Detecting endogenous effects by aggregate distributions: A case of lumpy investments

Chaoqun Lai
Makoto Nirei

IIR Working Paper WP#09-03

May 2009

Institute of Innovation Research
Hitotsubashi University
Tokyo, Japan
http://www.iir.hit-u.ac.jp

Detecting endogenous effects by aggregate distributions: A case of lumpy investments

Chaoqun Lai
Utah State University

Makoto Nirei*
Institute of Innovation Research
Hitotsubashi University

May 25, 2009

Abstract

This paper studies the effect of strategic complementarity among firms' lumpy investments on the fluctuations of aggregate investments. We investigate an extensive panel data set on Italian manufacturing firms. We first show that the fluctuations of fraction of firms that experience large investment rates in a region-year follow a double-exponential distribution. We then estimate the degree of the strategic complementarity within a region directly by estimating the firm's decision on lumpy investments. We propose a simple sectoral model which is capable of generating the double-exponential distribution for the aggregate fluctuations that arise from the strategic complementarity among firms' lumpy investments. We argue that the shape and magnitude of the aggregate fluctuations observed in the data are consistent with the degree of strategic complementarity estimated at the micro-level in the same data.

Keywords: Strategic complementarity, endogenous effect, non-Gaussian fluctuations

JEL codes: L16, E22

---

*Corresponding author, E-mail: nirei@iir.hit-u.ac.jp, Phone: +81-42-580-8417, Fax: +81-42-580-8410. We are grateful to Luigi Guiso and Fabiano Schivardi for generously sharing the data. This research is initiated during Nirei's visit at Ente Einaudi. We are benefitted by comments from Marc Henry, Ryo Kambayashi, Young-Gak Kim, Andrea Tiseno, and Hajime Tomura.

1 Introduction

This paper empirically studies the aggregate fluctuations that arise from the complementarity of firms' decisions. Our strategy is to focus on the distribution shape of the fluctuations in order to differentiate the aggregate consequence of the endogenous or exogenous mechanism of the fluctuations.

Models of interacting individuals have studied the possibility that the interaction gives rise to aggregate shifts endogenously (Brock and Durlauf [4], Glaeser, Sacerdote, and Scheinkman [15]). It has been recognized that the models of endogenous effects are often unidentified econometrically, as Manski [21] formulated as a reflection problem. In the context of sectoral comovements, it has been proposed to utilize the heterogeneous input-output relations to instrument the endogeneity by Shea [23, 24] and Bartelsman, Caballero, and Lyons [2]. In the context of information spillovers, Guiso and Schivardi [16] recently tackles the problem by utilizing additional information on the reference network of firms.

In this paper, we analyze the same data on Italian firms as Guiso and Schivardi, with a new concentration on large investment episodes. First, we establish that the aggregate fluctuation of the lumpy investments follows a non-normal, exponential distribution empirically. We show that the normality hypothesis is rejected statistically, and is also dominated statistically by an alternative double-exponential hypothesis. Secondly, we present a simple sectoral model with lumpy investments that generates the exponential distribution, while the model generates a normal distribution if it lacks the lumpy behaviors. The model predicts that the parameter of the distribution of the aggregate fluctuations is determined by the degree of strategic complementarity among firms' lumpy investments. Thirdly, we estimate the strategic complementarity from the firm-level data. We use the input-output matrix to identify the endogenous effects for the firm's decision on lumpy investments. In accordance with Bartelsman et al., we observe the externality effect through the output-weighted sectoral activities in the short run. Finally, we show that the complementarity estimated at the micro level is consistent with the fluctuation distribution estimated at the aggregate level. While Guiso and Schivardi investigate information spillovers of continuous investment decisions within an industry, we focus on interactions of lumpy investments within a region. By finding the exponential distribution, we argue that the lumpiness of the investment decision has an important consequence on aggregate fluctuations. Also, we use the input-output relations to identify the endogenous effects, while Guiso and Schivardi utilize the interview data that describe which firms observe which upon their investment decisions.

Aggregate consequences of micro-level lumpy investments have been a topic of extensive discussion. On the one hand, researchers traditionally emphasize the "law of large numbers" effect by which individual lumpy investments cancel out with each other in aggregation. For example, Long and Plosser's [19] sectoral business cycles model has met such arguments by Lucas [20] and Dupor [13]. On the other hand, researchers have investigated the effects of the density of firms around the threshold for lumpy adjustments on aggregate investments. If the capital level of many firms are positioned near to the threshold level for adjustments and far from the desired level of capital, then aggregate investment is likely to increase in near future. This mechanism has been studied empirically by Caballero and Engel [6] and Cooper, Haltiwanger, and Power [10]. However, theoretical investigations have shown that the distribution of firms' gaps between desired and actual capital has an invariant distribution and its convergence is fast (Caplin and Spulber [8] and Caballero and Engel [5]), implying that some aggregate shocks need to be present in order for the extensive margin to fluctuate, such as in the model of Caplin and Leahy [7]. In this line of research, the lumpy investments provide a mechanism of amplifying the exogenous aggregate shocks and of deferring the timing of the impacts.

There is a possibility that the interaction of firms' lumpy investments causes the aggregate fluctuations without aggregate shocks (Nirei [22]). Consider a one-sided (S,s) policy in which firms tend to be distributed uniformly over an inaction band. If there are a continuum of firms, we obtain the "neutrality" result in which lumpy investments do not cause aggregate fluctuations (Caplin and Spulber [8], Caballero and Engel [5]). Now, suppose that there are a large, finite number of firms which are distributed evenly in the inaction band with slight disturbances. Also suppose that the lumpy investments are strategic complements. Then, when a firm that is the closest to the threshold decides to invest, it can cause another lumpy investment of the firm that is positioned at the second closest to the threshold. The investment of the second closest can further cause an investment of the third closest, and so on. This chain reaction stops at the point where there is no firm that is positioned close enough to the threshold. The mechanism is similar to the domino game: the line of falling tiles stops at where two adjacent tiles are standing apart too far. Since the stopping point is altered greatly by a slight change in standing points, the domino effect generates a varying degree of amplification of individual shocks.

We investigate such domino effects in the interaction of lumpy investments in the Long-Plosser type model. The Long-Plosser model exhibits a multiplier effect of idiosyncratic shocks. If the individual behavior is lumpy, then the multiplier itself behaves stochastically, because the it depends on the initial random state just like the standing points of domino tiles. In the model, we show that the aggregate fluctuations generated by the domino effect follow an exponential tail, whereas the standard Long-Plosser model generates a normal distribution. We argue that the particular shape of the distribution of aggregate fluctuations may serve as a symptom that differentiates the endogenous shock model from the exogenous shock model. Often the sources of the exogenous common shock are driven by many factors. In those cases, the central limit theorem predicts that the common shock should follow a normal distribution even when the factors that comprise the common shock follow non-normal distributions. Conversely, when we find a normal distribution in the aggregate variable, it is reasonable to include an unspecified set of exogenous shocks in the model.

The central limit theorem does not characterize the aggregate distribution if there are endogenous effects, since the firms' actions are correlated in that case. Let us consider the simplest case of endogenous effects where the probability for agent i to act ($a_{i,r} = 1$) depends on the realized action of its neighbor $i-1$. Suppose that $\Pr(a_{i,r} = 1) = \beta a_{i-1,r}$. In this case, an action by $i = 1$ may cause a domino effect on the successive agents. The distribution of the number of agents who act, $\sum_{i=1}^{N} a_{i,r}$, follows an exponential distribution. The distribution shape depends on the precise structure of the interactions. In the case of the herd behavior model (Banerjee [1]), for example, the agent i's action is affected by the actions taken by any agent j who acted before i, i.e., $j = 1, 2, \ldots, i-1$. In this case, there is a probability mass on the event that all agents act. The exponential distribution generated by domino effects appears to be robust as long as the domino effect does not degenerate to such a deterministic cascade. We will show that this is in fact the case in our sectoral model.

The rest of the paper is organized as follows. Section 2 describes the data and presents an evidence that the fraction of firms that engage in large investments follows the exponential distribution rather than the normal distribution. Section 3 presents a model of lumpy investment that generates the exponential distribution. The model provides a testable prediction that relates the macro-level property, namely the distribution parameter of aggregate fluctuations, to the micro-level property, namely the degree of strategic complementarity. In Section 4, we numerically simulate the calibrated model, and show that the empirical distribution is fitted well by the simulated distribution. In Section 5, we estimate the firm's decision on lumpy investment directly, and match the estimated strategic complementarity with the distributional estimate. Section 6 concludes.

