# Supply chain formation and fragilities with imperfect information

Andrea Titton

February 27, 2023

**Abstract**

This paper investigates the role of supply chain unobservability in generating endogenously fragile production networks. In a simple production game, in which firms need to multisource to hedge against suppliers' risk under unobservability, firms underdiversify vis-à-vis the social optimum. The unobservability of suppliers' relations is the driver behind this. In production networks where upstream risk is highly correlated and supplier relationships are not observable, the marginal risk reduction of adding an additional supplier is low, because this additional supplier's risk is likely to be correlated to that of existing suppliers. This channel reduces firm incentives to diversify, which gives rise to inefficiently fragile production networks.

By solving the social planner problem, I show that, if the risk reduction experienced downstream resulting from upstream diversification were to be internalised by upstream firms, endogenous production networks would be resilient to most levels of risk. Despite its stylised form, the model identifies the trade-off firms face when diversifying risk and isolates the mechanism that aggregates these decisions into a production network. Furthermore, it maps the conditions of the trade-off, such as expected profits of the firm or the pairing costs, to the properties of the production network.

In August 2020, hurricane Laura hit one of the world's largest petrochemical districts, in the U.S. states of Louisiana and Texas. As polymer producers in the area were forced to halt production, up to 15% of the country's PE and PP producers were unable to source polymer inputs, which in turn caused shortages across the economy. This episode illustrates how correlation in suppliers' risk, in this case due to spatial proximity, can have yield sizeable downstream shutdowns and, hence, firms need to account for it in making sourcing decisions. Yet, the structure of the supply chain is often opaque: firms do not observe sourcing relations beyond their immediate suppliers (Williams et al., 2013). In face of this opacity, how do producers make sourcing decisions? And, should we expect these sourcing decisions to yield robust production networks?

In this paper, I study the role of supply chain opacity in determining firms' sourcing decisions and, in turn, the consequences on the resilience of the production network. A widespread approach to mitigate risk is to diversify it by multisourcing. This practice consists of procuring the same inputs from multiple suppliers, sometimes redundantly (Zhao and Freeman, 2019). Yet, when deciding how many suppliers from which to source, a firm faces decreasing marginal benefits in risk reduction, because each additional supplier's failure to deliver is increasingly likely to be correlated with that of the existing ones. In the presence of marginal costs of sourcing, for example, contractual costs or higher prices, the uncertainty behind the correlation of a firm's potential suppliers might induce it to diversify risk less than socially optimal. The wedge between endogenous firm decisions and social optimality arises because downstream firms would be willing to compensate their suppliers for increased diversification of inputs. This underdiversification can generate aggregate fragility in production networks. To understand the relationship between the opacity of the supply chain, firms' diversification decisions, and production network fragility, I study the properties of a stylised production game. In the equilibrium of the game, unobserved correlation among suppliers generates fragility via two channels. First, it directly introduces endogenous correlation in downstream firms' risk, which amplifies through the production network. This increases the probability of cascading failures, in which the entire production network is unable to produce. Second, it indirectly affects firms' decisions by reducing the expected marginal gain from adding a source of input goods. The latter channel leads to firms diversifying increasingly less, such that small shocks in the production of basal goods can generate cascading failures downstream.

The role that production networks play in determining economic outcomes has been long recognised. As far back as Leontief (1936), economists have studied how networks in production can act as aggregators of firm level activity. Following a foundational paper by Hulten (1978), which showed that the first order impact of a productivity shock to an industry is independent of the production network structure, macroeconomics has since de-emphasised this role (Baqaee and Farhi, 2019, p. 2). However, more recently, Baqaee and Farhi (2019) illustrated how the structure of the production network can aggregate micro shocks via second order effects.[(1)] Furthermore, the degree of competition in an

[(1)]These results build on a vast literature Gabaix, 2011; Acemoglu et al., 2012; Carvalho et al., 2016;

industry also interacts with the production network to aggregate shocks, which can lead to cascading failures (Baqaee, 2018). Once established that production networks play a central role in aggregating shocks, two natural questions arise. First, which networks can we expect to observe, given that firm endogenously and strategically choose suppliers? Second, are these endogenous network formations responsible for the growth or fragility that large economies display? These questions fuelled a number of recent papers studying endogenous production network formation. Focusing on growth, Acemoglu and Azar (2020) show that endogenous production networks can be a channel through which firms' increased productivity lowers costs throughout the supply chain and allows for sustained economic growth. In parallel, a vast literature dealt with studying the role of endogenous production networks and firm incentives in determining fragile or resilient economies. Erol and Vohra (2014) showed that in networks with strategic link formation, systemic endogenous fragility arises if the shocks experienced by firms are correlated. Later work, by Amelkin and Vohra (2020), shows that uncertainty in the time of production is crucial in determining whether production networks in equilibrium are sparse, hence fragile. Finally, Elliott et al. (2022) illustrate how complexity in the production process can also be a key driver of endogenous fragility in production networks. [(2)]

In this paper, I aim to study endogenous fragility in the presence of supply chain unobservability. Kopytov et al. (2021) studied the effect of uncertainty in endogenous production network formation on firms' productivity and business cycles. They find that higher uncertainty can lead to lower economic growth. In contrast, in this paper I study the role of uncertainty in generating endogenous fragility to cascading failures using a more stylised production network model, akin to that studied by Elliott et al. (2022). In line with the existing literature, in the model small idiosyncratic shocks can be massively amplified. The degree of amplification depends on the equilibrium behaviour of firms. This phenomenon holds true in vertical economies producing simple goods. The novel theoretical contribution of this paper is to extend the analysis of production network formation to an opaque environment in which firms aim to minimise risk while accounting for correlation between suppliers. To do so, I develop a tractable analytical framework that describes the propagation of idiosyncratic shocks through the supply chain when firms take sourcing decisions endogenously in an imperfect information environment. The model describes the evolution of risk through the supply chain as a dynamical system over the depth of the production network. Finally, I show that the model can be seen as an extension of the one developed by Elliott et al. (2022) for large, but finite, supply chains.

---

Baqaee and Farhi, 2019; Carvalho and Tahbaz-Salehi, 2019

[(2)] The literature on production networks is vast and it is unfortunately impossible to give a fair overview in this introduction. For a more comprehensive review of the literature I refer the reader to Carvalho and Tahbaz-Salehi (2019) and Amelkin and Vohra (2020)

## 1. Production Environment

Consider a vertical economy producing $K+1$ goods (as displayed in Figure 1), indexed by $k \in [K] := \{0, 1, \dots K\}$. Each firm produces a single good and each good is produced by $m$ firms. Good $k$ requires only good $k-1$ as input. I often refer to the set of firms producing a good $k$ as *layer* $k$.

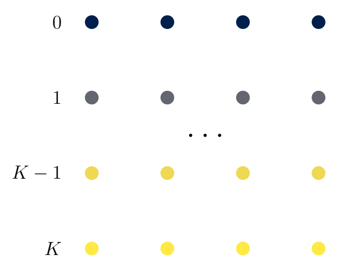



**Figure 1:** $K$-layers vertical economy

Each firm picks a set of suppliers in the previous layer. Establishing a relation with a supplier has a fixed cost $\kappa$. If no supplier is able to deliver the input good, then the firm is not "functional", and hence not able to deliver downstream. I assume that being functional yields an exogenous payoff $\pi$. This assumption can be relaxed by introducing a market structure to endogenise $\pi$, but this does not change the main model mechanics. Finally, I assume that firms know the structure of the economy but do not observe the realised supplier relationship in upstream layers. The only source of risk in the model is a non-necessarily idiosyncratic probability $\mu_0$ that basal firms in the zeroth layer are not able to carry out production.

A firm is identified by a tuple $(k, i)$, where $k \in [K]$ is its good or layer and $i \in [m]$ is the firm index. Each firm picks suppliers from which to source its input good among the producer of the previous layer. Let $\mathcal{S}_{k,i} \subseteq \{k-1\} \times [m]$ be this set of suppliers. For example, a possible supplier realisation of the network presented in Figure 1 can be seen in Figure 2.

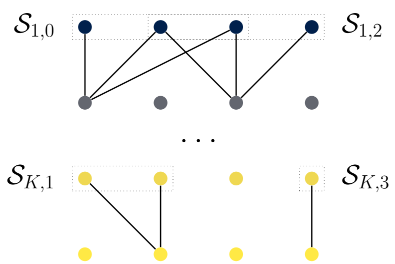



**Figure 2:** Supply chain realisation

The firm $(k, i)$ is then able to produce if at least one of its suppliers is able to deliver. Letting $\mathcal{F}_k$ be the functional firms in layer $k$ we can say that $(k, i) \in \mathcal{F}_k$ if and only if $\mathcal{F}_{k-1} \cap \mathcal{S}_{k,i}$ is not empty. Then, the probability that the firm is functioning can be written as

$$p_{k,i} := \mathbb{P}\big(i \in \mathcal{F}_k\big) = 1 - \mathbb{P}\big(\mathcal{F}_{k-1} \cap \mathcal{S}_{k,i} = \emptyset\big). \quad (1)$$

Before moving on with the model solution, it is useful to discuss the assumptions presented in this section. Clearly, the production game is highly stylised: first, firms do not adjust prices but only quantities, such that failure to produce only arises in the case that no input is sourced; second, they are able to obtain profits by simply producing, such that $\pi$ is exogenous and constant; third, they face fixed costs when establishing relations, such that $\kappa$ is constant. There are both theoretical and empirical aspect motivating this choice. Theoretically, a simpler model allows us to isolate the interplay between the variables of interest: correlation in the risk of suppliers, supply chain opacity, and the endogenous production network fragility. Introducing these aspects into a richer model such as the one proposed by Elliott et al. (2022) or Kopytov et al. (2021) can render the analysis intractable and prevent isolation of the desired causal mechanisms. In this sense, the model should not be seen as an alternative approach to the one adopted by these authors but rather a complementary one. Empirically, there is strong evidence that firms, first, when faced with supply chain shocks, adjust quantities rather than prices in the short run (di Giovanni and Levchenko, 2010; Macchiavello and Morjaria, 2015; Jiang et al., 2022; Lafrogne-Joussier et al., 2022) and, second, that production shutdowns can be extremely costly (Tan and Kramer, 1997; Hameed and Khan, 2014). These two aspects of supply chain disruptions are well captured by the model.

