I}_n(\eta, M_e) = \int_{F^{-1}(1-M/M_e)}^{\infty} a^\eta dF^n(a). \tag{1.7}$$

We will see that the properties of $\mathcal{I}_n$ play a crucial role, both in this economy and in the richer models of Section 2.

1.2 Optimal Firm Creation

The goal of this paper is to characterize optimal firm creation under disagreement. We introduce our notion of optimality, then derive a measure of the wedge between the competitive equilibrium and the efficient level of entry.

Welfare criterion. Evaluating efficiency in models with heterogenous beliefs requires making choices on the treatment of beliefs, and the literature has suggested various ways to do so. We study efficiency under the Pareto criterion, respecting each household's beliefs. This corresponds to evaluating the utility of each household under their own beliefs. Then an allocation is more efficient than another one if it makes all households better off. The main existing alternative is a more paternalistic approach, where the planner knows the true distribution and evaluates allocation under this distribution.[5]

The Pareto criterion is particularly well-suited for our object of study. In a situation where new firms are created, not much information is available about them. Households must rely on their priors to evaluate these firms, and there is no telling which of these is correct. Our criterion respects this difficulty. The planner is no better judge of firms' futures than any investor. In contrast, a paternalistic approach might be more suited where heterogenous beliefs come from a failure to communicate or the irrationality of some agent, and thereby viewed as inefficient on their own. Another more positive — as opposed to normative — reason to favor the Pareto criterion is that if households in our economy are offered a vote between two Pareto ranked allocations, they would all choose the more efficient one. Our characterization of efficiency is therefore

[5] Brunnermeier, Simsek, and Xiong (2014) propose a variation which avoid taking a stance on the true distribution by considering efficiency across any convex combination of agents' beliefs.

a practical one, answering the question of which policies agents in our economy would support.

We study the difference in implications between our particular choice and the paternalistic option in Section 3.3.

Policy instrument. We are interested in optimal entry, and thus study allocations where the planner freely chooses the level of entry, leaving financial markets unaffected. The symmetry we assumed across agents guarantees a unambiguous ranking of allocations. Further, this simple policy approach reflects the policy debate of whether firm creation should be subsidized or taxed.

We could also consider a more general constrained efficient planner problem. The constraint on the planner is to allocate date-1 consumption using positive linear combinations of firm profits, reflecting the limits on trading in financial markets in the competitive equilibrium. In the case of our simple model, because production is efficient, the allocation chosen by a planner who can only choose entry is Pareto efficient for the constrained planner. It is actually the most efficient allocation if we impose additionally that welfare weights are equal across agents.

The entry wedge. Under the assumptions stated above, the planner maximizes household's perceived utility by choosing the level of entry. This corresponds to maximizing $M_e V^{(n)}(M_e) - W(M_e)$, the perceived total expected profits minus the cost of effort for firm creation. The first-order condition of this problem characterizes optimal entry:

$$W'(M_e) = V^{(n)}(M_e) + M_e V^{(n)'}(M_e) \tag{1.8}$$

The planner equalizes the marginal cost of an additional firm, to its value $V^{(n)}(M_e)$ plus the effect it has on the value of the existing firms $M_e V^{(n)'}(M_e)$.

A simple way to implement this level of entry is to use an entry tax. Firm creators must pay out a fraction τ of their revenue if they sell a firm. The total equilibrium proceeds T from the tax are rebated lump-sum to households. Equilibrium entry with such a tax becomes $W'(M_e) = (1-\tau)V^{(n)}(M_e)$. This leads to an expression for the wedge that reconciles the allocation with optimal entry:

$$\tau = \tau_n(M_e) = \frac{-M_e V^{(n)'}(M_e)}{V^{(n)}(M_e)} \tag{1.9}$$

The wedge is the opposite of the elasticity of firm value to firm entry, $-\mathcal{E}_{\mathcal{I}_n} > 0$.[6] An entry tax is optimal because of the negative business stealing externality imposed on incumbents by new entrants. As new firms enter, existing firms

[6]Throughout the paper, we denote the elasticity of quantity X to firm entry M_e by $\mathcal{E}_X = d\log(X)/d\log(M_e)$.

are displaced as the cutoff for production increases, which lowers the value of existing firms. The wedge is the value of this change applied to all firms relative to the value of one firm, hence the elasticity.

Throughout the various models of the paper, a similar entry wedge τ exists to capture entry distortions. We will use this wedge to compare optimality across different economies with different levels of entry. We follow the commonly used practice of evaluating the tax at the competitive equilibrium outcome to simplify the analysis.

1.3 Speculation Dissipates Displacement

We now study how the entry wedge depends on disagreement. We first show that entry is always more efficient with disagreement than under agreement. Then, we show conditions under which entry becomes optimal in the large disagreement limit. We leave all proofs to Appendix A.

Agreement versus disagreement. Under agreement, the wedge is

$$\tau_1(M_e) = \frac{\gamma - \eta}{\gamma}. \tag{1.10}$$

The wedge is increasing with γ and decreasing in η. Recall that the wedge comes from value of displaced profits relative to expected profits. The former is determined by the quality of marginal firms, whereas the latter comes from the whole distribution of productivity above this cutoff. As the tail of the distribution becomes fatter, γ falls, and the tax decreases because the value of a marginal firm becomes small with respect to the average firm. Similarly, when η is large, small differences in productivity translate into large differences in profits and a large difference between the marginal and the average firm. The wedge does not depend on entry M_e. This property relies on the classic result that under a Pareto distribution, the ratio between marginal and average productivity is independent of the lower cutoff.

The following proposition shows that there is always less over-entry with speculation ($n > 1$) than with agreement ($n = 1$).

Proposition 1 (Disagreement lowers business stealing)**.** *The wedge is larger with agreement than with disagreement:*

$$\tau_n \leq \tau_1 \tag{1.11}$$

Under disagreement, the wedge is smaller even though there is a higher level of firm entry than with agreement. This seemingly contradictory result arises because the planner respects the beliefs of households. Under disagreement, households only invest in their favorite firms. Therefore, each household places a lower probability on being displaced by new entrants, thereby reducing the effect of the externality that gives rise to the wedge.

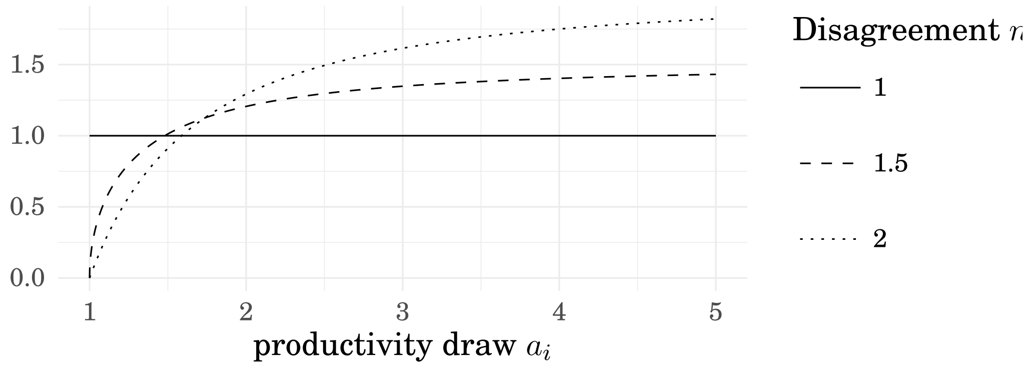


Figure 1
Distortion of productivity weights F'_n/F'

More formally, we can rewrite the wedge from (1.9) in its integral form:

$$\tau_n(M_e) = \frac{-M_e V^{(n)\prime}(M_e)}{V^{(n)}(M_e)} = \frac{\displaystyle\int_{\underline{a}}^{+\infty} \pi(\underline{a}) \frac{F'_n}{F'}(\underline{a})\, dF(a)}{\displaystyle\int_{\underline{a}}^{+\infty} \pi(a) \frac{F'_n}{F'}(a)\, dF(a)}.$$

Holding M_e and thus $\underline{a}$ constant, consider how this ratio changes from $n = 1$ and $n > 1$. The numerator, which is the expected value of displaced firms, is affected by the change in probability weights F'_n/F' at the threshold $\underline{a}$. In contrast, the denominator, which is the expected profits of a firm, is affected by changes in probability weights throughout the distribution above the threshold. Because increasing n corresponds to shifting the perceived distribution of productivities to the right, the change in probability weights is increasing as we move to higher productivities (see Figure 1). Such an increase therefore affects expected profits more strongly than the value of displaced firms, decreasing the wedge.

In Appendix A, we show that this result holds more generally with minimal assumptions on the productivity distribution $F(\cdot)$ and profit function $\pi(\cdot)$.

The high speculation limit. We now study what happens to the wedge in an economy with high speculation by considering the limit case when disagreement is large, $n \to \infty$. This situation delivers a sharp characterization of role of disagreement for the efficiency of entry. We will use the contrast of the wedge in the high speculation environment with the agreement case throughout the paper. The following proposition describes this asymptotic behavior for our model of business stealing.

Proposition 2 (Entry wedge with high speculation)**.** *In the high disagreement limit ($n \to \infty$), the entry wedge converges to a finite limit, which depends on the* *sign of $\gamma\theta - \eta$:*

- *If $\gamma\theta > \eta$, the wedge disappears:*

$$\lim_{n\to\infty} \tau_n = 0 \tag{1.12}$$

- *If $\gamma\theta < \eta$, then τ converges to the wedge in the agreement case ($n = 1$):*

$$\lim_{n\to\infty} \tau_n = \frac{\gamma - \eta}{\gamma} \tag{1.13}$$

- *In the knife-edge case of $\gamma\theta = \eta$,*

$$\lim_{n\to\infty} \tau_n = \check{\tau} < \frac{\gamma - \eta}{\gamma}, \tag{1.14}$$

  where $\check{\tau}$ is defined in appendix (A.15).

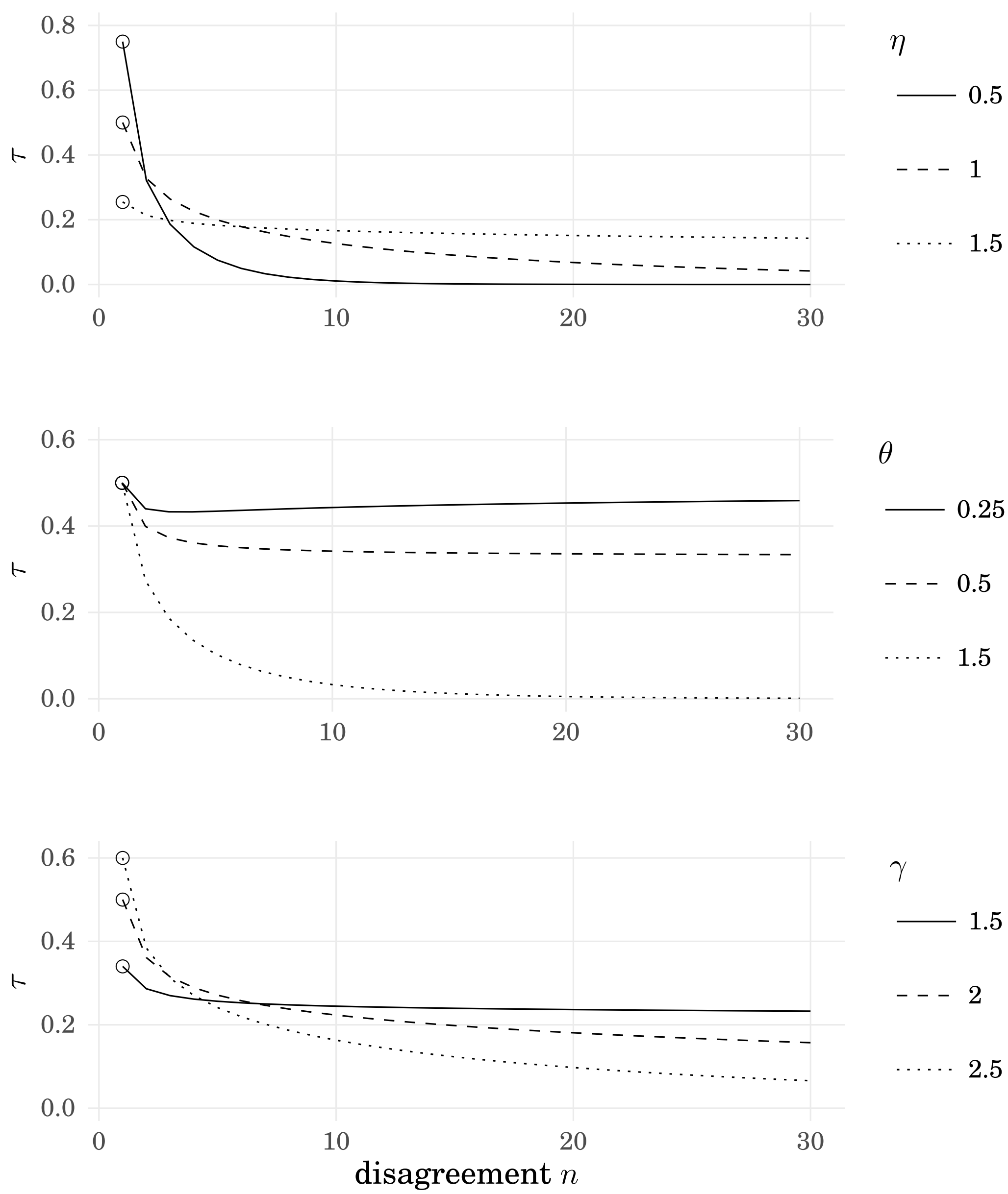
The proposition derives a condition under which entry becomes completely efficient as disagreement increases. In such a situation, as speculation picks up and prices and quantities of firms increase, the problem of over-entry actually vanishes. Interestingly, when this does not happen, the wedge goes back to the case of agreement. Figure 2 illustrates these cases.

Two forces determine the asymptotic behavior of the wedge. First, with more disagreement, investors increasingly believe that the firms they invest in are in the tail of the productivity distribution. Therefore, they are less concerned about the risk of being displaced by new entrants, as they expect that the mass of firms in their own portfolio that will be below the entry threshold is small. For a given level of entry M_e, this mass converges to 0 as n goes to infinity. This is an extreme case of the result we characterized in Proposition 1 where the wedge goes to zero.

Second, disagreement leads to more entry. For a given level of disagreement, n, as M_e converges to infinity, the M producing firms end up in the tail of both the population distribution, F, and the favorite-firm distribution, F_n. The tails of these two distributions have the same shape since $\lim_{x\to+\infty} F_n'(x)/F'(x) = n$. Therefore disagreement does not affect the relative position of the marginal and average valuation in the tail. This force brings the wedge back towards the agreement wedge.

Which of these to forces dominates depends of how fast firm creation increases with speculation. If $\theta > \eta/\gamma$, the latter force dominates. When θ is large, the marginal cost of creating rises more rapidly, which reduces equilibrium entry M_e and strengthens the first force. Such situations are characterized by high asset prices in addition to high firm entry. Similarly, as γ decreases, we have a fatter tailed firm productivity distribution. The size of the tail becomes more important than the relative ordering of firms, which strengthens the second force. As η increases, any difference in productivity translates into a larger difference in profit, which again diminishes the importance of the first force.

Figure 2
Wedge between the Competitive Equilibrium and the Planner Problem with increasing disagreement.



1.4 Zero Cutoff Profit Condition

We now investigate the robustness of our result to allowing for an alternative form of selection of producing firms. Our baseline model specifies that the M most productive firms will be allowed to produce. While this allows for tractability and a clear understanding of the model forces, this also results in marginal producers earning positive profits, $\pi(\underline{a}) > 0$. To make sure our main result does not rely on the profits of the marginal firms being positive, we augment the previous model with an intermediate stage where firms, after entering the market, compete to be among one of the M firms producing. We set this competition stage so that a business-stealing externality remains. We keep the belief and production structure of the model intact. We show how tax varies under this model that includes a zero cutoff profit (ZCP) condition for the marginal firm.