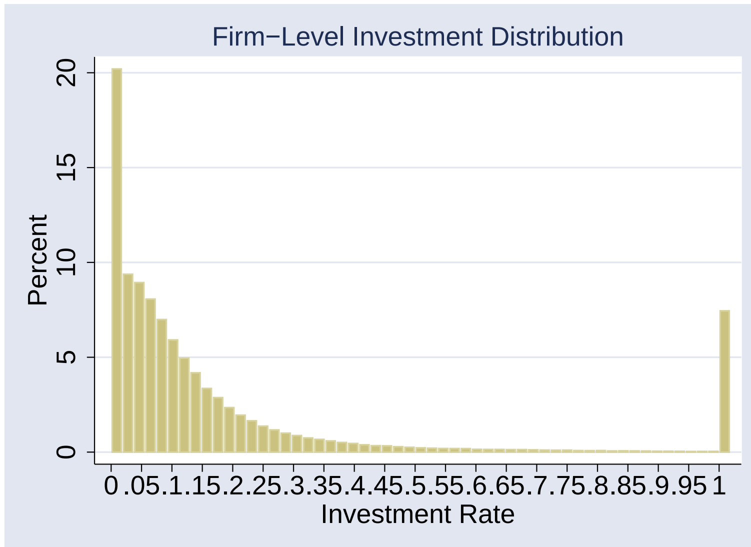

Figure 1: Histogram of firms' investment rates $IPK(i,t)$

2 Fluctuations of aggregate lumpy investments

2.1 Data and variables

We use longitudinal data of Italian firms drawn from the Company Accounts Data Service (CADS). The data set we use was compiled by Guiso and Schivardi [16]. The annual data set covers over 30,000 firms from 20 regions, 25 industries, and 15 years from 1982 to 1996.

We focus on the number of firms that experience large investment episodes in a given reference group in each year. First, we define an investment-capital ratio, $IPK(i,t)$, for each firm i and year t. Figure 1 depicts a histogram of $IPK(i,t)$. We can see that the distribution is highly skewed to the right. 24% of the samples have investment rates less than 2%. The tail is long, and it implies that a relatively small fraction of firms has large impacts on aggregate investments.

Such a long tail in investment rates is observed commonly as in Doms and Dunne [12] and Cooper et al. [10]. Doms and Dunne also found the salient spiking of investment behaviors, in which firms adjust capital actively in a short period and show little adjustments in other periods. Along with the literature following their findings, we are most concerned with the aggregate consequence of such spiking investments. Thus, we divide the heterogeneous behaviors of IPK into two polar groups, lumpy investments and inactions. That is, we convert the investment-capital ratio into a binary variable, $d(i,t)$, which takes 1 if $IPK(i,t) > \bar{d}$ and zero otherwise. We take the threshold $\bar{d}$ at 20% for our estimation, in order to be comparable with the literature on lumpy investments ([10] for example). The sum of the lumpy investments defined as such accounts for about 20% of the aggregate investments in our data. The results obtained in this paper are not affected very much when we change the threshold to 10% or 30%.

This paper is concerned with the *fraction of investing firms*, $X(G.,t) = \sum_{i \in G.} d(i,t)/\#G.$, where $G.$ is a reference group and $\#G.$ is the number of firms in $G.$.[1] We drop the groups that have firms less than 10, because the behavior of the fraction X can be overly volatile for the groups that have a small number of firms.[2] We then define a centered fraction variable $\tilde{X}(G.,t) \equiv X(G.,t) - \langle X(G.,t) \rangle_{G.}$. Namely, yearly common effects across reference groups is subtracted from $X(G.,t)$. We work with the region reference group G_r which is a set of firms that operate in the same region r, as well as the industry-region reference group $G_{l,r}$ where l is an industry. The variables of interest are summarized in Table 1.

Figure 2 plots the histogram of $X(G_{l,r},t)$ for each year. We observe a general pattern of the distributions: they are fairly centered and their center location fluctuates over years. In this paper, we regard the annual common shock as exogenous, and concentrate on the remaining fluctuations observed across reference groups. The left panel of Figure 3 shows the histogram of $\tilde{X}(G_{l,r})$ in which the yearly effect on X is subtracted. The histogram of $\tilde{X}(G_{l,r})$ is fairly symmetric. In the right panel is the centered histogram for region reference groups $\tilde{X}(G_r)$, which shows the similar pattern to the industry-region histogram.

Our goal is to characterize and explain the distribution of the fraction of investing firms $\tilde{X}$. Figure 4 shows semi-log plots of the cumulative distributions of $\tilde{X}(G_{l,r},t)$ and $\tilde{X}(G_r,t)$ for positive and negative sides. To produce the plot, we first divide $\tilde{X}$ into two groups depending on the sign of $\tilde{X}$, and for each positive or negative group we rank $|\tilde{X}|$ in a descending order. Then, the log of the rank divided by the total number of observations is plotted against $|\tilde{X}|$ for positive and negative groups.

[1] The number of observations in the reference $G.$ may change over time. The time subscript t is suppressed here only for simplicity.

[2] Original data file contains 306363 observations. We drop observations with missing industry or region codes, and one outlier of the variable "ioverk". Then we are left with 283210 observations. When we work with the fraction variables X, we also exclude small reference groups that contain less than 10 firms. In the regression analysis that involves a lag variable, we exclude observations in the initial year 1982.

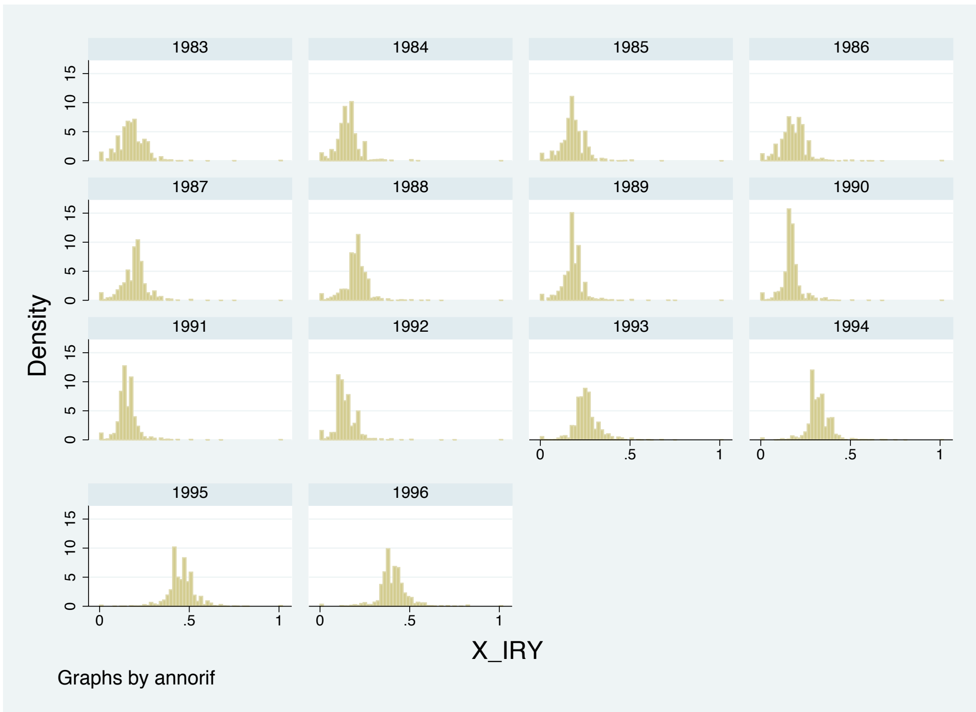

Figure 2: Histograms of $X(G_{l,r},t)$ for each year

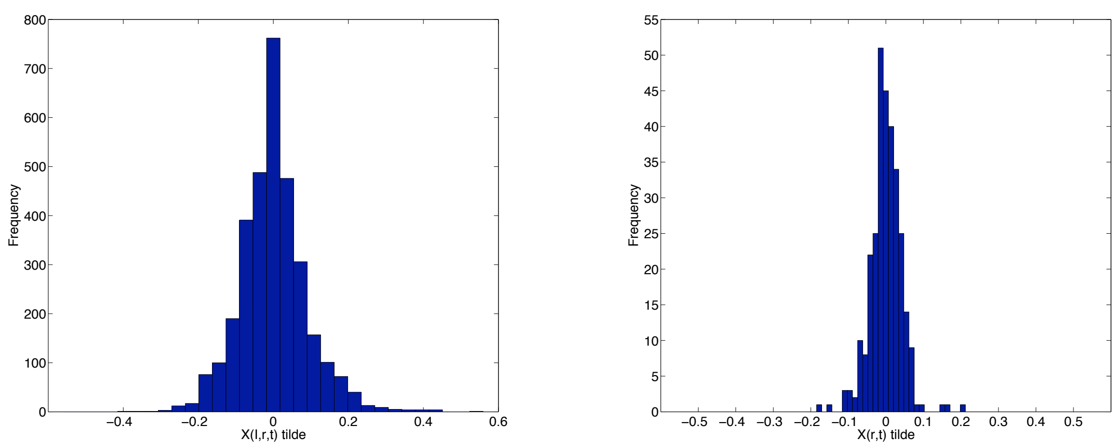

Figure 3: Histograms of $\tilde{X}(G_{l,r})$ (left) and $\tilde{X}(G_r)$ (right)

	Obs	Mean	Std. Dev.	Min	Max	Median
$d(i,t)$	283210	0.224	0.417	0	1	0
$X(G_{l,r},t)$	3234	0.213	0.141	0	0.830	0.185
$\#G_{l,r}$	3234	84	154	10	1465	33
$X(G_r,t)$	298	0.208	0.119	0	0.538	0.179
$\#G_r$	298	950	1396	10	7919	381
$X(G_l,t)$	362	0.210	0.116	0	0.571	0.183
$\#G_l$	362	782	936	10	4825	392

Table 1: Description of variables. X is the number of firms that engage in lumpy investments divided by the total number of firms in a reference group. Reference groups with less than 10 firms are dropped from the observation.

