Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 08-2022

Production function estimation controlling for endogenous productivity disruptions

Plutarchos Sakellaris and Dimitris Zaverdas

March 2022

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

Production function estimation controlling for endogenous productivity disruptions

Plutarchos Sakellaris* and Dimitris Zaverdas*,†

Athens University of Economics and Business

This version: March 27, 2022

Abstract

We modify the standard production function estimation framework to incorporate endogenous disruptions in the productivity process due to lumpy firm investment. We investigate the differences between the existing (baseline) approach and our own (disruption) on a large proprietary panel of Greek Manufacturing firms. We find significantly different production function estimates

*We thank the AUEB Laboratory of Econometrics and the Laboratory of Economic Policy Studies (EMOP) for their hospitality, the Hellenic Statistical Authority (ELSTAT) for providing us with 2-digit NACE rev. 2 Producer Price Indices for Greece, Prof. Vadym Volosovych for providing us with the NACE rev. 1 to (unique) NACE rev. 2 correspondence table used in Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych, and Yesiltas (2015), and Dr. Dimitris Smyrnakis for illuminating us on the topic of threshold model regressions.

†I acknowledge financial support received for this research, in the form of a scholarship by the Academy of Athens.

and different results for subsequent inference. The implied average levels of productivity and magnitude of endogenous disruptions are different. The baseline model cannot capture the full dynamics of disruption costs, which extend across time. The decomposition of Aggregate Productivity Growth is substantially different under the disruption model.

Keywords: Production function estimation, Investment spike, Productivity, Matching

JEL codes: D22, D24, E22, L60

1 Introduction

Studies on a variety of economic topics, such as productivity[1] and markups[2], as well as applications that use structural models of firms, all rely on the accurate estimation of production functions. The current standard for production function estimation is to assume that productivity is completely exogenous. However, the literature studying lumpy adjustment has shown that investment spikes are associated with disruptions in productivity. In this paper, we propose a modification to the standard production function estimation framework that incorporates endogenous disruptions in the productivity process due to lumpy firm investment. We investigate the differences between our novel (disruption) framework and the existing (baseline) approach on a large proprietary panel of Greek Manufacturing firms.

[1] See for example Allcott, Collard-Wexler, and O'Connell (2016), Bennett, Stulz, and Wang (2020), Brandt, Van Biesebroeck, and Zhang (2012), Cirera, Lederman, Máñez, Rochina, and Sanchis (2015), De Loecker (2011), Du and Wang (2020), Padilla-Pérez and Villarreal (2017), Pavcnik (2002), and Taştan and Gönel (2020).

[2] See for example Caselli, Schiavo, and Nesta (2018), De Loecker (2011), and Du and Wang (2020).

We find that estimates of production function coefficients differ significantly. The disruption model estimates the capital coefficient to be higher, the labour coefficient lower and the returns to scale lower. This, in turn, affects the results of subsequent inference. In particular, the two models imply different average levels of productivity. The disruption model estimates log TFP to be lower by 0.087 on average in the sample. The measured magnitude of endogenous disruptions is also different. A sample matching experiment shows that firms that choose to adjust via a spike suffer a decrease in TFP by 5.46 percentage points according to the baseline model, whereas for the disruption model, the decrease is even greater at 6.31 percentage points. In addition, the baseline model cannot capture the full pattern of disruption costs. In contrast, the disruption model allows for a persistent effect of the disruption on productivity that extends across time.

We also explore the different implications of the baseline and the disruption model contrasting a sub-sample of investment spike episodes with a matched sample that did not display lumpy investment. The sample matching is performed following one of the methods in Rosenbaum and Rubin (1985). We identify a spike observation as the case where the investment rate is above a large multiple of the expected investment rate conditional on current capital and a certain minimum threshold. The conditional expectation is defined as the fitted values of the regression of the investment rate on a polynomial of log capital and a set of control dummies. Both models produce evidence of disruptions, in the year after the spike (t=1). However, the disruption model gives an average drop in TFP that is 23.2% larger, than that of the baseline model. Two years on (at t=2), both models show that the average TFP of spike firms starts to recover towards that of the non-spike firms.

In addition, we are the first to extract the individual components of TFP, as they are produced by the estimation procedure, and repeat the above analysis per component. We find that disruptions affect the firm almost equally through both a part of TFP that is observable by the firm prior to production and another that is not. The observable part is solely responsible for the differences between the models (since the unobservable component is identical to both models). In the matching experiment, the baseline models shows a drop in observable productivity by 1.67 percentage points, while the disruption model shows a 2.53 point drop.

The discrepancies between models have an important effect when analyzing the decomposition of Aggregate Productivity Growth (APG). We construct APG for the inference sample and decompose it to its underlying components, following the methodology of Petrin and Levinsohn (2012). Our sample period extends from 2001 to 2017. This is a turbulent period for Greek manufacturing. It encompasses three distinct subperiods in terms of economic activity. The first subperiod is between 2001 and 2007, which is the pre-financial crisis period when the sector was booming. The second subperiod is from 2008 to 2015, the crisis period. Finally, from 2016 to 2017 the sector was under recovery.

Both models find technical efficiency as the more significant factor of APG, with reallocation efficiency being generally of the same sign but smaller. There is, however, the exception of the financial crisis period, where reallocation efficiency operates in a positive direction (demonstrating the cleansing effect of the crisis), but is overcome by the larger negative effect of technical efficiency. Although the qualitative conclusions are similar for both models, there is significant disagreement about the magnitudes of the components.

In the pre-crisis period the disruption model estimates average technical efficiency to be in absolute terms 10.11% higher and reallocation efficiency 38.82% lower, compared to the baseline model. During the crisis, the disruption model is more conservative, giving a 4.36% lower magnitude for average technical efficiency and 10% lower reallocation efficiency. During recovery, the differences in contributions are more extreme, with the disruption model giving a 17.17% lower technical efficiency and 232.67% higher reallocation efficiency contributions than the baseline model.

The two gross-output models we consider are based on the production model of Ackerberg, Caves, and Frazer (2015) (henceforth ACF). The use of gross-output in this paradigm may seem erroneous at first, as Bond and Söderbom (2005) have shown that such a combination does not produce uniquely identifiable production functions, in the presence of fully flexible inputs (i.e. the intermediate inputs present in gross-output specifications). This has lead many authors to estimate value-added functions, instead, whose implicit assumptions, however, have received their own skepticism[3]. In response to all this, we follow a new method developed by Gandhi, Navarro, and Rivers (2020), which enables the estimation of gross-output production functions in the ACF setting. In particular, we use a simplified version found in Collard-Wexler and De Loecker (2020).

Evidence of a spike induced disruption effect, has existed in the literature for decades now. Sakellaris (2004) finds that investment spikes on a sample of US Manufacturing plants are followed by TFP drops. Cooper and Haltiwanger (2006) estimate various models of capital adjustment on US Manufacturing plants and find that a model with mixed convex and non-convex adjustment costs, the latter being triggered by investment

[3] See Bruno (1978), Diewert (1978), and Basu and Fernald (1995).

spikes, to be the best fit for their data. More recently, Gradzewicz (2020) finds similar disruption patterns in a large panel of the Polish economy.

There exist other studies that introduce other endogenous variables in the productivity process[4], but they are only preoccupied with measuring the effects this has on productivity and not on the production function itself. This is much in the same way that the lumpy investment literature has been studying the effects of investment spikes on productivity, but not the implications its findings have for the underlying assumptions of the estimated production function, in the first place. We are the first to put the spotlight on this issue and align the production function model assumptions with the observed phenomena in the data. Our more realistic assumption of partly endogenous productivity, founded on the observation of productivity disruptions in our dataset, should allow for more accurate production function estimates and more reliability on results that rely on them.

The rest of this paper proceeds as follows. Section 2 presents the dataset we employ, discusses some variable definitions, and presents evidence of the disruption effect in our sample. Section 3 presents a disruption model, where the productivity process is affected by investment spikes. Section 4 discusses the differences in estimates and conclusions between this and the baseline model, and presents the empirical application on Aggregate Productivity Growth. Section 5 concludes. The appendices provide further useful details.

2 Data and Variable Definitions

We employ a large proprietary unbalanced dataset from the ICAP database on Greek firm annual financial statements and employment data from 1998 to 2017. After the

[4] See Doraszelski and Jaumandreu (2013), Cirera et al. (2015), and Khan and Khederlarian (2021).

basic dropping of duplicates and observations without financial data, the ICAP dataset contains about 750,000 unique firm-year observations for about 100,000 Greek firms. However, we restrict to a sample of roughly 125,000 observations for the Manufacturing sector, covering about 15,000 firms. We focus on Manufacturing because it reports the most reliable data and because there is better availability for the necessary deflators and other sub-sector specific variables from other sources.

Our key variables are output, the production inputs, and investment in physical capital. For output, we use the book values of net sales, other operating income, and the change in inventories of finished goods, which we deflate by a production price index provided to us by ELSTAT[5]. For intermediate inputs, we use the book values of the cost of goods sold and other operating expenses minus their depreciation and the cost of labour. Then, we deflate by the intermediate inputs price index reported by EUKLEMS. For investment and capital, we follow the perpetual inventory method (PIM) per asset class, and then add together the corresponding series for Buildings and Machinery & Equipment to get the aggregate measures.

More details on variable construction and data cleaning are given in Appendix B. After these tasks, we are left with a final sample of 47586 observations for 8087 Manufacturing firms. We will be referring to this as the inference sample, as it comprises of all of the observations that are involved in the various inference steps to follow.

Table 1 reports some summary statistics for output, capital, labour, and intermediate inputs in logs and for the investment rate, in the inference sample. All variables appear to be mostly symmetrically distributed without very fat tails, with the notable exception of the investment rate. Table 2 reports the cross correlations and autocorrelations of the

[5] In a few cases where that is not available, we use the Harmonized Consumer Price index for the whole economy, reported by Eurostat.

Table 1: Summary Statistics

	Output	Capital	Labour	Int. Inputs	Inv. Rate
Min	0.2838	-0.2007	0.0000	5.8815	-0.9670
Median	14.4992	13.5680	3.0828	14.0644	0.0616
Mean	14.5350	13.4053	3.0830	14.0550	0.1597
Max	19.2223	18.3132	6.9020	19.2564	5.3232
S.D.	1.3704	1.7479	1.1599	1.5428	0.4550
Skewness	-0.1178	-0.6963	-0.0771	-0.2280	4.5584
Kurtosis	4.0299	4.9691	3.0265	3.5800	35.7753

$N = 47586$

Output, capital, labour, and intermediate inputs are in logs. The investment rate is defined as total real investment in physical capital over total real physical capital.

same variables. The production function variables seem to display a significant positive relationship and strong autocorrelations. In particular, the large positive autocorrelation of log labour allows us to use its first lag as its instrument when appropriate, which is needed for the purposes of the estimation of the model presented in Section 3. Again, the investment rate is the odd one out, being generally independent of the rest of the variables and not particularly autocorrelated.

In table A.1 of Appendix A, we report the per year average growth rates of output and the production inputs, and the Aggregate Productivity Growth (APG), for the inference sample, the latter of which we calculate in Section 4.3 based on the methodology of Petrin and Levinsohn (2012). We observe that for the period between 2001 and 2007 the average growth of output is 1.64% and is mostly driven by the positive growth of its inputs. This

Table 2: Correlations and Autocorrelations

	Output	Capital	Labour	Int. Inputs	Inv. Rate
Output	1.0000				
Capital	0.6816	1.0000			
Labour	0.7427	0.5948	1.0000		
Int. Inputs	0.9469	0.6388	0.6361	1.0000	
Inv. Rate	0.0647	-0.0947	0.0310	0.0618	1.0000
Autocorrelation	0.9611	0.9835	0.8850	0.9317	0.0805

$N = 47586$

Output, capital, labour, and intermediate inputs are in logs. The investment rate is defined as total real investment in physical capital over total real physical capital.

is despite of average APG being at -0.72% for the same period. From 2008 to 2015, during the 2008 Greek financial crisis, the average output growth is -8.96%, owning to the decrease, and even reversal, of the growth of its inputs, while average APG falls even further to -1.60% in the same period.