In the new intermediate decision stage, firms can use some of their production as advertisement to reach consumers, a deadweight loss. Only the M firms that spend the most on advertising produce in equilibrium. Formally, each firm chooses how much of its production to use on advertisement, $h_i \leq \pi(a_i)$. In doing so, firms take as given the equilibrium level $\underline{h}$ of advertising necessary to attract consumers. Their profit function is therefore $\pi(a_i)\mathbf{1}\{h_i \geq \underline{h}\} - h_i$. Clearly, the optimal advertisement choice is $h_i = \underline{h}$ if $\pi(a_i) \geq \underline{h}$ and 0 otherwise. The equilibrium value of $\underline{h}$ of is such that exactly M firms choose to spend more on advertisement. Keeping the definition of the production cutoff $\underline{a}$ from earlier, this implies

$$\underline{h} = \pi(\underline{a}). \tag{1.15}$$

Firms must spend the profits of the marginal firm to be able to produce, resulting in zero profits for the marginal firm.

Firm value in this model with a ZCP condition is modified to account for the cost of advertisement paid by producing firms:

$$\tilde{V}^{(n)}(M_e) = \int_{\underline{a}}^{+\infty} (\pi(a) - \pi(\underline{a}))\, dF^n(a), \tag{1.16}$$

We can define the corresponding integral $\tilde{\mathcal{I}}_n$. With this new definition of firm value, the remainder of the competitive equilibrium and the planner problem are unchanged. In particular, the entry wedge is $\tau = -\mathcal{E}_{\tilde{\mathcal{I}}_n}$. We compare the wedge between the planner problem and the competitive equilibrium and we actually find the same result as in Proposition 2.

Proposition 3 (Entry wedge with advertising)**.** *In the model with a ZCP condition and agreement ($n = 1$), the entry wedge is*

$$\tau = \frac{\gamma - \eta}{\gamma} \tag{1.17}$$

Moreover in the high disagreement limit ($n \to \infty$), the wedge between the planner problem and the competitive equilibrium converges to a finite limit:

- *If $\gamma\theta < \eta$, then* $\lim_{n\to\infty} \tau_n = (\gamma - \eta)/\gamma$.
- *If $\gamma\theta > \eta$, then* $\lim_{n\to\infty} \tau_n = 0$.

Our conclusions are therefore robust to including competition to enter. Intuitively, this is because marginal firms drive the externality in both settings. In our baseline, the externality operates at the extensive margin: more entry displaces the profits of excluded marginal firms. In this model, the externality is at the intensive margin: more entry increases the productivity of the marginal firms and therefore advertisement costs for all producing firms.

2 Speculation and Optimal Entry in General Equilibrium

2.1 Model where Firms Compete for a Scarce Input

We now build on our model from Section 1 by considering a neoclassical benchmark where firms have a production technology with decreasing returns to scale, using an input in fixed supply. We refer to this input as labor. Because entrants do not internalize their impact on this market, this gives rise to additional sources of inefficiency. We characterize how speculation affect these inefficiencies.

2.1.1 Market for Labor

Setup. We maintain the same setup — preferences, beliefs and firm creation technology — as the model of Section 1 for date 0. We now endogenize the profit function $\pi(a)$ at date 1, by adding labor in fixed supply.

Households are endowed with a fixed quantity of labor L.[7] Firms use labor to produce an homogenous good according to a decreasing returns to scale technology. The production function for a given productivity level a is:

$$y(a) = a \cdot \frac{\sigma}{\sigma - 1} \ell^{\frac{\sigma-1}{\sigma}}, \tag{2.1}$$

where y is firm output, ℓ is firm labor input, and the parameter $\sigma \in [1, \infty]$ controls the returns to scale in labor.

Labor trades at a competitive wage w. The market clearing condition for labor is:

$$\int \ell_i di = L. \tag{2.2}$$

[7]Elastic input supply in Section 2.3.

We formally define an equilibrium of this economy in Appendix B. In equilibrium, the firm profit function for a given productivity level a becomes:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} a^{\sigma}. \tag{2.3}$$

Relative to Section 1, profits are still isoelastic with respect to productivity a. However, they now also depend on the equilibrium wage w, which is affected by equilibrium entry.

Equilibrium properties. We now characterize properties of the economy at date 1 which are sufficient to analyze the entry wedge. We derive the labor and profit shares, as well as the elasticity of the wage and aggregate consumption with respect to the quantity of firms, M_e, created at date 0.

First, because of the production technology, a fraction $(\sigma - 1)/\sigma$ of the revenue for each firm is used to pay for labor and the remaining $1/\sigma$ accrues to profits. Because this is true for each firm, aggregate labor income is equal to a fraction $(\sigma - 1)/\sigma$ of aggregate consumption $\mathcal{C}$, and aggregate profits are a fraction $1/\sigma$:

$$wL = \frac{\sigma - 1}{\sigma} \cdot \mathcal{C}.$$

Second, the aggregate production frontier is homogenous of degree 1 in the distribution of productivities. Because of the Pareto distribution, increasing M_e by one percent is equivalent to increasing all productivities by $1/\gamma$ percent. For the date 1 economy, the first welfare theorem holds and production is efficient. Therefore the elasticity of aggregate consumption with respect to firm creation $\mathcal{E}_{\mathcal{C}}$ is:

$$\mathcal{E}_{\mathcal{C}} = \frac{1}{\gamma} \tag{2.4}$$

Third, we can combine these two arguments to obtain the response of the wage to firm creation.[8] Recall that aggregate labor expenditure is $wL = \mathcal{C}(\sigma - 1)/\sigma$. This immediately gives the elasticity of wages with respect to firm creation $\mathcal{E}_w$:

$$\mathcal{E}_w = \frac{1}{\gamma}. \tag{2.5}$$

2.1.2 Comparing Planner and Competitive Equilibrium

We lay out the framework to derive the wedge between the planner and the competitive equilibrium allocation.

[8]Alternatively, we can aggregate individual labor demands to determine the equilibrium wage. A firm with productivity a uses $(w/a)^{-\sigma}$ units of labor. The market clearing condition gives $L = w^{-\sigma} M_e \mathcal{I}_1$. Using equation (B.1), we obtain $\mathcal{E}_w = \frac{1}{\gamma}$.

Expected profits. Recall the definition of the integral $\mathcal{I}_n$ to aggregate the productivity of firms under the distributions F_n:

$$\mathcal{I}_n(M_e, \sigma) := \int_{\underline{a}}^{+\infty} a^{\sigma} dF^n(a).$$

Under the population distribution, the average profit of a created firm is equal to:

$$V^{(1)}(M_e) := \int_{\underline{a}}^{+\infty} \pi(a) dF(a) = \frac{w^{1-\sigma}}{\sigma - 1} \cdot \mathcal{I}_1(M_e),$$

which is the product of the integral $\mathcal{I}_1$ studied in Section 1 and a term capturing equilibrium labor market conditions. Alternatively, because all firms are ex-ante identical, they represent the same share, $1/M_e$, of aggregate profits in expectation, thus:

$$V^{(1)}(M_e) := \frac{w^{1-\sigma}}{\sigma - 1} \cdot \mathcal{I}_1(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e}.$$

These expected profits correspond to the price of a firm in a competitive equilibrium with agreement. Under disagreement, profits are aggregated under the distribution of the favorite firm rather than the population distribution. We transform one into the other by multiplying the expected profit by $\mathcal{I}_n/\mathcal{I}_1$. The value of a firm with disagreement is therefore:

$$V^{(n)}(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\mathcal{I}_n(M_e)}{\mathcal{I}_1(M_e)}$$

Competitive equilibrium. In the competitive equilibrium, firm entry equalizes the marginal cost of entry with ex-ante firm valuation:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\mathcal{I}_n(M_e)}{\mathcal{I}_1(M_e)}, \tag{2.6}$$

Planner problem. The planner maximizes expected consumption while respecting the individual household's belief. Consumption for household j is the product of labor income and profits from its investment:

$$\mathcal{C}_j = \underbrace{\frac{\sigma - 1}{\sigma} \cdot \mathcal{C}}_{\text{labor income: } wL} + \underbrace{\frac{1}{\sigma} \cdot \frac{\mathcal{I}_n}{\mathcal{I}_1} \cdot \mathcal{C}}_{\text{firm profits: } V^{(n)}}$$

Hence the planner optimization sets the marginal cost of labor to equal its effect on perceived aggregate output $\mathcal{C}_j$:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{d\left(\mathcal{I}_n/\mathcal{I}_1 \cdot \mathcal{C}\right)}{dM_e} + \frac{\sigma - 1}{\sigma} \cdot \frac{d\mathcal{C}}{dM_e} \tag{2.7}$$

Wedge between planner and competitive equilibrium allocation. Taking the ratio of social and private benefits of firm creation, as in Section 1.3, we obtain the entry wedge.

Proposition 4 (Wedge in general equilibrium)**.** *The entry wedge, τ_n between the competitive equilibrium and the planner allocation is:*

$$\tau_n = \underbrace{-\mathcal{E}_{\mathcal{I}_n}}_{\textit{bus. stealing}} + \underbrace{1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}}}_{\textit{general equilibrium}} - \underbrace{(\sigma - 1)\mathcal{E}_{\mathcal{C}} \cdot \frac{\mathcal{I}_1}{\mathcal{I}_n}}_{\textit{labor surplus}}. \tag{2.8}$$

With agreement ($n = 1$), the wedge becomes:

$$\tau_1 = 1 - \sigma \mathcal{E}_{\mathcal{C}} \tag{2.9}$$

Three externalities give rise to the wedge. The *business stealing externality* is the only externality present in the model of Section 1. As before, the contribution to the wedge from the business stealing externality is $-\mathcal{E}_{\mathcal{I}_n}$, arising because entrants displace marginal firms. This is a negative externality that is weaker with disagreement than with agreement.

The *general equilibrium externality* corresponds to the change in profits due to the change in equilibrium wages, $\mathcal{E}_w \cdot d\log\pi/d\log w = -(1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}})$. As new firms are created, the set of producing firms is more productive. Aggregate demand for labor increases, pushing wages up. The negative effect on firm value depends on the wage elasticity and on the impact of wages on profits. Since a change in the wage has the same relative effect on all firms regardless of their productivity, the general equilibrium externality does not depend on the level of disagreement.

Thirdly, the *externality from the appropriability effect* arises because investors on financial markets do not take into account the surplus accruing to workers. When more firms enter, aggregate output increases and the surplus accruing to workers increases as well — the labor share is constant and equal to $(\sigma - 1)/\sigma$. Therefore this is a positive externality from firm creation. Unlike on the market for firms, there is no speculation in labor markets. Hence this part of the social value of a firm is evaluated under the population distribution, whereas the private value is evaluated under the favorite distribution. The difference accounts for the ratio $\mathcal{I}_1/\mathcal{I}_n$ in the externality.

2.1.3 The High Speculation Limit versus Agreement

We now contrast the economics of the optimal tax in the high speculation limit with the agreement case. Along the lines of Section 1, we focus on situation of high disagreement with inelastic entry, where speculation has a large impact on asset prices.[9] The following proposition characterizes the asymptotic wedge in this case.

[9] We fully characterize the cases with elastic entry in Appendix B.1.

Proposition 5 (General model with high disagreement)**.** *Assume $\theta > 1/\gamma$. In the high disagreement limit ($n \to \infty$), the entry wedge between the competitive equilibrium and the planner problem is:*

$$\tau_\infty := \lim_{n\to\infty} \tau_n = 1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}}$$

In the high speculation limit, only the general equilibrium effects from the tax formula in equation (2.8) remain. The business stealing externality, $-\mathcal{E}_{\mathcal{I}_n}$, disappears with high speculation in the case of inelastic entry. This result is analogous to Theorem 2 in Section 1, but with a weaker parameter restriction for the externality to disappear. With fixed labor supply, the impact of disagreement on firm entry is dampened, thus increasing the difference between the average and marginal valuation.

The externality from the appropriability effect also disappears in the high speculation limit. As disagreement increases, households believe that they are investing in increasingly better firms while their labor income stagnates at the equilibrium wage determined by the physical distribution of productivity. Therefore they expect and increasing share of their welfare to come from their capital investment profits rather than labor income. Formally, the ratio $\mathcal{I}_1/\mathcal{I}_n$ converges to zero.

Now we compare how the economic environment affects this tax relative to the case of agreement. From equation (2.9), we have the formula for the tax with agreement in the model with decreasing returns to scale:

$$\tau_1 = \frac{\gamma - \sigma}{\gamma}.$$

The tax with a high level of disagreement corresponds to the general equilibrium effect:

$$\tau_\infty = \frac{\sigma - 1}{\gamma}.$$

The following proposition shows that disagreement reverses the effect of the labor share and Pareto tail on the wedge:

Proposition 6 (Comparative statics of the wedge with and without disagreement)**.** *With agreement, the entry wedge is decreasing in the labor share and decreasing in the Pareto tail parameter, γ. With high disagreement ($n \to \infty$ and $\theta > 1/\gamma$), the entry wedge is increasing in the labor share and increasing in the Pareto tail parameter, γ.*

With agreement, as the labor share increases, the positive externality from not internalizing the labor surplus becomes more important, pushing the tax down. With disagreement, there are no externalities from the appropriability effect. Instead, only the general equilibrium effect survives. As the labor share increases, firms rely more on the labor input. Hence, firm profits become more

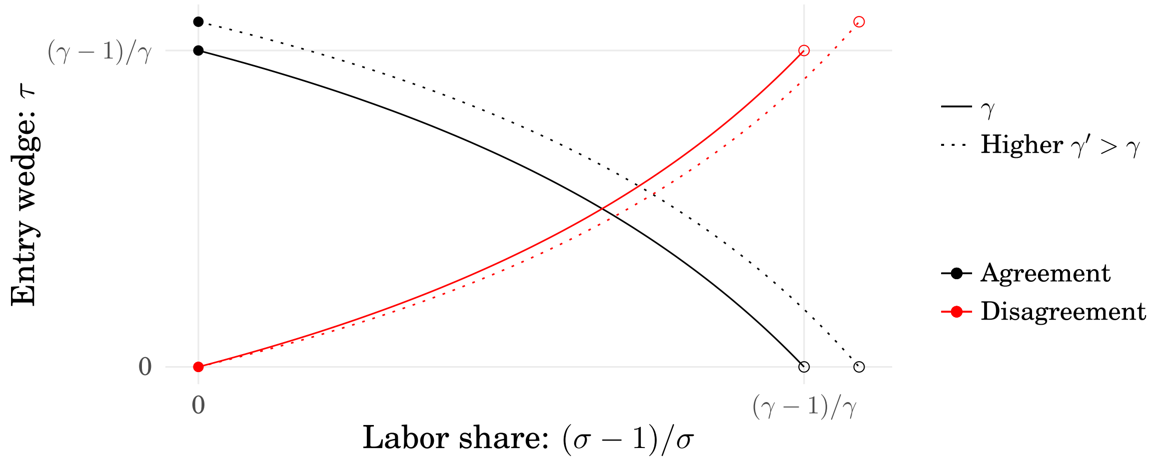


Figure 3
Comparisons of taxes with and without disagreement

sensitive to the wage, which increases the general equilibrium externality. The two opposite forces have similar strength. When the labor share moves through its range from zero to $(\gamma-1)/\gamma$, the wedge varies between zero and $(\gamma-1)/\gamma$ in both cases.[10]

Similarly, the tail of the productivity distribution affects the wedges in opposite directions. As the productivity distribution tail becomes thinner, γ is higher, and the elasticities of both aggregate consumption and the wage relative to firm entry decrease. This is because additional entry does not increase the number of superstar firms as much. With agreement, this implies that aggregate consumption exhibits more decreasing returns to scale. The growth rate of the economy in response to entry becomes even slower than implied by firms' private valuations. This leads to more over-entry and a larger wedge. With disagreement, this slowdown generates a smaller response of the wage, hence less negative externalities and a smaller wedge.

2.2 The Role of General Equilibrium Effects

Input markets are not the only mechanism through which firms interact in general equilibrium. We now study two other sources of firm interactions — aggregate demand and knowledge spillovers. We show that both Propositions 4 and 5 still hold. However the general equilibrium effects, and therefore the wedge with high speculation, depend on the nature of firms' interaction. In contrast, in the case with agreement, only the behavior of the economy's aggregates matters

[10]The labor share is bounded from above by $(\gamma-1)/\gamma$ and not one. This is the consequence of the integrability condition $\sigma < \gamma$.

and defines the wedge.