Thus, the vertical line shows the percentile in the descending order. We observe a steeper slope for the regional distribution compared to the industry-regional distribution.

We now fit parametric distributions to the empirical distribution. Our null hypothesis is the normal distribution. An alternative hypothesis is the double-exponential distribution, in which the distribution of $\tilde{X}(G_{l,r},t)$ follows exponential distributions for the positive and negative sides with possibly different means λ_+ and λ_-. Namely, the likelihood function is $\Pr(\tilde{X}=\tilde{x}|\tilde{x}>0)=(1/\lambda_+)e^{-(1/\lambda_+)\tilde{x}}$ and $\Pr(\tilde{X}=\tilde{x}|\tilde{x}<0)=(1/\lambda_-)e^{-(1/\lambda_-)(-\tilde{x})}$. The other alternative hypothesis is the Laplace distribution in which both positive and negative sides of $\tilde{X}$ follow an exponential distribution with the same mean λ: $\Pr(\tilde{X}=\tilde{x})=(1/(2\lambda))e^{-(1/\lambda)|\tilde{x}|}$.

Table 2 shows the results of the maximum likelihood estimation for each parametrization and for each reference group. We start from the industry-region reference group $\tilde{X}(G_{l,r},t)$. The estimated mean of the exponential distribution is 0.071 (standard error 0.002) for the positive side and 0.062 (standard error 0.002) for the negative side of $\tilde{X}$. The estimated λ for the Laplacian is 0.066, which is the middle value for the exponential slopes for the positive and negative sides. The exponential hypothesis has a larger log likelihood value than the Laplacian hypothesis, because the Laplacian distribution is equivalent to the double-exponential distribution with restriction $\lambda_+=\lambda_-$.

We test the normality hypothesis by a likelihood-based test. Let $L(i;H)$ denote the likelihood of

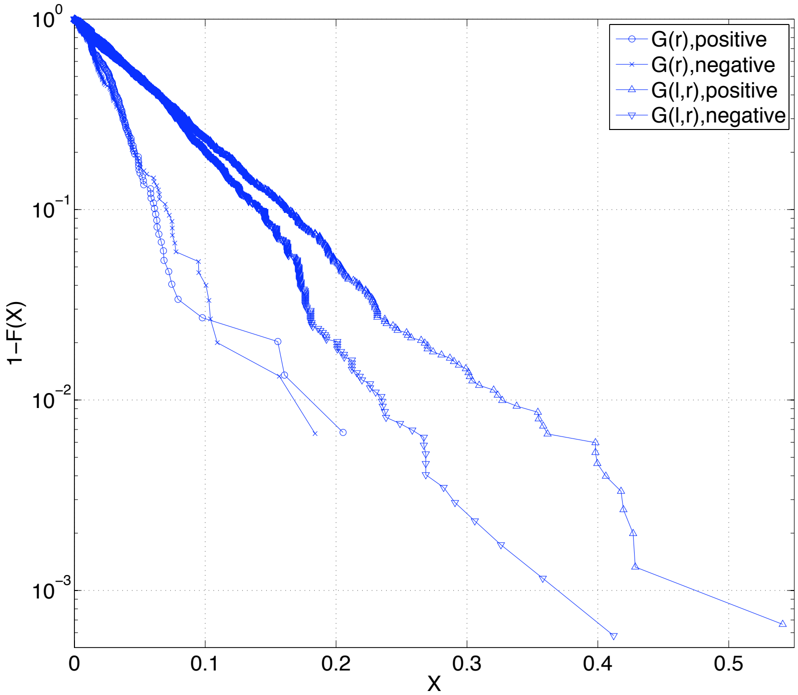

Figure 4: Semi-log plots of cumulative distributions of $|\tilde{X}(G_{l,r},t)|$ and $|\tilde{X}(G_r,t)|$ for positive and negative values. The distributions are cumulated from above.

	Industry-region $\tilde{X}(G_{l,r},t)$						Region $\tilde{X}(G_r,t)$					
	Exponential		Laplace		Normal		Exponential		Laplace		Normal	
ML estimates	λ_+	0.071	λ	0.066	μ	0.000	λ_+	0.031	λ	0.031	μ	0.000
		(0.002)		(0.001)		(0.002)		(0.003)		(0.002)		(0.003)
	λ_-	0.062			σ	0.090	λ_-	0.031			σ	0.042
		(0.002)				(0.000)		(0.003)				(0.000)
Log Likelihood		3337		3185		3200		531.9		518.4		519.7
Vuong's statistic		4.98				Null		1.30				Null
				-0.50		Null				-0.12		Null
		10.70		Null				3.06		Null		

Table 2: Parameter estimates for various distributions. Standard errors are in parentheses.

sample point i under the hypothesis H. Define the log likelihood ratio for each i as $l_i = \log L(i;H_1) - \log L(i;H_0)$. We use Vuong's statistic $V \equiv \sqrt{N}\bar{l}_i/\text{Std}(l_i)$ which follows a standard normal distribution if the hypotheses H_0 and H_1 are "equivalent" in the sense of Kullback-Leibler information. Thus, if V computed for H_1 against the null H_0 is greater than 1.96, then the null is rejected in favor of the alternative at the 5% significance level. Vuong's statistics are reported in Table 2 for the exponential and the Laplacian hypotheses against the Gaussian null hypothesis and for the exponential against the Laplacian. The exponential is favored against the Gaussian and the Laplacian at the 1% significance, while the Laplacian is not significantly different from the Gaussian. We repeat the same test for $\tilde{X}(G_r,t)$ on the right side of Table 2. Vuong's statistics suggest that the exponential hypothesis is favored against the Laplacian hypothesis at 1% significance level, but it is not significantly different from the Gaussian. Note that we have fewer observations for the region group compared to the industry-region group. Thus, we lose the testing power somewhat due to the limited number of observations for the case of regional distribution.

We now turn to another normality test which is based on higher moments. Table 3 shows the higher moments of $\tilde{X}$. The large kurtosis indicates that $\tilde{X}$ is leptokurtic. Standard normality tests such as skewness-kurtosis test, Shapiro-Wilk test, and Shapiro-Francia test reject the normality hypothesis overwhelmingly for both $\tilde{X}(G_{l,r},t)$ and $\tilde{X}(G_{r,t})$.

Variable	Observations	Mean	Std Deviation	Skewness	Kurtosis
$\tilde{X}(G_{l,r},t)$	3234	0	0.090	0.537	5.274
$\tilde{X}(G_r,t)$	298	0	0.042	0.692	6.703

Table 3: Moments of $\tilde{X}$

We also conduct a test that is only based on kurtosis but not on skewness. The concentration of the 4th moment is useful in order to test among a class of symmetric distributions, in our case the Laplacian and the Gaussian. The true kurtosis is equal to 3 under the Gaussian hypothesis and 6 under the Laplacian hypothesis. Consider a method of moment estimator $(\sum_{i=1}^{N} x_i^4/N)/(\sum_{i=1}^{N} x^2/N)^2$. Then, its asymptotic variance is $V(x_i^4)/\sigma^2 N$. Under the Gaussian hypothesis, the variance is equal to $96/N$. Thus, the sample kurtosis in Table 3 rejects the Gaussian hypothesis at 1% significance, even though the estimator is not efficient. In contrast, the asymptotic variance of the MOM estimator is $(8! - 4!^2)/2^8 N = 155.25/N$ under the Laplacian hypothesis. Thus, the MOM test permits the Laplacian hypothesis at 5% significance for $\tilde{X}(G_r,t)$.

The good fit of the exponential parametrization can be seen in Figure 4. In the semi-log scale, an exponential distribution function would show as a linear line. Our plot demonstrates that the positive side of the distribution of $\tilde{X}(G_{l,r},t)$ is well fitted by a linear line. The negative side has a kink beyond which the distribution shows a faster decline than the positive side. The faster decline in the negative side can be caused by the boundary effect of $X(G_{l,r},t)$ which takes only non-negative values. Since $X(G_{l,r},t)$ has mean 0.21 and standard deviation 0.14 as in Table 1, it is natural that the distribution of $\tilde{X}$ shows a boundary effect that is seen in the empirical distribution at around -0.17. This explanation is also consistent with the patterns of yearly distributions shown in Figure 2. The distribution is skewed in the years when the distributions are close to zero, such as in 1984 and 1992, whereas the distribution is not skewed in the boom years when the distributions are far from zero, such as in 1994, 1995, and 1996. The downward kink may be the reason why Vuong's test cannot reject the Gaussian hypothesis for $\tilde{X}(G_{l,r},t)$. The right panel of Figure 4 shows a distribution of $\tilde{X}(G_r,t)$. The distribution does not exhibit the kink for the negative side, since the zero-bound is not binding for $\tilde{X}(G_r,t)$ due to its small standard deviation.