2.1 Investment Spike Definition

The statistics in table 1 imply that the empirical distribution of the investment rate in our sample is not symmetric. Table 3 gives further insights on this by reporting the number of observations in the sample for which the first lag of the investment rate falls within certain intervals of interest. We report for the first lag because that is what is used at the estimation stage.

Out of the 47586 observations, only 7337 report disinvestment and 5123 are within the interval of -0.01 and 0.01, which is associated with investment inactivity in the literature.

Table 3: Investment Rate Statistics

Number of observations with:	
$I_{it-1}/K_{it-1} \leq -0.01$	7337
$\lvert I_{it-1}/K_{it-1} \rvert < 0.01$	5123
$I_{it-1}/K_{it-1} \geq 0.01$	35126
$I_{it-1}/K_{it-1} > 0.20$	12346

$N = 47586$

The rest report positive investment. This is evidence of the presence of non-convex adjustment costs in capital and justifies the need to investigate the effect of investment spikes on the firm. We see that 12346 observations lie above the 0.20 threshold for the investment rate, which is a typical lower bound used to characterize investment spikes in the literature.

We define investment spikes similarly to Gradzewicz (2020). We use a combination of the definitions in Power (1998) and Sakellaris (2004). Let K_{it} be the level of physical capital of firm i in period t, I_{it} its level of investment, and d^I_{it} be the spike dummy, then we define

$$d^I_{it} = \begin{cases} 1, & \dfrac{I_{it}}{K_{it}} > \max\left\{\alpha E\left(\dfrac{I_{it}}{K_{it}}\middle| K_{it}\right), \gamma\right\} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

with $\alpha = 2.75$ and $\gamma = 0.20$. $E\left(I_{it}/K_{it}\right| K_{it})$ is the expected investment rate for firm i in period t conditional on its level of capital. We construct it as the fitted value of the ordinary least squares regression of I_{it}/K_{it} on a third degree polynomial of the natural logarithm of K_{it} and a complete set of time dummies. To account for sector heterogeneity and the possibility of a structural break caused by the 2008 Greek financial crisis, we perform a different regression per 2-digit NACE rev. 2 sector and before and after 2008. Finally, if the criterion is satisfied for consecutive years for the same firm, we only identify the first instance as a spike.

This definition results in 4270 spike events being identified in the inference sample, which is about a third of the original 12346 candidates. Furthermore, table A.2 of Appendix A reports the average investment rate and percentage of identified investment spikes in the inference sample per year. We observe that average investment generally recedes during the 2008 financial crisis relative to the pre-crisis period. Because we control for the crisis, it does not affect the proportion of observations we identify as investment spikes each year, always being roughly between 5% - 10% of all contemporaneous observations in the sample.

2.2 Sample Matching Method

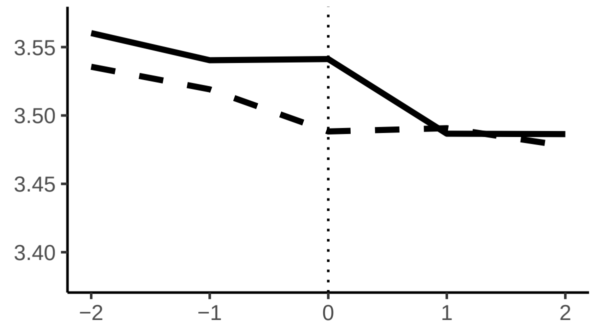
To study the effect investment spikes have on productivity, we define the sub-sample of spike observations, as the treatment group, and an equally sized matched sub-sample of non-spike firms with similar characteristics, as the control group. We follow a variant of the Mahalanobis distance within calipers approach described in Rosenbaum and Rubin (1985). Because we need our observations to have available data on productivity for five consecutive years, we defer the construction of the spike and matched sub-samples until after the models have been estimated (see Section 4) and the desired measures of productivity have been calculated. We, then, restrict to the observations with the data availability we require.

The matching process works with a set of approximate matching variables and a set of exact matching variables. We select firm age, log real sales, real sales growth (as the first difference in log sales), and lagged leverage (as total liabilities over total assets) as our approximate matching variables. We use the observation's year, the firm's legal form, its 2-digit NACE rev. 2 sector, and its trading status (public or private) for exact matching. We then regress a logit model with the spike dummy as the explained variable and the approximate matching variables, a set of time dummies, and a set of 2-digit sector dummies as the explanatory variables. The fitted log odds ratio is called the propensity score and is used to determine the closeness of the observations between samples.

Both the sample of spikes and candidate matches are sorted randomly. For each observation in the spike sample, we find its match by restricting to the candidates that exactly match the exact matching variables and are within a 0.5 range above or below of the propensity score of the spike. Then, we do not just match by the closest propensity score, but by the Mahalanobis distance between the spike and all of the candidate matches. This distance is calculated according to the approximate matching variables and the propensity score. The observation with the minimum Mahalanobis distance is designated a match and it is removed from the sample of candidates (matching without replacement). Once every spike has its match, we drop any pairs with a Mahalanobis distance greater than 3. The two final sub-samples cover 644 spike episodes.

Figure 1 illustrates the average performance of the TFP of a firm that undergoes an investment spike, for a time window of two periods before to two periods after a spike event, compared to its statistical match that does not. The measure of TFP used is derived based on the estimates of a baseline gross-output production function estimated according to GNR. Here, we have a superficial look at the figure and go into greater detail in Section 4.2. One can see that on average productivity drops by 5.46 percentage points, following an investment spike. Compared to the overall variability of the series, this drop is too large to be dismissed as a random occurrence. In addition, no similar

Figure 1: Baseline $log(TFP)$ averages around investment spikes



$N = 644$

The solid lines report for the spike sub-sample and the dashed lines report for the matched sub-sample.

behaviour is observed in the matched sample. This is evidence of endogenous productivity disruptions caused by lumpy adjustment. However, this phenomenon is not reflected in the standard assumptions of the ACF/GNR model. Thus, we find it necessary to align it with the evidence and develop a new version which allows for investment spikes to affect productivity, which is what we do in Section 3.

3 The Disruption Model

A lot of research that uses production functions focuses on value-added[6]. However, there has been skepticism regarding its validity and the validity of the necessary conditions for it to give equivalent results to gross-output[7]. There is also the structural real value-added specification, used for example in Ackerberg et al. (2015), where the production function is assumed to be separable in the primary and the intermediate inputs by a Leontief function. The practical difference between (deflated) value-added and structural

[6] See for example Bennett et al. (2020), Brandt et al. (2012), Ding, Kim, and Zhang (2018), Du and Wang (2020), Padilla-Pérez and Villarreal (2017), and Taştan and Gönel (2020).

[7] See Bruno (1978) and Diewert (1978).

value-added comes down to if and how to deduct the value of the intermediate inputs from output before regressing. In both cases, one estimates a model of some measure of output on the primary inputs alone.

Given that intermediate inputs are generally freely adjustable and the non-identifiability critique of Codd-Douglas functions with flexible inputs by Bond and Söderbom (2005), it would seem that value-added is the only option. Even with the added assumptions it requires. However, the new approach by GNR allows for the estimation of gross output production functions with fully flexible inputs. So, we here present a disruption model, which builds upon the baseline model of ACF and GNR, but includes investment spike decisions as a determinant of productivity. Its estimation is practically identical to that of the baseline. The model and its assumptions are as follows.

Consider a set of firms, with a gross-output Cobb-Douglas production function in logs

$$y_{it} = \beta_K k_{it} + \beta_L l_{it} + \beta_M m_{it} + \omega_{it} + \varepsilon_{it} \tag{2}$$

where the i subscript indicates the firm in question and t the time period. y_{it} is the natural logarithm of the firm's gross output and k_{it}, l_{it}, and m_{it} are the logs of capital, labour, and other intermediate inputs (raw materials, energy, etc), respectively. β_K, β_L, and β_M are the respective Cobb-Douglas parameters.

The production process is subject to two types of shocks, ω_{it} and ε_{it}, which combined give its total factor productivity. Both variables are unobservable to the econometrician. However, ω_{it} is observed by the firm at the start of the period, so we may also refer to it as the observable shock. ε_{it} is not observed by the firm until the end of the period, thus we may also refer to it as the unobservable shock. It may incorporate highly unpredictable factors that affect output or even the measurement errors of some variables, all of which may or may not display some serial correlation.

Assumption 1. *The current and all past levels of observable productivity, $\omega_{it-s} \forall s \in \{0,1,2,\dots\}$, are included in $\mathcal{I}_{it}$, but any future values, $\omega_{it+s} \forall s \in \{1,2,\dots\}$, are not. For the unobservable shock it holds that $E(\varepsilon_{it}|\mathcal{I}_{it}) = 0$.*

Assumption 2. *The observable shock, ω_{it}, follows a first order Markovian process*

$$\omega_{it} = g(\omega_{it-1}, d^I_{it-1}) + \xi_{it} \tag{3}$$

where d^I_{it-1} is the first lag of the disruption dummy defined in Section 2.1 and ξ_{it} is such that $E(\xi_{it}|\mathcal{I}_{it-1}) = 0$.

The choice of d^I_{it-1} instead of d^I_{it}, follows from our findings in Section 4 that disruption takes effect one period after the investment spike.

This is our point of differentiation to the baseline approach, which uses a restricted version of (3), where g is only a function of ω_{it-1}. We will be estimating both models in order to juxtapose them in Section 4.

For our particular needs, we parameterize g to be linear and have ω_{it} following an AR(1) process, in order to get model estimates for its long run average and its persistence. Thus, (3) becomes

$$\omega_{it} = c + \rho\omega_{it-1} + \delta d^I_{it-1} + \xi_{it} \tag{4}$$

and the baseline version becomes

$$\omega_{it} = c + \rho\omega_{it-1} + \xi_{it} \tag{5}$$

with $0 < \rho < 1$.

Assumption 3. *Capital is a fixed input and its level, k_{it}, is determined by the previous period's investment, at time $t-1$, and some law of motion. Labour is quasi-fixed and its level, l_{it}, is chosen at time $t-b$, with $0 < b < 1$*[8].

By assumptions 1 and 3, l_{it} is correlated with ω_{it} and thus ξ_{it}, whereas k_{it} is not. However, the relative rigidity of labour means that it should display some positive correlation with its first lag, which is independent of ξ_{it} and can be used as its instrument. Both k_{it} and l_{it} are uncorrelated with ε_{it}. From all this, we can get the following unconditional moment conditions

$$E(k_{it}\xi_{it}) = 0 \tag{6}$$

$$E(l_{it-1}\xi_{it}) = 0 \tag{7}$$

$$E(k_{it}\varepsilon_{it}) = 0 \tag{8}$$

$$E(l_{it}\varepsilon_{it}) = 0 \tag{9}$$

Assumption 4. *The other intermediate inputs are fully flexible and their level, m_{it}, is determined at the start of period t. There exists an intermediate input demand function*

$$m_{it} = f(k_{it}, l_{it}, \omega_{it}) \tag{10}$$

This means that m_{it} is chosen freely and optimally, up to the unobservable shock, ε_{it}.