## 2. Firm Problem

We can now examine and solve the multisourcing problem faced by firms. I derive an analytical expression that describes the risk propagation in the supply chain when firms' sourcing decisions are endogenous. The problem of the firm $(k,i)$ is to maximise the expected payoff, which, using equation (1), we can write as

$$\Big(1 - \mathbb{P}\big(\mathcal{F}_{k-1} \cap \mathcal{S}_{k,i} = \emptyset\big)\Big)\ \pi - \kappa\, |\mathcal{S}_{k,i}|\,, \tag{2}$$

by picking a set of suppliers from those producing the previous good $\mathcal{S}_{k,i} \subseteq \{k-1\} \times [m]$. To model supply chain uncertainty, I assume firms cannot observe the supplier decisions of the firms producing the necessary input good before making their supplier decision. Nevertheless, firms know the position they occupy within the supply chain $k$ and the number of firms in each layer $m$. Given this information, firms can derive the distribution of risk in each layer of the production network and make sourcing decisions based on it. This solution criterion captures the fact that firms, despite not knowing exactly the correlation of risk among producers of their input goods, have some information to estimate their suppliers' risk and make sourcing decisions. In particular, a firm producing good $k$ can, first, infer the distribution of the number $F_{k-1} \coloneqq |\mathcal{F}_{k-1}|$ of functioning firms in the previous layer; second, select the optimal number $s_{k,i} = |\mathcal{S}_{k,i}|$ of firms from which to source its input good; third, pick $s_{k,i}$ suppliers with equal probability. The randomness in this last step of the firms' choices arises since all potential suppliers producing good $k-1$ are ex-ante equal. Hence, each firm in layer $k$ might pick a different set of suppliers, that is, $\mathcal{S}_{k,i}$ is not necessarily equal to $\mathcal{S}_{k,j}$ for some $i$ and $j$, but they do pick the same number, $s_k$, of suppliers, that is

$$s_{k,i} = s_k \text{ for all } i. \tag{3}$$

If two firms $(k,i)$ and $(k,j)$ pick two sets of suppliers $\mathcal{S}_{k,i}$ and $\mathcal{S}_{k,j}$ with equal $|\mathcal{S}_{k,i}| = |\mathcal{S}_{k,j}| = s_k$, the probability that each of the firms is functional, $p_{k,i}$ and $p_{k,j}$, is a realization drawn from the same distribution. We can characterise this distribution by looking at the proportion of possible realisations in which, given a random set of suppliers $\mathcal{S}_{k,i}$

and a random set of functioning firms $\mathcal{F}_{k-1}$, there is overlap between them. This quantity only depends on the sizes $s_k$ and $F_{k-1}$ of the two sets. Let $P(s_k, F_{k-1}; m) \coloneqq p_{k,i}$ be the probability of the two sets overlapping, then

$$1 - \mathbb{P}\big(\mathcal{S}_{k,i} \cap \mathcal{F}_{k-1} = \emptyset\big) = P\left(s_k, F_{k-1}; m\right) = 1 - \overbrace{\binom{m - s_k}{F_{k-1}}}^{\substack{\text{non-overlapping} \\ \text{configurations}}} \Big/ \underbrace{\binom{m}{F_{k-1}}}_{\substack{\text{all possible} \\ \text{configurations}}} . \tag{4}$$

This function is identical for all firms in layer $k$, and depends on the layer size $m$ both indirectly, via the support of $F_{k-1} \in [m]$, and directly. Furthermore, the number of functioning firms among the producers of an input good $F_{k-1}$ is a random variable, hence the probability of functioning $P(s_k, F_{k-1}; m)$, conditional on a sourcing decision $s_k$, is a random variable too. We can now make a claim regarding the distribution of $P(s_k, F_{k-1}; m)$ and its relationship with that of $F_{k-1}$. First, I introduce the two key distributions we work with.

**Definition 1.** *A random variable $Y$ is said to follow a BetaPower distribution, with mean $1 - \mu$, overdispersion $\rho$, and power $s$ if it can be written as $Y = X^s$ where $X$ follows a Beta distribution with mean $\mu$ and overdispersion $\rho$.*

**Definition 2.** *If the probability of functioning of a firm producing good $k$ follows*

$$P(s_k, F_{k-1}) \sim \mathit{BetaPower}(1 - \mu_{k-1}, \rho_{k-1}, s_k), \tag{5}$$

*then the number of functioning firms in the same layer follows a compounded distribution*

$$F_k \mid P(s_k, F_{k-1}) = p \sim \mathit{Bin}(m, p). \tag{6}$$

*which we call BetaBinPower and denote as*

$$F_k \sim \mathit{BetaBinPower}(m, 1 - \mu_{k-1}, \rho_{k-1}, s_k). \tag{7}$$

Second we can link the distribution of functioning firms between the two layers $F_{k-1}$ and $F_k$.

**Proposition 1.** *If the number of functioning firms in layer $k-1$, $F_{k-1}$, follows a BetaBinPower distribution, then the probability of functioning of a firm producing good $k$, $P(s_k, F_{k-1}; m)$, converges to a BetaPower distribution as $m \to \infty$.*

The proof of can be found in Appendix B.2. This, combined with the Definitions 1 and 2 implies the distribution of functioning firms in the whole production network.

**Corollary 1.1.** *In the large $m$ limit, if for some layer or good $k$ the number of functioning firms $F_k$ follows a BetaBinPower distribution, then, in all downstream layers $l > k$, the number of functioning firms $F_l$ follows a BetaBinPower distribution.*

Corollary 1.1 asserts that the distribution of functioning firms will remain in the same distribution family as risk amplifies through the production network. This is a powerful result. It allows us to describe risk propagation by mapping the parameters $\mu_k$ and $\rho_k$ through the layers. Furthermore, a firm can compute $\mu_k$ and $\rho_k$ and get all the information required to determine its optimal sourcing decision $s_k$. The only missing piece is the distribution of functioning firms in the basal layer. As mentioned in the previous section, we assume that basal firms fail with a not necessarily independent probability $\mu_0$. We can model this by assuming that $F_0$ follows itself a BetaBin distribution with parameters $\mu_0$ and $\rho_0$.

**Assumption 1.** *The number of functioning firms producing the basal good follows a*

$$F_0 \sim BetaBin(m, 1 - \mu_0, \rho_0), \tag{8}$$

*for some initial condition $\mu_0, \rho_0 \in (0, 1)$. This generalises the case of idiosyncratic risk in the basal layer which can be retrieved by taking*

$$\lim_{\rho_0 \to 0} F_0 \sim Bin(m, 1 - \mu_0). \tag{9}$$

It is useful at this point to give an interpretation of $\mu_k$ and $\rho_k$, in the context of our model. The parameter $\mu_k$ is the fraction of firms that are expected not to deliver. Hence, I hereafter refer to $\mu_k$ as *risk*. The parameter $\rho_k$ tracks the degree of correlation in the risk of the firms operating in layer $k$. If $\rho_k \to 0$, then firms' risk is independent, $P_k$ concentrates at $1 - \mu_k$ (solid line, Figure 3), and $F_k$ degenerates into a binomial distribution. On the contrary, if $\rho_k \to 1$ then firms' risk is perfectly correlated, $P_k$ concentrates at 0 and 1 (dashed line, Figure 3), and either no firm is able to operate $F_k = 0$, with probability $\mu_k$, or all are able to operate $F_k = m$, with probability $1 - \mu_k$. Since $\rho_k$ controls the probability of tail events, I hereafter refer to it as *overdispersion* or *suppliers' risk correlation.*

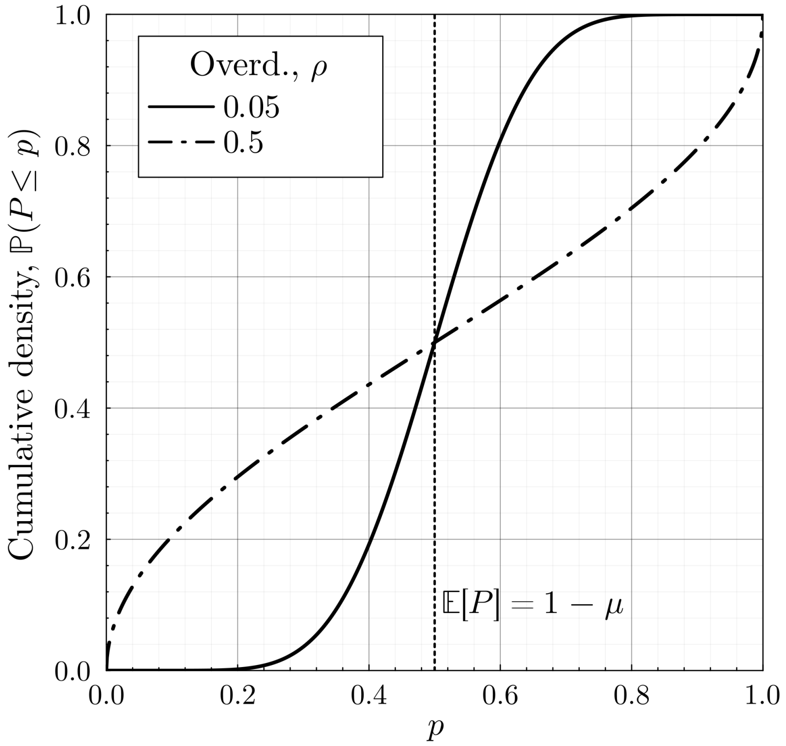



**Figure 3:** Cumulative distribution function for $P \sim \text{Beta}(\mu = 1/2, \rho)$

## 3. Dynamics of Risk and Overdispersion

As argued above, the system is fully described by the evolution of the distribution of functioning firms $\{F_k\}_{k=0}^{\infty}$ through the layers, and these two are entirely determined by the evolution of the parameters $\mu_k$ and $\rho_k$. Thus, we seek a function $G$ that maps the parameters of the distribution of the number of functioning firms from one layer to the next given a number of sources,

$$(\mu_{k+1}, \rho_{k+1}) = G(\mu_k, \rho_k;\ s_{k+1}). \tag{10}$$

Propositions 2 and 3 below specify this law of motion. For the derivation see Appendix C.

**Proposition 2.** *Risk $\mu_k$ amplifies through the production network following the law of motion*

$$\mu_{k+1} = G_1(\mu_k, \rho_k, s_{k+1}) = \begin{cases} \left(\mu_k \frac{1-\rho_k}{\rho_k}\right)^{\underline{s_{k+1}}} \Big/ \left(\frac{1-\rho_k}{\rho_k}\right)^{\underline{s_{k+1}}} & \text{if } \rho > 0 \\ \mu_k^{s_{k+1}} & \text{if } \rho_k = 0 \end{cases}, \tag{11}$$

*where* $x^{\underline{s}} := x(x-1)\dots(x-s+1)$ *is the falling factorial.*

**Corollary 2.1.** *The map* $s \mapsto G_1(\mu, \rho, s) - \mu$ *is monotonically decreasing in* $s$. *Namely, as the number of sources increases, risk decreases.*

Equation (11) captures the decreasing marginal benefit attained by firms when adding a further source. The relative marginal reduction in risk of attained by the $n$-th supplier is

$$\frac{G_1(\mu, \rho, n-1) - G_1(\mu, \rho, n)}{G_1(\mu, \rho, n-1)} = (1-\mu) \Bigg/ \left(1 + n\frac{\rho}{1-\rho}\right). \tag{12}$$

As the number of suppliers $n$ increases, the relative reduction in risk experienced by a downstream firm decreases, as its suppliers are increasingly likely to share inputs sources. For higher levels of suppliers' correlation $\rho$, this effect is exacerbated (see Figure 4). But, if there suppliers are not correlated, $\rho$, adding a supplier adds

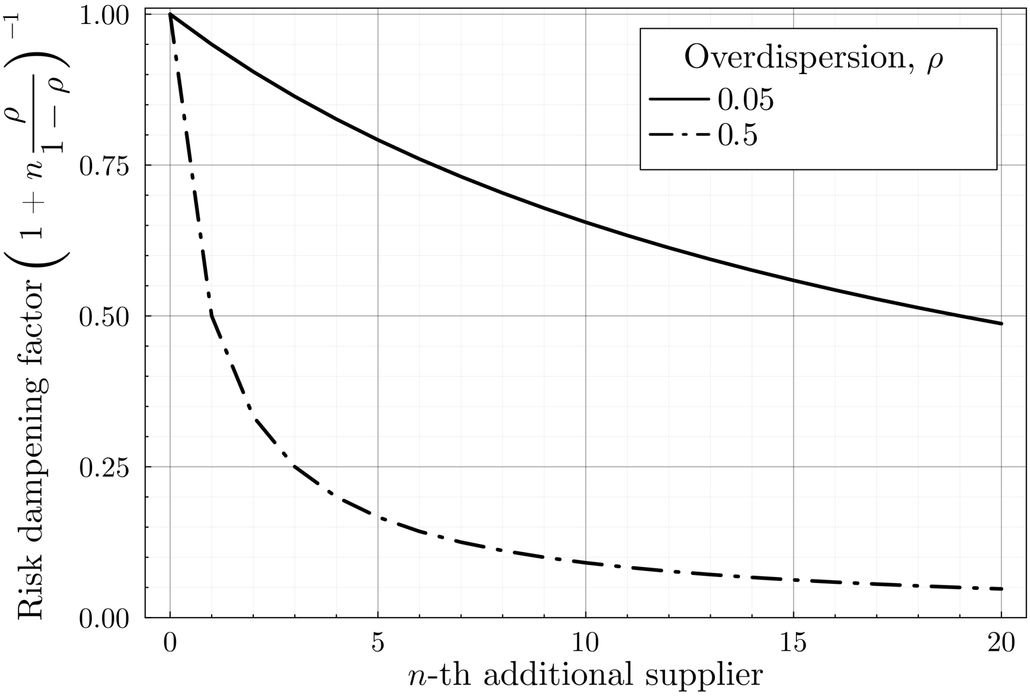



**Figure 4:** The marginal decrease in risk experienced by a firm adding an $n$-th supplier, in a low (solid) and a high (dash) supplier correlation environments.