In sum, we show that the distribution of the fraction of firms that engage in lumpy investments exhibits an exponential decay rather than a Gaussian decay at the industry-region level or the region level. Thus, the empirical distribution favors the model that generates an exponential pattern than the models that generate normal distributions for the aggregate fluctuation of investments. The fluctuation we investigated captures the extent of aggregate fluctuations that are not explained by the common annual shocks. For the case of the regional level, the variation of $\tilde{X}(G_r, t)$ accounts for a third of the total variation in $X(G_r, t)$. In the following sections, we try to understand the mechanism that generates the exponential pattern of the aggregate fluctuation.

3 A model of endogenous investment fluctuations

This section presents a simple model of endogenous investment fluctuations that generates an exponential tail for the distribution of X as we observed previously. Consider that there are N firms that produce differentiated goods with a production function:

$$y_i = A_i n_i^{\alpha}. \tag{1}$$

A_i is an exogenous productivity, and n_i is input goods. The returns-to-scale of production is $\alpha < 1$. We assume that the firm is constrained on the operation scale n_i to a discrete set:

$$n_i \in \{1, \lambda_i^{\pm 1}, \lambda_i^{\pm 2}, \ldots\} \tag{2}$$

where $\lambda_i > 1$ is the exogenous parameter for lumpiness. The constraint represents the situation where firms must choose an integer for the number of plants, although we restrict growth rates rather than levels for the sake of tractability.

Input n_i is the composite good produced by using all the goods in a CES manner:

$$n_i = \left(\sum_{j=1}^{N} \chi_{i,j}^{1/\sigma} z_{i,j}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}, \tag{3}$$

where $\sigma > 1$. Note that the input weight $\chi_{i,j}$ is different across i. By $\chi_{i,j}$ we incorporate the heterogeneous inter-industrial effects in the input-output relations. Define the aggregate index of input prices for industry i as a weighted sum:

$$P_i \equiv \left(\sum_{j=1}^{N} \chi_{i,j} p_j^{1-\sigma} \right)^{1/(1-\sigma)}. \tag{4}$$

Then, the derived demand for good j is given in an isoelastic form as: $z^*_{i,j} = (p_j/P_i)^{-\sigma}\chi_{i,j}n_i$. Thus, the derived demand $z^*_{i,j}$ changes proportionally to the weight $\chi_{i,j}$ when the input n_i varies. The minimized cost satisfies $\sum_j p_j z^*_{i,j} = P_i n_i$.

We assume that each firm i is owned by household i as in a backyard production economy. Household i's income is the profit π_i from producing good i. Households have a common utility function $U(C_i)$ over a composite consumption good: $C_i = \left(\sum_{j=1}^N \chi_{C,j}^{1/\sigma}(z^i_{C,j})^{(\sigma-1)/\sigma}\right)^{\sigma/(\sigma-1)}$. Aggregate variables across households are defined as $z_{C,j} = \sum_i z^i_{C,j}$, $C = \sum_i C_i$, and $\pi = \sum_i \pi_i$. Also define the aggregate index of consumer price:

$$P_C \equiv \left(\sum_{j=1}^N \chi_{C,j} p_j^{1-\sigma}\right)^{1/(1-\sigma)}, \tag{5}$$

and normalize it to 1. Then the derived demand for $z^i_{C,j}$ is: $\sum_i z^{i*}_{C,j} = (p_j/P_C)^{-\sigma}\chi_{C,j}\sum_i C_i$. The optimal expenditure satisfies $\sum_j p_j z^{i*}_{C,j} = P_C C_i = \pi_i$.

The equilibrium condition for good i is $y_i = \sum_j z_{j,i} + z_{C,i}$. Thus, the total demand function is isoelastic:

$$y_i = p_i^{-\sigma} Y_i, \tag{6}$$

where Y_i is the aggregate demand factor for good i:

$$Y_i \equiv \sum_j P_j^\sigma n_j \chi_{j,i} + C\chi_{C,i}. \tag{7}$$

Aggregate output is defined as $Y \equiv \sum_{i=1}^N y_i^{1-1/\sigma} Y_i^{1/\sigma}$.

Each firm is a monopolistic supplier of a differentiated good. Its profit is defined as $\pi_i = p_i y_i - \sum_j p_j z_{i,j}$. The second term, the cost of inputs, is equal to $P_i n_i$ at optimum by the result of cost minimization. The monopolist's problem is thus $\max_{n_i} p_i y_i - P_i n_i$ subject to (1), (2), and (6). We define an "inaction band" of n_i by the lower and upper thresholds $[n^*_i, \lambda_i n^*_i)$. We can find such thresholds by noting that the profits must be equalized at the two boundaries of the inaction band (see [22]). Then the bound is solved as:

$$n^*_i = \left(\frac{\lambda_i^{\alpha(1-1/\sigma)} - 1}{\lambda_i - 1} A_i^{1-1/\sigma} Y_i^{1/\sigma} P_i^{-1}\right)^{\frac{1}{1-\alpha(1-1/\sigma)}} \tag{8}$$

Consider the situation where the productivity A_i is a random variable independent across i. For each realization of the productivity profile, an equilibrium profile of n_i is determined by the threshold rule $n_i^* \leq n_i < \lambda_i n_i^*$ and the discreteness constraint (2). We are interested in the probability distribution of the fraction of investing firms (the fraction takes a negative value if firms adjust downward), when the exogenous productivities A_i's are slightly perturbed.

Suppose that we are at an equilibrium, and now a small perturbation on A_i affects the equilibrium. An increase in A_i raises n_i^*, and then induces i to increase n_i to $\lambda_i n_i$ if i was near the lower threshold n_i^*. The increase in n_i raises the demand for other goods Y_j's. The shift in demand increases n_j^*'s, and thus the firms near the lower threshold n_j^* may then choose to increase n_j by a factor of λ_j. This propagation process continues until the economy reaches a new equilibrium. We regard this process as a fictitious tatonnement, and an equilibrium selection algorithm as in Vives [25] and Cooper [9]. The fictitious tatonnement is defined in Appendix A.1 precisely. The selected equilibrium has the property that the equilibrium is the closest to the initial equilibrium in the direction of the search (namely, the direction of the initial adjustment).

In what follows we analytically show that the model generates the exponential distribution for the fluctuations of the fraction of investing firms. The analytical characterization is carried out by embedding the fictitious tatonnement in a stochastic process ex ante the random variables realize, following the method presented in a separate paper [22]. Here we consider the case when the input weight $\chi_{i,j}$ is the same across i: $\chi_j = \chi_{i,j}$ for all i. The numerical simulations of the heterogeneous case under calibrated parameters are presented in the next section. The aggregate output is simplified as $Y = (\sum_i \chi_i^{1/\sigma} y_i^{1-1/\sigma})^{\sigma/(\sigma-1)}$. Then we obtain the optimal threshold as:

$$n_i^* = a_i A_i^{\frac{1-1/\sigma}{1-\alpha(1-1/\sigma)}} Y^{\frac{1}{\sigma(1-\alpha(1-1/\sigma))}} \tag{9}$$

$$= a_i A_i^{\frac{1-1/\sigma}{1-\alpha(1-1/\sigma)}} \left(\sum_j \chi_j^{\frac{1}{\sigma}} A_j^{1-\frac{1}{\sigma}} n_j^{\alpha(1-1/\sigma)} \right)^{\frac{1}{(\sigma-1)(1-\alpha(1-1/\sigma))}} \tag{10}$$

where

$$a_i \equiv \frac{\lambda_i^{\alpha(1-1/\sigma)} - 1}{\lambda_i - 1} \chi_i^{\frac{1}{\sigma}}. \tag{11}$$

By the Taylor series expansion as in Appendix A.2, we obtain:

$$d \log n_i^* = \frac{\beta(\sigma-1)}{\alpha} d \log A_i + \beta \sum_{j=1}^{N} b_j \frac{d \log n_j}{\log \lambda_j} + O(N^{-2}) \tag{12}$$

where

$$\beta \equiv \frac{1}{\sigma(1/\alpha - 1) + 1} \tag{13}$$

$$b_j \equiv \frac{\chi_j^{\frac{1}{\sigma}} A_j^{1-\frac{1}{\sigma}} n_{j,0}^{\alpha(1-1/\sigma)}}{\sum_l \chi_l^{\frac{1}{\sigma}} A_l^{1-\frac{1}{\sigma}} n_{l,0}^{\alpha(1-1/\sigma)}} \frac{\lambda_j^{\alpha(1-1/\sigma)} - 1}{\alpha(1-1/\sigma)}. \tag{14}$$