Additionally, assumptions 1 and 4 imply that

$$E(m_{it}\varepsilon_{it}) = 0 \tag{11}$$

Assumption 5. *f is strictly increasing and invertible in ω_{it}.*

[8] Depending on the particularities of the data, the model, and the general context, the researcher may want to set b equal to exactly zero or exactly one. This can be accommodated by the GNR method with the necessary modifications to the estimation process, as described in their appendix.

The invertibility of f with respect to ω_{it} implies that there exists a function, f^{-1}, such that

$$\omega_{it} = f^{-1}(k_{it}, l_{it}, m_{it}) \tag{12}$$

The main argument of GNR, which we also follow, is that the flexibility of m_{it} permits the use of revenue share regressions to identify β_M. Using an estimate of β_M, we can define a measure of output net of the effect of intermediate inputs

$$\widetilde{y}_{it} = y_{it} - \widehat{\beta_M} m_{it} = \beta_K k_{it} + \beta_L l_{it} + \omega_{it} + \varepsilon_{it} \tag{13}$$

Given that the rest of the inputs are not fully flexible, one can use a revenue share regression for the flexible inputs in conjunction with the ACF framework for $\widetilde{y}_{it}$ and equation (13). Thus, one can retrieve estimates for every parameter in (2) and circumvent the unidentifiability issues underlined by Bond and Söderbom (2005).

The ACF two-stage procedure substitutes (12) in (13) to get

$$\widetilde{y}_{it} = \beta_K k_{it} + \beta_L l_{it} + f^{-1}(k_{it}, l_{it}, m_{it}) + \varepsilon_{it} = \Phi(k_{it}, l_{it}, m_{it}) + \varepsilon_{it} \tag{14}$$

where Φ is an unknown function of observables. At the first stage, Φ is approximated by least squares on a polynomial of the observables. At the second stage, the fitted values, $\widehat{\Phi}_{it}$, are used with candidate estimates of (β_K, β_L) to get estimates for ω with which to estimate (or approximate) g. The residuals, $\widehat{\xi}_{it}$, are combined with a set of instruments to give moment conditions for which $(\widehat{\beta_K}, \widehat{\beta_L})$ is chosen optimally.

We use the revenue share of intermediate inputs and the moment conditions derived above, following Collard-Wexler and De Loecker (2020), to estimate the production function in (2). However, we are not preoccupied with accounting for measurement errors, as they do, to avoid the systematic loss of a considerable number of observations. We employ a linear approximation for Φ and the linear functional form we have assumed for g. As a side-effect, we can also get estimates for the parameters in (4) and (5), as well as approximate the unobservable (to the econometrician) variables ω_{it} and ε_{it}. The estimation essentially follows the procedure described in the appendix of Collard-Wexler and De Loecker (2020). Nevertheless, we provide detailed estimation steps in Appendix C. LP, ACF, and GNR also provide their own detailed appendices.

4 Estimation Results

4.1 Model Estimates

We proceed by estimating the production function for the baseline and the disruption model described in Section 3. Table 4 reports the estimated parameters for both models. The standard errors for the production function parameters are calculated by 1000 bootstrap samples, following Levinsohn and Petrin (2003)[9]. The standard errors for the parameters in the productivity process, g, are heteroskedasticity and autocorrelation robust errors.

We first examine the estimates for the production function. Given how relatively small the estimated standard errors are, we can confidently say that all coefficients are statistically significant for both models. We also observe that the estimates for some coefficients change significantly. The coefficient for capital increases from 0.0759 to 0.0922, which using the standard deviation of the baseline model is 1.4554 standard deviations higher. The coefficient for labour decreases from 0.1888 to 0.1473, which is 1.4613 baseline

[9] According to Hahn and Liao (2021), standard error estimates produced by bootstrap can be more conservative (i.e. greater) than the true standard errors. This is not a problem for us, however, as it only means that our parameter estimates are more accurate than we report. For a more detailed analysis on the bootstrap, also see Horowitz (2001).

Table 4: Production Function and Productivity Process Estimates

	Production Function			Productivity Process		
	β_K	β_L	β_M	c	ρ	δ
Baseline	0.0759	0.1888	0.6730	0.5165	0.8516	
	(0.0112)	(0.0284)	(0.0015)	(0.0150)	(0.0043)	
Disruption	0.0922	0.1473	0.6730	0.5199	0.8474	-0.0217
	(0.0167)	(0.0427)	(0.0015)	(0.0154)	(0.0045)	(0.0011)

$N = 47586$

Values in parentheses report standard errors. For the Production Function, they are calculated using 1000 bootstrap samples. For the Productivity Process, they are heteroskedasticity and autocorrelation robust errors.

standard deviations lower. That these two coefficients adjust between models in the opposite direction is not surprising. According to Collard-Wexler and De Loecker (2020), because capital and labour are positively correlated, an upwards correction to the capital coefficient leads to a downwards adaptation of the coefficient for labour and the overall effect on returns to scale is unknown. In this case, returns to scale decrease, form 0.9377 to 0.9125. Moreover, the recent findings by Hahn and Liao (2021) suggest that second moment estimates using the bootstrap are biased upwards. So, the standard deviations for the production function coefficients may be smaller than estimated and the models are more precise and more statistically distinct than they appear to be. Of course, the estimate for the intermediate inputs coefficient remains unchanged, since it is calculated before the inclusion of the spike dummy takes place. The important conclusion is that including the spike dummy definitely has an effect on the production function estimates.

We now have a look at the estimates for the productivity process. Again, the relatively small standard errors imply that all coefficients are statistically significant. In both models, productivity is strongly persistent, with an autocorrelation coefficient of 0.8516 and 0.8474 for the baseline and the disruption model, respectively. Both models have a positive and statistically significant constant, implying a positive long-run average for productivity. In the disruption model, we can see that the investment spike dummy has a negative and statistically significant coefficient of -0.0217. This suggests a short-term negative effect of investment spikes on output of 2.17 percentage points. Furthermore, the persistent nature of the productivity process means that investment spikes also have their mark in later periods, and productivity needs more time to recover. We return to this idea in the next subsection.

Finally, although the constant and persistence parameters between both models are statistically very close, the implied average productivity is different. The implied average for the baseline model is given by

$$\bar{\omega} = \frac{c}{1-\rho} \tag{15}$$

which gives an estimate of 3.4812. For the disruption model, the calculation becomes

$$\bar{\omega} = \frac{c + \delta P(d^I_{it-1} = 1)}{1-\rho} \tag{16}$$

In the inference sample, 8.97% of observations report an investment spike (in first lags). Using this percentage as the probability of an investment spike, the estimated $\bar{\omega}$ for the disruption model is 3.3942. So, between models, ω_{it} is on average calculated to be smaller by 0.0870 (or 2.5% in terms of the baseline estimate). Since the average value of the unobservable shock is zero for both models, the difference in averages for log TFP is also 0.0870. The baseline model generally predicts a TFP for this sample that is biased upwards.

4.2 Productivity Component Performance

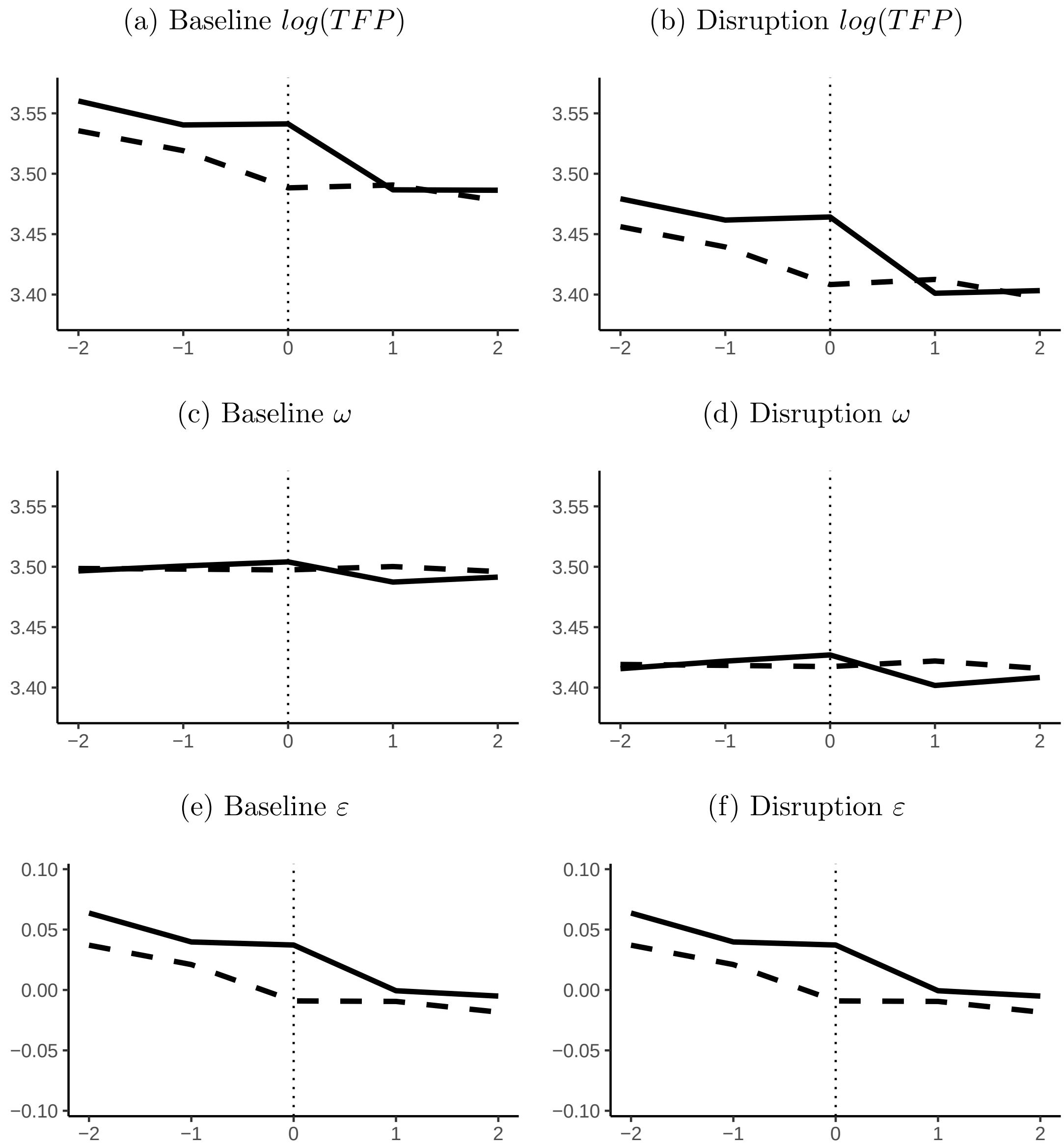
We now study the effects of investment spikes on log TFP and its components, ω_{it} and ε_{it}. We retrieve the three variables for each model as follows: for $log(TFP_{it})$ we use the estimated production function parameters with the production function in logs to get $log(TFP_{it}) = y_{it} - \widehat{\beta_K} k_{it} - \widehat{\beta_L} l_{it} - \widehat{\beta_M} m_{it}$, for ω_{it} we use the values produced in the second stage of the estimation, and we get ε_{it} as the difference between $log(TFP_{it})$ and ω_{it}[10].