**Proposition 3.** *Overdispersion $\rho$ amplifies following the law of motion*

$$\rho_{k+1} = G_2(\mu_k, \rho_k, s_{k+1}) = \begin{cases} \frac{\left(\mu_k \frac{1-\rho_k}{\rho_k}+2s_{k+1}\right)^{\overline{2s_{k+1}}} \Big/ \left(\frac{1-\rho_k}{\rho_k}+2s_{k+1}\right)^{\overline{2s_{k+1}}} - \mu_{k+1}^2}{\mu_{k+1}\left(1-\mu_{k+1}\right)} & \text{if } \rho_k > 0 \\ 0 & \text{if } \rho_k = 0 \end{cases}, \tag{13}$$

*where $\mu_{k+1} \coloneqq G_1(\mu_k, \rho_k, s_{k+1})$.*

Equation (13) controls the evolution of suppliers' correlation and overdispersion in the production network. Figure 5 illustrates the evolution of suppliers' correlation $G_2$, when firms multisource (solid line) and single sourcing (dashed line). If the suppliers in layer $k$ are weakly correlated, multisourcing dampens the propagation of overdispersion. But, if upstream correlation is high, multisourcing exacerbates overdispersion. This mechanism anticipates an asymmetry between the firms' incentives and the social optimum: a firm which does not internalise the correlation between its own risk and that of its competitors will not invest sufficiently in diversification and will hence allow overdispersion. This type of externality is more extensively explored below, in Section 6.

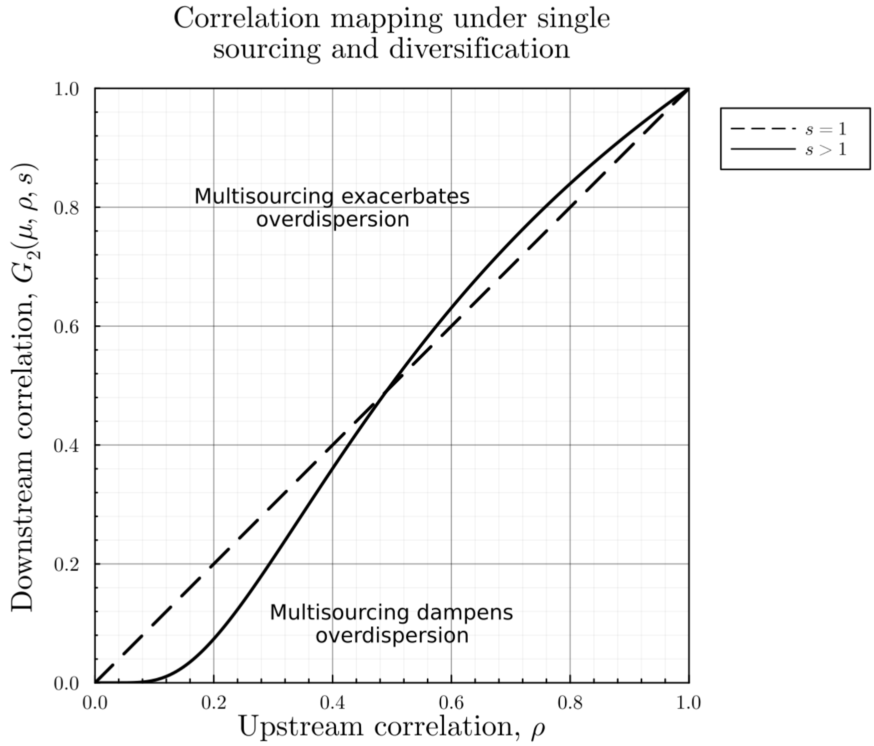



**Figure 5:** Suppliers' correlation $\rho_k$ evolution for multisourcing $s_k = 5$ and single sourcing $s_k = 1$

The dynamical system $G = (G_1, G_2)$ fully describes the propagation of risk in the production network, given a sequence of sourcing decisions of firms $\{s_k\}_{k=0}^K$. Before looking at competitive sourcing and social optimum sourcing $\{s_k\}_{k=0}^K$, it is useful to establish some properties of the dynamical system and give their economic interpretation.

**Corollary 3.1.** *If all firms pick a single supplier $s = 1$, every non degenerate level of risk and overdispersion $\mu, \rho \in (0, 1)$ is a fixed point of the risk dynamics, that is,*

$$(\mu, \rho) = G(\mu, \rho, 1). \tag{14}$$

*Proof.* This follows immediately from the definition of $G$. It states that an industry in which firms single source shares the same fate as the industry supplying its input goods. □

**Corollary 3.2.** *If firms do not single source, that is $s > 1$, the only fixed points are the degenerate points $(\mu, \rho)$ in which risk is fully diversified, $(0, 0)$, or in which all firms fail to operate, $(1, 0)$.*

*Proof.* Proof of corollary 3.2 follows from the definition of $G_1$ (equation 11). Notice that for non degenerate values of $\mu$ and $\rho$ and non-single sourcing $s \neq 1$, risk is never constant, $G_1(\mu, \rho, s) \neq \mu$ if $\mu \in (0, 1)$. □

These results (2.1, 3.1, and 3.2) together highlight how the production network must converge to either a finite stable distribution of functioning firms or to a degenerate distribution in which all or no firms fail. In the next section we will see that the economic environment and endogenous choices of the firms play a crucial role in determining which of these scenarios arises.

## 4. Firm Optimal Diversification and Competitive Equilibrium

Now that we have characterised risk and overdispersion dynamics $G$, we can derive the optimal firms' sourcing strategies and describe how these interact with risk propagation. In this section, I will first analyse a limit case in which suppliers' risk is not correlated ($\rho = 0$), in which analytical derivations and interpretation are more straight forward, before turning towards the general framework ($\rho > 0$).

### 4.1. A Special Case: No Correlation Risk

The dynamical system without supplier correlation can be retrieved by evaluating the function $G$ at $\rho = 0$,

$$G(\mu, 0, s) = (\mu^s, 0). \tag{15}$$

Studying the dynamical system without supplier correlation, given by the one-dimensional map $g_s(\mu) \coloneqq G_1(\mu, 0, s)$, allows us to derive properties that can be later generalised to the non-degenerate dynamics $G$ with $\rho > 0$. Under optimal firm sourcing, risk's evolution equation is

$$\mu_{k+1} = g(\mu_k) = -\frac{\kappa/\pi}{\log(\mu_k)}. \tag{16}$$

**Proposition 4.** *If relative sourcing costs are too high, $\kappa/\pi > -1/e$, firms have no incentives to diversify risk and the supply chain fails to produce $\mu_k \to 1$. I will refer to this as the no-production regime.*

**Proposition 5.** *In a production regime, if the basal risk is smaller than the relative pairing costs,*

$$\mu_0 < \mu_c \coloneqq \kappa/\pi, \tag{17}$$

*the production network convergence to a stable distribution of risk, $\mu_k < 1$ for all $k$. Otherwise, firms fail to diversify the basal risk and eventually, for some $k$, firms will fail with probability $\mu_k = 1$.*

To make sense of Propositions 4 and 5, we can plot the function $g$ (Figure 6). If profits $\pi$ are larger than pairing costs $\kappa$, then there exists a critical risk level $\mu_c$ below which firms diversify risk and downstream firms are able to produce, $\mu_k < 1$ for all $k$. In this low relative cost parameter region the system has two steady states, a low risk stable steady state (indicated by a black dot in Figure 6) and a high risk unstable steady state (indicated by a white dot). The stability condition of the fixed points (equation 17) immediately highlights the qualitative difference between them. At the stable fixed point the expected loss in profits, $\mu\pi$, is smaller than the marginal diversification cost. Hence, it is optimal for the firms to respond to an increase in risk, $\mu$, by diversifying. On the other hand, at the unstable fixed point, the expected loss in profits is not worth recuperating. Hence the optimal response to an increase in risk is to cease production. Furthermore, the figure highlights how, for sufficiently high levels of relative costs, diversifying is never optimal. In this case, arbitrarily small levels of risk in the basal layer will amplify and the downstream risk will be such that no firm operates in equilibrium.

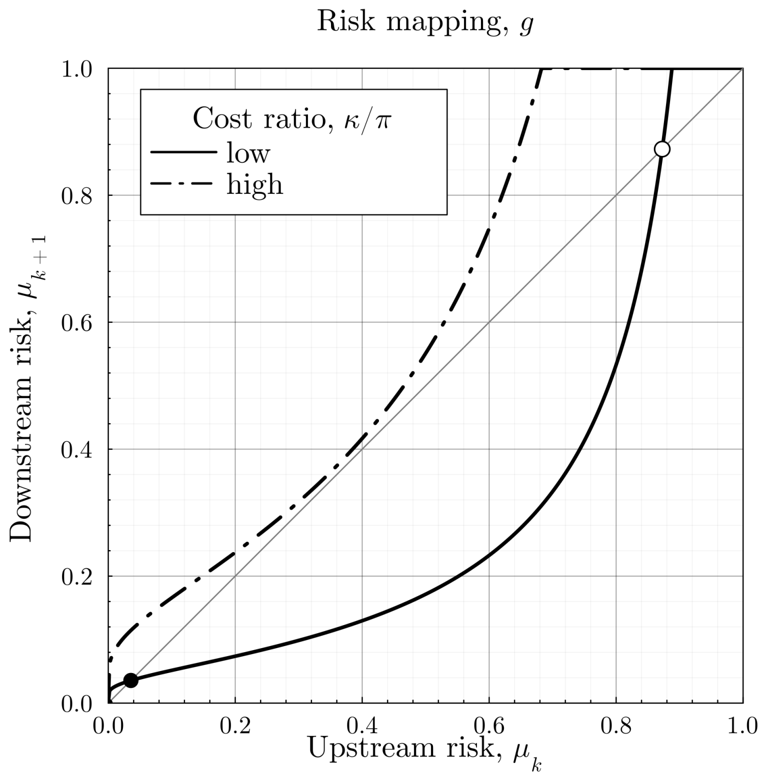



**Figure 6:** The function $g$ for high (dashed) and low (solid) relative costs, $\kappa/\pi$.

## 4.2. Introducing correlation risk

### 4.2.1. Optimal Firm Sourcing

We can now extend the analysis to the case in which suppliers' risk is correlated. Consider the problem of a firm in layer $k$, that is, given a level of risk and overdispersion in the previous layer $(\mu_{k-1}, \rho_{k-1})$, to choose a number of suppliers $s_k \in [m]$ to maximise profits. The profit $\Pi_k : \mathbb{R}_{\geq 0} \to \mathbb{R}$ of a firm in layer $k+1$ is

$$\Pi_{k+1}(s) = \Big(1 - G_1(\mu_k, \rho_k, s)\Big)\pi - \kappa\ s. \tag{18}$$

Then the optimal sourcing decision is

$$s_{k+1}(\mu_k, \rho_k) = \arg \max_{s \in \{0,1,\ldots m\}} \Pi_{k+1}(s). \tag{19}$$