Now define the initial gap between the actual input and the lower threshold as follows:

$$s_i \equiv \frac{\log n_i - \log n_i^*}{\log \lambda_i} \tag{15}$$

Here n_i is the initial level of input that is determined by the productivity and the aggregate demand before the perturbation. Thus s_i is the state associated to the initial equilibrium before the perturbation. Then, the optimal rule is to invest upon the perturbation if:

$$s_i < -\frac{d \log n_i^*}{\log \lambda_i} = -\frac{1}{\log \lambda_i} \left(\frac{\beta(\sigma-1)}{\alpha} d \log A_i + \beta \sum_{j=1}^{N} b_j \frac{d \log n_j}{\log \lambda_j} \right) \tag{16}$$

for large N. From (16), we see that the investment decision depends on the exogenous shock on A_i and the endogenous shock on Y. An investment by firm j increases Y, which reduces the gap and increases the likelihood of firm i to invest. The magnitude of this impact of $\log \lambda_j$ on the gap of i through the strategic complementarity is:

$$-\frac{\beta b_j}{\log \lambda_i} \frac{d \log n_j}{\log \lambda_j} \tag{17}$$

Note that for an upward lumpy adjustment, $d \log n_j / \log \lambda_j = 1$. Also, the first fraction in (14) is $1/N$ if firms are homogeneous, and the second fraction is approximately $\log \lambda_j$ when the lumpiness is small (λ_j is close to 1). Thus, the impact on the gap variable is roughly equal to β/N in the homogeneous case. Consider a random variable of s_i unconditional on λ_i and A_i that is distributed on the unit circumference of a circle. We suppose, as a first-order approximation of the distribution, that the density around $s_i = 0$ is constant q. Then the probability that j's adjustment induces i to adjust is equal to $q\beta b_j / \log \lambda_i$. The probability is reduced to $q\beta/N$ for the homogeneous case with small lumpiness.

Consider the fictitious tatonnement process that starts from the initial state (s_i). Suppose that there is one firm j that adjusts capital at the first round of the tatonnement upon the perturbation. Then, the number of firms that are induced to adjust in the second step of the tatonnement conditional on the adjustment of firm j follows a convolution of a Bernoulli trial with probability $q\beta b_i / \log \lambda_j$ across i. Thus the number of firms that are induced to adjust in the second step unconditionally follows a mixture across j of the convolution above, which is an integer random variable with mean $\phi \equiv$

$q\beta E(\sum_{i=1}^{N} b_i)E(1/\log \lambda_j)$. Then, the number of firms m_u that adjust in each step u of the tatonnement conditional to m_{u-1} follows a distribution that is represented as a m_{u-1}-times convolution of a certain integer random variable with mean ϕ. The integer random variable is identically distributed across u as long as N is so large that $\sum_{v=1}^{u} m_v/N$ is small and thus the density of s_i of the affected firms is constant at q (see [22] for further details and generalizations). Then we can utilize the following theorem.

Theorem 1 (Otter) *Consider a branching process m_u, $u = 1, 2, \dots$, that starts from $m_1 = 1$. Let ϕ denote the mean number of children in $u + 1$ born by each parent in u. Define the total number $M = \sum_{u=1}^{T} m_u$. Then:*

$$\Pr(|M| = m|m_1 = 1) = C_0(e^{\phi-1}/\phi)^{-m} m^{-1.5} \tag{18}$$

for a large integer m, where C_0 is a constant.

When N tends to infinity, m_u approximately follows a Poisson distribution with mean ϕm_{u-1}. By using this, the analytical result can be sharpened when the lumpiness λ_i, the productivity A_i, and the input weight $\chi_{i,j}$ are homogeneous across i, j.

Proposition 1 *When λ_i, A_i and $\chi_{i,j}$ are common across i, j, M follows a symmetric probability distribution function:*

$$\Pr(|M| = m \mid m_1) = m_1 e^{-\phi(m+m_1)} \phi^m (m + m_1)^{m-1}/m!. \tag{19}$$

The tail of the distribution function is approximated as:

$$\Pr(|M| = w \mid m_1) \sim (m_1 e^{(1-\phi)m_1}/\sqrt{2\pi})(e^{\phi-1}/\phi)^{-m} m^{-1.5}. \tag{20}$$

Proof: This result draws on Nirei [22]. See Appendix A.3.

Proposition 1 shows that M follows a distribution that is a mixture of power and exponential functions as seen in (20). Since the exponential function declines faster than the power function, the tail of the distribution is dominated by the exponential part. We will argue that this corresponds to our empirical findings that X follows an exponential distribution. Proposition 1 further shows that the slope of the exponential distribution is determined by ϕ. By Equation (18), the exponential slope is $\phi - 1 - \log \phi$, and the mean and standard deviation of the exponential part is determined by the inverse of the slope. Since the slope is decreasing in ϕ in the region $\phi < 1$, a large ϕ implies a large fluctuation of M. This is intuitive, because ϕ is the mean number of firms that are induced to invest by a firm's lumpy investment in the fictitious tatonnement.

In the model, ϕ is directly affected by β, which summarizes the impact of a lumpy investment on the gap variable s_i of other firms in the micro-level investment decision (16). In this sense, ϕ is the parameter of strategic complementarity among firms' lumpy investments. From (13), a small returns-to-scale (small α) or a low mark-up (large σ) implies a small β and ϕ, and thus a steep slope. This prediction will be confirmed in a calibrated simulation in the next section.

Since β appears in the single firm's decision rule (16), we can estimate it from the panel data independently from the estimation of the distribution of the aggregate fluctuations. Thus, Proposition 1 provides a testable implication that connects the two estimates at the micro level and the macro level. In Section 5, we estimate the firm's decision at the micro level, and use this connection to test the validity of the model.

4 Numerical simulations

In this section, we simulate the model under calibrated parameters. We calibrate the distribution of λ_i to an exponential distribution with lower bound 1.1 and mean 1.2, which mimics the histogram of investment rate in our data (shown in Figure 1). We set $\chi_{i,j}$ as a transpose of the Italian input-output matrix normalized to 1 column-wise.[3] We also set A_i to mimic the distribution of total input demand across industry in the input-output matrix. The returns to scale is set at $\alpha = 0.8$ and 0.9. The σ is set at 6 and 11, which corresponds to the mark-up rate μ at 20% and 10%, respectively (the mark-up rate satisfies $\mu = 1/(\sigma - 1)$).

We compute the distribution of the number of investing firms numerically under the calibration. We first draw the profiles of λ_i and A_i, and compute the initial equilibrium for the realization. Then we add a small shock to A_i (the variance of the shock is set so that the mean number of firms that adjust to the shock initially is 1), and compute a new equilibrium. The left panel of Figure 5 shows the histogram of the fraction of firms that adjust their capital upward in the transition from the old to the

[3] We use the 1985 version of the Italian transaction table to identify the industrial relation. We convert the original 4-digit industry code of the Italian firms in our data set to 2-digit industry code that is consistent with the OECD classification. With the data available we end up with 25 industries.

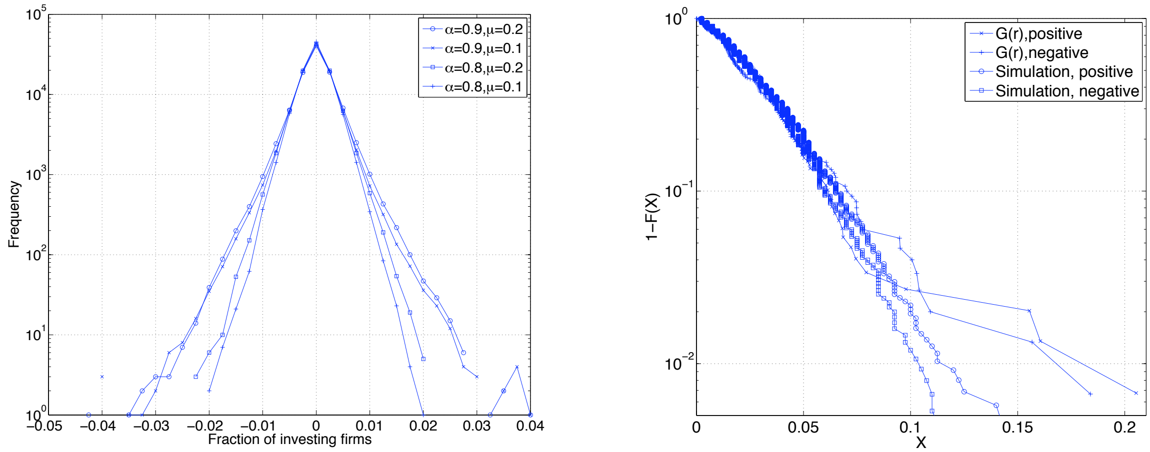

Figure 5: Left: simulated histogram of the fraction of investing firms when the sizes of lumpiness, industry, and input weights are heterogeneous. Right: the cumulative distribution of $|\tilde{X}(G_r, t)|$ and the simulated distribution.

new equilibrium (a negative number means that those firms adjust capital downward). The vertical axis is log-scaled, and thus a linear line implies an exponential decay. The simulated histogram clearly conforms to the exponential pattern.