In the sample used for the estimation, the disruption model estimates a smaller contribution of TFP to output, compared to the baseline model. Based on the data, both the average logarithm of TFP and the average of ω_{it} of the baseline model are 3.4760 (compared to the parameter based estimate of 3.4812, calculated in Section 4.1). For the disruption model, they are both at 3.3857 (compared to 3.3942 in Section 4.1), which is less by 0.0903 (or 2.6%, compared to 0.0870 and 2.5% in Section 4.1). These data based averages and their difference are relatively close to the corresponding parameter based estimates of Section 4.1. This is providing comfort that both the baseline and the disruption model can adequately fit the data. For ε_{it} in both models the average is zero, which is to be expected since it is equivalent to the OLS residual from the first stage of the estimation.

We select a sub-sample of 644 cases, in our sample, that experience an investment spike and have available data for a time window of two periods before and after the spike. Following the methodology described in Section 2.2, we select an equally sized sub-sample of cases of firms that are statistically similar, but do not exhibit a spike, and have the

[10] Equivalently, we could use the residuals from the first stage of the estimation. We have found that the two alternatives only differ at the very low digits, which could be attributed to representational rounding errors.

Figure 2: Productivity component averages around investment spikes

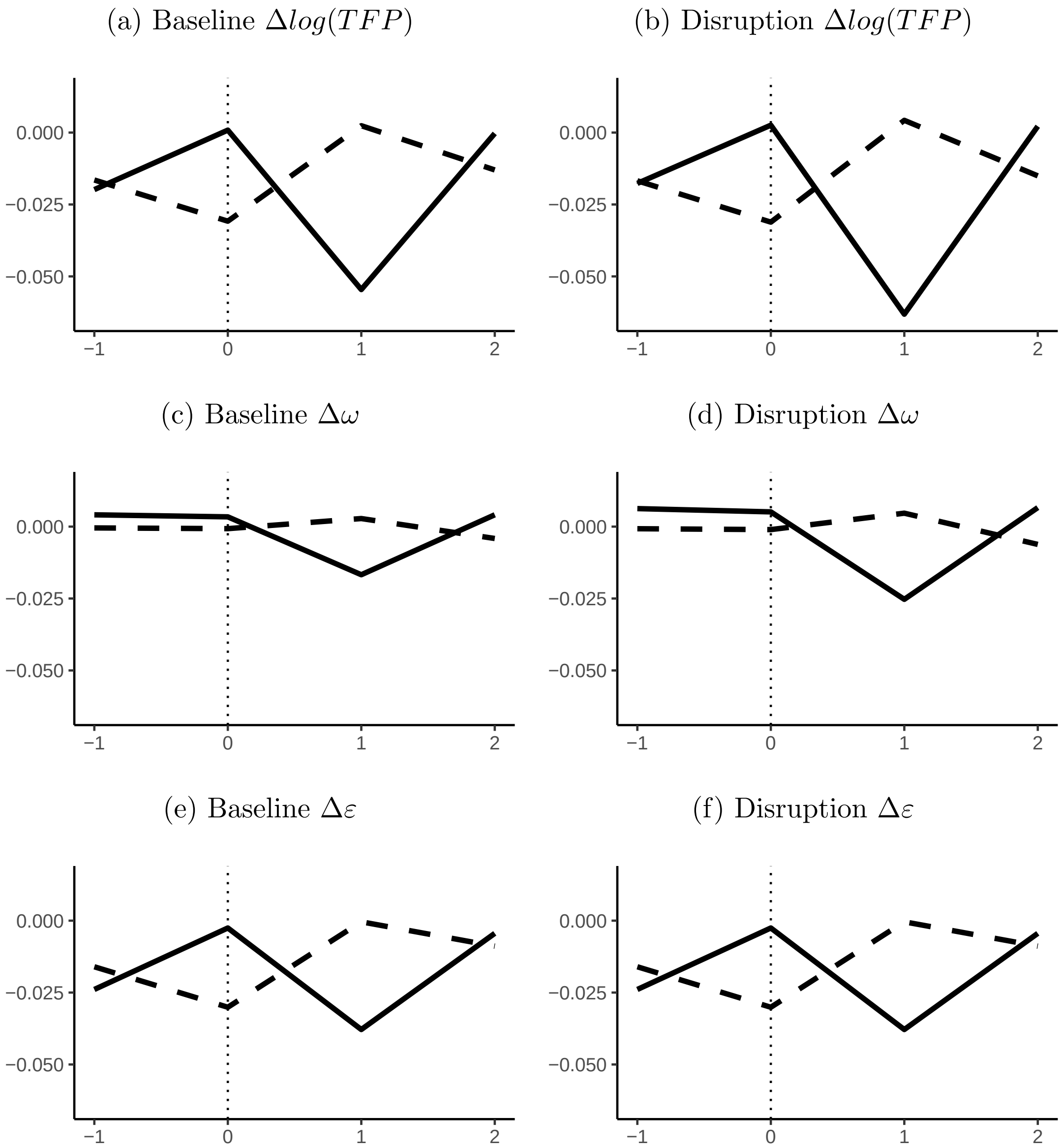


$N = 644$

The solid lines report for the spike sub-sample and the dashed lines report for the matched sub-sample.

same time window of available data. We draw a series of plots for these sub-samples, regarding the performance of TFP and its components around investment spike episodes. All investment spike episodes are grouped together, regardless of the time period they happened, thus mitigating any time effects.

Figure 3: Productivity component first difference averages around investment spikes



$N = 644$

The solid lines report for the spike sub-sample and the dashed lines report for the matched sub-sample.

Figure 2, plots the average values of $log(TFP_{it})$, ω_{it}, and ε_{it} separately for the spike and the matched sub-samples, for a time window of two periods before to two periods after an investment spike. Figure 3 does the same for the corresponding first differences, which helps to better illustrate the presence of productivity disruptions following investment spikes, as well as any differences between the models. Observe that, also in these subsamples, the averages for both $log(TFP_{it})$ and ω_{it} are at visibly lower levels for the disruption model compared to the baseline model, same as it is found by the parameter based and the data based estimates above. That is not the case for ε_{it}, however, which is identical for both models. This makes sense, since the residuals of the first stage, with which ε_{it} is equivalent, are calculated before any assumption on the process of ω_{it} is involved.

In figure 2, for both models (subfigures (a) and (b)), the averages of log TFP and those of its components, for both spike and non-spike firms, start out relatively close two periods before the spike (at $t = -2$). Their difference is around 2 to 3 percentage points[11], in both cases. They also move mostly in parallel towards the period before the spike ($t = -1$), maintaining that difference. This is also evident by the corresponding subfigures of figure 3, where the starting first differences essentially coincide, being at around the 1.6 to 2 percentage points region in all cases. This reassures that the matching process did a good job at finding appropriate statistical twins of the spike firms, even though productivity and its components are not targeted variables of the matching algorithm.

During the spike period (at $t = 0$), in both models, the average spike firm enjoys a relatively higher TFP than its non-spike counterpart. For the baseline model the average log TFP of a spike firm is 3.5412 compared to 3.4883 for the non-spike firms, increasing the gap between them to about 5 percentage points. For the disruption model, the difference is slightly greater at about 5.5 percentage points (average log TFP is 3.4642 for the spike firms and 3.4083 for the non-spike firms). The same picture can also be seen in their growth rates from before the spike to the spike period. In the baseline model,

[11] A difference of 0.01 in log TFP approximately implies a difference in TFP in levels of 1 percentage point.

non-spike TFP drops by 3 percentage points, whereas spike TFP stays essentially the same. Roughly the same happens in the disruption model.

The first period after the spike (at $t = 1$), we can see a stark difference in productivity growth rates, between spike and non-spike firms. In the baseline model, the average TFP of the spike firms drops by 5.46 percentage points, compared to about zero for the non-spike firms, closing the gap between them. In the disruption model, the difference is more pronounced. Spike firms fall by 6.31 percentage points, while non-spike firms increase by 0.43. Using the baseline model as the reference point, in this limited sample of 644 cases, it can be calculated that the disruption model predicts a productivity differential that is 23.22% greater. So, it seems that both models show signs of a disruption effect, the baseline model underestimates its magnitude, since it is trying to fit the data in a specification without disruption.

Such productivity drops, which are correlated with lumpy adjustment behaviour, constitute an extra adjustment cost that the firm has to face for choosing spiky adjustment, compared to smoother adjustment, on top of any usual convex costs. These costs may appear because significant adjustment may require production units to halt operations while the new capital is installed, or may be the result of the production workers slowly adapting to the new capital they work with (learning-by-doing). As we will see below, such productivity disruption have a lingering effect on the firm's future output, due to the persistent nature of ω_{it}.

Two periods after the spike (at $t = 2$), both types of firms appear to co-move again, in both models. However, the first difference plots in figure 3 reveal that the spike firms have a slightly higher growth rate, possibly while they recover from last period's disruption shock. All of the above observations are in line with the general findings of the lumpy adjustment literature.

We now move on to dissecting TFP into its individual components, the observable ω_{it} component and the unobservable ε_{it} component (subfigures (c), (d), (e), and (f), for each figure, respectively). Firstly, we observe in both figure 2 and 3 that ω_{it} is much less variable for both models, compared to ε_{it}, which appears to drive a lot of the variability in TFP. Before the spike (at $t = -2$ and $t = -1$), ω_{it} and ε_{it} are relatively close between sub-samples for both models, same as with TFP.

In both models, the relatively increased productivity of spike firms at the spike period ($t = 0$) is almost exclusively due to the unobservable shock, ε_{it}. Since the spike and the non-spike firms are otherwise similar, one interpretation of this is that the relative productivity advantage of the spike firms is a random occurrence, which cannot be captured by the models. This boosted productivity helps counteract the disruption costs that an investment spike would cause, and incentivizes the firm to go through with it. Alternatively, this productivity spread could also be a sign of the spike firms proactively intensifying the usage of physical capital to get the most out of it before it is replaced and to help build a buffer that will protect them from the ensuing productivity disruptions. This may be evidence of variable input utilization. However, the models used here could not account for it, since the dataset we employ provides no such useful information.

In the first period after the spike ($t = 1$), the situation is different. Now, both components respond to the spike with a significant drop. The observable ω_{it} drops by 1.67 percentage points for the baseline model and 2.53 for the disruption model, 0.86 percentage points more. The unobservable ε_{it} drops by 3.79 in both models. So, the disparity in the decrease between models is entirely due to differences in the observable shock. At $t = 2$, ω_{it} appears to be the main contributor towards the recovery of TFP we have observed.

Even though they are useful in providing with some insight into the performance of productivity and its components, the spike and the matched non-spike sub-samples are rather limited in size, and thus any measurements of the size of disruption may be significantly inaccurate. For the remainder of this subsection, we will be using the estimate of the δ parameter of the disruption model, which is based on the entire inference sample, as the true size of disruption that acts through ω_{it}.

We are interested in examining whether our estimate of the disruption effect on the observable shock, $\widehat{\delta}$, is backed by the data. We do this via a paired Wilcoxon signed rank test[12]. We take 2452 cases for which ω_{it} is available during and after an investment spike and consider two samples, one containing the observations during the spike and one for the next period. We use the ω_{it} that is produced by the disruption model.

We conduct a two-sided Wilcoxon test, where the null hypothesis is that the location shift between the two samples is equal to $-\widehat{\delta}$. Our estimate for δ is -0.0217, implying a location shift of 0.0217. The test's p-value is 0.0197, which rejects the null at the 5% confidence level. Interestingly, however, the reported 95% confidence interval is [0.0220, 0.0249], which has two implications. Firstly, we see that our estimate is rejected by a small margin. Secondly, the confidence interval excludes zero, which reinforces the evidence for the presence of a disruption effect in the examined sample. We also conduct a one-sided test, where the null is that the shift is at least equal to $-\widehat{\delta}$ or greater. Unsurprisingly, the test strongly accepts the null at the 5% confidence level with a p-value of 0.9901.