Solving the integer problem (19) analytically is not trivial. Yet, we can relate the properties of the optimal sourcing strategy $s_{k+1}$ to the difference between the relative marginal cost of pairing and the marginal risk reduction

$$\Delta\Pi_{k+1}(s) := \frac{1}{\pi}\frac{\partial \Pi_{k+1}(s)}{\partial s} = \underbrace{\kappa/\pi}_{\text{relative pairing costs}} - \underbrace{-\frac{\partial G_1}{\partial s}(\mu_k, \rho_k, s)}_{\text{marginal risk reduction}}. \tag{20}$$

This quantity captures the incentive a firm has to increase diversification at a given diversification level $s$.

**Proposition 6.** *The profit function $\Pi_{k+1}$ is strictly concave in the number of sources $s$.*

Given Proposition 6, we can establish two properties of the optimal sourcing strategy $s_{k+1}$.

**Corollary 6.1.** *If $\Delta\Pi_{k+1}(s) = 0$ at $s$, then $|s_{k+1} - s| < 1$. That is, firms multisource as long as relative pairing costs are smaller than the marginal risk reduction.ß*

**Corollary 6.2.** *If $\Delta\Pi_{k+1}(s) = 0$ at $s < 1$, then $s_{k+1} = 1$. That is, if relative pairing costs are sufficiently high, firms choose to single source.*

Proposition 6 and Corollaries 6.1, 6.2 are proven in Appendix (C.3). These give us a number of insights. First, firms will add number of input sources as long as the relative cost of doing so $\kappa/\pi$ is larger than the marginal reduction in risk they can expect $\partial G_1/\partial s$. Figure 7 illustrates the optimal sourcing strategy of firms $s_k$, when facing suppliers' risk $\mu_k$ and overdispersion $\rho_k$. This optimal sourcing function extends previous results from the literature by accounting for supplier correlation $\rho_k$. Namely, consistently with previous literature and the no-overdispersion case analysed above there exist a critical level of risk $\mu_c$ above which firms are better off not producing and stop sourcing goods, that is, if $\mu_k > \mu_c$ for some $k$ then $s_{k+1}(\mu_k, \rho_k) = 0$. Yet, given a level of risk in the previous layer $\mu_k$, an increase in correlation among suppliers in the previous layer $\rho_k$ disincentivises firms' diversification (i.e. moving upwards in the contour plot). In Section 5, I show how this new channel can lead to underdiversification and represents a new form of externality that upstream producers impose on downstream firms.

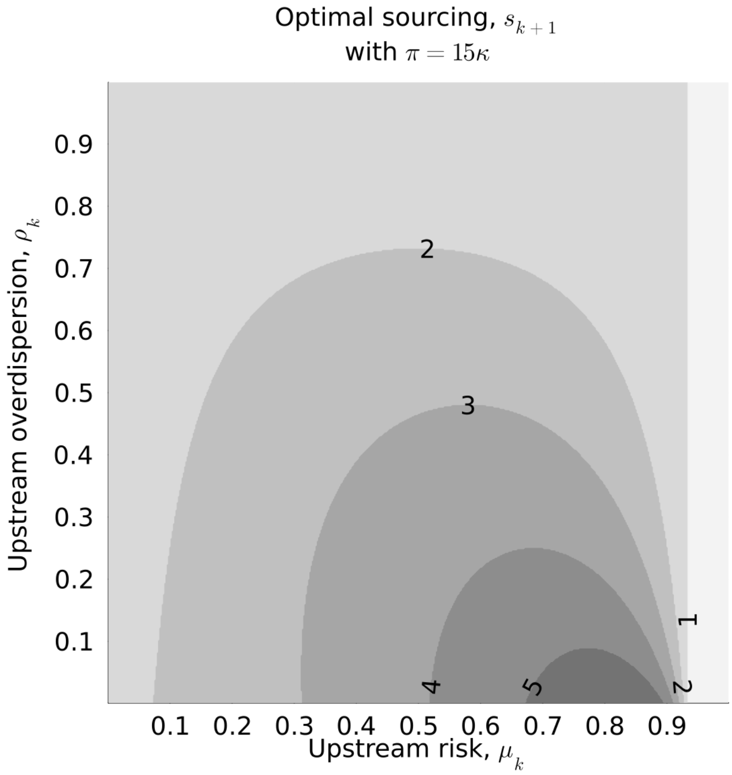



**Figure 7:** Contour plot of $s_{k+1}(\mu_k, \rho_k)$ and $\pi = 20\kappa$.

#### 4.2.2. Risk and Overdispersion Dynamics with Firms' Optimal Diversification

Now that I have characterised the optimal sourcing strategy of firms, in this section I look at the dynamics of risk this implies,

$$(\mu_{k+1}, \rho_{k+1}) = G(\mu_k, \rho_k, \underbrace{s_k(\mu_k, \rho_k)}_{\text{optimal sourcing}}). \quad (21)$$

This system is analysed by, first, deriving an implicit equation for the steady state levels of risk and overdispersion. Second, I compute numerically the downstream levels of risks implied by each basal condition, $\mu_0$ and $\rho_0$. Finally, implications are derived about sensitivity to basal conditions of cascading failures.

**Definition 3.** *A stable downstream distribution is defined by a constant level of risk $\bar{\mu}$ and overdispersion $\bar{\rho}$, such that, after a certain layer $\overline{k}$, the distribution of failing firms is stable, namely $(\mu_k, \rho_k) = (\bar{\mu}, \bar{\rho})$ for all $k \geq \overline{k}$.*

Corollaries 3.1 and 3.2 require that a distribution of failure among suppliers is stable if and only if downstream firms single source when facing it, namely $s_k(\bar{\mu}, \bar{\rho}) = 1$. This condition allows us to derive boundaries on the set of downstream stable distributions $(\bar{\mu}, \bar{\rho})$ (see Appendix C.5 for derivation).

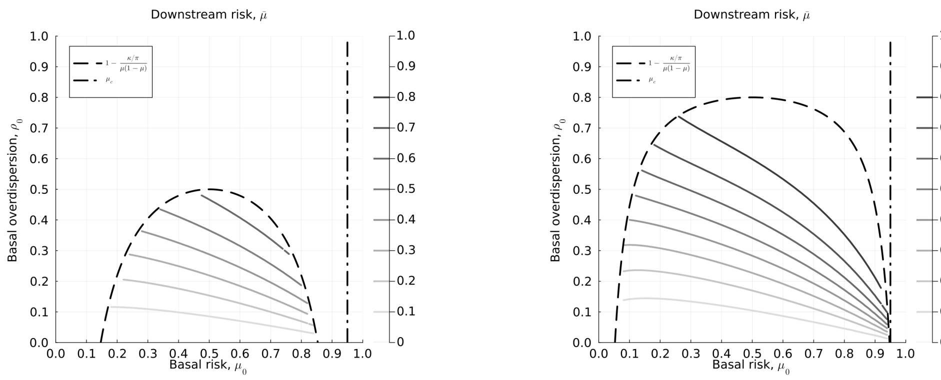
**Proposition 7.** *A downstream stable distribution $(\bar{\rho}, \bar{\mu})$ satisfies*

$$\bar{\rho} > 1 - \frac{\kappa/\pi}{\bar{\mu}(1-\bar{\mu})} \tag{22}$$

*and*

$$\bar{\mu} < 1 - \frac{\kappa}{\pi} = \mu_c. \tag{23}$$

The latter condition (23) is not surprising. For a non-degenerate equilibrium, the risk experienced by firms needs to be below the critical level $\mu_c$. The former (22) determines which levels of suppliers' correlation $\bar{\mu}$ are compatible with a given level of suppliers' risk $\bar{\rho}$ in equilibrium. As equilibrium risk $\bar{\mu}$ increase, optimal diversification imposes higher levels of correlation on the production network. After $\bar{\mu} = 1/2$, this effect reverses, and correlation starts decreasing. This effect is more pronounced, the higher the relative sourcing costs are. These results are graphically illustrated in Figure 8. I plot, for each basal level of risk and overdispersion $(\mu_0, \rho_0)$, the downstream risk $\bar{\mu}$, in the case of high (8a) and low (8b) relative costs. These dynamics can be traced back to the firms' choices. Faced with higher level of correlation, firms employ higher levels of multisourcing. This exacerbates downstream correlation, which reduces downstream sourcing incentives. Hence, the downstream levels of risk are higher. This observation has powerful implications: correlation in suppliers' risk increases equilibrium supply chain risk by distorting the incentives of firms to diversify.




**(a)** High relative costs, $\pi = 8\kappa$

**(b)** Low relative costs, $\pi = 20\kappa$

**Figure 8:** Level curves of the downstream level of risk $\mu_k$ for a given basal distribution $(\mu_0, \rho_0)$.

# 5. Social Planner Problem

To establish a benchmark to which one can compare the competitive equilibrium analysed in the previous section, I now solve the model from the perspective of a social planner. The social planner faces the trade-off of, on the one hand, minimising the number of firms expected to fail, and, on the other, minimising the number of sourcing relations, which have fixed cost for the firm. We can write the expected welfare as a function of the sequence of suppliers $\mathcal{S}_{k,i}$ by averaging the individual firms' expected payoffs, namely,

$$W(\mathcal{S}_{1,1}, \mathcal{S}_{1,1}, \ldots \mathcal{S}_{k,i}, \ldots \mathcal{S}_{K,m}) \coloneqq \frac{1}{mK} \sum_{(k,i)=(1,0)}^{(K,m)} \Big(1 - \mathbb{P}\big(\mathcal{F}_{k-1} \cap \mathcal{S}_{k,i} = \emptyset\big)\Big)\ \pi - \kappa \left|\mathcal{S}_{k,i}\right|. \tag{24}$$

Maximising the function $W$ with respect to the sequence $\{\mathcal{S}_{k,i}\}_{(k,i)} \subseteq [m]^{mK}$ requires the social planner to dictate to each firm which suppliers to source from. This requires a lot of regulatory power and information hidden from the firm. A less demanding social planner problem is the one in which the planner is only able to coerce the firm into picking a given number of suppliers and the firm is free to pick which suppliers to source from. We can write this constrained social planner problem in the "language" of our model as maximising

$$W_c(s_1, s_2, \ldots, s_K) \coloneqq \frac{1}{K} \sum_{k=1}^{K} \Big(1 - G_1(\mu_{k-1}, \rho_{k-1}, s_k)\Big)\pi - \kappa\ s_k, \tag{25}$$

with respect to the sequence of number of sources in each layer, $(s_1, s_2, \ldots, s_K) \in [m]^K$. By solving the unconstrained problem, I show that for a risk-neutral social planner (as described by 24 and 25) the two formulations yield the same solution. This result follows only if the social planner is exclusively concerned with the expected number of firms functioning and not with the probability of extreme events, in which a large number of firms is unable to produce. In what follows, I solve the latter case and then refine the solution to introduce a selection criterion based on the risk of tail events.