Next, we try to match with the empirical exponential distribution of $\tilde{X}(G_r, t)$ shown in Figure 4. The variance of $\tilde{X}(G_r, t)$ is much larger than the variances of the simulated distributions shown in the cases $\alpha = 0.8, 0.9$. Thus, we repeat the simulation with a higher returns to scale at $\alpha = 0.95$, and also computed the sum of the fraction of firms across 60 periods. We compare the resulting simulated distribution with the distribution of $\tilde{X}(G_r, t)$ in the right panel of Figure 5. We observe a good fitting. The summation across periods has a similar effect to increasing the perturbation size. The time-accumulation does round the shape of exponential distribution toward the normal distribution near the mean 0, as predicted by the central limit theorem, but it still preserves the exponential tail far from the mean. Moreover, it improves the fitting around 0, because the empirical distribution shows a slight deviation from the exponential distribution that is well captured by the rounding effect.

The left panel of Figure 6 shows the distribution of the log-deviation in aggregate output $\sum_i p_i y_i$. The plots show that the aggregate output largely inherits the exponential pattern in the distribution

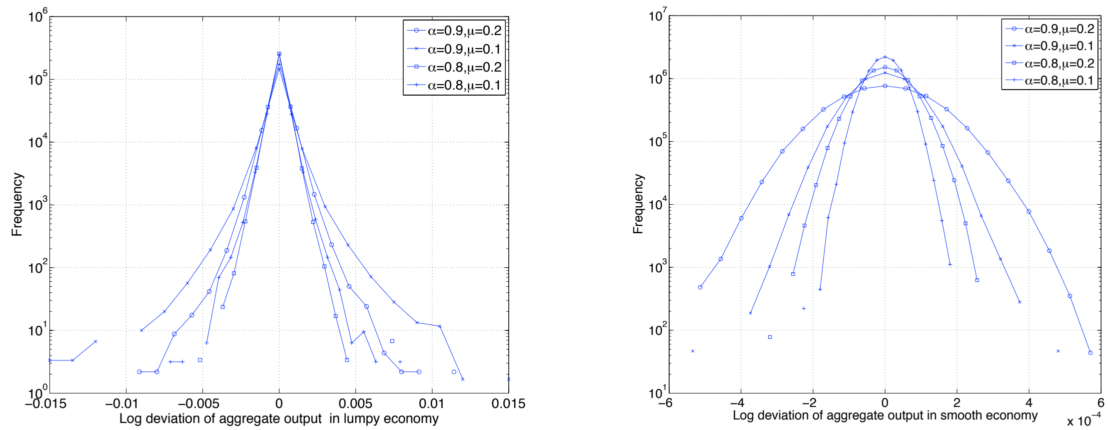

Figure 6: Simulated histograms of the log-deviations of aggregate output when adjustments are lumpy (left) and when adjustments are smooth (right)

of the fraction of investing firms, although we observe a fatter tail than the exponential for the case $\alpha = 0.9$ and $\mu = 0.2$. We can analytically show that the inheritance of the exponential decay from the fraction of investing firms to the aggregate output is exact if the input weights and the lumpiness are homogeneous.

In contrast to the exponential pattern we observe so far, the aggregate fluctuation follows a normal distribution if there is no discreteness constraint (2). The right panel of Figure 6 shows that the log deviation of aggregate output Y when the same realizations of the random variables occur as in the lumpy economy we have observed so far. In the semi-log plot, a normal distribution looks as a parabola. The plot clearly shows the parabola for the case of continuous adjustments.

We can derive the normal distribution in the smooth economy analytically when λ_i and $\chi_{i,j}$ are common across i. By taking the first order condition with respect to n_i and combining with the equilibrium conditions, we obtain the equilibrium aggregate output as:

$$Y = \left(\left(\alpha \left(1 - \frac{1}{\sigma} \right) \right)^{\left(1-\frac{1}{\sigma}\right)\frac{\alpha}{1-\alpha(1-1/\sigma)}} \sum_i \left(A_i^{1-\frac{1}{\sigma}} \chi_i^{\frac{1}{\sigma}} \right)^{\frac{1}{1-\alpha(1-1/\sigma)}} \right)^{\frac{\sigma(1-\alpha(1-1/\sigma))}{(\sigma-1)(1-\alpha(1-1/\sigma))-\alpha}}. \tag{21}$$

As we can see above, the equilibrium output Y is a weighted sum of idiosyncratic shocks A_i, and follows a normal distribution as N tends to infinity.

Another salient difference between the distributions in Figure 6 is the magnitude of fluctuations. The standard deviation in the smooth economy is one-digit smaller than that in the lumpy economy even though they incur the same realization of random shocks. Thus, the micro-level lumpy adjustments have a strong effect of magnifying exogenous independent shocks. It is not that the smooth economy is completely lacking the magnifying mechanism. Consider a simple case where the micro decision is positively dependent on the aggregate such as $x_i = \phi \sum_{j=1}^{N} x_j/N + \epsilon$. Then the exogenous shock will be magnified by the multiplier $1/(1-\phi)$. In the case of lumpy adjustments, however, the multiplier effect at the micro-level is non-linear because of the threshold rule. Thus, the aggregation mechanism works as if the multiplier itself is stochastic, because the multiplier depends on the profile of underlying heterogeneous productivity before the perturbation. This is just like the small deviations in the standing points of domino tiles greatly affect the resulting number of falling tiles. The exponential tail we have observed in the aggregates captures the nature of this stochastic multiplier effect, and thus it is a symptom of the endogenous fluctuation mechanism.

5 Consistency with micro-level estimate

In this section, we test the model in the previous section by examining whether the strategic complementarity observed in the micro data is strong enough to generate the exponential distribution observed in the aggregate data. In the model, the magnitude of fluctuations is largely determined by β, the degree of strategic complementarity, which is determined by σ and α. The σ was calibrated to a reasonable markup rate 10%, but α was taken as a free parameter. In the simulations, α was set so that β was at around 0.67. In this section, we provide an empirical estimate that justifies this choice of β.

We start by directly estimating the individual firm's decision rule on lumpy investment. Each firm faces a binary choice $d(i,t) \in \{0,1\}$ whether or not to engage in a lumpy investment. We would like to estimate the binary decision rule in the model (16) by a probit model. According to the optimal decision rule, the lumpy investment is more likely to occur when the perturbation in productivity $\Delta \log A_i$ and/or the perturbed derived demand ΔY is large. (Note that, in (16), n_i is the equilibrium input before the perturbation, and thus predetermined by the levels of A_i and Y.) By using (17), ΔY is rewritten as $\beta \sum_l b_l X_l N_l$ where l is an industry, b_l is b_j for any firm j in industry l (by abuse of notation), X_l is the fraction of firms in industry l that engage in lumpy investments after the perturbation, and N_l is the number of firms in industry l. As seen in the definition (14), b_l is the weight that is determined by the importance of industry l as an input good and the relative size of the industry l.

Our goal is to estimate the strategic complementarity in lumpy investments within a region, namely, the impact of a lumpy investment on the likelihood of lumpy investment of other firms within the same region. Thus, we are most concerned with β, the impact of the fraction of firms that engage in lumpy investment in G_{r_i} on i's likelihood for investment spiking. A simple regression of $d_{i,t}$ on $X_{r_i,t}$ would pick up a common shock effect that affects all the firms in a region-year. Thus, we include a region-year dummy to control for the common shock. Because $X_{r_i,t}$ is constructed by the left-hand-side variable $d(i,t)$ in the same region, however, a direct estimation would still cause an endogeneity bias in the estimate of β. We avoid this issue by the structure of the model in which the impact derived demand is heterogeneous across industries. Namely, we will identify the endogenous effects within region by using the industrial dimension as an instrument. This method follows Shea [23] and Bartelsman et al. [2]. Our estimator satisfies the conditions in Brock and Durlauf [4] for the identification of this type of estimator.

We follow a two-step procedure to construct a proxy variable that captures the portion of the derived demand in a region-year that is affected by industrial exogenous shocks. First, we regress $X_{l_i,r_i,t}$ on a constant and year-dummies, and construct $\tilde{X}_{l_i,r_i,t}$ by subtracting the year effect (captured by the year-dummy) from $X_{l_i,r_i,t}$. Repeat the same procedure for the industry-year fraction $X_{l_i,t}$ and construct $\tilde{X}_{l_i,t}$ that is the industry-year fraction less common year effects. Thus, $\tilde{X}_{l_i,t}$ captures the industry-wide common shock on lumpy investments, which we deem exogenous to the firm in industries other than l_i. We regress $\tilde{X}_{l_i,r_i,t}$ on $\tilde{X}_{l_i,t}$ and obtain linear predictors. The predictors $\hat{X}_{l_i,r_i,t}$ are then used as proxies for the exogenous portion of the variation in $\tilde{X}_{l_i,r_i,t}$.