We set up both models and then we study the peculiarities of the wedge in both environments. We leave details of the models and derivations to Appendix B.2 and Appendix B.3 for the model of aggregate demand and knowledge spillovers respectively.

Aggregate demand. To capture the role of aggregate demand, we study an economy with differentiated goods where firms operate under monopolistic competition at date 1, and date 0 is unchanged from Section 2.1. Each firm produces a differentiated variety and household utility over the set of goods produced is:

$$\mathcal{C}=\left(\int_0^{M_e}\int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty}c\left(a,i\right)^{\frac{\sigma-1}{\sigma}}\,dF(a)di\right)^{\frac{\sigma}{\sigma-1}}$$

Firms operate a linear technology in labor and output for a firm with productivity a is $y=a\ell$.

The economy is similar to the previous model at the macroeconomic level. The profit share is $1/\sigma$. The aggregate production function is also homogeneous of degree one in the distribution of productivities and the relative labor allocations are efficient.[11] The macroeconomic elasticity of aggregate consumption to firm entry is therefore $\mathcal{E}_{\mathcal{C}}=1/\gamma$, which is equal to the wage elasticity, $\mathcal{E}_w=1/\gamma$.

However the microeconomics of firms' interactions is different. Profits are:

$$\pi(a)=\frac{1}{\sigma}\cdot\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}w^{1-\sigma}\cdot\mathcal{C}\cdot a^{\sigma-1}.$$

Profits are less sensitive to individual firm productivity, with an elasticity of $\sigma-1$ instead of σ. However, profits are now increasing in aggregate demand, $\mathcal{C}$, the consequence of imperfect substitution across goods.

Knowledge spillovers. We capture the role of knowledge spillovers by assuming a firm's productivity combines its own type, a, and an aggregate of all the active firms' productivity, A. We assume the aggregator is homogenous of degree one in the productivity distribution.[12] The production function is:

$$y=\frac{\sigma}{\sigma-1}a^{1-\alpha}A^{\alpha}\ell^{\frac{\sigma-1}{\sigma}},$$

where α is the intensity of the knowledge spillovers. This implies aggregate productivity has an elasticity $\mathcal{E}_A=1/\gamma$ with respect to firm creation.

[11]This result was first shown in Lerner (1934). It is the consequence of the homogeneous distortions at the firm level when markups are constant.

[12]In Appendix B.3 we derive the case of Hölder mean of degree q, $A=\left(M_e/M\int_{\underline{a}}^{\infty}a^q dF(a)\right)^{1/q}$, where the parameter $q<\gamma$ controls how aggregate knowledge comes from the top firms.

All the macroeconomic features of the economy with decreasing returns to scale are preserved: the labor share is $(\sigma - 1)/\sigma$ and $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = 1/\gamma$. However, the microeconomics of firms' interactions differ from the model with decreasing returns to scale. Profits are:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot A^{\alpha\sigma} \cdot a^{(1-\alpha)\sigma}$$

As the intensity of knowledge spillovers increases, larger α, firm profits become less sensitivite to individual types and more sensitive to the aggregate distribution.

Behavior of the wedge with externalities. In both of these economies, the labor share is $(\sigma - 1)/\sigma$ and profits are isoelastic in firms' productivity. These two conditions are sufficient to obtain both proposition 4 and 5.[13] The following proposition stresses how the economics of the wedge with and without disagreement differs sharply.

Proposition 7 (Role of the microeconomic structure with disagreement)**.** *With agreement, our baseline economy (decreasing returns to scale) and the economies with aggregate demand or knowledge spillovers have the same wedge:*

$$\tau_1 = \frac{\gamma - \sigma}{\gamma}$$

With high disagreement ($n \to \infty$) and $\theta > 1/\gamma$, the wedges differ, taking the forms:

$$\tau_\infty^{DRS} = \frac{\sigma - 1}{\gamma}$$
$$\tau_\infty^{AD} = \frac{\sigma - 2}{\gamma}$$
$$\tau_\infty^{KS} = \frac{(1-\alpha)\sigma - 1}{\gamma}$$

for the decreasing returns to scale, aggregate demand, and knowledge spillovers models respectively.

The wedge with agreement only depends on the macroeconomic aggregates, and not on the microeconomic structure of the economy. The size of the appropriability effect relative to profits is determined by the labor share. The private value of profits assume a linear growth with firm creation M_e, whereas social value takes into account the elasticity of aggregate consumption $\mathcal{E}_{\mathcal{C}}$.

The irrelevance of microeconomic structure does not hold in an economy with disagreement. With high levels of speculation, only the general equilibrium externality contributes to the wedge, rather than the business stealing or

[13]The condition for convergence $\theta > 1/\gamma$ is unchanged as long as $\mathcal{E}_{\mathcal{C}} = 1/\gamma$.

appropriability effects. The microeconomic structure of the economy is therefore crucial to understanding the wedge. In the baseline model, the general equilibrium effect operates through the wage, giving rise to a tax of $-(1-\sigma)\mathcal{E}_w$. In the aggregate demand model, aggregate demand influences profits, leading to an additional positive externality that lowers the tax by $\mathcal{E}_{\mathcal{C}} \cdot d\log\pi/d\log\mathcal{C} = 1/\gamma$. Similarly, knowledge spillovers cause increases in aggregate productivity to induce a positive externality. The tax is lower than in the baseline model by $\mathcal{E}_A \cdot d\log\pi/d\log A = \alpha\sigma/\gamma$. The stronger is the spillover, the lower the wedge.

In both cases, the labor share continues to have opposite effects on the tax with and without disagreement. With agreement, the tax is decreasing in the labor share. In contrast, it is increasing under disagreement. Furthermore, the additional positive externalities arising from the aggregate demand and knowledge spillover channels can lower the wedge below zero, leading to an optimal policy of subsidies to firm entry rather than taxes. This gives rise to another sharp difference in the optimal tax between the case with agreement and disagreement. We state the conditions under which the wedge changes sign under agreement or disagreement in the following proposition:

Proposition 8 (Sign reversal for the wedge)**.** *With demand externalities or knowledge spillovers, if the labor share is close to zero, the wedge is positive with agreement and negative with large disagreement. The converse happens when the labor share is close to its upper-bound.*

The proposition implies that situations calling for an entry tax with agreement require a subsidy with disagreement and vice-versa — see Figure 4. The reversal of the sign is true throughout the range of the labor share whenever $\gamma = 2$ with demand externalities or $\alpha = 1 - 1/\gamma$ with knowledge spillovers. For instance when the labor share is low, the labor surplus is relatively small, and the dominant force for the tax is that firms do not internalize the aggregate decreasing returns to scale of the economy. This issue is addressed by a positive tax on firm entry. With disagreement however, since firms do not rely much on labor, the general equilibrium externality is small, hence the positive demand or knowledge externality dominates. An entry subsidy encourages firm creation to take advantage of this dominant positive externality.

2.3 Extensions

To highlight the flexibility of our framework, we now discuss several extensions. In particular, we study the role of an elastic supply of labor input, a variable number of producing firms M, and a setup when firms compete to participate as in Section 1.4. Finally, we consider a setting where fixed costs determine the set of producing firms, as in Melitz (2003). For all these models, we obtain simple generalizations of the wedge formula of Proposition 4 and find that the

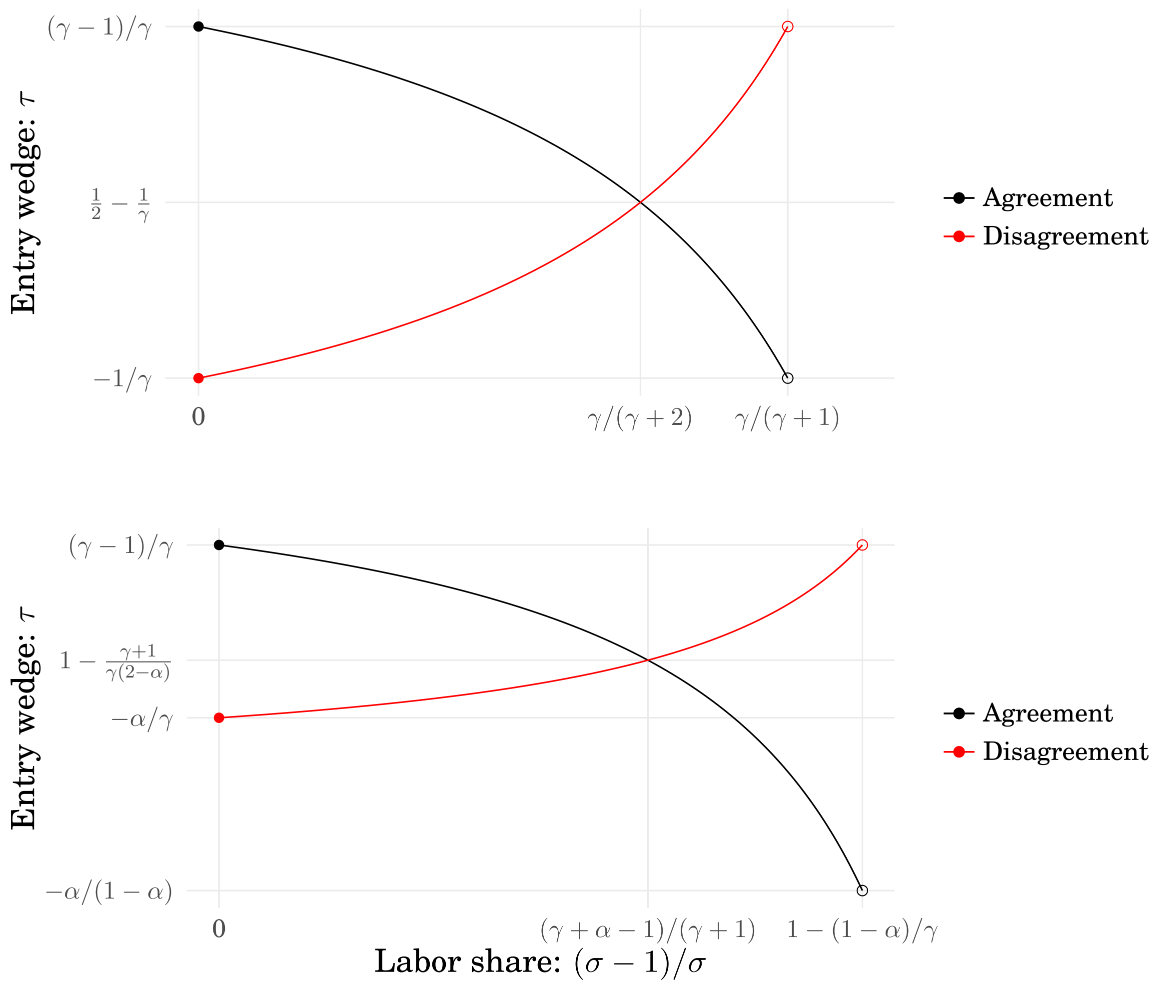


Figure 4
Comparisons of taxes with and without disagreement

comparative statics of Proposition 6 are still valid. Details and derivations of the models are in Appendix C and D.

Production factor in variable supply. We consider the case of a variable labor supply by allowing households to provide labor L by exerting an effort cost $S(L)$. We assume

$$S'(L) = f_l \left(\frac{L}{L_0}\right)^{1/\kappa},$$

where κ is the Frisch elasticity of labor supply. As κ converges to 0, the model converges to a constant labor supply L_0. The remainder of the model is unchanged.

The labor share is still constant and equal to $(\sigma-1)/\sigma$, but the elasticities of both the wage and consumption to firm entry change: $\mathcal{E}_w = \sigma/\left(\gamma(\kappa+\sigma)\right)$ and $\mathcal{E}_{\mathcal{C}} = (\sigma+\kappa\sigma)/\left(\gamma(\sigma+\kappa)\right)$. As we increase the elasticity of labor supply, the wage becomes less elastic to firm entry and consumption becomes more elastic to firm entry as it becomes less costly to expand labor.

The wedge is similar to Proposition 4. However, the labor surplus term now accounts for the utility cost of expanding the labor supply

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma-1)\underbrace{(\mathcal{E}_{\mathcal{C}} - \mathcal{E}_L)}_{\mathcal{E}_w}\frac{\mathcal{I}_1}{\mathcal{I}_n}$$

The wedges for the cases with agreement and with high disagreement (and high θ) are:

$$\tau_1 = \frac{\gamma-\sigma}{\gamma},$$
$$\tau_\infty = (\sigma-1)\cdot\mathcal{E}_w = \frac{\sigma-1}{\gamma}\frac{\sigma}{\kappa+\sigma}.$$

The wedge with agreement is unchanged from the baseline model. The wedge with high disagreement is increasing in σ and decreasing in γ, as is the case in Proposition 6. The elasticity of labor supply does not affect the wedge with agreement, but with disagreement a higher κ lowers the tax. As labor supply becomes more elastic, the wage becomes less responsive to entry, and firms have less influence on each other through general equlibrium effects. With perfectly elastic labor supply, there are no general equilibrium effects, and the economy with disagreement is efficient, i.e. $\tau_\infty = 0$.

Variable number of participating firms. We study a model where the number of participating firms, M, responds to firm creation M_e. For instance, households' consumption bundles may become more or less concentrated as more firms enter the economy. We assume

$$M = \frac{1}{M_0^{\chi-1}}\cdot M_e^{\chi},$$

where χ is the elasticity of firms producing to firms created.

The elasticities of consumption and the wage are still equal to each other:

$$\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = \frac{1}{\gamma} + \chi\cdot\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)$$

The equation for the wedge from Proposition 4 still applies with a generalized version of $\mathcal{I}_n$ that accounts for firm participation. The wedges for the case with agreement and the case with high disagreement (and high θ) are:

$$\tau_1 = 1 - \sigma \mathcal{E}_{\mathcal{C}} = \frac{\gamma - \sigma}{\gamma} \cdot (1 - \chi)$$
$$\tau_\infty = (\sigma - 1) \cdot \mathcal{E}_w = \frac{\sigma - 1}{\gamma} + \chi (\sigma - 1) \left(\frac{1}{\sigma} - \frac{1}{\gamma} \right)$$

As long as $|\chi| < 1$, Proposition 6 still holds with the generalized version of $\mathcal{I}_n$. A higher elasticity of firm participation with respect to firm entry leads to opposite results with or without disagreement. A large elasticity χ dampens the wedge with agreement because it diminishes business stealing. However it leads to a higher wedge with disagreement: in response to firm entry, labor demand responds at the intensive margin with more productive firms, and at the extensive margin with more participating firms.

Advertisement costs. In Section 1.4, we discussed the behavior of the business-stealing externality in a model where firms compete to produce, and found identical results to the baseline model. We revisit this analysis in the context of our general equilibrium model. The wedge is:

$$\tau_n = -\mathcal{E}_{\tilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_{\mathcal{C}} \frac{\mathcal{I}_1}{\tilde{\mathcal{I}}_n},$$

where the adjusted integral $\tilde{\mathcal{I}}_n$ aggregates productivity net of advertisement costs. The aggregate economic quantities are the same as in the baseline model, with $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = 1/\gamma$. However, the profit share is now lower because of advertisement costs. A fraction $(\sigma - 1)/\sigma$ of output accrues to labor but only a fraction $1/\gamma$ ($< 1/\sigma$) is collected as profits. The larger importance of labor relative to profits gives rise to a lower tax with agreement, $\tau_1 = 1/\sigma - 1/\gamma$. The tax with disagreement is unchanged, with $\tau_\infty = (\sigma - 1)/\gamma$. Proposition 6 still holds. In Appendix C.3 we confirm that it also holds with aggregate demand and knowledge spillovers.