[12] Because the exact Wilcoxon test can be computationally demanding even for relatively small samples, we use a continuity corrected normal approximation.

Furthermore, both tests report a pseudomedian for the location shift of 0.0234, which is the tests' estimate of the shift. So, our estimate is relatively more conservative about the size of the disruption effect, compared to what the paired signed rank test finds. In practice, they both roughly agree and, at the very least, the disruption effect is at least as prominent as we estimate it to be and should not be ignored when conducting any kind of inference.

Finally, using model parameter estimates and sample statistics, we can get an idea of the overall effect a single investment spike has on output across time. As we have discovered, disruption acts through two channels, by affecting both the observable and the unobservable shock. Due to the persistence of ω_{it}, the overall effect of an investment spike on it from the time disruptions take effect and across all future periods can be calculated as

$$\delta + \rho\delta + \rho^2\delta + \rho^3\delta + \cdots = \frac{\delta}{1-\rho} \tag{17}$$

Using the parameter estimates from the disruption model in (17) we get -0.1421. By this estimate alone, we see that, ceteris paribus, the firm losses 14.21% of its future output just through the observable shock channel, when it chooses lumpy adjustment as opposed to smooth adjustment. Notice here how the baseline model provides no structural way of measuring this long-term effect of lumpy investment. We have seen that ε_{it} also drops in response to investment spikes. In the inference sample, the average drop is 0.0288, or 2.88%, which we will use in place of a lacking model estimate. Adding everything together, we find that investment spikes result in the overall loss of 17.10% of the firm's output in forgone output, 5.05% of which in the very next year (calculated as 2.17% plus 2.88%) and the remaining 12.05% in the subsequent years. Thus, the extra cost the firm has to bear, when choosing to adjust in spikes rather than smoothly, is not limited to one period.

4.3 Aggregate Productivity Growth

We further illustrate the importance of the difference in estimates between the baseline and the disruption model with an empirical application. We construct Aggregate Productivity Growth (APG) for the inference sample and decompose it to its underlying components, following the methodology of Petrin and Levinsohn (2012). We deviate slightly from their definition, to reflect our gross-output assumption, and calculate the discrete time version of APG in each time period as

$$APG_t = \sum_i \overline{D}_{it}\Delta y_{it} - \sum_i \overline{D}_{it}\overline{s}_{ikt}\Delta k_{it} - \sum_i \overline{D}_{it}\overline{s}_{ilt}\Delta l_{it} - \sum_i \overline{D}_{it}\overline{s}_{imt}\Delta m_{it} \tag{18}$$

where summation is across all firms present in the sample in period t, $\overline{D}_{it}$ is the firm's Domar weight, which we calculate as total revenue over the sum of all revenues across all firms in period t, and $\overline{s}_{ikt}$ is the revenue share of capital and similarly with $\overline{s}_{ilt}$ and $\overline{s}_{imt}$ for labour and intermediate inputs, respectively. The bars over $\overline{D}_{it}$, $\overline{s}_{ikt}$, $\overline{s}_{ilt}$, and $\overline{s}_{imt}$ indicate average values between periods $t-1$ and t. Δy_{it}, Δk_{it}, Δl_{it}, and Δm_{it} are the first differences of log output and the respective log inputs. Before constructing APG and any of its components, we drop observations that lie at the top and bottom 0.5% of the empirical distribution of the sum of revenue shares.

For the revenue shares of the inputs, we need data on their costs. For labour and intermediate inputs, we simply multiply the real variables in levels with the average cost of labour and the intermediate inputs price index, respectively (see Appendix B). Of course, data on capital use costs are not reported in any of the firm's financial statements. Basu and Fernald (1995) and Balk (2021) provide some measures one can construct, but the necessary data for Greece are not easily available. Petrin and Levinsohn (2012), who face the same problem, opt to simply not account for capital in their data and provide an APG measure which is accurate "up to an adjustment for capital expenditures". We prefer to use a simple, but attainable, approximation used in Bitros and Panas (2001) for Greek Manufacturing. We define the cost of capital as

$$r_{it}^{K} = P_{it}^{I}(r_{it} + \delta_{it}) \tag{19}$$

where P_{it}^{I} is the investment deflator, r_{it} is the intertemporal interest rate, and δ_{it} the capital depreciation rate. We have data for the investment deflator and the capital depreciation rates, which we use for our variable construction. We calibrate the intertemporal interest rate at a fixed 0.025 for all periods, taken from table 9 of the online appendix of Fakos, Sakellaris, and Tavares (2021), who study the effects of credit supply shocks on Greek Manufacturing firms' investment before and during the Greek 2008 financial crisis. We calculate the unit cost both for Buildings and Machinery & Equipment and then appropriately calculate the revenue share of aggregate capital.

Notice how the calculation of APG relies only on observable data and does not depend on any estimated parameters. Production function estimates are needed, however, to get a breakdown of the forces acting on APG. Petrin and Levinsohn (2012) split APG into changes in Technical Efficiency (TE), Reallocation Efficiency (RE) for each input, and fixed and sunk costs (F). Thus,

$$APG_t = TE_t + RE_t + F_t \tag{20}$$

Table 5: Aggregate Productivity Growth breakdown according to the each model. Averages per subperiod.

Period	Model	APG	TE	RE	RE_K	RE_L	RE_M	N
2001-2007	Baseline	-0.0072	-0.0057	-0.0015	0.0042	-0.0023	-0.0035	19799
	Disruption	-0.0072	-0.0063	-0.0009	0.0057	-0.0031	-0.0035	
2008-2015	Baseline	-0.0160	-0.0284	0.0124	0.0014	0.0081	0.0028	10295
	Disruption	-0.0160	-0.0272	0.0111	0.0019	0.0065	0.0028	
2016-2017	Baseline	-0.0050	-0.0046	-0.0003	0.0018	0.0002	-0.0023	5345
	Disruption	-0.0050	-0.0038	-0.0011	0.0024	-0.0012	-0.0023	

Averages of the annual values of APG and its breakdown for the specified subperiods (annual values are found in tables A.3 and A.4 in Appendix A).

The estimates for fixed and sunk costs (F) are not reported, because they are zero everywhere.

N is the number of observations.

with

$$TE_t = \sum_i \overline{D}_{it} \Delta ln(TFP_{it}) \tag{21}$$

$$RE_t = \sum_i \overline{D}_{it}(\beta_K - \overline{s}_{ikt})\Delta k_{it} + \sum_i \overline{D}_{it}(\beta_L - \overline{s}_{ilt})\Delta l_{it} + \sum_i \overline{D}_{it}(\beta_M - \overline{s}_{imt})\Delta m_{it} \tag{22}$$

$$F_t = -\sum_i \overline{D}_{it} \Delta ln(F_{it}) \tag{23}$$

and F_{it} being the fixed and sunk costs of firm i at time t. F_t is calculated residually as $F_t = APG_t - TE_t - RE_t$. TFP is calculated as in Section 4.2.

Table 5 shows the evolution of APG and its components, during three time periods in our sample, according to both the baseline and the disruption model. The values reported are the averages of the corresponding annual values, given by the Petrin and Levinsohn (2012) methodology, for each subperiod we define. The per year values are reported in

Appendix A (tables A.3 and A.4). Our first subperiod is between 2001 and 2007, which is the pre-financial crisis period. The second subperiod is from 2008 to 2015, which we consider to be the crisis period. Finally, from 2016 to 2017 is our recovery period. The values of APG are common for both models, since its calculation is model independent. We observe that the average of APG is slightly decreasing in every subperiod, except for the crisis period where the decrease is more prominent at -0.0160, implying an average decrease in aggregate productivity in our sample of 1.60% per year during the crisis.

Moving on to the breakdowns by component, at first glance, it appears that both models mostly agree on the signs of each part of APG. Both models find technical efficiency as the more significant factor of APG, with reallocation efficiency being generally of the same sign but smaller. There is, however, the exception of the financial crisis period, where reallocation efficiency operates in a positive direction (demonstrating the cleansing effect of the crisis), but is overcome by the larger negative effect of technical efficiency.

Upon closer inspection, one can also see that the generally smaller size of reallocation efficiency is misleading. In reality, the reallocation effects from each production input are comparable in size to that of technical efficiency, but usually cancel each other out. Efficiency growth due to capital reallocation is positive throughout time, but the other inputs are ultimately the deciding factor of the effect of total reallocation. Finally, both models report the exact same values for the intermediate inputs reallocation efficiency, because the intermediate inputs production function coefficient estimate is identical between models.

Although the qualitative conclusions are similar for both models, there is significant disagreement about the magnitudes of the components. Table 6 reports the ratio of the

Table 6: Aggregate Productivity Growth breakdown as percentage of annual level and ratios between models. Averages per subperiod.

	Baseline		Disruption		Ratio		N
Period	TE	RE	TE	RE	TE	RE	
2001-2007	1.8322	-0.8322	2.0562	-1.0562	1.1011	0.6118	19799
2008-2015	2.0766	-1.0766	1.9978	-0.9978	0.9564	0.9000	10295
2016-2017	0.9694	0.0306	0.8426	0.1574	0.8283	3.3267	5345

The columns under Baseline and Disruption report the ratios of TE/APG and RE/APG in each time period for the baseline and the disruption model, respectively. The columns under Ratio report the ratio of the disruption estimate over the corresponding baseline estimate.

N is the number of observations.

corresponding values of TE and RE over APG for the baseline and the disruption model, as well as the ratio between the models (as disruption over baseline). Notice that for each grouping the values of TE and RE sum up to one. We can observe that the models assign significantly different contributions of TE and RE to APG in all subperiods. In the pre-crisis period the disruption model estimates average technical efficiency to be in absolute terms 10.11% higher and reallocation efficiency 38.82% lower[13], compared to the baseline model. During the crisis, the disruption model is more conservative, giving a 4.36% lower magnitude for average technical efficiency and 10% lower reallocation efficiency. During recovery, the differences can be more extreme, with a 17.17% lower technical efficiency and 232.67% higher reallocation efficiency (or equivalently, the baseline model

[13] These values are found by subtracting one from the corresponding columns under Ratio, which report the value estimated by the disruption model over the corresponding value of the baseline model.

is 69.95%[14] lower than the disruption model). However, this last large difference could be due to the more limited data for the recovery period. Overall, however, the evidence we have reviewed suggests that the choice of model can paint a substantially different picture of the underpinnings of APG.

5 Summary and Conclusions

In this paper, we study the effects of investment spikes on firm-level productivity on a panel of Greek Manufacturing firms from the ICAP database. Motivated by the evidence from the lumpy investment literature, we argue that production function estimation needs to take into account endogenous productivity disruptions caused by investment spikes. We contribute to the production function estimation literature, by providing a modified structural model and an estimation method similar to the unobservable variable proxy approach that incorporates spike induced disruptions to production.

We find that our modification is meaningful and produces significantly different production function estimates, compared to the standard exogeneity assumption for productivity, on a large proprietary panel of the Greek Manufacturing sector. Furthermore, we break down TFP in an observable (by the firm) part and an unobservable part. We find that the disruptions examined have an effect on output both in the next period and a gradually waning effect in subsequent periods, the latter of which can only be structurally estimated using our model. These constitute an additional adjustment cost the the firm incurs for lumpy adjustment, in contrast to smoother adjustment where only the typical convex costs may apply. Additionally, we calculate Aggregate Productivity Growth and its implied components according to each

[14] This is found by taking the reciprocal of 3.3267 and subtracting one.

model. Again, we find significant differences.