Next, we normalize the input-output matrix Π so that the sum for each row is equal to 1. We call it Π^{OW} which stands for "output-weighted" following Bartelsman et al. [2]. $\Pi^{OW}(j,k)$ represents a "demand-pull" or "customer" impact from k-th column to j-th row. This is seen as follows. Suppose that there is one percent increase in the production in industry k. Now $\Pi^{OW}(j,k)$ represents the fraction of industry j's output demanded by k. Thus the demand for j is increased by $\Pi^{OW}(j,k)$, other things being equal. Then we define $X_{i,t}^{OW} \equiv \sum_{k\neq l_i} \Pi^{OW}(l_i,k)\hat{X}_{k,r_i,t}\#G_{k,r_i}/\#G_{r_i}$. $X_{i,t}^{OW}$ represents the total exogenous demand shock coming from all the firms in the same region but in the different industries. Likewise, we construct an input-weighted matrix Π^{IW}, which represents a "supply-push"

	Output-weighted		Input-weighted		Hybrid	
	$d\Phi/dx$	(s.e.)	$d\Phi/dx$	(s.e.)	$d\Phi/dx$	(s.e.)
ΔTFP	0.0021***	(0.0006)	0.0021***	(0.0006)	0.0021***	(0.0006)
Lagged Cash Flow	-3.3e-07	(2.1e-06)	-3.3e-07	(2.1e-06)	-3.3e-07	(2.1e-06)
Investment in (l_i, r_i)	0.0023*	(0.0012)	0.0020*	(0.0012)	0.0022*	(0.0012)
X^{OW}	1.42***	(0.37)			1.38***	(0.38)
X^{IW}			0.79	(0.63)	0.41	(0.64)
Loglikelihood	-87376.19		-87382.72		-87375.99	

Table 4: Probit estimation of firm's lumpy investment

or "supplier" impact from j-th row to k-th column, by normalizing Π column-wise. Using Π^{IW} as a weight, we obtain $X_{i,t}^{IW}$.

Then, we obtain a probit estimation equation as follows:

$$\Pr(d(i,t) = 1) = \Phi(D_{i,t}\gamma_D + Z_{i,t}\gamma_Z + \beta^{OW} X_{i,t}^{OW} + \beta^{IW} X_{i,t}^{IW}). \tag{22}$$

D's are dummy variables for industry, region, year, and region-year pair. $Z_{i,t}$ includes the measure of productivity growth for firm i, the cash flow (normalized by capital) in the previous period, and the aggregate investment in the industry-region which firm i belongs to, $I_{l_i,r_i,t}$. The construction of these variables in $Z_{i,t}$ is explained in Appendix B. The region-year dummy is included to address the region-year common shocks. $X_{i,t}^{OW}$ and/or $X_{i,t}^{IW}$ are output- or input- weighted averages of the exogenous industrial shocks. $X_{i,t}^{OW}$ and $X_{i,t}^{IW}$ are the exogenous portion of the impact of the derived demand that arises from lumpy investment of other firms in the same region.

The estimation result is summarized in Table 4. The effect of total factor productivity is significantly positive, which is consistent with the decision rule (16). We observe a positive and significant estimate of β^{OW}, whereas the estimate of β^{IW} is nonsignificant. This result is consistent with the "within" estimate of Bartelsman et al. [2]. In the short-run, the inter-industrial effect seems to flow from the downstream to upstream. The aggregate investment in the industry-region has a weakly significant mild effect. The lagged cash flow does not show a significant effect on the lumpy investments.

Now we examine if the strategic complementarity estimated above can explain the exponential distribution estimated in Section 2. The mean of the exponential distribution of $|\tilde{X}|$ is estimated at

0.03 for the region-wide fluctuations in Table 2. The median number of firms in G_r is 381 by Table 1. Hence, the median *number* of affected firms is about 12. By solving $1/12 = \phi - 1 - \log \phi$, we indirectly obtain the estimate of ϕ as $\hat{\phi} = 0.65$. This number, estimated from aggregated data, falls within the 90% confidence interval of ϕ, $(0.64, 2.12)$, which is directly estimated from the micro-level investment decisions in Table 4. Thus, our model of endogenous fluctuations have passed a test of internal consistency with the data we investigated.

This is not to say that our estimation of the strategic complementarity has explained all of the exponential fluctuation we observed. First, the connection between the complementarity and the exponential slope was analytically established only for the case when the input weight $\chi_{i,j}$ are common across i. The actual input-output matrix shows a great deal of heterogeneity across rows, and hence it may alter the relationship between ϕ and the slope in the model. The importance of the heterogeneity of the input weights in propagation mechanism has been emphasized by Horvath [17]. Secondly, the confidence interval of ϕ is rather wide, and it will pass any exponential distribution that has a sufficiently flat slope. The meaningful content of the internal test we passed lies on the lower bound of the estimated ϕ: the observed exponential distribution has a standard deviation that is larger than what the minimum estimate of the strategic complementarity at the micro-level implies.

6 Conclusion

This paper argues that the distribution of the fraction of firms that engage in large investments provides a useful test for the existence of endogenous effects among the firms' investment decisions. Testing of the endogenous effects is often a challenge because it is observationally equivalent to the model with exogenous common shocks. We argue that, in a binary choice model, the common shock model results in the normal distribution of the number of investing firms when the exogenous common shock consists of many independent factors, whereas the endogenous effects leads to a non-normal distribution that is better characterized by an exponential distribution.

We investigate the panel of investment data for Italian firms. We construct the fraction of firms that engage in the investment with which the investment-capital ratio of the firm exceeds certain threshold. The fraction samples are constructed for region-year groups and industry-region-year groups. Those samples constitute a distribution of the fraction of investing firms for each definition of the reference group. The results show that the normal distribution hypothesis is rejected by the empirical distributions, whereas the exponential distribution hypothesis is consistent with the empirical distributions.

We present a simple model of lumpy investments when the firms' investment decisions are interrelated by strategic complementarity. The complementarity stems from the fact that firms are linked by input-output relations. When a firm increases production, it generates an increased factor demand, and thus provides an incentive for the firms in the upstream of the input-output relations to produce more. This model generates a non-normal distribution with exponential tail for the fraction of investing firms, while it generates a normal distribution if there is no lumpy investments in the model.

Finally we estimate the degree of strategic complementarity by a micro-level estimation of the firm's decision of the lumpy investment. We utilize the input-output relations to instrument the endogeneity in estimation. Then, we compute the degree of strategic complementarity that is required by the model so that the model explains the empirical magnitude of fluctuations in the fractions of investing firms. We confirmed that the result is consistent with the model prediction.

Appendix

A Proofs

A.1 Fictitious tatonnement

The tatonnement that converges to an equilibrium is a time series of the input profile $n_u = (n_{1,u}, n_{2,u}, \dots, n_{N,u})$ for $u = 1, 2, \cdots$. It starts from the initial equilibrium that is defined for a realization of $(A_{i,0}, \lambda_i)$. The initial point is denoted by n_0. There is a corresponding threshold rule at the initial equilibrium. Its lower bound profile is denoted by n_0^*. The gap variable is accordingly defined as $(s_{i,0})$.

Now a perturbation ϵ_i is drawn independently across i. The productivity is perturbed such that $\log A_{i,1} = \log A_{i,0} + \epsilon$. Then the threshold is altered even if no firms change their input level at n_0. The new lower threshold under (A_1, n_0) is denoted by n_1^*. The gap variable changes to s_1. We apply the threshold rule, and some firms find it optimal to adjust their input. The best reply is denoted by n_1.

Then the second round starts. Under (A_1, n_1), we compute a new profile of lower threshold n_2^*, a gap variable s_2, and a best reply n_2. We iterate this process until n_u converges. The stopping time is denoted by T. Thus, $n_{i,T} = n_{i,T-1}$ for all i.

A.2 Impact of a lumpy investment on total demand

We start from:

$$Y_0 = \left(\sum_j \chi_j^{\frac{1}{\sigma}} A_j^{1-\frac{1}{\sigma}} n_{j,0}^{\alpha(1-1/\sigma)} \right)^{\frac{\sigma}{(\sigma-1)}} . \tag{23}$$

Suppose that firm i chooses to invest. Then $\log n_{i,1} - \log n_{i,0} = \log \lambda_i$. We calculate the new aggregate output by the Taylor series expansion around Y_0 as follows:

$$\log Y_1 - \log Y_0 = \alpha \frac{\chi_i^{\frac{1}{\sigma}} A_i^{1-\frac{1}{\sigma}} n_{i,0}^{\alpha(1-1/\sigma)}}{\sum_j \chi_j^{\frac{1}{\sigma}} A_j^{1-\frac{1}{\sigma}} n_{j,0}^{\alpha(1-1/\sigma)}} \sum_{k=1}^{\infty} \frac{(\alpha(1-1/\sigma))^{k-1} (\log \lambda_i)^k}{k!} + O(N^{-2}) \tag{24}$$

$$= \alpha \frac{\chi_i^{\frac{1}{\sigma}} A_i^{1-\frac{1}{\sigma}} n_{i,0}^{\alpha(1-1/\sigma)}}{\sum_j \chi_j^{\frac{1}{\sigma}} A_j^{1-\frac{1}{\sigma}} n_{j,0}^{\alpha(1-1/\sigma)}} \frac{\lambda_i^{\alpha(1-1/\sigma)} - 1}{\alpha(1-1/\sigma)} + O(N^{-2}). \tag{25}$$

Note that the first term is of order N^{-1}, and the higher order term results from the derivative of the summation in the denominator of the first term with respect to n_i. From this we obtain the expression (17) and (12).