The only source of risk in the model is represented by the shutdowns experienced by firms in the basal layer with probability $\mu_0$. This implies that the risk faced by a firm depends only on how many firms in the first layer it is connected to by any path. Namely, if $n$ basal firms are involved in a firm's production, its risk is $1 - \mu_0^n$. This further imposes a lower bound of $\mu_0^m$ on the risk experienced by a firm. Hence, to find the welfare maximising sequence of suppliers, we can first look for the most edge parsimonious way to achieve a given level of risk $\mu_0^n$, such that

$$p_{k,i} = \mathbb{P}\big(\mathcal{F}_{k-1} \cap \mathcal{S}_{k,i} = \emptyset\big) \equiv \mu_0^n \text{ for all } (k,i), \tag{26}$$

and then find the optimum $n$.

**Definition 4.** *Let min-max($n$) be the class of networks in which all firms in layer 1 have $n$ suppliers and thereafter, each firm, is connect to only one supplier. That is any sequence $\mathcal{S}_{k,i} \subseteq [m]$ such that*

$$s_k = |\mathcal{S}_{k,i}| = \begin{cases} n & \text{if } k = 1, \\ 1 & \text{otherwise.} \end{cases} \tag{27}$$

**Proposition 8.** *The min-max($n$) networks are the networks with fewest edges that achieve $\mu_0^n$ risk.*

*Proof.* The proof can be best illustrated graphically. Consider Figure (9). If the planner wants to achieve risk $\mu_0^n$ in layer $k$, any further branching (right) from the *min-max*($n$) network (left), requires at least an additional link to close the branching, hence it has strictly more edges than the *min-max*($n$) network. Furthermore, note that if a firm in layer $k$ is connected, the marginal benefit of connecting it $\pi(1 - \mu_0^n)$ must have been bigger than the marginal cost $\kappa$. This implies that it must be profitable to connect a firm in layer $k+1$, since both marginal benefit and cost are equal. Namely, $s_k = 1$ implies $s_{k+1} = 1$, for $k \geq 2$. This allows us to exclude "truncated" *min-max*($n$) networks.

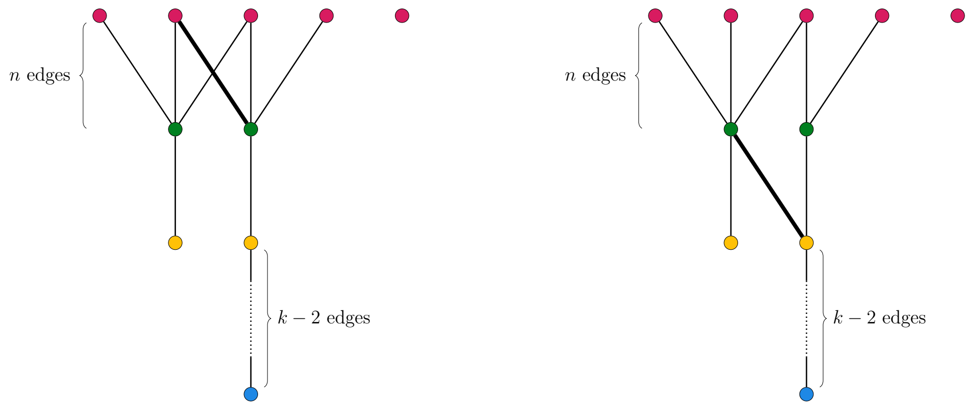



**Figure 9:** Two production networks that achieve risk $\mu^n$. On the left a *min-max*$(n)$ network and on the right a deviation where a supplier relationship has been moved from the first to second layer (thick black line).

□

Now that we know that *min-max*$(n)$ are the most edge parsimonious networks that achieve risk $\mu_0^n$, we can look for the optimal $n$. Conditional on using a *min-max*$(n)$ network, we write the welfare function (24) as a function of $n$,

$$W_{mm}(n) := W(\textit{min-max}(n)) = \frac{1}{K}\overbrace{(\pi\ (1-\mu^n) - \kappa n)}^{\text{first layer}} + \frac{K-1}{K}\overbrace{(\pi\ (1-\mu^n) - \kappa)}^{\text{downstream layers}} \\ = \pi(1-\mu^n) - \frac{n+K-1}{K}\kappa \tag{28}$$

**Proposition 9.** *Let*

$$n^* := \left\lfloor \frac{\log(\kappa/\pi) - \log(K) - \log(1-\mu_0)}{\log(\mu_0)} \right\rfloor \in [m]. \tag{29}$$

*If $W_{mm}(n^*) < 0$, then the social optimum is achieved by the empty production network, $\mathcal{S}_{k,i} = \emptyset$ for all firms $(k,i)$.*

*If $W_{mm}(n^*) > 0$, then the social optimum is achieved by a min-max$(n^*)$ network.*

*Proof.* The optimal number of basal firms involved in production

$$n^* = \arg\max_{n\in[m]} W_{mm}(n) \tag{30}$$

must be such that adding a new supplier yields a lower payoff $W_{mm}(n^*+1) - W_{mm}(n^*) < 0$. Furthermore, welfare must be positive, since shutting down production and not establishing any connection yields a payoff of zero, $W(\emptyset) = 0$. This immediately determines $n^*$. □

The fact that the *min-max* network is socially optimal is quite intuitive: the optimal way to mitigate risk is to diversify as close as possible to the source of risk in the production network (i.e. layer $k = 1$) and keep the risk constant thereafter. This intuition relates closely to the idea that in network games with negative spillovers (risk) a planner should target low eigenvector centrality nodes (basal firms) (Galeotti et al., 2017). The number of sources the social planner mandates in the first layer $n^*$ is increasing in the basal risk $\mu$ and decreasing in the marginal cost $\kappa/\pi$ (as seen in Figure 10).

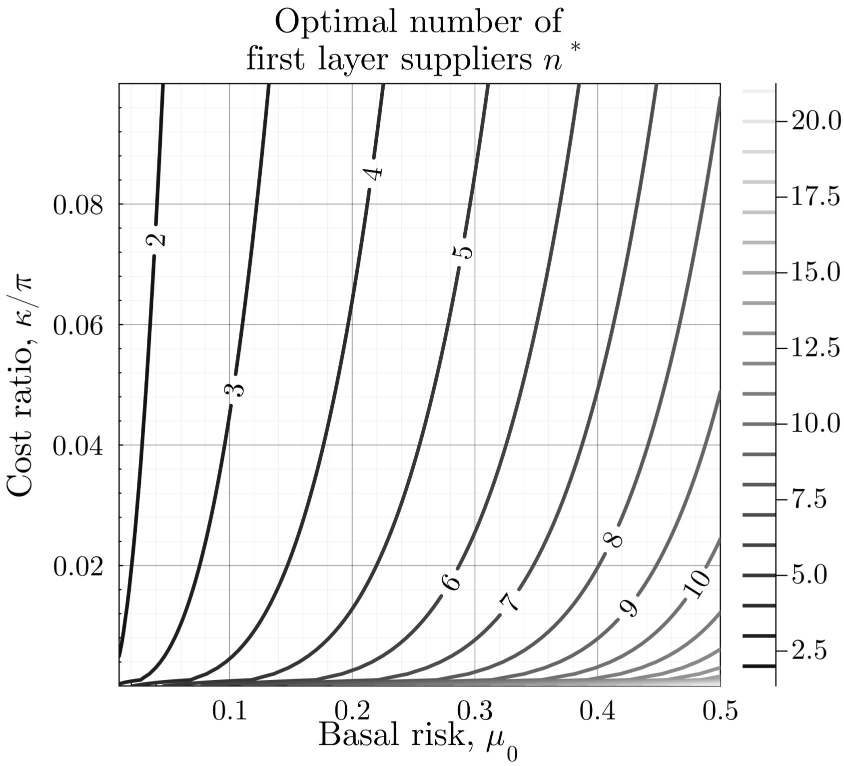



**Figure 10:** The socially optimal number of sources in the first layer, given different levels of basal risk and cost ratios. The color bar is in logarithmic scale.

## 6. Welfare Analysis

Now that we have derived the difference in sourcing strategies employed endogenously by the firms and by the social planner, we can study the differences these bring in terms of outcomes for the production network, and derive policy implications from these. Particularly, in this section, I analyse the gain in social welfare the social planner achieves with respect to the competitive equilibrium.

First, by definition, the social planner solution maximises expected welfare, hence it achieves higher welfare than the competitive equilibrium. Yet, a natural question is:

under which basal conditions, $\mu_0$ and $\rho_0$, is the welfare gain maximised, and what sourcing decision the planner makes allow her to achieve it? In Figure 11 I plot the welfare gain for different basal conditions $(\mu_0, \rho_0)$. The gain is maximised in parameters regions (area enclosed by the level curve 0.7 in 11) for which the competitive equilibrium is unstable. This suggests that the social planner is able to diversify higher levels of risk and correlation than the firms acting competitively. This occurs because firms do not internalise, when making their sourcing decision, the increase in risk downstream firms suffer. Another insight into the mechanism by which the social planner achieves first best, can be derived by focusing on the welfare gain achieved in the stable competitive region (under the $s_k = 1$ curve, in Figure 11). Particularly, notice that for higher levels of suppliers' correlation (i.e., higher $\rho$), the welfare gain of the social planner is higher. This is because, even a risk neutral social planner, unlike the firms, incorporates into its optimisation problem the overdispersion ($\rho$) and is hence able to minimise suppliers' correlation, which affects downstream risk ($\mu$).

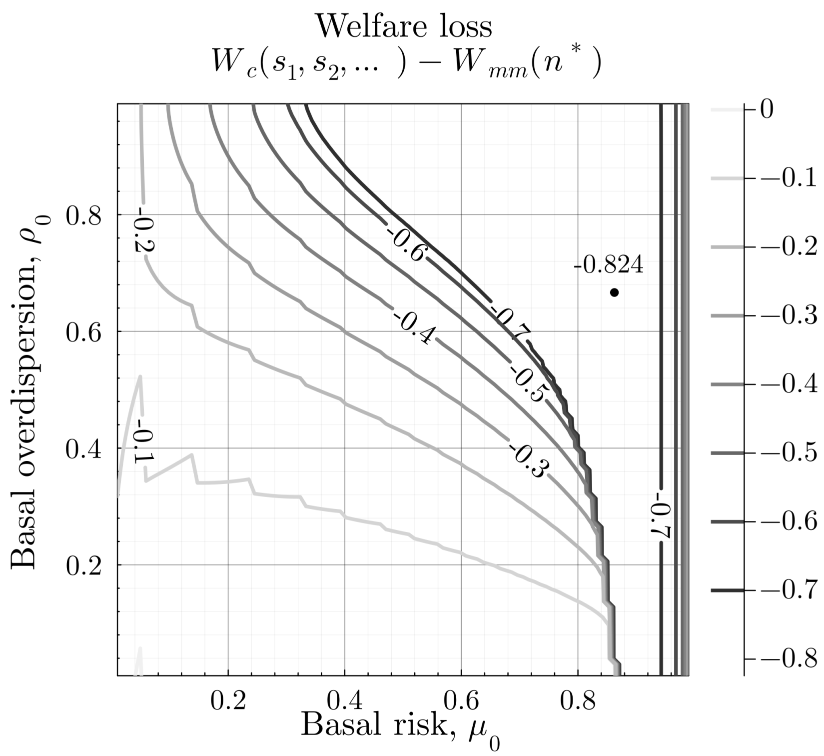



**Figure 11:** Contour plot of the gain in welfare achieved by the social planner, given a basal condition $(\mu_0, \rho_0)$.