We suggest that other firm-level datasets may display patterns of endogenous productivity disruptions. If ignored, they can skew production function or other estimates and produce erroneous inference results, as we have seen here. They need not be present in every dataset, but researchers ought to not exclude the possibility and try to control for it.

References

Ackerberg, D., Benkard, L. C., Berry, S., & Pakes, A. (2007). Econometric Tools for Analyzing Market Outcomes. *Handbook of econometrics* (pp. 4171–4276). https://doi.org/10.1016/S1573-4412(07)06063-1

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, *83*(6), 2411–2451.

Allcott, H., Collard-Wexler, A., & O'Connell, S. (2016). How Do Electricity Shortages Affect Industry? Evidence from India. *American Economic Review*, *106*(3), 587–624. https://doi.org/10.1257/aer.20140389

Balk, B. M. (2021). *Productivity* (1st ed.). Springer, Cham. https://doi.org/10.1007/978-3-030-75448-8

Basu, S., & Fernald, J. G. (1995). Are apparent productive spillovers a figment of specification error? *Journal of Monetary Economics*, *36*(1), 165–188. https://doi.org/10.1016/0304-3932(95)01208-6

Bennett, B., Stulz, R., & Wang, Z. (2020). Does the stock market make firms more productive? *Journal of Financial Economics*, *136*(2), 281–306. https://doi.org/10.1016/j.jfineco.2019.09.006

Bitros, G. C., & Panas, E. E. (2001). Is there an inflation-productivity trade-off? Some evidence from the manufacturing sector in Greece. *Applied Economics*, *33*(15), 1961–1969. https://doi.org/10.1080/00036840110043730

Bond, S., & Söderbom, M. (2005). Adjustment Costs and the Identification of Cobb Douglas Production Functions. *The Institute for Fiscal Studies*, *Working Paper*(WP05/04).

Brandt, L., Van Biesebroeck, J., & Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics*, *97*(2), 339–351. https://doi.org/10.1016/j.jdeveco.2011.02.002

Bruno, M. (1978). Duality, Intermediate Inputs and Value-Added. In M. Fuss & D. McFadden (Eds.), *Production economics: A dual approach to theory and applications* (pp. 3–16). McMaster University Archive for the History of Economic Thought. https://eml.berkeley.edu/prodecon/apps/ch8.pdf

Caselli, M., Schiavo, S., & Nesta, L. (2018). Markups and markdowns. *Economics Letters*, *173*, 104–107. https://doi.org/10.1016/j.econlet.2018.09.026

Cirera, X., Lederman, D., Máñez, J. A., Rochina, M. E., & Sanchis, J. A. (2015). The Export-productivity Link for Brazilian Manufacturing Firms. *Economics*, *9*(1), 20150022. https://doi.org/10.5018/economics-ejournal.ja.2015-22

Collard-Wexler, A., & De Loecker, J. (2020). Production Function Estimation and Capital Measurement Error. *(under revision)*, *0*(0), 0. https://drive.google.com/file/d/14IPUIxG3kOMVrDuniwUTosZyRXF2qzgT/view

Cooper, R. W., & Haltiwanger, J. C. (2006). On the Nature of Capital Adjustment Costs. *Review of Economic Studies*, *73*(3), 611–633.

De Loecker, J. (2011). Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Production. *Econometrica*, *79*(5), 1407–1451. https://doi.org/10.3982/ECTA7617

de Loecker, J., Eeckhout, J., & Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications. *Quarterly Journal of Economics*, *135*(2), 561–644.

Diewert, W. (1978). Hick's Aggregation Theorem and the Existence of a Real Value Added Function. In M. Fuss & D. McFadden (Eds.), *Production economics: A dual approach to theory and applications*. McMaster University Archive for the History of Economic Thought. https://eml.berkeley.edu/prodecon/apps/ch9.pdf

Ding, S., Kim, M., & Zhang, X. (2018). Do firms care about investment opportunities? Evidence from China. *Journal of Corporate Finance*, *52*, 214–237. https://doi.org/10.1016/j.jcorpfin.2018.07.003

Doraszelski, U., & Jaumandreu, J. (2013). R&D and Productivity: Estimating Endogenous Productivity. *The Review of Economic Studies*, *80*(4 (285)), 1338–1383. http://www.jstor.org/stable/43551561

Du, P., & Wang, S. (2020). The effect of minimum wage on firm markup: Evidence from China. *Economic Modelling*, *86*, 241–250. https://doi.org/10.1016/j.econmod.2019.10.012

Fakos, A., Sakellaris, P., & Tavares, T. (2021). *Investment Slumps During Financial Crises: The Real Effects of Credit Supply* (Working Paper). SSRN. https://dx.doi.org/10.2139/ssrn.3197369

Gandhi, A., Navarro, S., & Rivers, D. A. (2020). On the Identification of Gross Output Production Functions. *Journal of Political Economy*, *128*(8), 2973–3016. https://doi.org/10.1086/707736

Gradzewicz, M. (2020). What Happens After an Investment Spike—Investment Events and Firm Performance. *Journal of Business and Economic Statistics*, *00*(0), 1–16. https://doi.org/10.1080/07350015.2019.1708369

Hahn, J., & Liao, Z. (2021). Bootstrap Standard Error Estimates and Inference. *Econometrica*, *89*(4), 1963–1977. https://doi.org/10.3982/ECTA17912

Horowitz, J. L. (2001). The Bootstrap. In J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics* (1st ed., pp. 3159–3228). Elsevier.

Hu, Y., Huang, G., & Sasaki, Y. (2020). Estimating production functions with robustness against errors in the proxy variables. *Journal of Econometrics*, *215*(2), 375–398. https://doi.org/10.1016/j.jeconom.2019.05.024

İmrohoroğlu, A., & Tüzel, Ş. (2014). Firm-Level Productivity, Risk, and Return. *Management Science*, *60*(8), 1861–2109.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., & Yesiltas, S. (2015). *How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications* (Working Paper No. 21558). National Bureau of Economic Research. https://doi.org/10.3386/w21558

Keller, W., & Yeaple, S. R. (2009). Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States. *The Review of Economics and Statistics*, *91*(4), 821–831.

Khan, S. Y., & Khederlarian, A. (2021). *Inventories, Input Costs, and Productivity Gains from Trade Liberalizations* (Working Paper No. 9564). World Bank. https://openknowledge.worldbank.org/handle/10986/35213

Kim, K. i., Petrin, A., & Song, S. (2016). Estimating production functions with control functions when capital is measured with error. *Journal of Econometrics*, *190*(2), 267–279. https://doi.org/10.1016/j.jeconom.2015.06.016

Levinsohn, J., & Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, *70*(2), 317–342.

Manjón, M., & Mañez, J. (2016). Production function estimation in Stata using the Ackerberg–Caves–Frazer method. *The Stata Journal*, *16*(4), 900–916. https://doi.org/10.1177/1536867X1601600406

Olley, S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*(6), 1263–1297.

Padilla-Pérez, R., & Villarreal, F. G. (2017). Structural change and productivity growth in Mexico, 1990–2014. *Structural Change and Economic Dynamics*, *41*, 53–63. https://doi.org/10.1016/j.strueco.2017.02.002

Pakes, A., & Olley, S. (1995). A limit theorem for a smooth class of semiparametric estimators. *Journal of Econometrics*, *65*(1), 295–332.

Pavcnik, N. (2002). Trade Liberalization, Exit, And Productivity Improvements: Evidence From Chilean Plants. *Review of Economic Studies*, *69*(1), 245–276. https://doi.org/10.1111/1467-937X.00205

Petrin, A., & Levinsohn, J. (2012). Measuring aggregate productivity growth using plant-level data. *RAND Journal of Economics*, *43*(4), 705–725. https://doi.org/10.1111/1756-2171.12005

Power, L. (1998). The Missing Link: Technology, Investment, and Productivity. *The Review of Economics and Statistics*, *80*(2), 300–313. https://doi.org/10.1162/003465398557393

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, *39*(1), 33–38. https://doi.org/10.2307/2683903

Rovigatti, G. (2017). Production Function Estimation in R: The prodest Package. *The Journal of Open Source Software*, *2*(18). https://doi.org/10.21105/joss.00371

Rovigatti, G., & Mollisi, V. (2018). Theory and Practice of TFP Estimation: the Control Function Approach Using Stata. *The Stata Journal*, *18*(3), 618–662. https://doi.org/10.1177/1536867X1801800307

Sakellaris, P. (2004). Patterns of Plant Adjustment. *Journal of Monetary Economics*, *51*(2), 425–450.

Stehrer, R., Bykova, A., Jäger, K., Reiter, O., & Schwarzhappel, M. (2019). *Industry Level Growth and Productivity Data with Special Focus on Intangible Assets: Report on methodologies and data construction for the EU KLEMS Release 2019* (No. 8). The Vienna Institute for International Economic Studies (wiiw).

Taştan, H., & Gönel, F. (2020). ICT labor, software usage, and productivity: firm-level evidence from Turkey. *Journal of Productivity Analysis*, *53*(2), 265–285. https://doi.org/10.1007/s11123-020-00573-x

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*(3), 112–114. https://doi.org/10.1016/j.econlet.2009.04.026

A Tables

Table A.1 reports for the inference sample, per year for the pre-crisis, crisis, and recovery periods: Aggregate Productivity Growth (APG), calculated according to Petrin and Levinsohn (2012), and the average growth rates of the output and input variables (capital, labour, and other intermediate inputs), calculated as the average of their first differences in logs.

Table A.2 reports, per year for the inference sample, the average investment rate (i.e. average I_{it}/K_{it} across i for each t) and the percentage of observations identified as investment spikes (i.e. when $d_{it} = 1$). Observe the noticeable drop in average investment behaviour after the onset of the Greek financial crisis in 2008. Our definition of lumpy investment appears unaffected, however, because it takes the financial regime into account (pre 2008 or not). In fact, one could notice some increase in lumpy behaviour when approaching the recovery period (circa post 2015).

Table A.3 reports Aggregate Productivity Growth (APG) and its components, technical efficiency (TE) and reallocation efficiency (RE), for the baseline model. RE is further broken down into its summands, RE due to capital ($\mathrm{RE_K}$), RE due to labour ($\mathrm{RE_L}$), and RE due to other intermediate inputs ($\mathrm{RE_M}$). Table A.4 reports the same for the disruption model.

Table A.5 reports, per year for the inference sample, technical efficiency and reallocation efficiency as a percentage of APG (TE/APG and RE/APG, respectively) for the baseline and the disruption model, and their ratio between models (as disruption over baseline). Notice how for each grouping TE and RE sum up to one.