A.3 Proof of Proposition 1

In the homogeneous set-up, the probability of the Bernoulli trial $q\beta b_i / \log \lambda_j$ becomes $q\beta/N$. Thus, the number of firms m_u that are induced to adjust upward in u by m_{u-1} firms who adjusted in the previous step of the tatonnement $u-1$ follows a Poisson distribution with mean $q\beta m_{u-1}$ asymptotically as $N \to \infty$.

Since a Poisson distribution is infinitely divisible, the Poisson variable with mean ϕm_{u-1} is equivalent to a m_{u-1}-times convolution of a Poisson variable with mean ϕ. Thus the process m_u for $u \geq 2$ conditional to m_{u-1} is a branching process with a step random variable being a Poisson with mean ϕ for $u \geq 3$, and m_2 follows a Poisson with mean ϕm_1. Since $\phi \leq 1$, the process m_u reaches 0 by a finite stopping time with probability one (see [14]). Thus the best response dynamics is a valid algorithm of equilibrium selection in a sense that the convergence is achieved by a finite stopping time T. By using the property of the Poisson branching process [18], we obtain an infinitely divisible distribution called Borel-Tanner distribution for the accumulated sum M conditional to m_2 as:

$$\Pr(M = m \mid m_2) = (m_2/m) e^{-\phi m} (\phi m)^{m-m_2} / (m - m_2)! \tag{26}$$

for $m = m_2, m_2+1, \ldots$. Using that m_2 follows the Poisson distribution with mean ϕm_1 (note that m_1 is not necessarily an integer), we obtain (19) in the Proposition as follows:

$$\begin{aligned}
\Pr(M = m \mid m_1) &= \sum_{m_2=0}^{m} ((m_2/m)e^{-\phi m}(\phi m)^{m-m_2}/(m-m_2)!)e^{-\phi m_1}(\phi m_1)^{m_2}/m_2! \\
&= (\phi m_1 e^{-\phi(m+m_1)}/m) \sum_{m_2=1}^{m} (\phi m)^{m-m_2}(\phi m_1)^{m_2-1}/((m-m_2)!(m_2-1)!) \\
&= (\phi m_1 e^{-\phi(m+m_1)}/m)(\phi m + \phi m_1)^{m-1}/(m-1)! \\
&= m_1 e^{-\phi(m+m_1)}\phi^m (m+m_1)^{m-1}/m!. \qquad (27)
\end{aligned}$$

Approximation (20) is obtained by applying the Stirling's formula $m! \sim \sqrt{2\pi}e^{-m}m^{m+0.5}$:

$$\begin{aligned}
& m_1 e^{-\phi(m+m_1)}\phi^m (m+m_1)^{m-1}/m! \\
\sim\;& m_1 e^{-\phi m_1}(e^{-\phi}\phi)^m (m+m_1)^{m-1}/(\sqrt{2\pi}e^{-m}m^{m+0.5}) \\
=\;& (m_1 e^{-\phi m_1}/\sqrt{2\pi})(e^{1-\phi}\phi)^m m^{-1.5}(1+m_1/m)^m(1+m_1/m)^{-1} \\
\sim\;& (m_1 e^{(1-\phi)m_1}/\sqrt{2\pi})(e^{\phi-1}/\phi)^{-m}m^{-1.5}. \qquad (28)
\end{aligned}$$

This completes the proof. □

B Construction of the productivity measure and aggregate investments

To estimate the total factor productivity A_i, we follow the procedure similar to Cooper and Haltiwanger [11] and Bayer [3]. We start with the static optimization problem of a firm under a Cobb-Douglas production function:

$$Y = AL^{\alpha_L}K^{\alpha_K}. \qquad (29)$$

The first order condition for the wage-taking firm implies $wL = \alpha_L Y$. Then we obtain:

$$Y = \left[A(\frac{\alpha_L}{w})^{\alpha_L}\right]^{\frac{1}{1-\alpha_L}} K^{\frac{\alpha_K}{1-\alpha_L}}. \qquad (30)$$

Using this, we estimate the productivity of capital as:

$$\log(Y) - \frac{\alpha_K}{1-\alpha_L}\log(K), \qquad (31)$$

where Y is the value added.

The expenditure shares for both labor and capital are heterogeneous across firms, and it is impossible to directly estimate α_L and α_K due to the dynamic structure of the panel data as Bayer [3] pointed out. We instead calculate the average expenditure shares of firms for each industry and use them for α_L and α_K above.

We also use an aggregate investment variable in order to capture its direct effect on $d(i,t)$ in the regression analysis. The aggregate investment is defined as:

$$I(G_{l_i,r_i},t) = \frac{\sum_{i \in G_{l_i,r_i}} I_{i,t}}{\sum_{i \in G_{l_i,r_i}} \sum_{t=1}^{T} I_{i,t}/(\#G_{l_i,r_i}T)}, \tag{32}$$

where $I_{i,t}$ represents the real investment of firm i in year t and T is the total years in observation. We take logarithm of $I(G_{l_i,r_i},t)$ and then subtract its yearly trend. We use the residual, i.e. the log-deviation of the aggregate investment in the industry-region, as a regressor in $Z_{i,t}$.

References

[1] Abhijit V. Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107:797–817, 1992.

[2] Eric J. Bartelsman, Ricardo J. Caballero, and Richard K. Lyons. Customer- and supplier-driven externalities. *American Economic Review*, 84:1075–1084, 1994.

[3] Christian Bayer. Investment dynamics with fixed capital adjustment cost and capital market imperfections. *Journal of Monetary Economics*, 53:1909–1947, 2006.

[4] William A. Brock and Steven N. Durlauf. Discrete choice with social interactions. *Review of Economic Studies*, 68:235–260, 2001.

[5] Ricardo J. Caballero and Eduardo M. R. A. Engel. Dynamic (S,s) economies. *Econometrica*, 59:1659–1686, 1991.

[6] Ricardo J. Caballero and Eduardo M. R. A. Engel. Explaining investment dynamics in U.S. manufacturing: A generalized (S,s) approach. *Econometrica*, 67:783–826, 1999.

[7] Andrew Caplin and John Leahy. Aggregation and optimization with state-dependent pricing. *Econometrica*, 65:601–625, 1997.

[8] Andrew S. Caplin and Daniel F. Spulber. Menu cost and the neutrality of money. *Quarterly Journal of Economics*, 102:703–726, 1987.

[9] Russell Cooper. Equilibrium selection in imperfectly competitive economies with multiple equilibria. *Economic Journal*, 104:1106–1122, 1994.

[10] Russell Cooper, John Haltiwanger, and Laura Power. Machine replacement and the business cycles: Lumps and bumps. *American Economic Review*, 89:921–946, 1999.

[11] Russell W. Cooper and John C. Haltiwanger. On the nature of capital adjustment costs. *Review of Economic Studies*, 73:611–633, 2006.

[12] Mark Doms and Timothy Dunne. Capital adjustment patterns in manufacturing plants. *Review of Economic Dynamics*, 1:409–429, 1998.

[13] Bill Dupor. Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics*, 43:391–409, 1999.

[14] William Feller. *An Introduction to Probability Theory and Its Applications*, volume I. Wiley, NY, second edition, 1957.

[15] Edward L. Glaeser, Bruce I. Sacerdote, and José A. Scheinkman. The social multiplier. *Journal of the European Economic Association*, 1:345–353, 2003.

[16] Luigi Guiso and Fabiano Schivardi. Spillovers in industrial districts. *Economic Journal*, 117:68–93, 2007.

[17] Michael Horvath. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45:69–106, 2000.

[18] J. F. C. Kingman. *Poisson Processes*. Oxford, NY, 1993.

[19] John B. Long, Jr and Charles I. Plosser. Real business cycles. *Journal of Political Economy*, 91:39–69, 1983.

[20] Robert E. Lucas, Jr. *Studies in business-cycle theory*. MIT Press, 1981.

[21] Charles F. Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60:531–542, 1993.

[22] Makoto Nirei. Threshold behavior and aggregate fluctuation. *Journal of Economic Theory*, 127:309–322, 2006.

[23] John Shea. The input-output approach to instrument selection. *Journal of Business and Economic Statistics*, 11:145–155, 1993.

[24] John Shea. Complementarities and comovements. *Journal of Money, Credit and Banking*, 34:412–433, 2002.

[25] Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19:305–321, 1990.