These two channels highlight the inherent fragility of the competitive equilibrium of the production game. First, as well established in the literature, when faced with a simple sourcing problem, firms tend to organise critically and the production network is susceptible to cascading failures as a consequence of small shocks in risk. Second, the unobservability of the production network introduces a high probability of tail risk and exacerbates the fragility of the production network, which is a novel feature of the model presented in this paper. The social planner solution shows that such fragility arises endogenously, due to the firm sourcing decision, and does not follow as a feature of the structure of the production network. Particularly, downstream externalities imposed by the firms' sourcing decision have to be addressed to move the competitive equilibrium towards the first best.

# 7. Conclusion

Risk diversification is a crucial determinant of firms' sourcing strategies. In this paper, I show that firms endogenously under-diversify risk when they have incomplete information about upstream sourcing relations. I do so by deriving an analytical solution to a simple production game in which firms' sole objective is to minimise the risk of failing to source input goods. Despite its simple structure, the game identifies an important externality firms impose on the production network when making sourcing decisions: upstream multisourcing introduces correlation in firms' risk, which disincentivises multisourcing downstream. This externality exacerbates the risk of fragile production networks. Furthermore, I show that a risk neutral social planner can not only design production networks that mitigate fully this externality, but, unexpectedly, can do so with the same information as the firms. Particularly, it is sufficient for the planner to enforce a sourcing strategy on firms that does not depend on the realisation of the production network. Such a result no longer holds true for a risk averse social planner. A remarkable consequence of this result is that, in principle, it is possible to design a transfer mechanism that allows downstream firms to compensate upstream firms to internalise such externality.

As mentioned before, the approach presented here is just one of the many points of view that can be taken when studying endogenous production network and imperfect information (others being that of cite). Simple production games can be helpful in isolating mechanisms, but it is often equally, if not more, valuable to embed such mechanisms into more comprehensive and complex models and study how they interact with each other. In this spirit, it would be of great interest to develop a general equilibrium model with endogenous production network with growth and risk diversification motives. In addition, the social planner solution presented here serve as a natural steppingstone for the study of insurance or taxation schemes to mitigate production network externalities. How can we make funds flow upstream in the supply chain to incentives or disincentives diversification? Can such a transfer mechanism be setup without knowledge of the production network structure? The result presented in this paper suggest that this is possible.

## References


Leontief, W. (1936). Quantitative Input and Output Relations in the Economic Systems of the United States [MAG ID: 1977856930]. *The Review of Economics and Statistics*, *18*(3), 105. https://doi.org/10.2307/1927837

Hulten, C. R. (1978). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies*, *45*(3), 511–518. https://doi.org/10.2307/2297252

Billingsley, P. (1995). *Probability and Measure* (Third). John Wiley; Sons.

Tan, J. S., & Kramer, M. A. (1997). A general framework for preventive maintenance optimization in chemical process operations. *Computers & Chemical Engineering*, *21*(12), 1451–1469. https://doi.org/10.1016/S0098-1354(97)88493-1

di Giovanni, J., & Levchenko, A. A. (2010). Putting the Parts Together: Trade, Vertical Linkages, and Business Cycle Comovement [MAG ID: 1980074098]. *American Economic Journal: Macroeconomics*, *2*(2), 95–124. https://doi.org/10.1257/mac.2.2.95

Gabaix, X. (2011). The Granular Origins of Aggregate Fluctuations [MAG ID: 3122444308]. *Econometrica*, *79*(3), 733–772. https://doi.org/10.3386/w15286

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica*, *80*(5), 1977–2016. https://doi.org/10.3982/ECTA9623

Williams, B. D., Roh, J., Tokar, T., & Swink, M. (2013). Leveraging supply chain visibility for responsiveness: The moderating role of internal integration. *Journal of Operations Management*, *31*(7-8), 543–554. https://doi.org/10.1016/j.jom.2013.09.003

Erol, S., & Vohra, R. (2014). Network Formation and Systemic Risk [MAG ID: 599238147 S2ID: 5ca3a071ddd810c2759c8aba0e153aad071f58ad]. *European Economic Review*. https://doi.org/10.2139/ssrn.2500846

Hameed, A., & Khan, F. (2014). A framework to estimate the risk-based shutdown interval for a processing plant. *Journal of Loss Prevention in the Process Industries*, *32*, 18–29. https://doi.org/10.1016/j.jlp.2014.07.009

Macchiavello, R., & Morjaria, A. (2015). The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports. *American Economic Review*, *105*(9), 2911–2945. https://doi.org/10.1257/aer.20120141

Carvalho, V. M., Nirei, M., Saito, Y., & Tahbaz-Salehi, A. (2016). Supply Chain Disruptions: Evidence from the Great East Japan Earthquake [MAG ID: 1524026022]. https://doi.org/10.2139/ssrn.2883800

Galeotti, A., Golub, B., & Goyal, S. (2017). Targeting Interventions in Networks [ARXIV_ID: 1710.06026 MAG ID: 3101673019 S2ID: 25acf76bef3ccbde408873523b760f957d25c137]. *arXiv: Computer Science and Game Theory*. https://doi.org/10.2139/ssrn.3054353

Baqaee, D. R. (2018). Cascading Failures in Production Networks [MAG ID: 2188284008]. *Econometrica*, *86*(5), 1819–1838. https://doi.org/10.3982/ecta15280

Baqaee, D. R., & Farhi, E. (2019). The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem. *Econometrica*, *87*(4), 1155–1203. https://doi.org/10.3982/ECTA15202

Carvalho, V. M., & Tahbaz-Salehi, A. (2019). Production Networks: A Primer. *Annual Review of Economics*, *11*(1), 635–663. https://doi.org/10.1146/annurev-economics-080218-030212

Zhao, M., & Freeman, N. K. (2019). Robust Sourcing from Suppliers under Ambiguously Correlated Major Disruption Risks. *Production and Operations Management*, *28*(2), 441–456. https://doi.org/10.1111/poms.12933

Acemoglu, D., & Azar, P. D. (2020). Endogenous Production Networks. *Econometrica*, *88*(1), 33–82. https://doi.org/10.3982/ECTA15899

Amelkin, V., & Vohra, R. (2020). Strategic Formation and Reliability of Supply Chain Networks [arXiv: 1909.08021]. *arXiv:1909.08021 [cs, econ, eess, math, q-fin]*. Retrieved March 1, 2022, from http://arxiv.org/abs/1909.08021

Kopytov, A., Bineet Mishra, Nimark, K., & Taschereau-Dumouchel, M. (2021). Endogenous Production Networks under Uncertainty [MAG ID: 3212064397]. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3936969

Elliott, M., Golub, B., & Leduc, M. V. (2022). Supply Network Formation and Fragility [arXiv: 2001.03853]. *arXiv:2001.03853 [physics]*. Retrieved February 27, 2022, from http://arxiv.org/abs/2001.03853

Jiang, B., Rigobon, D., & Rigobon, R. (2022). From Just-in-Time, to Just-in-Case, to Just-in-Worst-Case: Simple Models of a Global Supply Chain under Uncertain Aggregate Shocks. *IMF Economic Review*, *70*(1), 141–184. https://doi.org/10.1057/s41308-021-00148-2

Lafrogne-Joussier, R., Martin, J., & Mejean, I. (2022). Supply Shocks in Supply Chains: Evidence from the Early Lockdown in China. *IMF Economic Review*. https://doi.org/10.1057/s41308-022-00166-8

# A. Notation and Distributions

To study firms' decision in an opaque production network, I assume firms reason probabilistically about what is, at the core, a combinatorial problem. To talk about such a problem, it is useful to introduce some notation and distributions that play a central role in describing the reasoning and decisions of firms.

For a real number $x$ and a positive integer $n$, I denote the rising factorial as

$$x^{\overline{n}} := \underbrace{x(x+1)(x+2)\dots(x+n-1)}_{n \text{ terms}}. \tag{31}$$

and the falling factorial

$$x^{\underline{n}} := \underbrace{x(x-1)(x-2)\dots(x-(n+1))}_{n \text{ terms}}. \tag{32}$$

Abusing notation, I often extend this function to non-integer exponents $n$ by using the gamma function $\Gamma$, such that

$$x^{\overline{n}} := \frac{\Gamma(x+n)}{\Gamma(x)}. \tag{33}$$

Two distributions that arise naturally in combinatorial problems and in the framework presented below are the beta and beta-binomial distributions. The former, which I denote $\text{Beta}(\alpha, \beta)$, has two positive parameters. In this paper, I use a more convenient parametrisation, namely, I write

$$\begin{gathered}\text{Beta}(1-\mu, \rho), \text{ where} \\ \mu = \frac{\beta}{\alpha+\beta} \text{ and } \rho = \frac{1}{1+\alpha+\beta}.\end{gathered} \tag{34}$$

This distribution is used here to model uncertainty around the value of a probability distribution of a binomial process. Consequently, we can define a beta-binomial distribution as a compounded binomial distribution with a beta distributed probability parameter. Namely, by letting $X \sim \text{Bin}(m, p)$, where $m \in \mathbb{N}$ and $p \sim \text{Beta}(1-\mu, \rho)$, we can write $X \sim \text{BetaBin}(m, 1-\mu, \rho)$. The random variable $X$ takes values in $[m] := \{0, 1, \dots, m\}$. Informally, one can think of $X$ as a distribution with fatter tails, vis-à-vis a simple binomial, induced by the added uncertainty around the probability parameter $p$. For example, consider the limit case where $\mu = 1/2$ and $\rho = 1/3$. In this case, $p$ is uniformly distributed on the interval $[0, 1]$. Then, $X$ is uniformly distributed on $[m]$. On the other hand, if $\rho = 0$, then $X$ follows a binomial distribution with $p = 1/2$. The first two moments of the beta-binomial distribution can be written as

$$\mathbb{E}[F] = m(1-\mu) \text{ and } \mathbb{V}\text{ar}[F] = m\mu(1-\mu)(1+(m-1)\rho). \tag{35}$$

If $\rho = 0$, the beta-binomial degenerates into a binomial distribution, hence $\rho$ can be interpreted as an "overdispersion" vis-à-vis the binomial distribution. The probability mass function of the beta-binomial distribution is

$$f_F(k) = \binom{m}{k} \frac{B(k+\alpha, m-k+\beta)}{B(\alpha, \beta)} \tag{36}$$

where $B$ is the beta function. The moment generating function of the beta-binomial is

$$M_F(t) = \mathbb{E}\left[e^{tF}\right] = {}_2F_1(-m, \alpha, \alpha+\beta; 1-e^t) = \sum_{n=0}^{m} (-1)^n \binom{m}{n} \frac{B(\alpha+n, \beta)}{B(\alpha, \beta)} (1-e^t). \tag{37}$$

Consider a r.v. $P_k \sim \text{Beta}(1-\mu_k, \rho_k)$, for some parameters $(\mu_k, \rho_k)$ and a r.v. defined conditionally on a realisation $p_k$ of $P_k$ as

$$\big(F_k \mid P_k = p_k\big) \sim \text{Bin}(m, p_k). \tag{38}$$

Then it unconditionally follows a beta-binomial distribution,

$$F_k \sim \text{BetaBin}(m, 1-\mu_k, \rho_k), \tag{39}$$

with mean

$$\mathbb{E}\left[F_k\right] = m\mathbb{E}\left[P_k\right] = m(1-\mu_k) \tag{40}$$

and variance

$$\mathbb{V}\text{ar}\left[F_k\right] = m\mu_k(1-\mu_k)\big(1+(m-1)\rho_k\big). \tag{41}$$

Note that $F_k$ "inherits" its parameters from $P$.