Entry costs. To ensure the marginal firm makes zero profits, we consider an alternative mechanism: firms invest in infrastructure to produce. We assume that upon entry all firms can participate on the goods market. Firms must acquire one unit of infrastructure to reach all of their customers. Households produce infrastructure competitively at a cost of effort Φ. In an equilibrium with M producing firms the price of infrastructure is

$$\Phi'(M) = \varphi(M) = \varphi_0 \cdot M^\nu,$$

with $\nu > 0$, such that the higher the mass of producing firms M, the larger the cost of infrastructure. Participation is now a good traded on a competitive market, so the first welfare theorem holds and the wedge with agreement is zero: $\tau_1 = 0$. More generally the wedge is

$$\tau_n = -\mathcal{E}_{\tilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_{\mathcal{C}} \cdot \frac{\mathcal{I}_1}{\tilde{\mathcal{I}}_n} - \underbrace{\frac{\gamma - \sigma}{\gamma}(\mathcal{E}_{\mathcal{C}} - \mathcal{E}_M) \cdot \frac{\mathcal{I}_1}{\tilde{\mathcal{I}}_n}}_{\substack{\text{surplus from} \\ \text{participation costs}}},$$

where the last term accounts for the surplus from infrastructure creation. With high disagreement (and θ large enough) we have the following limit:

$$\tau_\infty = (\sigma - 1) \cdot \mathcal{E}_w = \frac{\sigma - 1}{\gamma} \cdot \left(\frac{1 + \nu}{1 + \nu + 1/\gamma - 1/\sigma} \right).$$

Again, Proposition 6 holds. Moreover τ_∞ is decreasing in ν. As the cost of producing infrastructure becomes steeper, the sensitivity of firm participation to firm creation is smaller, and the wedge is less responsive. In the limit with $\nu \to \infty$, a fixed number of firms produces, we are back to our baseline $\tau_\infty = (1 - \sigma)/\gamma$.

These types of entry costs are often used in conjunction with monopolistic competition and CES demand, for instance in Melitz (2003) and related work. We study this combination in Appendix D.2.

3 Implications

In this section we discuss situations where our analysis is likely to be empirically relevant. We point out that it is difficult to separate the role of speculation from that of productivity using macroeconomic aggregates. However their normative implications differ sharply. The model suggests directions to measure speculation using disaggregated data. We discuss the extant empirical literature along the lines of our framework. Finally we compare our welfare approach to a paternalistic criterion, highlighting the different implications for tax policy.

3.1 High Entry Episodes

An unambiguous consequence of high speculation is high firm creation. Across all of our settings, higher disagreement n goes with a higher level of firm entry M_e. Large entry in a sector usually coincides with disruptive innovation, either through a large technological change or the introduction of new products. The intrinsic novelty of these breakthroughs is conducive to an environment with low information where investors must rely on their priors. This makes our model of speculation, with heterogeneous beliefs, well suited to represent these situations. Further supporting this view, these episodes of high firm entry also tend to coincide with effervescent financial markets — Janeway (2012) and Brunnermeier and Oehmke (2013) review the evidence on this relation.

The current condition of the Tech sector is representative of this nexus of facts. The development of large-scale social networks and new forms of communication technologies has fostered the creation of a slew of creative firms. These firms are financed at unusually high valuations in a decentralized manner through venture capital. It is difficult to predict which of these ideas will be successful in the long run and different financiers bet on different firms.

However, it is important to acknowledge other explanations of a joint observation of high valuations and high entry. They could also be the consequence of the high future productivity of these new ideas, actual or perceived. Indeed, if all agents in the economy think that the new firms will be highly productive and will collect higher profits than usual, this will result in a higher demand for new firms. This pushes the price and quantity of new firms up.

Within our theory, we can see that aggregate observations of the price and quantity of entry cannot distinguish these two interpretations. To see this, we include an aggregate productivity component A to firm profits in the model of Section 1, and note the entry condition becomes

$$p_i = W'(M_e) = A\,\mathcal{I}_n(M_e).$$

Any combination of price and quantity consistent with the entry technology $W(.)$ can be rationalized by adjusting either the aggregate productivity A or the level of disagreement n.

While these two explanations are impossible to disentangle in real time based on aggregate observations alone, they have radically different implications for optimal entry policy. For instance, innovative firms often have a relatively high profit share because their value stems from their original idea. If one interprets the high entry and prices as high productivity expectations, this pushes towards enforcing a maximal entry tax. These firms displace each other, but do not create much surplus for workers. With disagreement, however, investors do not care about business stealing, and because such firms do not affect wages, there is no negative externality, pushing towards a low entry tax — see Proposition 6. The analysis in Section 2.2 shows that with speculation, the microeconomic structure of the economy has a further impact on the optimal tax. Innovative firms often learn from each other and facilitate the development of new sectors of consumption, resulting in additional positive externalities that can lead to a sign reversal in the entry tax between the two interpretations: there is over-entry under the productivity narrative, but under-entry with speculation — see Proposition 8.

3.2 Detecting Speculation

While aggregate quantities do not allow one to separate the two interpretations of high entry episodes, disaggregated data can help. A direct way to measure speculation is by getting at the primitive of the model: beliefs. Because investors rank firms differently, we should observe a larger dispersion in forecasts of future profits and stock returns in the presence of speculation. Data on individual investors' beliefs is scarce, but the empirical literature has been using dispersion in analysts' forecasts, which is more readily available. For example, Diether, Malloy, and Scherbina (2002) find that stocks with dispersed analysts forecasts experience low subsequent returns, consistent with our model. Yu (2011) aggregates this measure to portfolios such as the market and finds a similar result. Our particular model focuses on a more specific form of disagreement, where the identity of relatively optimistic investors is different across different stocks. To our knowledge, there are no empirical studies of this joint distribution of beliefs.

Another approach to detect speculation is to study outcomes: positions on the stock market. In the model, as n increases, only a fraction $1/n$ of investors buys each firm. Different investors buy different firms such that all investors have the same total position. Chen, Hong, and Stein (2002) measure the breadth of ownership for individual stocks as the fraction of the population of mutual funds investing in a given stock. They find this measure predicts low stock returns, again in line with disagreement models.

To implement the optimal entry tax, it is instrumental to not only focus on aggregate stock prices or firm quantities, but also to detect speculation. Measures of belief or ownership data are the direct empirical counterpart of speculation.

3.3 Role of the Welfare Criterion

Even if the decision-maker has confirmed the presence of speculation, one last issue remains: the choice of a welfare criterion. A crucial issue for our analysis is how to handle heterogenous beliefs. In this paper, we take a non-paternalistic approach. We use the Pareto criterion, respecting the beliefs of each investor. The main alternative in the literature is to use a more paternalistic approach, assuming the planner knows the true distribution of productivity types and evaluate allocations on behalf of households using the physical distribution. Brunnermeier, Simsek, and Xiong (2014) take a more subtle paternalistic approach and focus on allocations that are efficient across any convex combination of agents' beliefs. Because all agents in the economy agree on the aggregate productivity distribution, this criterion reduces to the standard paternalistic one in our setting.

In Section 1.1.2, we discussed reasons we favor the Pareto criterion. We now contrast the implications of this choice with the paternalistic approach, for the family of models of Section 2. Under the physical distribution, the planner maximizes aggregate consumption net of entry costs, $\mathcal{C} - W(M_e)$. This immediately gives an entry wedge of:

$$\tau_n^{\text{pater}} = 1 - \sigma \mathcal{E}_{\mathcal{C}} \frac{\mathcal{I}_1}{\mathcal{I}_n}.$$

This optimal tax is similar to the case of agreement, except that the relative social value is weighted by the ratio $\mathcal{I}_1/\mathcal{I}_n$, reflecting the distortion in the entry decision due to heterogenous beliefs — a friction with this welfare criterion. As disagreement increases, more and more firms enter even though these firms have a low value according to the planner evaluating them under the population distribution. Therefore, the tax increases. In the high disagreement ($n \to \infty$) limit, the market values firms infinitely more than their social value, $\mathcal{I}_n/\mathcal{I}_1 \to \infty$. The planner can only rein in this exhuberance by taxing entry completely, $\tau_n^{\text{pater}} \to 1$.

This result is in sharp contrast to our analysis, where the additional entry generated by disagreement is not always unesirable. Instead, we showed that high disagreement often results in less over-entry than agreement, and even under-entry overall in some cases.

4 Concluding Remarks

In this paper we propose a framework to study the role of speculation on firm creation. We find that under a non-paternalistic planner, classic results on the efficiency of firm entry are reversed in an economy with speculation. In an economy where there would be over-entry with agreement, we show that there is under-entry with speculation, despite the higher level of firm creation. We find that speculation plays a distinct role on different externalities in models of firm creation. With high speculation, direct externalities like business stealing and the distortion of households' labor surplus disappear, leading to a more efficient economy. However, indirect externalities on aggregate quantities such as labor wages, aggregate demand or knowledge spillovers are left intact. The labor share in the economy affects the efficient level of firm creation distinctly under agreement or under a speculative environment. Hence, an increase in the labor share would have dramatically different implications for efficiency under each situation.

Our paper suggests that macroeconomic aggregates are sufficient to establish the efficiency of firm creation under agreement. However, under speculation, a planner relying solely on aggregates will deliver the wrong solutions. Therefore, gathering information about the microeconomic structure of firms' interaction is crucial to determine the optimal level of firm creation.

References

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60 (2):323–351.

Borjas, George J. and Kirk B. Doran. 2012. "The Collapse of the Soviet Union and the Productivity of American Mathematicians." *Quarterly Journal of Economics* 127 (3):1143–1203.

Brunnermeier, Markus K. and Martin Oehmke. 2013. "Bubbles, Financial Crises, and Systemic Risk." In *Bubbles, Financial Crises, and Systemic Risk*, *Handbook of the Economics of Finance*, vol. 2, edited by George M. Constantinides, Milton Harris, and Rene M. Stulz, chap. 18. Elsevier, 1221 – 1288.

Brunnermeier, Markus K., Alp Simsek, and Wei Xiong. 2014. "A Welfare Criterion For Models With Distorted Beliefs." *Quarterly Journal of Economics* 129 (4):1753–1797.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein. 2002. "Breadth of ownership and stock returns." *Journal of Financial Economics* 66 (2):171 – 205. Limits on Arbitrage.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina. 2002. "Differences of Opinion and the Cross Section of Stock Returns." *Journal of Finance* 57 (5):2113–2141.

Grossman, Gene M and Elhanan Helpman. 1992. *Innovation and growth in the global economy*. MIT Press.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5):1127–1150.

Janeway, W.H. 2012. *Doing Capitalism in the Innovation Economy: Markets, Speculation and the State*. Cambridge University Press.

Lerner, A P. 1934. "The Concept of Monopoly and the Measurement of Monopoly Power." *Review of Economic Studies* 1 (3):157–175.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6):1695–1725.

Romer, Paul M. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94 (5):1002–1037.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5):71–102.

Van den Steen, Eric. 2004. "Rational Overoptimism (and Other Biases)." *American Economic Review* 94 (4):1141–1151.

Yu, Jialin. 2011. “Disagreement and return predictability of stock portfolios.” *Journal of Financial Economics* 99 (1):162 – 183.

Contents

Appendix

A Simple Model

A.1 General Results

We first express the wedge in integral form, then prove Proposition 1 for general functions F and π.

A.1.1 General Formulas

The wedge between the competitive equilibrium and the planner problem is $\tau_n(M_e) = -[M_e V^{(n)'}(M_e)]/[V^{(n)}(M_e)]$. The numerator and denominator have interpretable expressions. First rewrite the denominator in the following integral form:

$$\begin{aligned} V^{(n)} &= \int_{F^{-1}}^{\infty} \pi(x) dF_n(x) \\ &= \int_{F^{-1}}^{\infty} \pi(x) \frac{F_n'}{F'}(x)\, dF(x). \end{aligned} \tag{A.1}$$

Now the numerator can be written:

$$-M_e \frac{dV^{(n)}}{dM_e} = -\frac{M}{M_e} \cdot \pi \left[F^{-1} \left(1 - \frac{M_e}{M} \right) \right] \cdot \frac{F_n'}{F'} \left[F^{-1} \left(1 - \frac{M_e}{M} \right) \right]. \tag{A.2}$$

Denoting $F^{-1}(1 - M/M_e)$ by F^{-1} for convenienc, we can rewrite this in the integral form:

$$-M_e \frac{dV^{(n)}}{dM_e} = \int_{F^{-1}}^{\infty} \pi \left[F^{-1} \right] \frac{F_n'}{F'} \left[F^{-1} \right] dF(x). \tag{A.3}$$

This leads to the following formula for the wedge:

$$\tau_n = \frac{\int_{F^{-1}}^{\infty} \pi \left[F^{-1} \right] \frac{F_n'}{F'} \left[F^{-1} \right] dF(x)}{\int_{F^{-1}}^{\infty} \pi(x) \frac{F_n'}{F'}(x)\, dF(x)}. \tag{A.4}$$

A.1.2 Comparing the wedges

Lemma A.1. *Holding constant the number of firms, the wedge is larger with agreement than with disagreement.*

Proof. First recall the wedge $\tau_n(M_e) = \frac{-M_e V^{(n)'}(M_e)}{V^{(n)}(M_e)}$. We have the derivative:

$$V^{(n)'} = -\frac{1}{M_e} \frac{M}{M_e} \pi \left[F^{-1} \right] \cdot n \left(1 - \frac{M}{M_e} \right)^{n-1},$$

and we can bound $V^{(n)}$:

$$\begin{aligned} V^{(n)} &= \int_{F^{-1}}^{\infty} \pi(x) n F^{n-1}(x) dF(x) \\ &\geq \int_{F^{-1}}^{\infty} \pi(x) n F^{n-1} \left[F^{-1} \right] dF(x) \\ &\geq n \left(1 - \frac{M}{M_e} \right)^{n-1} \int_{F^{-1}}^{\infty} \pi(x) dF(x). \end{aligned}$$

Therefore we are able to bound the wedge for a given M_e and n:

$$\tau_n(M_e) \leq \frac{\int_{F^{-1}}^{\infty} \pi\left[F^{-1}\right] dF(x)}{\int_{F^{-1}}^{\infty} \pi(x) dF(x)} \leq \tau_1(M_e), \tag{A.5}$$

where the second inequality comes from the definition of $\tau_1(M_e)$. ∎

A.2 Power case derivations

We now outline the derivations for the case we focus on in the main text, with $F(a) = 1 - a^{-\gamma}$ and $\pi(a) = a^{\eta} \cdot \mathbf{1}\{a \geq \underline{a}\}$.

First, define

$$\underline{a} = F^{-1}(1 - \frac{M_e}{M}) = \left(\frac{M_e}{M}\right)^{1/\gamma}.$$

The ex-ante value of a firm, $V^{(n)}(M_e)$, is:

$$V^{(n)}(M_e) = \int_{\left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}}^{\infty} x^{\eta} \gamma n x^{-\gamma-1} \left(1 - x^{-\gamma}\right)^{n-1} dx \tag{A.6}$$

$$= \gamma n \underline{a}^{\eta-\gamma} \int_{1}^{\infty} t^{\eta-\gamma-1} \left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt, \tag{A.7}$$

The first derivative with respect to entrants is:

$$\frac{dV^{(n)}}{dM_e} = -\frac{1}{M_e} \cdot \frac{M}{M_e} \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}} \cdot n \left(1 - \frac{M}{M_e}\right)^{n-1}. \tag{A.8}$$

It is convenient to express $-M_e V^{(n)'}(M_e)$ as:

$$-M_e V^{(n)'}(M_e) = \underline{a}^{\eta-\gamma} \cdot n \left(1 - \underline{a}^{-\gamma}\right)^{n-1}. \tag{A.9}$$

A.2.1 Wedge and firm entry under agreement

Lemma A.2. *With agreement ($n = 1$), the wedge does not depend on the level of entry:*

$$\tau_1(M_e) = \frac{\gamma - \eta}{\gamma}. \tag{A.10}$$

The level of entry is:

$$\frac{M_e}{M} = \left(f_e \frac{\gamma - \eta}{\gamma}\right)^{-\frac{\gamma}{\gamma(\theta+1)-\eta}}. \tag{A.11}$$

Proof. Under agreement, $n = 1$, and we can derive an exact solution for the mass of firms entering in equilibrium, M_e. The value of a firm is:

$$V^{(1)}(M_e) = \gamma \underline{a}^{\eta-\gamma} \int_{1}^{\infty} t^{\eta-\gamma-1} dt$$

$$= \frac{\gamma}{\gamma - \eta} \underline{a}^{\eta-\gamma} = \frac{\gamma}{\gamma - \eta} \left(\frac{M_e}{M}\right)^{\frac{\eta-\gamma}{\gamma}}. \tag{A.12}$$

From equation (A.9) with $n = 1$, we have the numerator of the wedge:

$$-M_e \frac{dV^{(1)}}{dM_e} = \left(\frac{M_e}{M}\right)^{\frac{\eta-\gamma}{\gamma}}, \tag{A.13}$$

which leads directly to the desired formula (A.10) for the wedge. Finally, we can rewrite equation (1.6):

$$f_e \left(\frac{M_e}{M}\right)^{\theta} = \frac{\gamma}{\gamma-\eta}\left(\frac{M_e}{M}\right)^{\frac{\eta-\gamma}{\gamma}},$$

which reduces to (A.11) as desired. ∎

A.2.2 Disagreement asymptotics

Lemma A.3. *If $\theta \geq 0$, then as disagreement increases ($n \to \infty$) the mass of entrants also increases and goes to infinity:* $\lim_{n\to\infty} M_e = \infty$.