Table A.1: Aggregate Productivity Growth and Average Growth Rates

Year	APG	Output	Capital	Labour	Int. Inputs
2001	-0.0362	-0.0005	0.1157	0.1221	0.0389
2002	-0.0206	0.0155	0.0767	0.0845	0.0162
2003	-0.0152	0.0175	0.0814	0.0647	0.0226
2004	-0.0149	0.0295	0.1142	0.0608	0.0459
2005	0.0177	-0.0299	0.0747	-0.0500	-0.0665
2006	-0.0002	0.0325	0.0827	-0.0031	0.0257
2007	0.0190	0.0503	0.0844	0.0252	0.0218
2008	-0.0226	-0.0374	0.0720	-0.0083	0.0223
2009	-0.0236	-0.2191	0.0076	-0.0912	-0.2243
2010	-0.0181	-0.1534	0.0070	-0.0843	-0.1122
2011	-0.0219	-0.1864	-0.0196	-0.1024	-0.1282
2012	-0.0065	-0.1709	-0.0111	-0.1956	-0.1180
2013	-0.0069	-0.0253	0.0189	-0.0158	0.0272
2014	-0.0170	0.0409	0.0122	0.2757	0.0522
2015	-0.0115	0.0349	0.0754	0.0504	0.0480
2016	-0.0035	0.0020	0.0084	-0.0034	0.0413
2017	-0.0064	0.0188	0.0286	0.0568	0.0088
Average	-0.0111	-0.0342	0.0488	0.0109	-0.0164
SD	0.0142	0.0890	0.0440	0.1053	0.0817
Average 2001-2007	-0.0072	0.0164	0.0900	0.0435	0.0149
Average 2008-2015	-0.0160	-0.0896	0.0203	-0.0214	-0.0541
Average 2016-2017	-0.0050	0.0104	0.0185	0.0267	0.0250

Aggregate Productivity Growth is calculated based on Petrin and Levinsohn (2012). The average growth rates for the rest of the variables are calculated as the average of

Table A.2: Average Investment Rate and Proportion of Investment Spikes per year

Year	Investment Rate	Investment Spike Percentage
1999	0.2551	0.0650
2000	0.2518	0.0755
2001	0.1856	0.0773
2002	0.1814	0.0796
2003	0.2326	0.0690
2004	0.1796	0.0886
2005	0.1944	0.0862
2006	0.2003	0.0791
2007	0.1819	0.0681
2008	0.0916	0.0681
2009	0.0868	0.0642
2010	0.0576	0.0673
2011	0.0621	0.0695
2012	0.1096	0.0627
2013	0.0862	0.0780
2014	0.2160	0.0939
2015	0.1074	0.0990
2016	0.1255	0.0796

Table A.3: Aggregate Productivity Growth breakdown according to the Baseline Model

Year	APG	TE	RE	RE_K	RE_L	RE_M
2001	-0.0362	-0.0398	0.0035	0.0060	0.0002	-0.0026
2002	-0.0206	-0.0192	-0.0014	0.0040	-0.0013	-0.0041
2003	-0.0152	-0.0176	0.0023	0.0036	-0.0001	-0.0013
2004	-0.0149	-0.0108	-0.0041	0.0060	-0.0057	-0.0044
2005	0.0177	0.0279	-0.0102	0.0033	-0.0086	-0.0050
2006	-0.0002	-0.0014	0.0011	0.0035	0.0024	-0.0047
2007	0.0190	0.0206	-0.0017	0.0032	-0.0027	-0.0022
2008	-0.0226	-0.0483	0.0257	0.0031	0.0204	0.0023
2009	-0.0236	-0.0256	0.0020	0.0011	-0.0052	0.0061
2010	-0.0181	-0.0373	0.0192	0.0006	0.0156	0.0030
2011	-0.0219	-0.0152	-0.0067	0.0003	-0.0066	-0.0004
2012	-0.0065	-0.0216	0.0151	0.0003	0.0052	0.0096
2013	-0.0069	-0.0230	0.0161	0.0011	0.0124	0.0025
2014	-0.0170	-0.0314	0.0144	0.0011	0.0121	0.0012
2015	-0.0115	-0.0248	0.0133	0.0037	0.0113	-0.0018
2016	-0.0035	-0.0038	0.0003	0.0019	0.0006	-0.0022
2017	-0.0064	-0.0054	-0.0010	0.0017	-0.0003	-0.0025
Average	-0.0111	-0.0163	0.0052	0.0026	0.0029	-0.0004
SD	0.0142	0.0199	0.0101	0.0018	0.0085	0.0040
Average 2001-2007	-0.0072	-0.0057	-0.0015	0.0042	-0.0023	-0.0035
Average 2008-2015	-0.0160	-0.0284	0.0124	0.0014	0.0081	0.0028
Average 2016-2017	-0.0050	-0.0046	-0.0003	0.0018	0.0002	-0.0023

The estimates for fixed and sunk costs (F) are not reported, because they are zero everywhere.

Table A.4: Aggregate Productivity Growth breakdown according to the Disruption Model

Year	APG	TE	RE	RE_K	RE_L	RE_M
2001	-0.0362	-0.0373	0.0011	0.0079	-0.0042	-0.0026
2002	-0.0206	-0.0177	-0.0029	0.0052	-0.0041	-0.0041
2003	-0.0152	-0.0166	0.0013	0.0048	-0.0022	-0.0013
2004	-0.0149	-0.0124	-0.0025	0.0083	-0.0064	-0.0044
2005	0.0177	0.0217	-0.0040	0.0045	-0.0035	-0.0050
2006	-0.0002	-0.0018	0.0016	0.0048	0.0015	-0.0047
2007	0.0190	0.0199	-0.0009	0.0042	-0.0030	-0.0022
2008	-0.0226	-0.0457	0.0231	0.0041	0.0167	0.0023
2009	-0.0236	-0.0296	0.0060	0.0015	-0.0016	0.0061
2010	-0.0181	-0.0343	0.0162	0.0007	0.0125	0.0030
2011	-0.0219	-0.0194	-0.0025	0.0004	-0.0025	-0.0004
2012	-0.0065	-0.0249	0.0184	0.0002	0.0085	0.0096
2013	-0.0069	-0.0207	0.0138	0.0015	0.0097	0.0025
2014	-0.0170	-0.0213	0.0043	0.0014	0.0016	0.0012
2015	-0.0115	-0.0214	0.0099	0.0051	0.0066	-0.0018
2016	-0.0035	-0.0038	0.0003	0.0024	0.0001	-0.0022
2017	-0.0064	-0.0039	-0.0026	0.0024	-0.0024	-0.0025
Average	-0.0111	-0.0158	0.0047	0.0035	0.0016	-0.0004
SD	0.0142	0.0181	0.0085	0.0024	0.0067	0.0040
Average 2001-2007	-0.0072	-0.0063	-0.0009	0.0057	-0.0031	-0.0035
Average 2008-2015	-0.0160	-0.0272	0.0111	0.0019	0.0065	0.0028
Average 2016-2017	-0.0050	-0.0038	-0.0011	0.0024	-0.0012	-0.0023

The estimates for fixed and sunk costs are not reported, because they are zero everywhere.

Table A.5: Aggregate Productivity Growth breakdown as percentage of annual level and ratios between models

	Baseline		Disruption		Ratio	
Year	TE	RE	TE	RE	TE	RE
2001	1.0978	-0.0978	1.0292	-0.0292	0.9375	0.2982
2002	0.9307	0.0693	0.8573	0.1427	0.9212	2.0589
2003	1.1520	-0.1520	1.0870	-0.0870	0.9436	0.5725
2004	0.7250	0.2750	0.8306	0.1694	1.1456	0.6162
2005	1.5790	-0.5790	1.2260	-0.2260	0.7765	0.3904
2006	6.2520	-5.2520	8.3160	-7.3160	1.3301	1.3930
2007	1.0888	-0.0888	1.0471	-0.0471	0.9617	0.5303
2008	2.1360	-1.1360	2.0202	-1.0202	0.9458	0.8980
2009	1.0843	-0.0843	1.2543	-0.2543	1.1567	3.0152
2010	2.0587	-1.0587	1.8917	-0.8917	0.9189	0.8423
2011	0.6933	0.3067	0.8861	0.1139	1.2780	0.3714
2012	3.3189	-2.3189	3.8212	-2.8212	1.1514	1.2166
2013	3.3260	-2.3260	3.0009	-2.0009	0.9022	0.8602
2014	1.8445	-0.8445	1.2506	-0.2506	0.6780	0.2967
2015	2.1513	-1.1513	1.8577	-0.8577	0.8635	0.7450
2016	1.0983	-0.0983	1.0833	-0.0833	0.9864	0.8481
2017	0.8405	0.1595	0.6019	0.3981	0.7161	2.4966
Average	1.8457	-0.8457	1.8859	-0.8859	0.9772	1.0264
SD	1.3976	1.3976	1.8563	1.8563	0.1825	0.7929
Average 2001-2007	1.8322	-0.8322	2.0562	-1.0562	1.1011	0.6118
Average 2008-2015	2.0766	-1.0766	1.9978	-0.9978	0.9564	0.9000
Average 2016-2017	0.9694	0.0306	0.8426	0.1574	0.8283	3.3267

B Variable Construction

We try to recover any missing information in our financial statement data based on accounting identities[15]. We then construct the variables we need as explained below. Finally, we take some additional steps to reach the final dataset we use for inference.

The inference variables are constructed as follows:

Firm Age, as the difference between the current year of each observation and the Year of Establishment reported for each firm.

Physical Capital and Investment Deflators (base year 2010), as the Implicit Price Index for Gross Fixed Capital Formation, provided by the World Bank.

Geometric Depreciation Rates by Asset Class per Year, δ_{jit}, as the weighted average of the geometric depreciation rates for each subclass of capital, provided by the EU KLEMS 2019 survey, weighted by the gross capital stocks reported by Eurostat (nama10 file) every year and for every 2-digit NACE 2 division. If data for a division are not reported, the data for the immediate parent grouping available are used. If some δ_{jit} is undefined due to a zero over zero division, this means that the amount particular asset class for that particular time and sector is zero, and thus δ_{jit} set to zero. Then firms are assigned the proper δ_{jit} according to their NACE 2 classification.

[15] We set to missing any value that is reported as negative where negative values do not apply. The only financial variables that we allow to be negative are: Cash and Deposits, Shareholders' Equity, Retained Earnings/Accumulated Losses, Gross Margin, Operating Margin, Profits Before Taxes, Income Taxes, and EBITDA. To account for rounding errors, we set every instance of the value of negative one to zero.

Deflated Book Value of Capital per Asset Class[16], as

$$K_{Book\ jit} = \frac{Book_{jit} - AccumulatedDepreciation_{jit}}{IPIGFCF_t}, \forall i, j, t \tag{B.1}$$

where $j \in \{Buildings,\ Machinery\ and\ Equipment\}$ and $IPIGFCF$ is the Implicit Price Index for Gross Fixed Capital Formation for Greece provided by the World Bank per asset class per year.

Real Investment per Asset Class[17], as

$$I_{jit} = \frac{\begin{array}{l}\Delta\left(Book_{jit+1} - AccumulatedDepreciation_{jit+1}\right) \\ + CurrentDepreciation_{jit}\end{array}}{IPIGFCF_t}, \forall i, j, t \tag{B.2}$$

where $j \in \{Buildings,\ Machinery\ and\ Equipment\}$.

Real Capital per Asset Class, according to the Perpetual Inventory Method (PIM) where for consecutive years

$$K_{jit} = I_{jit} + (1 - \delta_{jit})K_{jit-1}, \forall i, j, t \tag{B.3}$$

and K_{jit} is initialized using the corresponding Deflated Book Value, $K_{Book\ jit}$, when the firm first appears in the sample or if there is a skip in the series.

Aggregate Physical Capital and Aggregate Investment in Physical Capital, as the respective sums of Buildings, and Machinery and Equipment, $\sum_{j\in\{Buildings, Machinery\ and\ Equipment\}} K_{jit}$ and $\sum_{j\in\{Buildings, Machinery\ and\ Equipment\}} I_{jit}$.

[16] Here, we could have lagged the deflator by the average age of each asset class, as in İmrohoroğlu and Tüzel (2014). However, in our case the estimates for capital age were impossibly high for a considerable number of cases, so we opted not to use it.

[17] Our data do not report Current Depreciation for each asset class, but only in total. We decompose this to its per asset class components as best as possible, using a series of reasonable assumptions (e.g. setting current depreciation to zero when the associated book value is zero, etc).

Real Gross Output, as Net Sales plus Other Operating Income and the change in Inventories of Finished Goods deflated by the Producer Price Index (PPI) of the corresponding 2-digit NACE rev. 2 division (base year 2010) provided to us by ELSTAT. For a few cases where PPI is not available, we use Eurostat's Harmonized Consumer Price Index.