# B. Omitted Proofs

## B.1. The Normalised Beta-Binomial Converges To A Beta Distribution

**Proposition 10.** *Take a r.v. $F_m \sim BetaBin(m, \alpha, \beta)$ and let the normalised beta-binomial be $R_m \coloneqq F_m/m$. Then*

$$R_m \xrightarrow{d} R \sim Beta(\alpha, \beta) \text{ as } m \to \infty. \tag{42}$$

*Proof.* The idea of the proof is to show that the moment generating function of $R_m$ converges pointwise to that of a beta distribution. This implies that the sequence $R_m$ converges in distribution to a beta, since the latter is determined by its moments (Theorem 30.2 in Billingsley, 1995). The moment generating function of $R_m$ can be written in terms of that of $F_m$ as

$$
\begin{aligned}
M_{R_m}(t) &= \mathbb{E}\left[e^{tR_m}\right] \\
&= \mathbb{E}\left[e^{(t/m)F}\right] = M_{F_m}(t/m) \\
&= {}_2F_1(-m, \alpha, \alpha+\beta; 1 - e^{t/m}).
\end{aligned} \tag{43}
$$

We seek to prove that this converges pointwise to

$$
M_R(t) = {}_1F_1(\alpha, \alpha+\beta, t). \tag{44}
$$

Consider the Taylor series of

$$
1 - e^{t/m} = -\frac{t}{m} - \sum_{k=2}^{\infty} \frac{t^k}{m^k\ k!}. \tag{45}
$$

This and the fact that ${}_2F_1$ is continuous in the radius $|t| < m$, allows us to write

$$
\begin{aligned}
\lim_{m\to\infty} M_{R_m}(t) &= \lim_{m\to\infty} {}_2F_1(-m, \alpha, \alpha+\beta; 1 - e^{t/m}) \\
&= \lim_{m\to\infty} {}_2F_1(-m, \alpha, \alpha+\beta; -t/m) \\
&= \lim_{m\to\infty} \sum_{n=0}^{\infty} \frac{\alpha^{(n)}}{(\alpha+\beta)^{(n)}} t^n\ \frac{(-m)^{(n)}}{(-m)^n}.
\end{aligned} \tag{46}
$$

Consider the last term in the summand

$$
\begin{aligned}
\frac{(-m)^{(n)}}{(-m)^n} &= \frac{(-m)(-m+1)(-m+2)\ldots(-m+n-1)}{(-1)^n m^n} \\
&= \frac{(-1)^n m^n + o((-m)^{m-1})}{(-1)^n m^n} \xrightarrow{m\to\infty} 1.
\end{aligned} \tag{47}
$$

Hence

$$
\lim_{m\to\infty} M_{R_m}(t) = \lim_{m\to\infty} \sum_{n=0}^{\infty} \frac{\alpha^{(n)}}{(\alpha+\beta)^{(n)}} t^n\ \frac{(-m)^{(n)}}{(-m)^n} = \sum_{n=0}^{\infty} \frac{\alpha^{(n)}}{(\alpha+\beta)^{(n)}} t^n = M_R(t) \tag{48}
$$

such that $R_m \xrightarrow{d} R$. □

### B.2. Proof of mapping between $F_{k-1}$ and $F_k$ for large $m$

#### B.2.1. Definitions of distribution

We first characterise the two distribution we need (Definitions 5 and 6).

**Definition 5.** *A random variable $P$ follows a BetaPower$(\alpha, \beta, s)$ distribution if it can be written as $P = X^s$ where $X \sim$ Beta$(\alpha, \beta)$.*

**Proposition 11.** *The density of $P$ is*

$$h(p) = \frac{p^{\frac{\alpha}{s}-1}(1 - p^{1/s})^{\beta-1}}{sB(\alpha, \beta)}. \tag{49}$$

*Proof.* The proof follows from letting $g(p) = p^s$, noticing that $g(X) = P$, and using the chain rule to show that

$$h(p) = f_X(g^{-1}(p)) \left| \frac{\mathrm{d}}{\mathrm{d}p} g^{-1}(p) \right| \tag{50}$$

. □

**Corollary 11.1.** *If $P$ follows a BetaBinPower distribution, then $1-P$ follows a BetaBinPower distribution.*

*Proof.* This follows from the symmetry between $h(p)$ and $h(1-p)$. □

**Definition 6.** *A random variable $F_k$ follows a BetaBinPower$(m, \alpha, \beta, s)$ distribution if, given a $P \sim$ BetaPower$(\alpha, \beta, s)$,*

$$\big(F_k | P = p\big) \sim Bin(m, p). \tag{51}$$

*In other words, $F_k$ is a count of the number of successful trials if the probability of success is sampled at each trial from a BetaBinPower distribution.*

#### B.2.2. Objective

I seek to prove that, for countably infinite number of firms $m$, if $F_{k-1}$ follows BetaBinPower distribution, then $F_k$ follows a BetaPower distribution. The strategy is quite straight forward. The number of downstream functioning firms follows

$$\big(F_k | P(s, F_{k-1}) = p\big) \sim \mathrm{Bin}(m, p) \tag{52}$$

Abusing notation, we can redefine $P$ to be a function of $R_m \coloneqq F_m/m$, particularly,

$$1 - P(s, R_m) = \frac{\big(m\ (1 - R_m)\big)_{(s)}}{\big(m\big)_{(s)}}. \tag{53}$$

By considering only the leading term in the falling factorial we have

$$\lim_{m \to \infty} 1 - P(s, R_m) = (1 - R)^s. \tag{54}$$

The number of functioning suppliers is a random variable $F_{k-1}$ that takes value in $[m] := \{0, 1, 2 \ldots m\}$. A firm picks $s \in [m]$ suppliers, based on the expectation of $\mathbb{E}\big[F_{k-1}\big]$. Writing the proportion of firms functioning as $R_m := F_m/m$, probability of functioning of the downstream firm is

$$P(s, F_{k-1}) = 1 - \frac{\big(m(1 - R_m)\big)_{(s)}}{\big(m\big)_{(s)}} \xrightarrow{m \to \infty} 1 - (1 - R)^s \tag{55}$$

where $R = \lim_{m \to \infty} R_m$.

Hence, combining definition 6 and equation (55), if we manage to prove that $R \sim$ BetaBinPower we are done, since reflecting about 1/2 and exponentiating preserves the distributional family. The strategy is to simply prove that the probability mass function of $R_m$ converges to a probability density function of the BetaBinPower family.

### B.2.3. Ingredients

The main building blocks of the proof are two definitions involving the beta function (7 and 8) and a result due to Laplace 12.

**Proposition 12.** *Let $g$ be a twice continuously differentiable on the interval $[0, 1]$, on which it attains a unique maximum, $x_0$. Furthermore, let $h$ be a positive function. Then*

$$\lim_{m \to \infty} \frac{\sqrt{\frac{2\pi}{m|g''(x_0)|}}\, h(x_0) e^{m g(x_0)}}{\int_0^1 h(x) e^{m g(x)}\, dx} = 1 \tag{56}$$

**Definition 7.** *The beta function is defined as*

$$B(\alpha + 1, \beta + 1) = \int_0^1 p^\alpha (1 - p)^\beta\, dp. \tag{57}$$

**Definition 8.** *For an integer $m$ and $r = k/m$ for some integer $k$, we can use definition 7 to write*

$$\binom{m}{mr} = \left((m + 1) \int_0^1 p^{mr} (1 - p)^{m(1-r)}\, dp\right)^{-1}. \tag{58}$$

### B.2.4. Main Result

We are all set.

**Proposition 13.** *If $F \sim$ BetaBinPower$(m, \alpha, \beta, s)$ distribution, then*

$$R = \lim_{m\to\infty} F/m \sim \textit{BetaPower}(\alpha, \beta, s). \tag{59}$$

*Proof.* First we can define the probability mass function of $R_m$

$$f_m(x) = \frac{\int_0^1 h(p)\ p^{mx}(1-p)^{m(1-x)}\mathrm{d}p}{(m+1)\int_0^1 y^{mx}(1-y)^{m(1-x)}\mathrm{d}y} \tag{60}$$

where I rewrote the binomial coefficient as 8 and where $h$ is the density of the BetaPower distribution (11). Consider an $r \in [0, 1]$. Define $\delta_\varepsilon(r)$ to be all the numbers in the support of $f_m$ with distance at most $\varepsilon$ from $r$,

$$\delta_\varepsilon(r) = \left\{x \in \mathrm{supp}(f_m) : |x - r| < \varepsilon\right\} \tag{61}$$

.

Then we can define the probability density function of $R$ as

$$\begin{aligned} f(r) &= \lim_{\varepsilon\to 0}\lim_{m\to\infty} \frac{\sum_{x\in\delta_\varepsilon(r)} f_m(x)}{2\varepsilon} \\ &= \lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\lim_{m\to\infty}\sum_{x\in\delta_\varepsilon(r)} \frac{\int_0^1 h(p)\ p^{mx}(1-p)^{m(1-x)}\mathrm{d}p}{(m+1)\int_0^1 y^{mx}(1-y)^{m(1-x)}\mathrm{d}y} \end{aligned} \tag{62}$$

I will tackle this limit by using the Laplace's method (12). First, we can rewrite the numerator

$$\int_0^1 h(p)\ p^{mx}(1-p)^{m(1-x)}\mathrm{d}p = \int_0^1 h(p) e^{m(x\log p + (1-x)\log(1-p))}\mathrm{d}p, \tag{63}$$

s and notice that the function $p \mapsto x\log p + (1-x)\log(1-p)$ is uniquely maximised at $x$ and has second derivative $-1/x(1-x)$ at the optimum. Proposition (12) implies that the the term in (63) has the same behaviour as

$$h(x)\ x^{mx}(1-x)^{m(1-x)}\sqrt{\frac{2\pi}{m}x(1-x)} \text{ as } m\to\infty. \tag{64}$$

A similar exercise allows us to establish that

$$\lim_{m\to\infty} \frac{\int_0^1 y^{mx}(1-y)^{m(1-x)}\mathrm{d}y}{x^{mx}(1-x)^{m(1-x)}\sqrt{\frac{2\pi}{m}x(1-x)}} = 1. \tag{65}$$

Now we can relate the two limits by picking $\varepsilon = 1/\sqrt{m+1}$ such that we can rewrite the probability density function of $R$ (62) as

$$f(r) = \lim_{m\to\infty} \frac{\sqrt{m+1}}{m+1} \sum_{x\in\delta_{\sqrt{m+1}}(r)} h(x) = h(r). \tag{66}$$