Proof. We define $\underline{a}_n = (M_e/M)^{1/\gamma}$, where M_e now depends on n, and show that $\underline{a}_n \to \infty$. Equation (1.6) implies an implicit definition of the sequence $\underline{a}_n$:

$$f_e \underline{a}_n^{\gamma(\theta+1)-\eta} = \gamma n \int_1^{\infty} t^{\eta-\gamma-1}\left(1 - \underline{a}_n^{-\gamma} t^{-\gamma}\right)^{n-1} dt.$$

Suppose $\underline{a}_n$ has a finite limit that is strictly larger than zero, i.e. $\underline{a}_\infty > 0$.[14] Then there exists N large enough such that $\forall n > N$, $\underline{a}_n > A = \underline{a}_\infty - \epsilon > 0$. We obtain a lower bound for the right-hand side of the implicit equation above:

$$\begin{aligned} I_n &= \gamma n \int_1^{\infty} t^{\eta-\gamma-1}\left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt \\ &> \gamma n \int_1^{\infty} t^{\eta-\gamma-1}\left(1 - A^{-\gamma} t^{-\gamma}\right)^{n-1} dt. \end{aligned}$$

Consider an arbitrary threshold T_n that depends on n and satisfies:

$$\begin{aligned} I_n &> \gamma n \int_{T_n}^{\infty} t^{\eta-\gamma-1}\left(1 - A^{-\gamma} t^{-\gamma}\right)^{n-1} dt \\ &> \gamma n \left(1 - A^{-\gamma} T_n^{-\gamma}\right)^{n-1} \int_{T_n}^{\infty} t^{\eta-\gamma-1} dt \\ &= \frac{\gamma}{\gamma-\eta} \cdot n \cdot T_n^{\eta-\gamma}\left(1 - A^{-\gamma} T_n^{-\gamma}\right)^{n-1}. \end{aligned}$$

Choose the threshold $T_n = n^{1/\gamma}$. The bound becomes:

$$\begin{aligned} I_n &> \frac{\gamma}{\gamma-\eta} \cdot n^{\frac{\eta}{\gamma}} \exp\left(-(n-1)\log(1 - A^{-\gamma} n^{-1})\right) \\ &> \frac{\gamma}{\gamma-\eta} \cdot n^{\frac{\eta}{\gamma}} \exp\left(-(n-1)A^{-\gamma} n^{-1} + \mathcal{O}(n^{-1})\right). \end{aligned}$$

Since $\gamma(\theta+1) - \eta \geq \gamma - \eta > 0$, this implies $I_n \to \infty$, contradicting $\underline{a}_\infty < \infty$. ∎

[14] Since the mass of firms producing cannot be higher than the mass of firms created, $\underline{a}_n \geq 1$.

Lemma A.4 (Asymptotics for firm creation)**.** *In the high disagreement limit ($n \to \infty$), we have the following asymptotics for the mass of firms created, M_e:*

- *If $\gamma\theta < \eta$, then $M_e/M = \left(\frac{1}{f_e}\frac{\gamma}{\gamma-\eta} \cdot n\right)^{\frac{\gamma}{\gamma(\theta+1)-\eta}}$.*
- *If $\gamma\theta = \eta$, then $\lim_{n\to\infty} M_e/M = \alpha_\infty n$, where α_∞ is a constant defined below.*

Proof. Substituting $\underline{a}$ into (1.6), we have:

$$\begin{aligned} f_e &= \gamma \underline{a}^{\eta-\gamma(\theta+1)} n \int_1^\infty t^{\eta-\gamma-1}\left(1-\underline{a}^{-\gamma}t^{-\gamma}\right)^{n-1} dt \\ &\simeq \gamma \underline{a}^{\eta-\gamma(\theta+1)} n \int_1^\infty t^{\eta-\gamma-1} \exp\left(-(n-1)\underline{a}^{-\gamma}t^{-\gamma}\right) dt, \end{aligned}$$

where we have used the fact that $\underline{a} \to \infty$ from Lemma A.4, and $\log(1-x) = -x + \mathcal{O}(x^2)$. To find a solution, we guess the asymptotics of $\underline{a}(n)$. We rewrite $\underline{a} = \alpha(n) n^{1/(\gamma(1+\theta)-\eta)}$ and show that $\alpha(n)$ converges to a finite limit α. The above equation becomes:

$$f_e = \gamma\alpha(n)^{\eta-\gamma(\theta+1)} \int_1^\infty t^{\eta-\gamma-1} \exp\left(-\alpha(n)^{-\gamma} \frac{n-1}{n^{\frac{\gamma}{\gamma(1+\theta)-\eta}}} t^{-\gamma}\right) dt.$$

Suppose $\gamma\theta < \eta$. Then the exponential term converges to zero and we have:

$$f_e = \gamma\alpha^{\eta-\gamma(\theta+1)} \int_1^\infty t^{\eta-\gamma-1} = \alpha^{\eta-\gamma(\theta+1)} \frac{\gamma}{\gamma-\eta},$$

such that we have the following asymptotics for firm entry:

$$\frac{M_e}{M} = \left(\frac{1}{f_e}\frac{\gamma}{\gamma-\eta} \cdot n\right)^{\frac{\gamma}{\gamma(\theta+1)-\eta}}. \tag{A.14}$$

Suppose $\gamma\theta = \eta$. Then $\underline{a}$ is defined by:

$$f_e = \gamma \underline{a}_n^{-\gamma} n \int_1^\infty t^{\eta-\gamma-1}\left(1-\underline{a}^{-\gamma}t^{-\gamma}\right)^{n-1} dt.$$

Since $\underline{a} = (M_e/M)^{1/\gamma}$, it is sufficient to guess and verify that $\underline{a}_n = \alpha(n)^{-1/\gamma} n^{1/\gamma}$, and $\alpha(n)$ has a finite limit, α_∞ defined by:

$$\begin{aligned} f_e &= \gamma\alpha(n) \int_1^\infty t^{\eta-\gamma-1} \exp\left(-(n-1)(\alpha(n)n^{-1} + \mathcal{O}(\alpha(n)^2 n^{-2}))t^{-\gamma}\right) dt \\ &\to \gamma\alpha_\infty \int_1^\infty t^{\eta-\gamma-1} e^{-\alpha_\infty t^{-\gamma}} dt, \end{aligned}$$

where we take the limit when $n \to \infty$. The wedge with agreement implies:

$$\begin{aligned} f_e &> \gamma\alpha_\infty e^{-\alpha_\infty} \int_1^\infty t^{\eta-\gamma-1} dt \\ &> \alpha_\infty e^{-\alpha_\infty} \frac{\gamma}{\gamma-\eta} \end{aligned}$$

and thus

$$\frac{\alpha_\infty e^{-\alpha_\infty}}{f_e} < \frac{\gamma - \eta}{\gamma}, \tag{A.15}$$

which implies a finite bound on α_∞. ∎

Theorem A.5 (Asymptotics for the wedge, τ)**.** *In the high disagreement limit ($n \to \infty$), the wedge between the planner problem and the competitive equilibrium converges to a finite limit.*

- *If $\gamma\theta < \eta$, then $\tau \to (\gamma - \eta)/\gamma$.*
- *If $\gamma\theta > \eta$, then $\tau \to 0$.*
- *If $\gamma\theta = \eta$, then $\tau \to \alpha_\infty e^{-\alpha_\infty}/f_e$.*

Proof. Using the asymptotics derived in lemma A.4, we consider the three cases separately.

Suppose $\gamma\theta < \eta$. Substitute the asymptotics derived in equation (A.14) into the formula for the wedge:

$$\tau_n(M_e) = \frac{\frac{M}{M_e}\left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}} n \left(1 - \frac{M}{M_e}\right)^{n-1}}{f_e \left(\frac{M_e}{M}\right)^{\theta}} \tag{A.16}$$

$$\simeq \frac{1}{f_e} \cdot f_e \frac{\gamma - \eta}{\gamma}\frac{1}{n} \cdot n \left(1 - \frac{M}{M_e}\right)^{n-1} \to \frac{\gamma - \eta}{\gamma}, \tag{A.17}$$

where we have used the fact that $(1 - M/M_e)^{n-1} \to 1$.[15] The wedge therefore converges to the wedge with agreement in this case.

Now suppose $\gamma\theta > \eta$. We write the wedge directly:

$$\tau_n(M_e) = \frac{n\underline{a}^{\eta-\gamma}(1 - \underline{a}^{-\gamma})^{n-1}}{f_e \underline{a}^{\gamma\theta}}.$$

First suppose $\underline{a} \to \infty$. We rewrite the competitive equilibrium condition (1.6):

$$n\underline{a}^{-\gamma} = \frac{f_e \underline{a}^{\gamma\theta - \eta}}{\gamma \int_1^\infty t^{\eta-\gamma-1}(1 - \underline{a}^{-\gamma}t^{-\gamma})^{n-1} dt}.$$

The denominator is bounded from above by $\gamma \int_1^\infty t^{\eta-\gamma-1} dt$, which implies $n\underline{a} \to \infty$. Using a first-order approximation, we have:

$$(1 - \underline{a}^{-\gamma}t^{-\gamma})^{n-1} \simeq \exp\left(-n\underline{a}^{-\gamma}\right).$$

Therefore, the wedge in the limit is:

$$\tau \simeq \frac{n\underline{a}^{-\gamma} \exp\left(-n\underline{a}^{-\gamma}\right)}{f_e \underline{a}^{\gamma\theta - \eta}} \to 0, \tag{A.18}$$

[15] This follows from $(1 - M/M_e)^{n-1} = \exp\left[-(n-1)\log(M_e/M)\right]$ and using the asymptotics derived above for $\gamma\theta < \eta$: $(1 - M/M_e)^{n-1} = \exp\left[-(n-1)\left(f_e^{-1}\frac{\gamma}{\gamma-\eta} n\right)^{-\frac{\gamma}{\gamma(1+\theta)-\eta}}\right] \to 1$.

since the numerator goes to zero and the denominator goes to infinity. Suppose instead that $\underline{a}$ has a finite limit. We obtain the expression for τ:

$$\tau = \frac{n(1-\underline{a}^{-\gamma})^{n-1}}{f_e \underline{a}^{\gamma(1+\theta)-\eta}} = \frac{n\exp\left((n-1)\log(1-\underline{a}^{-\gamma})\right)}{f_e \underline{a}^{\gamma(1+\theta)-\eta}} \to 0, \tag{A.19}$$

since the denominator has a finite limit and the numerator goes to 0.

Lastly, consider the case where $\gamma\theta = \eta$. The tax expression simplifies to:

$$\tau = \frac{1}{f_e} \cdot n\underline{a}^{-\gamma}\left(1-\underline{a}^{-\gamma}\right)^{n-1}.$$

Using lemma A.4 and the result that $\underline{a}^{-\gamma} = \alpha(n)/n$, and $\alpha(n) \to \alpha_\infty$ we have:

$$\tau \simeq \frac{1}{f_e}\alpha(n)\exp\left(-(n-1)\alpha(n)/n\right) \tag{A.20}$$

$$\simeq \frac{1}{f_e}\alpha(n)\exp\left(-\alpha(n)\right) \to \frac{1}{f_e}\alpha_\infty e^{-\alpha_\infty}. \tag{A.21}$$

Moreover using lemma A.4, this also proves that in the limit τ is below the wedge with agreement $(\gamma-\eta)/\gamma$. ∎

A.3 Results with a Zero Cutoff Profit Condition

We now derive the results for the model of advertisement in Section 1.4.

A.3.1 General derivations

We decompose firms' valuations into two the revenue (from (1.16)) and advertising cost components:

$$V^{(n)}(M_e) = \int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty} \pi(a)dF_n(a) - \left(\frac{M}{M_e}\right)^{1-\frac{\eta}{\gamma}} \cdot n\left(1-\frac{M}{M_e}\right)^{n-1}. \tag{A.22}$$

The first derivative of $V^{(n)}$ is:

$$-M_e \cdot \frac{dV^{(n)}(M_e)}{dM_e} = \frac{\eta}{\gamma}\left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}} \cdot \left[1-\left(1-\frac{M}{M_e}\right)^n\right].$$

Using the free-entry condition, $V^{(n)}(M_e) = W'(M_e)$, we have following formula for the wedge between planner problem and competitive equilibrium:

$$\tau_n(M_e) = \frac{\eta}{\gamma} \cdot \frac{1}{f_e} \cdot \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}-\theta} \cdot \left[1-\left(1-\frac{M}{M_e}\right)^n\right]. \tag{A.23}$$

A.3.2 Wedge and firm entry

Lemma A.6. *In the model with a ZCP condition and agreement ($n=1$), the wedge between the competitive equilibrium and the planner problem is:*

$$\tau = \frac{\gamma-\eta}{\gamma}.$$

Proof. The free-entry condition with $n = 1$ gives us:

$$\left(\frac{M_e}{M}\right)^{\frac{\gamma(1+\theta)-\eta}{\gamma}} = \frac{1}{f_e} \cdot \frac{\eta}{\gamma - \eta}.$$

Given the derivation of the wedge in (A.23), we have:

$$\tau_1(M_e) = \frac{\eta}{\gamma} \cdot \frac{1}{f_e} \cdot \left(\frac{M_e}{M}\right)^{\frac{\eta}{\gamma}-\theta} \cdot \frac{M}{M_e} = \frac{\gamma - \eta}{\gamma}, \tag{A.24}$$

where we have used our equilibrium solution for M_e/M. ∎

Lemma A.7. *In the high disagreement limit ($n \to \infty$), the mass of entrants also increases and goes to infinity:* $\lim_{n\to\infty} M_e = \infty$.

Proof. We adapt the proof from Lemma A.3, again defining the sequence $\underline{a}_n = (M_e/M)^{1/\gamma}$ and showing that $\underline{a}_n \to \infty$. Equation (A.22) implies the implicit definition of the sequence $(\underline{a}_n)_n$ in this case:

$$f_e \underline{a}_n^{\gamma(\theta+1)-\eta} = \gamma n \int_1^{\infty} t^{\eta-\gamma-1} \left(1 - \underline{a}_n^{-\gamma} t^{-\gamma}\right)^{n-1} dt - n(1 - \underline{a}_n^{-\gamma})^{n-1}.$$

Assume that $\underline{a}_n$ has a finite limit that is strictly larger than zero, $\underline{a}_\infty > 0$. Then there exists N large enough such that $\forall n > N, \underline{a}_n > A = \underline{a}_\infty - \epsilon > 0$. For any arbitrary threshold T_n we have

$$I_n > n \left[\frac{\gamma}{\gamma - \eta} \cdot T_n^{\eta-\gamma} \left(1 - A^{-\gamma} T_n^{-\gamma}\right)^{n-1} - 1\right].$$

As in Lemma A.3, we conclude by considering the threshold $T_n = n^{1/\gamma}$. ∎

Lemma A.8 (Asymptotics for firm creation). *In the high disagreement limit ($n \to \infty$), we have the following asymptotics for the mass of firms created M_e:*

- *If $\gamma\theta < \eta$, then* $\lim_{n\to\infty} M_e/M = \alpha_\infty^{\gamma} n^{\frac{\gamma}{\gamma(1+\theta)-\eta}}$.
- *If $\gamma\theta = \eta$, then* $\lim_{n\to\infty} M_e/M = \alpha_\infty^{-1} n$.