Labour, as the reported number of employees times the year and sector specific ratio of persons employed to number of employees reported by Eurostat. If possible, we fill any missing ratios by linear interpolation. We fill the rest with the ratio of persons employed to employees reported by the EU KLEMS survey. Because self employment may be reported as zero employees (as opposed to missing), we first set observations reporting zero employees to one for non S.A. firms.

Intermediate Inputs, according to Keller and Yeaple (2009) and de Loecker, Eeckhout, and Unger (2020), as Cost of Goods Sold and Other Operating Expenses less Depreciation in Cost of Finished Goods and wage expenditures. Wage expenditures are calculated as the Average Cost of Labour for the corresponding division and year times the original number of employees reported by the firm at that year. We get the average wage by dividing the total cost of employees by the number of employees, reported by Eurostat. If possible, we fill any missing average cost by linear interpolation. We fill the rest with the ratio of the EU KLEMS item COMP over EMPE.

Leverage, as $1 - \dfrac{TotalEquity}{TotalAssets}$.

We then drop any firm associated with a consolidated group in any way, to control for consolidation effects. Then, we drop observations with negative values for age, output, real Buildings, real Machinery and Equipment, and intermediate inputs. Following Gradzewicz (2020), we also set to missing any investment rate above 12, as unlikely. Finally, we treat as outliers and drop observations below the 0.5-th and above the 99.5-th percentile for real output, real aggregate capital, labour, intermediate inputs, and the investment rate.

C Estimation

The estimation procedure of the models in Section 3 is a simplified version of Gandhi, Navarro, and Rivers (2020), as described in the appendix of Collard-Wexler and De Loecker (2020) (less the instrumentation of log capital by log investment to control for measurement errors). The estimation process is identical for both models, with the difference that in the disruption model, g takes a not only lagged productivity as an input, but also a lagged investment spike dummy. We base our computer code for the estimation on the `prodest` R package by Rovigatti (2017). Here we describe the precise steps that yield the results in table 4.

Consider the Cobb-Douglas production function in logs in equation (2)

$$y_{it} = \beta_K k_{it} + \beta_L l_{it} + \beta_M m_{it} + \omega_{it} + \varepsilon_{it} \tag{2}$$

Produce an estimate of β_M as the median of the empirical distribution for the revenue share of intermediate inputs

$$\widehat{\beta_M} = median\left(\frac{p_{it}^M M_{it}}{P_{it} Y_{it}}\right) \tag{C.1}$$

where the capitalized variables, Y_{it} and M_{it}, are output and intermediate inputs in levels instead of logs. P_{it} is the price firm i receives at time t for its output and p_{it}^M the corresponding price it pays for the inputs. In practice, they are the corresponding deflators used when constructing the variables.

Subtract the effect of intermediate inputs from output to get

$$\widetilde{y}_{it} = y_{it} - \widehat{\beta_M} m_{it} \tag{C.2}$$

and use $\widetilde{y}_{it}$ in the Ackerberg, Caves, and Frazer (2015) framework to get estimates for β_K and β_L in the following specification

$$\widetilde{y}_{it} = \beta_K k_{it} + \beta_L l_{it} + \omega_{it} + \varepsilon_{it} \tag{13}$$

This allows for the unique identification of all Cobb-Douglas coefficients.

The Ackerberg et al. (2015) framework is as follows.

For the first stage, proxy for the unobservable ω_{it} by inverting f and substituting in equation (2)

$$\begin{aligned}\widetilde{y}_{it} &= \beta_K k_{it} + \beta_L l_{it} + \omega_{it} + \varepsilon_{it} \\ &= \beta_K k_{it} + \beta_L l_{it} + f^{-1}(k_{it}, l_{it}, m_{it}) + \varepsilon_{it} \\ &= \Phi(k_{it}, l_{it}, m_{it}) + \varepsilon_{it}\end{aligned} \tag{C.3}$$

so that the production function can be written as an unknown function of observable variables, k_{it}, l_{it}, and m_{it}, and the unobservable shock, ε_{it}. To deal with the unknown function, Φ, approximate it with a linear function of its inputs. Because $E(k_{it}\varepsilon_{it}|\mathcal{I}_{it}) = E(l_{it}\varepsilon_{it}|\mathcal{I}_{it}) = E(m_{it}\varepsilon_{it}|\mathcal{I}_{it}) = 0$, this approximation can be estimated by OLS.

Then, take the fitted values of the regression, $\widehat{\Phi}_{it}$, to use in the second stage. This, essentially, purges ε_{it} from output.

In the second stage, choose candidates for $\widehat{\beta_K}$ and $\widehat{\beta_L}$ and get an estimate measure of ω_{it}

$$\widehat{\omega}_{it} = \widehat{\Phi}_{it} - \widehat{\beta_K} k_{it} - \widehat{\beta_L} l_{it} \tag{C.4}$$

The next step is to estimate g. For our AR(1) process for the baseline model we get

$$\omega_{it} = c + \rho\omega_{it-1} + \xi_{it} \tag{5}$$

and for the disruption model we get

$$\omega_{it} = c + \rho\omega_{it-1} + \delta d^{I}_{it-1} + \xi_{it} \tag{3}$$

The model assumptions in Section 3 imply that $E(k_{it}\xi_{it}) = E(l_{it-1}\xi_{it}) = 0$.

Thus, use $\widehat{\omega}_{it}$ (and any other data necessary) to estimate either (5) or (3) by OLS, depending on the chosen model, and use the residuals, $\widehat{\xi}_{it}$ to form the above moment conditions. Essentially, k_{it} and l_{it-1} act as instruments of a non-linear GMM specification given by (C.4) and (5) or (C.4) and (3), respectively. Form the GMM objective as

$$J = \widehat{\xi}' Z' W Z \widehat{\xi} \tag{C.5}$$

where $Z = [k\, l_{-1}]$ and W a weighting matrix. We set $W = (Z'Z)^{-1}$.

We optimize the objective using the Newton-Raphson algorithm, but any appropriate optimization routine can equally be used.

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2020

01-20 Incubated Early Stage Startuppers in Athens and their Initiatives during the Crisis (2010-2016), Ioannis Besis and Ioanna Sapfo Pepelasis
02-20 Alternative Monetary Policy Rules in an Imperfectly Competitive Dynamic Stochastic General Equilibrium Model, George Alogoskoufis and Stelios Giannoulakis
03-20 Quantifying Qualitative Survey Data: New Insights on the (Ir)Rationality of Firms' Forecasts, Alexandros Botsis, Christoph Görtz and Plutarchos Sakellaris
04-20 Can Country-Specific Interest Rate Factors Explain the Forward Premium Anomaly?, Efthymios Argyropoulos, Nikolaos Elias, Dimitris Smyrnakis and Elias Tzavalis
05-20 Signaling product quality through prices in an oligopoly model with costly consumer search, Minghua Cheny, Konstantinos Serfes and Eleftherios Zacharias
06-20 Thinking ahead of the next big crash: Clues from Athens in classical times, George C. Bitros
07-20 Financial crises, firm-level shocks, and large downturns: Evidence from Greece, Stelios Giannoulakis and Plutarchos Sakellaris
08-20 Notes on the Demand Side Secular Stagnation, George D. Demopoulos and Nicholas A. Yannacopoulos
09-20 Institutions and Macroeconomic Performance: Core vs Periphery Countries in the Eurozone, Tryfonas Christou, Apostolis Philippopoulos and Vanghelis Vassilatos
10-20 Macroeconomic Policy Lessons for Greece, George Economides, Dimitris Papageorgiou and Apostolis Philippopoulos
11-20 Energy Transitions and the role of the EU ETS: The case of Greece, Andriana Vlachou and Georgios Pantelias
12-20 Measuring the Default Risk of Small Business Loans: Improved Credit Risk Prediction using Deep Learning, Yiannis Dendramis, Elias Tzavalis and Aikaterini Cheimarioti
14-20 Saving Democracy by Switching to Corporate-like Governance, George C. Bitros
15-20 The impact of the lockdown on the Greek economy and the role of the Recovery Fund, George Economides, Apostolis Philippopoulos and Vanghelis Vassilatos

2021

01-21 Historical Cycles of the Economy of Modern Greece From 1821 to the Present, George Alogoskoufis
02-21 Greece Before and After the Euro: Macroeconomics, Politics and the Quest for Reforms, George Alogoskoufis
03-21 Commodity money and the price level, George C. Bitros
04-21 Destabilizing asymmetries in central banking: With some enlightenment from money in classical Athens, George C. Bitros
05-21 Exploring the Long-Term Impact of Maximum Markup Deregulation, Athanasios Dimas and Christos Genakos

06-21 A regularization approach for estimation and variable selection in high dimensional regression models, Y. Dendramis, L. Giraitis, G. Kapetanios
07-21 Tax Competition in the Presence of Environmental Spillovers, Fabio Antoniou, Panos Hatzipanayotou, Michael S. Michael, Nikos Tsakiris
08-21 Firm Dynamics by Age and Size Classes and the Choice of Size Measure, Stelios Giannoulakis and Plutarchos Sakellaris
09-21 Measuring the Systemic Importance of Banks, Georgios Moratis, Plutarchos Sakellaris
10-21 Firms' Financing Dynamics Around Lumpy Capacity Adjustments, Christoph Görtz, Plutarchos Sakellaris, John D. Tsoukalas
11-21 On the provision of excludable public goods General taxes or user prices? George Economides and Apostolis Philippopoulos
12-21 Asymmetries of Financial Openness in an Optimal Growth Model, George Alogoskoufis
13-21 Evaluating the impact of labour market reforms in Greece during 2010-2018 , Georgios Gatopoulos, Alexandros Louka, Ioannis Polycarpou, Nikolaos Vettas
14-21 From the Athenian silver to the bitcoin standard: Private money in a state-enforced free banking model, George C. Bitros
15-21 Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities. Stelios Arvanitis and Thierry Post
16-21 Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka
17-21 Competition and Pass-Through: Evidence from Isolated Markets, Christos Genakos and Mario Pagliero
18-21 Exploring Okun's Law Asymmetry: An Endogenous Threshold LSTR Approach, Dimitris Christopoulos, Peter McAdam and Elias Tzavalis
19-21 Limit Theory for Martingale Transforms with Heavy-Tailed Multiplicative Noise, Stelios Arvanitis and Alexandros Louka
20-21 Optimal taxation with positional considerations, Ourania Karakosta and Eleftherios Zacharias
21-21 The ECB's policy, the Recovery Fund and the importance of trust: The case of Greece, Vasiliki Dimakopoulou, George Economides and Apostolis Philippopoulos

2022

01-22 Is Ireland the most intangible intensive economy in Europe? A growth accounting perspective, Ilias Kostarakos, KieranMcQuinn and Petros Varthalitis
02-22 Common bank supervision and profitability convergence in the EU, Ioanna Avgeri, Yiannis Dendramis and Helen Louri
03-22 Missing Values in Panel Data Unit Root Tests, Yiannis Karavias, Elias Tzavalis and Haotian Zhang
04-22 Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities, Stelios Arvanitis and Thierry Post
05-22 Concentration Inequalities for Kernel Density Estimators under Uniform Mixing, Stelios Arvanitis
06-22 Public Sector Corruption and the Valuation of Systemically Important Banks, Georgios Bertsatos, Spyros Pagratis, Plutarchos Sakellaris
07-22 Finance or Demand: What drives the Responses of Young and Small Firms to Financial Crises? Stelios Giannoulakis and Plutarchos Sakellaris

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/