□

# C. Derivations

## C.1. Risk mapping, $G_1$

We can derive analytically the mapping $\mu_k = G_1(\mu_{k-1}, \rho_{k-1}, s_k)$. The first moment of $P_k(s_k, F_{k-1})$ implies that

$$\mu_k = 1 - \mathbb{E}\big[P_k(s_k, F_{k-1})\big] = \frac{\mathbb{E}\big[(m - F_{k-1})_{(s_k)}\big]}{(m)_{(s_k)}}. \tag{67}$$

$F_{k-1} \sim \text{BetaBin}(m, 1-\mu_{k-1}, \rho_{k-1})$ implies

$$(m - F_{k-1}) \sim \text{BetaBin}(m, \mu_{k-1}, \rho_{k-1}). \tag{68}$$

and that $\mathbb{E}\big[(m - F_{k-1})^{\underline{s_k}}\big]$ is the factorial moment of a beta-binomial distribution, which has a known analytical form

$$\mathbb{E}\big[(m - F_{k-1})^{\underline{s_k}}\big] = m^{\underline{s_k}} \frac{B\Big(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}} + s_k, (1-\mu_{k-1})\frac{1-\rho_{k-1}}{\rho_{k-1}}\Big)}{B\Big(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}}, (1-\mu_{k-1})\frac{1-\rho_{k-1}}{\rho_{k-1}}\Big)}, \tag{69}$$

we can write

$$G_1(\mu_{k-1}, \rho_{k-1}, s_k) = \frac{B\Big(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}} + s_k, (1-\mu_{k-1})\frac{1-\rho_{k-1}}{\rho_{k-1}}\Big)}{B\Big(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}}, (1-\mu_{k-1})\frac{1-\rho_{k-1}}{\rho_{k-1}}\Big)}. \tag{70}$$

where $B(x, y)$ is the beta function. If $s_k \in [m]$, we can write

$$G_1(\mu_{k-1}, \rho_{k-1}, s_k) = \frac{\Big(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}}\Big)^{\overline{s_k}}}{\Big(\frac{1-\rho_{k-1}}{\rho_{k-1}}\Big)^{\overline{s_k}}} \tag{71}$$

Furthermore, it is easy to see that

$$\lim_{\rho_{k-1}\to 0} G_1(\mu_{k-1}, \rho_{k-1}, s_k) = \mu_{k-1}^{s_k} \tag{72}$$

which is the limit case we expect if there is no correlation among suppliers and, hence, $F$ follows a binomial distribution.

## C.2. Overdispersion mapping, $G_2$

Unfortunately, we are not as lucky with the mapping for $\rho_k = G_2(\mu_{k-1}, \rho_{k-1}, s_k)$. The first link we can make is using the definition of $\mathbb{V}\text{ar}[P_k(s_k, F_{k-1})]$ to see that

$$\rho_k = \frac{\mathbb{V}\text{ar}[P_k(s_k, F_{k-1})]}{\mu_k\ (1-\mu_k)}. \tag{73}$$

where

$$\mathbb{V}\text{ar}[P_k(s_k, F_{k-1})] = \mathbb{E}[P_k(s_k, F_{k-1})^2] - \underbrace{\mathbb{E}[P_k(s_k, F_{k-1})]^2}_{(1-\mu_k)^2}. \tag{74}$$

Using the definition of $P_k$ we can write

$$\begin{aligned} \mathbb{E}\big[P_k(s_k, F_{k-1})^2\big] &= \mathbb{E}\left[1 - 2\frac{(m-F_{k-1})_{(s_k)}}{(m)_{(s_k)}} + \left(\frac{(m-F_{k-1})_{(s_k)}}{(m)_s}\right)^2\right] \\ &= 1 - 2\ \mu_k + \mathbb{E}\left[\left(\frac{(m-F_{k-1})_{(s_k)}}{(m)_s}\right)^2\right]. \end{aligned} \tag{75}$$

We can derive an analytical expression for this if $m$ is sufficiently large. In particular, let $R_m = F_{k-1}/m$ we can rewrite the last term as

$$\begin{aligned} \mathbb{E}\left[\left(\frac{(m(1-R_m))_{(s_k)}}{(m)_{(s_k)}}\right)^2\right] &= \mathbb{E}\left[\frac{m^2(1-R_m)^2(m(1-R_m)-1)^2\dots(m(1-R_m)-s_k+1)^2}{m^2(m-1)^2\dots(m-s_k+1)^2}\right] \\ &= \mathbb{E}\left[\frac{m^{2s_k}(1-R_m)^{2s_k} + o(m^{2s_k-1})}{m^{2s_k} + o(m^{2s_k-1})}\right] \xrightarrow{m\to\infty} \mathbb{E}\Big[(1-R)^{2s_k}\Big] \end{aligned} \tag{76}$$

where $R = \lim_{m\to\infty} R_m$. As shown in B.2, $R \sim \text{Beta}(1-\mu_k, \rho_k)$, such that $(1-R) \sim \text{Beta}(\mu_k, \rho_k)$, we can write the $2s$-th moment of $1-R$ as

$$\mathbb{E}[(1-R)^{2s_k}] = \left.\frac{\mathrm{d}^{2s_k} M_{1-R}}{\mathrm{d}t^{2s_k}}\right|_{t=0} \tag{77}$$

hence, for large $m$,

$$\begin{aligned}\mathbb{E}\left[p(s,R_m)^2\right] &\approx 1-2\ \mu_k+\left.\frac{\mathrm{d}^{2s_k}\,{}_1F_1}{\mathrm{d}t^{2s_k}}\right|_{(\beta,\alpha+\beta,0)} \\ &= 1-2\ \mu_k+\frac{\beta(\beta+1)\dots(\beta+2s_k-1)}{(\alpha+\beta)(\alpha+\beta+1)\dots(\alpha+\beta+2s_k-1)} \\ \text{re-parametrising,}\quad &= 1-2\ \mu_k+B\left(\frac{1-\rho_k}{\rho_k},2s_k\right)\Big/B\left(\mu_k\ \frac{1-\rho_k}{\rho_k},2s_k\right)\end{aligned} \tag{78}$$

where we have used $\frac{\mathrm{d}}{\mathrm{d}t}\,{}_1F_1(a;b;t)=\frac{a}{b}\,{}_1F_1(a{+}1;b{+}1;t)$ and ${}_1F_1(\cdot,\cdot,0)=1$. This implies that, for large $m$ and letting $\mu_k=G_1(\mu_{k-1},\rho_{k-1},s_k)$, we can write the overdispersion mapping as

$$G_2(\mu_{k-1},\rho_{k-1},s_k)=\frac{B\left(\frac{1-\rho_k}{\rho_k},2s_k\right)\Big/B\left(\mu_k\ \frac{1-\rho_k}{\rho_k},2s_k\right)-(\mu_k)^2}{\mu_k(1-\mu_k)} \tag{79}$$

As a sanity check, we can do the same exercises as above (71) and compute the limit as $\rho_{k-1}\to 0$. We know already that $\lim_{\rho_{k-1}\to 0}\mu_k=\mu_{k-1}^{s_k}$. We are left to compute

$$\lim_{\rho_{k-1}\to 0}\frac{B\left(\frac{1-\rho_{k-1}}{\rho_{k-1}},2s_k\right)}{B\left(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}},2s_k\right)}=\lim_{\rho_{k-1}\to 0}\frac{\left(\mu_{k-1}\frac{1-\rho_{k-1}}{\rho_{k-1}}+2s_k\right)^{2s_k}}{\left(\frac{1-\rho_{k-1}}{\rho_{k-1}}+2s_k\right)^{2s_k}}=\mu_{k-1}^{2s_k}. \tag{80}$$

Then, as expected

$$\lim_{\rho_{k-1}\to 0}G_2(\mu_{k-1},\rho_{k-1},s_k)=0. \tag{81}$$

### C.3. Proof of proximity between $G$ and $\tilde{G}$

First, I prove corollary 6.1

*Proof.* Let $\bar{s}$ be such that $\Pi(\bar{s})=\Pi(\bar{s}+1)$. Since $\Pi$ admits a maximum and is strictly concave, $\bar{s}$ is guaranteed to exist and be unique. This implies that

$$s_k\in[\bar{s},\bar{s}+1]\cap\left\{\lceil\tilde{s}_k\rceil,\lfloor\tilde{s}_k\rfloor\right\} \tag{82}$$

This condition implies that $|s_k-\tilde{s}_k|<1$. □

Second, I prove corollary 6.2

*Proof.* This follows again from Equation 82. In particular, if $\tilde{s}_k > 1$, then

$$s_k \geq \inf\Big\{[\bar{s}, \bar{s}+1] \cap \{\, \lceil \tilde{s}_k \rceil, \lfloor \tilde{s}_k \rfloor \,\}\Big\} \geq \inf\{\, \lceil \tilde{s}_k \rceil, \lfloor \tilde{s}_k \rfloor \,\} = \lfloor \tilde{s}_k \rfloor \geq 1. \tag{83}$$

The analogous procedure can be done with $\tilde{s}_k < 1$. □

## C.4. One-dimensional system

The agents' first order condition implies that

$$\mu^{\tilde{s}} \log(\mu) = \kappa/\pi. \tag{84}$$

This immediately allows to see that $\tilde{g}(\mu) = -\frac{\kappa/\pi}{\log(\mu)}$. The stability of a steady state $\bar{\mu}$ can be derived by

$$\begin{aligned} 1 &> \frac{\partial \tilde{g}}{\partial \mu}\bigg|_{\bar{\mu}} \\ &> \frac{\kappa/\pi}{\log(\bar{\mu})^2 \bar{\mu}} \\ &> \frac{\kappa/\pi}{\bar{\mu}}, \text{ by using } \bar{\mu}\log(\bar{\mu}) = -\frac{\kappa}{\pi}, \\ \bar{\mu}\pi &> \kappa \end{aligned} \tag{85}$$

## C.5. Boundaries of Stable Locus

Consider the profit function

$$\begin{aligned} &\Pi_{k+1} : \mathbb{N}_0 \to \mathbb{R}, \\ &\Pi_{k+1}(s) = \Big(1 - G_1(\mu_k, \rho_k, s)\Big)\pi - \kappa\ s. \end{aligned} \tag{86}$$

We seek to find the boundaries of the set

$$S = \{(\mu, \rho) \in [0,1]^2 :\ s_{k+1}(\mu, \rho) = 1\}. \tag{87}$$

Given that $s_{k+1} = \arg\max_{s \in [m]} \Pi_{k+1}(s)$, the boundaries between the loci with diversification $s$ and $s+1$ are such that the agent is indifferent between $\Pi_{k+1}(s)$ and $\Pi_{k+1}(s+1)$.

This gives us a condition

$$\begin{aligned} 0 &= \Pi_{k+1}(s+1) - \Pi_{k+1}(s) \\ &= \Big(G_1(\mu_k, \rho_k, s) - G_1(\mu_k, \rho_k, s+1)\Big)\pi - \kappa \\ \frac{\kappa}{\pi} &= G_1(\mu_k, \rho_k, s)(1-\mu_k)\left(\frac{\frac{1-\rho}{\rho}}{\frac{1-\rho}{\rho} + s}\right). \end{aligned} \tag{88}$$

If we seek to find the boundary between $s = 0$ and $s + 1 = 1$, we obtain $\mu_k = 1 - \kappa/\pi$. Otherwise, if we seek the boundary between $s = 1$ and $s + 1 = 2$ we obtain

$$\rho_k = 1 - \frac{\kappa/\pi}{\mu_k(1-\mu_k)}. \tag{89}$$