In each case, α_∞ is a constant defined below.

Proof. We adapt the proof from Lemma A.4. Starting from (1.6), and using $\underline{a}$:

$$f_e \simeq \gamma \underline{a}^{\eta-\gamma(\theta+1)} n \int_1^{\infty} \left(t^{\eta} - 1\right) t^{-\gamma-1} \exp\left(-(n-1)\underline{a}^{-\gamma} t^{-\gamma}\right) dt,$$

where we have used the fact that $\underline{a} \to \infty$ and $\log(1-x) = -x + \mathcal{O}(x^2)$. To find a solution, we guess that asymptotically $\underline{a} \simeq \alpha(n) n^{\frac{1}{\gamma(1+\theta)-\eta}}$ and show that $\alpha(n)$ converges to a finite limit α. The first-order condition becomes:

$$f_e \simeq \gamma \alpha(n)^{\eta-\gamma(\theta+1)} \int_1^{\infty} \left(t^{\eta} - 1\right) t^{-\gamma-1} \exp\left(-\alpha(n)^{-\gamma} \frac{n-1}{n^{\frac{\gamma}{\gamma(1+\theta)-\eta}}} t^{-\gamma}\right) dt.$$

If $\gamma\theta < \eta$, then the exponential term converges to zero and we have:

$$\alpha_\infty = \left(\frac{1}{f_e} \cdot \frac{\eta}{\gamma - \eta}\right)^{\frac{1}{\gamma(1+\theta)-\eta}}. \tag{A.25}$$

If $\gamma\theta = \eta$, then $\underline{a}$ is defined by:

$$f_e = \gamma \underline{a}_n^{-\gamma} n \int_1^\infty t^{\eta-\gamma-1}\left(1 - \underline{a}^{-\gamma} t^{-\gamma}\right)^{n-1} dt.$$

We guess and verify that $\underline{a}_n = \alpha(n)^{-1/\gamma} n^{1/\gamma}$, and $\alpha(n)$ has a finite limit, α_∞:

$$f_e = \gamma \alpha_\infty \int_1^\infty (t^\eta - 1)\, t^{-\gamma-1} e^{-\alpha_\infty t^{-\gamma}} dt,$$

where we took the limit when $n \to \infty$. Moreover we are able to bound the wedge above the wedge with agreement using a bound on α_∞:

$$\frac{\alpha_\infty e^{-\alpha_\infty}}{f_e} < \frac{\gamma - \eta}{\eta}, \tag{A.26}$$

which verifies that α_∞ is finite. ∎

Theorem A.9. *In the high disagreement limit ($n \to \infty$), the wedge between the planner problem and the competitive equilibrium converges to a finite limit.*

- *If $\gamma\theta < \eta$, then $\tau \to (\gamma - \eta)/\gamma$.*
- *If $\gamma\theta > \eta$, then $\tau \to 0$.*
- *If $\gamma\theta = \eta$, then $\tau \to \frac{\eta}{\gamma}\frac{1}{f_e} e^{-\alpha_\infty}$.*

Proof. If $\gamma\theta > \eta$, then given equation (A.23), we use that $M_e \to \infty$ to conclude that $\lim_{n\to\infty} \tau = 0$.

If $\gamma\theta < \eta$, then we can use the asymptotics from A.8 and the formula for the wedge from (A.23):

$$\tau_n(M_e) \simeq \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\theta\gamma} \cdot n^{\frac{\eta-\theta\gamma}{\gamma(1+\theta)-\eta}} \cdot \left[1 - \left(1 - \alpha_\infty^{-\gamma} n^{\frac{-\gamma}{\gamma(1+\theta)-\eta}}\right)^n\right] \tag{A.27}$$

$$\simeq \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\theta\gamma} \cdot n^{\frac{\eta-\theta\gamma}{\gamma(1+\theta)-\eta}} \cdot \left[1 - \exp\left(-\alpha_\infty^{-\gamma} n^{\frac{\gamma\theta-\eta}{\gamma(1+\theta)-\eta}}\right)\right] \tag{A.28}$$

$$\simeq \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\theta\gamma} \cdot n^{\frac{\eta-\theta\gamma}{\gamma(1+\theta)-\eta}} \cdot \alpha_\infty^{-\gamma} n^{\frac{\gamma\theta-\eta}{\gamma(1+\theta)-\eta}} \to \frac{\eta}{\gamma}\frac{1}{f_e}\alpha_\infty^{\eta-\gamma(1+\theta)}. \tag{A.29}$$

Using the definition of α_∞ from the proof above, we conclude $\lim_{n\to\infty} \tau_n(M_e) = (\gamma - \eta)/\gamma$.

In the knife-edge case with $\gamma\theta = \eta$, we have:

$$\tau_n(M_e) \simeq \frac{\eta}{\gamma}\frac{1}{f_e} \cdot \left[1 - \left(1 - \alpha_\infty n^{-1}\right)^n\right] \to \frac{\eta}{\gamma}\frac{1}{f_e} \cdot e^{-\alpha_\infty}. \tag{A.30}$$

We can bound the wedge in the limit: $\lim_{n\to\infty} \tau_n(M_e) < \alpha_\infty^{-1} \cdot (\gamma - \eta)/\gamma$. ∎

B General Equilibrium Models

Recall the definition of the average of a power function in productivity under the physical or distorted beliefs:

$$\mathcal{I}_n(M_e,\sigma) = \int_{\underline{a}}^{\infty} a^{\sigma} dF^n(a).$$

The integral with no disagreement is $\mathcal{I}_1$. We will use $\mathcal{I}_n$ when the dependence of the integral to M_e or σ is unambiguous. Under the Pareto distribution with parameter γ, we have the following result:

$$\mathcal{I}_1 = \frac{\gamma}{\gamma-\sigma} \cdot \left(\frac{M_e}{M}\right)^{\frac{\sigma}{\gamma}-1}. \tag{B.1}$$

B.1 Model with decreasing returns to scale

B.1.1 Equilibrium

The firm optimization problem given the production function and the competitive input price w is:

$$\max_{\ell(a)} \pi(a) = a \cdot \frac{\sigma}{\sigma-1} \ell(a)^{\frac{\sigma-1}{\sigma}},$$

The first-order condition leads to demand for labor at the firm level:

$$\ell(a) = \left(\frac{w}{a}\right)^{-\sigma}.$$

Output and profit at the firm level are:

$$y(a) = \frac{\sigma}{\sigma-1} \cdot w^{1-\sigma} \cdot a^{\sigma}$$
$$\pi(a) = \frac{1}{\sigma-1} \cdot w^{1-\sigma} \cdot a^{\sigma}.$$

Market clearing on the input market yields:

$$L = M_e \cdot w^{-\sigma} \int_{\underline{a}}^{\infty} a^{-\sigma} dF(a) = M_e \cdot w^{-\sigma} \cdot \mathcal{I}_1, \tag{B.2}$$

which, given (B.1), leads to the following wage in equilibrium under a Pareto distribution for F:

$$w = \left(\frac{\gamma}{\gamma-\sigma}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{L}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}-\frac{1}{\sigma}}.$$

Given the equilibrium quantities, we can decompose aggregate output into the profit and labor shares. First, observe that aggregate ouput is:

$$\mathcal{C} = M_e \cdot \int_{\underline{a}}^{\infty} y(a) dF(a) = M_e \cdot \frac{\sigma}{\sigma-1} w^{1-\sigma} \mathcal{I}_1,$$

From this expression we immediately conclude that:

$$w^{1-\sigma} \cdot \mathcal{I}_1 = \frac{\sigma-1}{\sigma} \cdot \frac{\mathcal{C}}{M_e},$$

and we are able to simplify the ex-ante valuation of firms:

$$\begin{aligned} V^{(n)}(M_e) &= \int_{\underline{a}}^{\infty} \pi(a) dF^n(a) = \frac{1}{\sigma-1} \cdot w^{1-\sigma} \cdot \mathcal{I}_n(M_e) \\ &= \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\mathcal{I}_n(M_e)}{\mathcal{I}_1(M_e)}. \end{aligned}$$

Finally, we express the wage as function of the equilibrium mass of firms:

$$w = \left(\frac{\gamma-\sigma}{\gamma} L\right)^{-\frac{1}{\sigma}} \cdot M^{\frac{\gamma-\sigma}{\gamma\sigma}} \cdot M_e^{\frac{1}{\gamma}}.$$

The equilibrium condition that determines entry in equilibrium is:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\mathcal{I}_n(M_e)}{\mathcal{I}_1(M_e)}. \tag{B.3}$$

B.1.2 Entry wedge

We have shown that the optimal entry tax is:

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n} + (1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}}) - (\sigma-1)\mathcal{E}_C \cdot \frac{\mathcal{I}_1}{\mathcal{I}_n}. \tag{B.4}$$

The asymptotic environment is similar to that of Section 1. We start by studying the asymptotics of the business stealing effect.

Lemma B.1 (Asymptotics for business stealing distortion)**.** *In the high disagreement limit ($n \to \infty$), the business stealing distortion converges to a limit that depends on the marginal cost of firm creation θ:*

- *If $\theta\gamma < 1$, then $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n} = \mathcal{E}_{\mathcal{I}_1} = \frac{\sigma}{\gamma} - 1$.*
- *If $\theta\gamma > 1$, then $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n} = 0$.*
- *If $\theta\gamma = 1$, then $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n} = \alpha_\infty e^{-\alpha_\infty}/f_e$.*

Proof. The free entry condition equation (B.3) leads to:

$$(\sigma-1)\left(\frac{\gamma-\sigma}{\gamma} L\right)^{\frac{1-\sigma}{\sigma}} \cdot f_e = M^{\theta + \frac{\sigma-\gamma}{\gamma}\frac{\sigma-1}{\sigma}} \cdot M_e^{\frac{1-\sigma}{\gamma} - \theta} \cdot \mathcal{I}_n$$

We recast the free entry condition using $\underline{a}$ to be able to use the asymptotic results from lemma A.4

$$\text{constant} = \underline{a}^{1-\sigma-\gamma\theta} \int_{\underline{a}}^{\infty} x^\sigma dF_n(x).$$

Writing $\tilde{\theta} = \theta + (\sigma-1)/\gamma$ and $\tilde{\eta} = \sigma$, we recognize the first-order condition from equation Lemma A.4 and use Theorem A.5. ∎

For the labor surplus term, we study the behavior of $\mathcal{I}_1/\mathcal{I}_n$.

Lemma B.2 (Asymptotics for labor surplus distortion)**.** *In the high disagreement limit ($n \to \infty$), the labor surplus distortion disappears:*

$$\lim_{n\to\infty}(\sigma-1)\mathcal{E}_{\mathcal{C}}\frac{\mathcal{I}_1}{\mathcal{I}_n}=0$$

Proof. Since $\tilde{\theta}>0$, Lemma A.3 gives $\lim_{n\to\infty}M_e=\infty$. The proof of Lemma A.3 implies $\lim_{n\to\infty}\mathcal{I}_n=\infty$. Finally, because $\sigma<\gamma$,

$$\mathcal{I}_1=\frac{\gamma}{\gamma-\sigma}\left(\frac{M_e}{M}\right)^{\frac{\sigma-\gamma}{\gamma}}\to 0$$

as $n\to\infty$. Therefore $\mathcal{I}_1/\mathcal{I}_n$ converges to 0. ∎

B.2 Differentiated goods

B.2.1 Date 1 economy

The introduction of differentiated goods in Section 2.2 changes the production stage. We therefore focus on the equilibrium conditions in date 1.

Firms produce a mass M of differentiated goods, indexed by (a,i), where a is firm productivity and i indexes the firms. We drop the i index when unambiguous. Household utility aggregates consumption of these goods with constant elasticity of substitution σ across goods. At date 1, household j with total expenditure E_j solves:

$$\mathcal{C}(E_j)=\max_{\{c(a,i)\}}\left(\int_0^{M_e}\int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty}c\,(a,i)^{\frac{\sigma-1}{\sigma}}\,dF(a)di\right)^{\frac{\sigma}{\sigma-1}}$$
$$\text{s.t.}\int_0^{M_e}\int_{F^{-1}\left(1-\frac{M}{M_e}\right)}^{\infty}p\,(a,i)\,c\,(a,i)\,dF(a)di\le E_j.$$

For reasons that will soon be clear, we denote by $1/\mathcal{P}$ the Lagrange multiplier on the budget constraint. Because the objective function is homogenous of degree 1 in consumption and the budget constraint is linear, $\mathcal{C}(E_j)$ is linear in E_j. Thus we have $\mathcal{C}(E_j)=E_j/\mathcal{P}$. Therefore $\mathcal{P}$ is the price of one unit of the consumption basket. We use this consumption basket as the numeraire at date 1 by normalizing $\mathcal{P}=1$. The linearity also implies that to aggregate individual demands, it is sufficient to know the aggregate expenditure in the economy, and not the whole distribution of individual expenditures.

The first-order condition in the problem above implies the demand curve:

$$c(p)=\mathcal{C}p^{-\sigma}.$$

Output for a firm with productivity a is $y=a\ell$. Firms face monopolistic competition. They maximize profits by setting prices, taking as given the demand curve from each household. Each firm solves

$$\max_{p(a)}\;p(a)y\left(p\left(a\right)\right)-\frac{wy\left(p\left(a\right)\right)}{a}=\mathcal{C}\left[p(a)^{1-\sigma}-\frac{w}{a}p(a)^{-\sigma}\right].$$

The optimal price is therefore

$$p(a) = \frac{\sigma}{\sigma - 1}\frac{w}{a}.$$

Firms charge a markup $\sigma/(\sigma-1)$ over their marginal cost w/a.

We can then compute output y, revenue py, labor expenditure $w\ell$ and profits π as functions of productivity:

$$\begin{aligned}
y &= \mathcal{C}w^{-\sigma}a^{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \\
py &= \mathcal{C}w^{1-\sigma}a^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \\
w\ell &= \frac{\sigma-1}{\sigma}\mathcal{C}w^{1-\sigma}a^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \\
\pi &= \frac{1}{\sigma}\mathcal{C}w^{1-\sigma}a^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}
\end{aligned}$$

We see that labor expenditure is a fraction $(\sigma-1)/\sigma$ of revenues, and profits make up the remaining $1/\sigma$ share.

Labor market clearing gives $\mathcal{C}(\sigma-1)/\sigma = wL$. In equilibrium, aggregate expenditure is equal to aggregate consumption, so we have:

$$\begin{aligned}
\mathcal{C} &= \mathcal{C}\left(\frac{\sigma}{\sigma-1}w\right)^{1-\sigma} M_e\mathcal{I}_1(M_e,\sigma-1) \\
&= M_e^{\frac{1}{\sigma-1}}\left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\sigma-1}}\left(\frac{M_e}{M}\right)^{\frac{(\sigma-1)-\gamma}{(\sigma-1)\gamma}} \cdot L \\
&= \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\sigma-1}} M^{\frac{1}{\sigma-1}-\frac{1}{\gamma}} M_e^{\frac{1}{\gamma}} \cdot L.
\end{aligned}$$

Therefore we have $\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = 1/\gamma$. Alternatively, notice that the labor allocation is efficient with monopolistic competition and the aggregate production function is homogenous of degree 1 in the distribution of productivity. Because an increase in M_e increases all productivities with an elasticity $1/\gamma$, this results in an elasticity of aggregate consumption of $1/\gamma$.

B.2.2 Entry wedge

All arguments behind Proposition 4 apply, so the propostion is still valid, but with $\mathcal{I}_1$ and $\mathcal{I}_n$ now evaluated with parameter $\sigma-1$.

With agreement, because the aggregate consumption elasticity is unchanged, the entry wedge is unchanged: $\tau_1 = (\gamma-\sigma)/\gamma$.

With speculation, the free entry condition is:

$$\frac{M_e}{M}^{\theta} = \frac{1}{\sigma}\mathcal{C}\left(\frac{\mathcal{C}}{L}\right)^{1-\sigma}\mathcal{I}_n,$$

which we can rewrite as:

$$KM_e^{\theta-(1-(\sigma-1))/\gamma} = \mathcal{I}_n,$$

where K does not depend on M_e and n. This is again the same condition as the homogenous goods model, with σ replaced by $\sigma - 1$. The condition for the convergence of $\mathcal{E}_{\mathcal{I}_n}$ from Lemma B.1, still applies as well. In the high disagreement limit with $\theta > 1/\gamma$, the tax becomes:

$$\tau_\infty = 1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}} = \frac{\sigma-2}{\gamma}. \tag{B.5}$$

B.3 Knowledge externalities

B.3.1 Date 1 economy

We again focus on the date 1 economy and return to a setting with decreasing return to scale.

Now firm productivity depends on the productivity of other firms producing. In particular, consider the production function:

$$y = \frac{\sigma}{\sigma-1} a^{1-\alpha} A^\alpha \ell^{\frac{\sigma-1}{\sigma}},$$

where a is firm productivity and A is an aggregator of all producing firms' productivities. $\alpha > 0$ captures the intensity of knowledge spillovers. We use a Hölder mean of the productivity of all firms producing:

$$A = \left(\frac{M_e}{M} \int_{\left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}} a^q dF(a) \right)^{\frac{1}{q}}.$$

Imposing $q < \gamma$ so that the integral is well defined, we have:

$$A = \left(\frac{\gamma}{\gamma - q} \right)^{\frac{1}{q}} \left(\frac{M_e}{M} \right)^{\frac{1}{\gamma}}.$$

Our results generalize to any aggregator which is homogenous of degree 1 in the productivity distribution of producing firms. Any such aggregator would similarly yield an elasticity $1/\gamma$ with respect to M_e.

Firms maximize their profits taking the wage as given:

$$\max_\ell \frac{\sigma}{\sigma-1} a^{1-\alpha} A^\alpha \ell^{\frac{\sigma-1}{\sigma}} - w\ell.$$

The demand for labor is therefore

$$\ell = \left(\frac{w}{a^{1-\alpha} A^\alpha} \right)^{-\sigma},$$

and we have:

$$y(a) = \frac{\sigma}{\sigma-1} \left(a^{1-\alpha} A^\alpha\right)^\sigma w^{1-\sigma}$$

$$w\ell(a) = \left(a^{1-\alpha} A^\alpha\right)^\sigma w^{1-\sigma} = \frac{\sigma-1}{\sigma} y(a)$$

$$\pi(a) = \frac{1}{\sigma-1} \left(a^{1-\alpha} A^\alpha\right)^\sigma w^{1-\sigma} = \frac{1}{\sigma} y(a)$$

The labor share is still $(\sigma-1)/\sigma$.

The market clearing condition for labor is:

$$\begin{aligned}
L &= w^{-\sigma} A^{\alpha\sigma} M_e \mathcal{I}_1\left(M_e, (1-\alpha)\sigma\right) \\
&= \left(\frac{\gamma}{\gamma-q}\right)^{\frac{\alpha\sigma}{q}} \left(\frac{M_e}{M}\right)^{\frac{\alpha\sigma}{\gamma}} M_e \frac{\gamma}{\gamma-(1-\alpha)\sigma} \left(\frac{M_e}{M}\right)^{\frac{(1-\alpha)\sigma}{\gamma}-1} w^{-\sigma} \\
&= \left(\frac{\gamma}{\gamma-q}\right)^{\frac{\alpha\sigma}{q}} \frac{\gamma}{\gamma-(1-\alpha)\sigma} M \left(\frac{M_e}{M}\right)^{\frac{\sigma}{\gamma}} w^{-\sigma} \\
w &= \left(\frac{M}{L}\right)^{\frac{1}{\sigma}} \left(\frac{\gamma}{\gamma-q}\right)^{\frac{\alpha}{q}} \left(\frac{\gamma}{\gamma-(1-\alpha)\sigma}\right)^{\frac{1}{\sigma}} \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}.
\end{aligned}$$

We still have $(\sigma-1)/\sigma\mathcal{C} = wL$, and the same elasticities: $\mathcal{E}_w = \mathcal{E}_{\mathcal{C}} = \mathcal{E}_A = 1/\gamma$.

B.3.2 Entry wedge

Proposition 4 and Lemma B.1 still apply, with $\mathcal{I}_1$ and $\mathcal{I}_n$ evaluated with parameter $(1-\alpha)\sigma$. The wedge with agreement is unchanged: $\tau_1 = (\gamma-\sigma)/\gamma$. The wedge in the high disagreement limit with $\theta > 1/\gamma$ becomes:

$$\tau_\infty = 1 + \mathcal{E}_{\mathcal{I}_1} - \mathcal{E}_{\mathcal{C}} = \frac{(1-\alpha)\,\sigma - 1}{\gamma}. \tag{B.6}$$

C Extensions

C.1 Elastic labor supply

We now consider the case of variable labor supply. Households can provide labor L by exerting an effort cost $S(L)$. We assume

$$S'(L) = f_l \left(\frac{L}{L_0}\right)^{1/\kappa},$$

where the parameter κ is the Frisch elasticity of labor supply. As κ converges to 0, the model converges to a constant labor supply L_0. The remainder of the model is unchanged.

C.1.1 Equilibrium.

In equilibrium, we have: $S'(L) = w$. This implies that $\mathcal{E}_L = \kappa\mathcal{E}_w$. Using the observation of a constant labor share, $(\sigma-1)/\sigma\mathcal{C} = wL$, we also have: $\mathcal{E}_C = \mathcal{E}_w + \mathcal{E}_L = (1+\kappa)\mathcal{E}_w$.

Market clearing for labor yields:

$$L = M_e \cdot w^{-\sigma} \int_{\underline{a}}^{\infty} a^{-\sigma} dF(a) = M_e \cdot w^{-\sigma} \cdot \mathcal{I}_1$$

Which given the expression for $\mathcal{I}_1$ in (B.1) under a Pareto leads to the following restriction:

$$wL^{1/\sigma} = \left(\frac{\gamma}{\gamma-\sigma}\right)^{\frac{1}{\sigma}} \cdot M_e^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}-\frac{1}{\sigma}}$$

Using the labor supply equation, we obtain:

$$\mathcal{E}_w = \frac{1}{\gamma}\frac{\sigma}{\kappa+\sigma}$$
$$\mathcal{E}_{\mathcal{C}} = \frac{1}{\gamma}\frac{\sigma+\kappa\sigma}{\sigma+\kappa}$$

C.1.2 Entry wedge

Asymptotics. To study the asymptotic behavior of the wedge, recall the free entry condition:

$$f_e\left(\frac{M_e}{M}\right)^{\theta} = \frac{1}{\sigma-1} w^{1-\sigma}\mathcal{I}_n.$$

We define

$$\tilde{\theta} = \theta + \frac{\sigma-1}{\gamma}\frac{\sigma}{\kappa+\sigma} \geq 0$$

and recognize the free entry condition from the pure business stealing model. This guarantees that $\lim_{n\to\infty} \mathcal{I}_1/\mathcal{I}_n = 0$. In addition, $\mathcal{E}_{\mathcal{I}_n}$ converges to 0 if:

$$\gamma\tilde{\theta} > \sigma \tag{C.1}$$
$$\Leftrightarrow \gamma\theta > 1 + (\sigma-1)\frac{\kappa}{\kappa+\sigma} \tag{C.2}$$

Planner problem and wedge. The planner objective is now:

$$\max_{M_e} \frac{1}{\sigma}\mathcal{C}\frac{\mathcal{I}_n}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C} - S(L) - W(M_e)$$

The first-order condition is:

$$W'(M_e) = \frac{1}{\sigma}\mathcal{C}\frac{\mathcal{I}_n}{\mathcal{I}_1}\frac{\mathcal{I}_n'}{\mathcal{I}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\mathcal{I}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\mathcal{I}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C}' - \underbrace{S'(L)L}_{\frac{\sigma-1}{\sigma}\mathcal{C}}\frac{1}{L}\frac{dL}{dM_e}.$$

The wedge is therefore:

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma-1)\underbrace{(\mathcal{E}_{\mathcal{C}} - \mathcal{E}_L)}_{\mathcal{E}_w}\frac{\mathcal{I}_1}{\mathcal{I}_n}.$$

With agreement, we have:

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - (\sigma-1)\mathcal{E}_w \tag{C.3}$$

$$= 1 - (\sigma+\kappa)\mathcal{E}_w = \frac{\gamma-\sigma}{\gamma}. \tag{C.4}$$

With high speculation and large θ, we have:

$$\tau_\infty = (\sigma-1)\mathcal{E}_w = \frac{\sigma-1}{\gamma}\frac{\sigma}{\kappa+\sigma}. \tag{C.5}$$

C.2 Variable number of participating firms

C.2.1 Setting and equilibrium

We introduce a variable number of firms M. We assume that M varies exogenously with the level of firm entry M_e:

$$M = \frac{1}{M_0^{\chi-1}} \cdot M_e^{\chi},$$

where χ is the elasticity of firms producing to firms created and M_0 a normalization constant. We assume that $\chi \leq 1$ such that we always have $M \leq M_e$.

The cost of creating a firm only depends on M_0 through:

$$W'(M_e) = f_e\left(\frac{M_e}{M_0}\right)^{\theta}.$$

The productivity threshold to produce is now:

$$\underline{a} := F^{-1}\left(1 - \frac{M}{M_e}\right) = \left(\frac{M_e}{M_0}\right)^{\frac{1-\chi}{\gamma}}.$$

The model still features a constant labor share and firm profits are still isoelastic in the productivity:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot a^{\sigma} = \frac{1}{\sigma} \frac{\mathcal{C}}{M_e} \cdot \frac{a^{\sigma}}{\mathcal{I}_1},$$

where we have redefined the integrals $\mathcal{I}_1$ and $\mathcal{I}_n$ to adjust for the new expressions for the productivity threshold $\underline{a}$:

$$\mathcal{I}_n(\chi) = \int_{\left(\frac{M_e}{M_0}\right)^{\frac{1-\chi}{\gamma}}}^{\infty} a^{\sigma} dF_n(a)$$

$$\mathcal{I}_1(\chi) = \frac{\gamma}{\gamma - \sigma} \cdot \left(\frac{M_e}{M_0}\right)^{(\chi-1)\frac{\gamma-\sigma}{\gamma}}.$$

The market-clearing condition $L = M_e w^{-\sigma} \mathcal{I}_1$ implies the equilibrium wage:

$$w = \left(\frac{\gamma}{\gamma - \sigma}\right)^{\frac{1}{\sigma}} \cdot L^{-\frac{1}{\sigma}} M_0^{(1-\chi)\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)} \cdot M_e^{\frac{\chi}{\sigma} + \frac{1-\chi}{\gamma}},$$

so that the labor elasticity is :

$$\mathcal{E}_w = \frac{1}{\gamma} + \chi \cdot \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right).$$

We obtain aggregate consumption by aggregating individual output $\mathcal{C} = M_e \sigma/(\sigma - 1) w^{1-\sigma} \mathcal{I}_1$, which yields equilibrium aggregate consumption and elasticity:

$$\mathcal{C} = \frac{\sigma}{\sigma - 1} \frac{\gamma}{\gamma - \sigma} \cdot L^{\frac{\sigma-1}{\sigma}} \cdot M_0^{(1-\chi)\left(\frac{1-\sigma}{\sigma} + \frac{\gamma-1}{\gamma}\right)} \cdot M_e^{\frac{1}{\gamma} + \chi \cdot \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)}$$

$$\mathcal{E}_{\mathcal{C}} = \mathcal{E}_w = \frac{1}{\gamma} + \chi \cdot \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)$$

C.2.2 Entry wedge

Given the constant labor share and isoelastic profits, we can apply Proposition 4 and obtain the wedge:

$$\tau_n = -\mathcal{E}_{\mathcal{I}_n(\chi)} + 1 + \mathcal{E}_{\mathcal{I}_1(\chi)} - \mathcal{E}_{\mathcal{C}} - (\sigma - 1)\mathcal{E}_C \cdot \frac{\mathcal{I}_1(\chi)}{\mathcal{I}_n(\chi)}. \tag{C.6}$$

From the expression for $\mathcal{I}_n$ we have the following change in the elasticities:

$$\mathcal{E}_{\mathcal{I}_n(\chi)} = (1 - \chi)\mathcal{E}_{\mathcal{I}_n(\chi=0)} = (1 - \chi)\mathcal{E}_{\mathcal{I}_n} \tag{C.7}$$

Asymptotics. We now turn to the high disagreement limit. The first-order condition for firm creation is:

$$f_e \left(\frac{M_e}{M_0}\right)^{\theta} = \frac{1}{\sigma - 1} w^{1-\sigma} \mathcal{I}_n(\chi)$$

$$\iff \quad \text{constant} = \underline{a}^{-\theta\frac{\gamma}{1-\chi} + (1-\sigma)\frac{\gamma}{1-\chi}\left(\frac{1}{\gamma} + \chi\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\right)} \int_{\underline{a}}^{\infty} a^{\sigma} dF_n(a).$$

We define:

$$\tilde{\theta} = \frac{1}{1-\chi}\left(\theta + \frac{\sigma-1}{\gamma} + \chi(\sigma-1)\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\right),$$

and recognize the entry condition of the baseline model. We apply our previous results, changing the condition for $\mathcal{E}_{\mathcal{I}_n(\chi)} \to 0$ to $\gamma\tilde{\theta} > \sigma$, which reduces to:

$$\gamma(\theta+\chi) > 1 + \chi\left(\frac{\gamma}{\sigma} - 1\right). \tag{C.8}$$

If this condition is satisfied, then $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n(\chi)} = 0$. When the inequality is reversed, $\lim_{n\to\infty} \mathcal{E}_{\mathcal{I}_n(\chi)} = \mathcal{E}_{\mathcal{I}_1(\chi)} = (1-\chi)(\sigma-\gamma)/\gamma$. With equality, the elasticity admits a finite limit between these two values.

Behavior of the wedge. The wedge with agreement is

$$\tau_1 = 1 - \sigma\mathcal{E}_{\mathcal{C}} = \frac{\gamma-\sigma}{\gamma} - \chi\sigma\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right). \tag{C.9}$$

The wedge with high disagreement, when θ is large enough, is:

$$\tau_\infty = 1 + \mathcal{E}_{\mathcal{I}_1(\chi)} - \mathcal{E}_{\mathcal{C}} = \frac{\sigma-1}{\gamma} + \chi\,(\sigma-1)\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) \tag{C.10}$$

The wedge with agreement is decreasing in σ (for $|\chi| < 1$) and increasing in γ.

C.3 Advertising to Participate

We augment the previous model with an intermediate stage after market entry when firms compete to be one of M firms producing, as in Section 1.4.

C.3.1 Setting and equilibrium

In the new intermediate decision stage, firms choose to spend on advertisement to reach consumers. We assume that only the M firms that spend the most on advertising produce in equilibrium. Formally if a firm with productivity a_i spends h_i in advertising, its profit is: $\pi(a_i)\mathbf{1}\{h_i \geq \underline{h}\} - h_i$. Hence there is a threshold level of advertising, $\underline{h}$, below which firms cannot reach any consumers and above which firms do produce. Firms take the threshold as given and decide on their choice of advertising. Thus the advertising equilibrium is such that the threshold matches the profit of the marginal firm: $\underline{h} = \pi(\underline{a})$.

Profits are modified with respect to the standard model of Section B.1 to incorporate the advertisement payments:

$$\pi(a) = \frac{1}{\sigma} w^{1-\sigma}\left(a^\sigma - \underline{a}^\sigma\right),$$

The ex-ante firm valuation is therefore:

$$V^{(n)}(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1},$$

where we define the adjusted integral $\tilde{\mathcal{I}}_n$ as

$$\tilde{\mathcal{I}}_n(M_e) = \int_{\left(\frac{M_e}{M}\right)^{\frac{1}{\gamma}}}^{\infty} \left(a^{\sigma} - \left(\frac{M_e}{M}\right)^{\frac{\sigma}{\gamma}}\right) dF_n(a).$$

The entry condition in the competitive equilibrium is now:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}.$$

C.3.2 Entry wedge

The planner maximizes expected utility:

$$\max_{M_e} \; - W(M_e) + \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C}$$

The corresponding optimality condition is:

$$W'(M_e) = \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\tilde{\mathcal{I}}_n'}{\tilde{\mathcal{I}}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C}'.$$

The wedge is therefore

$$\tau_n = -\mathcal{E}_{\tilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - (\sigma-1)\mathcal{E}_{\mathcal{C}}\frac{\mathcal{I}_1}{\tilde{\mathcal{I}}_n}. \tag{C.11}$$

With agreement, the wedge is:

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - (\sigma-1)\mathcal{E}_{\mathcal{C}}\frac{\gamma}{\sigma} = \frac{1}{\sigma} - \frac{1}{\gamma}, \tag{C.12}$$

where we have used that $\mathcal{E}_{\mathcal{C}} = 1/\gamma$. The wedge with agreement, τ_1, is equal to its value in the standard model at its lower and upper bound with $\sigma = 1$ and $\sigma = \gamma$ respectively. As in the standard model, τ_1 is decreasing in σ and increasing in γ.

To derive the wedge with high disagreement, notice that $\mathcal{I}_1/\tilde{\mathcal{I}}_n \to 0$ and therefore, if $\theta > 1/\gamma$, then:

$$\tau_\infty = \frac{\sigma}{\gamma} - \frac{1}{\gamma} \tag{C.13}$$

τ_∞ is unchanged from the standard model.

C.3.3 Advertising with Demand or Knowledge Spillovers.

We now introduce advertisement costs to the aggregate demand and knowledge externalities models from Sections B.2 and B.3.

The formula for τ_n in both of these models is the same as in Equation (C.11), so that we have the limit with high disagreement $\tau_\infty = \sigma/\gamma - 1/\gamma$.

The main difference between wedge in the standard model and the models in Section B.2 and B.3 arises through the differences in the profit function, which affects the ratio $\mathcal{I}_1/\tilde{\mathcal{I}}_n$.

With aggregate demand externalities the ratio is $\gamma/(\sigma-1)$ and the wedge with agreement is:

$$\tau_1^{\mathrm{AD}} = -\frac{1}{\gamma}. \tag{C.14}$$

For the model with knowledge spillovers the ratio is $\gamma/(1-\alpha)\sigma$ and the wedge with agreement is:

$$\tau_1^{\mathrm{KS}} = 1 - \frac{1}{\gamma} - \frac{\sigma-1}{(1-\alpha)\sigma}. \tag{C.15}$$

D Participation Costs

We study models where firms have to invest in infrastructure to participate in the goods market.

D.1 Participation Costs in the Baseline Model

We assume that upon entry all firms can participate on the goods market, but firms must buy one unit of infrastructure to reach all of their customers. Households produce infrastructure competitively at a cost of effort Φ. In an equilibrium with M producing firms, the price of infrastructure is:

$$\Phi'(M) = \varphi(M) = \varphi_0 \cdot M^{\nu}$$

such that the higher the mass of producing firms M, the larger the cost of infrastructure.

D.1.1 Participating firms

Given M_e and M, profits before the infrastructure costs are unchanged from the standard model with decreasing returns to scale:

$$\pi(a) = \frac{1}{\sigma - 1} \cdot w^{1-\sigma} \cdot a^{\sigma}.$$

The equilibrium wage is also unchanged:

$$w = \left(\frac{\gamma}{\gamma - \sigma}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{L}\right)^{\frac{1}{\sigma}} \cdot \left(\frac{M_e}{M}\right)^{\frac{1}{\gamma} - \frac{1}{\sigma}}.$$

The marginal firm has productivity $\underline{a}$ and spends all of its profit in infrastructure. Therefore, we have the zero cutoff profit condition $\Phi'(M) = \pi(\underline{a})$, which implies:

$$M^{\nu + \frac{1}{\gamma} + \frac{\sigma-1}{\sigma}} = \frac{1}{\varphi_0} \frac{1}{\sigma - 1} \left(\frac{\gamma - \sigma}{\gamma}\right)^{\frac{\sigma-1}{\sigma}} L^{\frac{\sigma-1}{\sigma}} \cdot M_e^{\frac{1}{\gamma}},$$

where we use the fact that $\underline{a} = (M_e/M)^{1/\gamma}$. In Section C.2, we specified an exogenous set of producing firms $M = M_e^{\chi}/M_0^{\chi-1}$. This arises endogenously through our cost of infrastructure with

$$\chi = \frac{1}{\gamma}\left(\nu + \frac{1}{\gamma} + \frac{\sigma - 1}{\sigma}\right)^{-1}$$

$$M_0^{1-\chi} = \left(\frac{1}{\varphi_0} \frac{1}{\sigma - 1} \left(\frac{\gamma - \sigma}{\gamma}\right)^{\frac{\sigma-1}{\sigma}} L^{\frac{\sigma-1}{\sigma}}\right)^{\left(\nu + \frac{1}{\gamma} + \frac{\sigma-1}{\sigma}\right)^{-1}},$$

where the exponent satisfies $\chi \leq 1$.

We can also compute the elasticity $\mathcal{E}_{\mathcal{C}}$:

$$\mathcal{E}_{\mathcal{C}} = \frac{1}{\gamma} + \chi\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) = \frac{1}{\gamma} \cdot \frac{1 + \nu}{1 + \nu + 1/\gamma - 1/\sigma} = \chi \cdot (1 + \nu).$$

D.1.2 Equilibrium

The equilibrium condition in the competitive equilibrium is:

$$W'(M_e) = \frac{1}{\sigma} \cdot \frac{\mathcal{C}}{M_e} \cdot \frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1},$$

where we define the modified $\tilde{\mathcal{I}}_n$ integral to account for the infrastructure expenditures of the firm:

$$\tilde{\mathcal{I}}_n(M_e, \chi) = \int_{\left(\frac{M_e}{M_0}\right)^{\frac{1-\chi}{\gamma}}}^{\infty} \left(a^\sigma - \left(\frac{M_e}{M_0}\right)^{\sigma \frac{1-\chi}{\gamma}}\right) dF_n(a).$$

With $n = 1$, we have:

$$\tilde{\mathcal{I}}_1(M_e, \chi) = \frac{\sigma}{\gamma - \sigma} \cdot \left(\frac{M_0}{M_e}\right)^{(1-\chi)\frac{\gamma-\sigma}{\gamma}} = \frac{\sigma}{\gamma} \cdot \mathcal{I}_1(M_e, \chi).$$

Aggregate profits therefore represent a fraction σ/γ of aggregate revenue after labor costs, while aggregate infrastructure costs account for the other $(\gamma - \sigma)/\gamma$. Therefore aggregate profits represent a share $1/\gamma$ of consumption and aggregate infrastructure costs $1/\sigma - 1/\gamma$.

D.1.3 Entry wedge

The planner maximizes expected utility:

$$\max_{M_e} \; - W(M_e) + \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C} + \underbrace{\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\mathcal{C}}_{\text{consumption from infrastructure}} - \underbrace{\Phi(M)}_{\text{cost of infrastructure}} .$$

The corresponding optimality condition is:

$$\begin{aligned} W'(M_e) =& \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\tilde{\mathcal{I}}_n'}{\tilde{\mathcal{I}}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} \\ &+ \frac{\sigma-1}{\sigma}\mathcal{C}' + \left(\frac{1}{\sigma} - \frac{1}{\gamma}\right)\mathcal{C}' - \underbrace{\Phi'(M)M}_{(\frac{1}{\sigma}-\frac{1}{\gamma})\mathcal{C}} \cdot \frac{1}{M}\frac{dM}{dM_e}. \end{aligned}$$

The wedge is therefore:

$$\tau_n = -\mathcal{E}_{\tilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma-1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\mathcal{I}_1}{\tilde{\mathcal{I}}_n}.$$

In particular, the wedge with agreement ($n = 1$) is:

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma-1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\gamma}{\sigma}.$$

Using the values of $\mathcal{E}_\mathcal{C}$ and χ, we obtain:

$$\tau_1 = 1 - \chi\gamma\left(1 + \nu - \frac{1}{\sigma} + \frac{1}{\gamma}\right) = 0, \tag{D.1}$$

given the formula above for χ. The competitive equilibrium with agreement is efficient since the first welfare theorem applies.

Now we apply the reasoning from Section C.1 to find the condition for convergence when θ is large. The condition for convergence of $\mathcal{E}_{\tilde{\mathcal{I}}_n}$ is the same as for $\mathcal{E}_{\mathcal{I}_n}$:

$$\gamma(\theta + \chi) > 1 + \chi\left(\frac{\gamma}{\sigma} - 1\right) \tag{D.2}$$

$$\iff \quad \gamma\theta > 1 + \frac{1}{\nu + \frac{1}{\gamma} + \frac{\sigma-1}{\sigma}}\left(\frac{1}{\sigma} - \frac{1}{\gamma} - 1\right) \tag{D.3}$$

As $n \to \infty$, we have that $\tilde{\mathcal{I}}_n \to \infty$ and $\mathcal{I}_1 \to 0$, and therefore $\mathcal{I}_1/\tilde{\mathcal{I}}_n \to 0$.

For the high disagreement wedge we have:

$$\tau_\infty = (\sigma - 1)\cdot\mathcal{E}_w = \frac{\sigma-1}{\gamma} + (\sigma-1)\chi\left(\frac{1}{\sigma} - \frac{1}{\gamma}\right) \tag{D.4}$$

$$= \frac{\sigma-1}{\gamma}\cdot\left(\frac{1+\nu}{1+\nu+1/\gamma-1/\sigma}\right). \tag{D.5}$$

Since $1/\gamma < 1/\sigma$, τ_∞ is decreasing in ν. Furthermore we have $\lim_{\nu\to\infty}\tau_\infty = (1-\sigma)/\gamma$ which is our baseline result. Finally, τ_∞ is increasing in σ.

D.2 Melitz (2003) Model: Participation Costs and Dixit-Sitglitz

We now turn to the case with Dixit-Stiglitz preferences and linear technology.

D.2.1 Participating firms

Given M_e and M, profits before the infrastructure costs are unchanged from the standard model with decreasing returns to scale:

$$\pi(a) = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\cdot\mathcal{C}\cdot w^{1-\sigma}\cdot a^{\sigma-1}.$$

The equilibrium consumption is also unchanged:

$$\frac{\mathcal{C}}{L} = \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\sigma-1}} M^{\frac{1}{\sigma-1}-\frac{1}{\gamma}} M_e^{\frac{1}{\gamma}}.$$

The marginal firm has productivity $\underline{a}$ and spends all of its profit in infrastructure. Therefore, we have the zero cutoff profit condition $\Phi'(M) = \pi(\underline{a})$, which implies:

$$M^{\nu+\frac{1}{\gamma}+\frac{\sigma-2}{\sigma-1}} = \frac{1}{\varphi_0}\frac{1}{\sigma}\left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{\frac{\sigma-2}{\sigma-1}} L\cdot M_e^{\frac{1}{\gamma}},$$

where we use the fact that $\underline{a} = (M_e/M)^{1/\gamma}$. In Section C.2, we specified an exogenous set of producing firms $M = M_e^\chi / M_0^{\chi-1}$. This arises endogenously through our cost of infrastructure with

$$\chi = \frac{1}{\gamma}\left(\nu + \frac{1}{\gamma} + \frac{\sigma-2}{\sigma-1}\right)^{-1},$$

$$M_0^{1-\chi} = \left(\frac{1}{\varphi_0}\frac{1}{\sigma}\left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{\frac{\sigma-2}{\sigma-1}} L\right)^{\left(\nu+\frac{1}{\gamma}+\frac{\sigma-2}{\sigma-1}\right)^{-1}},$$

where the exponent satisfies $\chi \leq 1$ if and only if $\nu + \frac{\sigma-2}{\sigma-1} \in (-\infty, -1/\gamma) \cup [0, \infty)$. Otherwise, all firms participate as M_e grows to infinity.

Finally, we derive the elasticity $\mathcal{E}_{\mathcal{C}}$:

$$\mathcal{E}_{\mathcal{C}} = \frac{1}{\gamma} + \chi\left(\frac{1}{\sigma-1} - \frac{1}{\gamma}\right) = \frac{1}{\gamma}\cdot\frac{1+\nu}{1+\nu+1/\gamma-1/(\sigma-1)} = \chi\cdot(1+\nu).$$

D.2.2 Equilibrium

The equilibrium condition in the competitive equilibrium is:

$$W'(M_e) = \frac{1}{\sigma}\cdot\frac{\mathcal{C}}{M_e}\cdot\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}.$$

Aggregate profits represent a fraction $(\sigma-1)/\gamma$ of aggregate revenue after labor costs, and aggregate infrastructure costs account for the other $(\gamma-(\sigma-1))/\gamma$. Therefore aggregate profits represent a share $(\sigma-1)/(\sigma\gamma)$ of consumption and aggregate infrastructure costs $(\gamma-(\sigma-1))/(\sigma\gamma)$.

D.2.3 Entry wedge

The planner maximizes expected utility:

$$\max_{M_e} \; -W(M_e) + \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1} + \frac{\sigma-1}{\sigma}\mathcal{C} + \left(\frac{\gamma-(\sigma-1)}{\sigma\gamma}\right)\mathcal{C} - \Phi(M).$$

The corresponding optimality condition is

$$\begin{aligned} W'(M_e) = {} & \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\tilde{\mathcal{I}}_n'}{\tilde{\mathcal{I}}_n} + \frac{1}{\sigma}\mathcal{C}'\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1} - \frac{1}{\sigma}\mathcal{C}\frac{\tilde{\mathcal{I}}_n}{\mathcal{I}_1}\frac{\mathcal{I}_1'}{\mathcal{I}_1} \\ & + \frac{\sigma-1}{\sigma}\mathcal{C}' + \left(\frac{\gamma-(\sigma-1)}{\sigma\gamma}\right)\mathcal{C}' - \Phi'(M)M\frac{1}{M}\frac{dM}{dM_e}. \end{aligned}$$

The wedge is therefore

$$\tau_n = -\mathcal{E}_{\tilde{\mathcal{I}}_n} + \mathcal{E}_{\mathcal{I}_1} + 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma-1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma-1}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\mathcal{I}_1}{\tilde{\mathcal{I}}_n}. \tag{D.6}$$

In particular, the wedge with agreement ($n = 1$):

$$\tau_1 = 1 - \mathcal{E}_{\mathcal{C}} - \left[(\sigma-1)\mathcal{E}_{\mathcal{C}} + \left(1 - \frac{\sigma-1}{\gamma}\right)(\mathcal{E}_{\mathcal{C}} - \chi)\right]\frac{\gamma}{\sigma-1} \tag{D.7}$$

$$= -\frac{1}{\sigma-1} \cdot \frac{1+\nu}{1+\nu+1/\gamma-1/(\sigma-1)}. \tag{D.8}$$

Now we apply the reasoning from Section C.1 to find the condition for convergence when θ is large. The condition for convergence of $\mathcal{E}_{\tilde{\mathcal{I}}_n}$ is the same as for $\mathcal{E}_{\mathcal{I}_n}$:

$$\gamma(\theta+\chi) > 1 + \chi\left(\frac{\gamma}{\sigma-1} - 1\right) \tag{D.9}$$

$$\iff \quad \gamma\theta > 1 + \frac{1}{\nu + \frac{1}{\gamma} + \frac{\sigma-2}{\sigma-1}}\left(\frac{1}{\sigma-1} - \frac{1}{\gamma} - 1\right) \tag{D.10}$$

As $n \to \infty$, we have $\tilde{\mathcal{I}}_n \to \infty$ and $\mathcal{I}_1 \to 0$, and therefore $\mathcal{I}_1/\tilde{\mathcal{I}}_n \to 0$.

For the high disagreement wedge we have:

$$\tau_\infty = (\sigma-1)\cdot\mathcal{E}_w - \mathcal{E}_{\mathcal{C}} = \frac{\sigma-2}{\gamma}\cdot\left(\frac{1+\nu}{1+\nu+1/\gamma-1/(\sigma-1)}\right). \tag{D.11}$$

When $\nu \to \infty$, there is a fixed supply of infrastructure and therefore a fixed number of firms, which implies:

$$\tau_1 = -\frac{1}{\sigma-1},$$
$$\tau_\infty = \frac{\sigma-2}{\gamma}.$$