Production Function Estimation Controlling for Endogenous Productivity Disruptions

Plutarchos Sakellaris and Dimitris Zaverdas

Athens University of Economics and Business

July 14, 2022

Motivation and Contribution

The current standard for production function estimation assumes productivity is completely exogenous. However, the lumpy adjustment literature has shown that investment spikes can have a systematic negative effect on productivity.

We propose a modification to the standard framework that incorporates endogenous firm decisions for spiky behaviour in the productivity process. We call this the disruption model and juxtapose with a baseline model, which assumes exogenous productivity.

We investigate the differences between models on a large proprietary panel of Greek Manufacturing firms from the ICAP database.

Findings

The two models lead to significantly different production function and productivity process estimates, which in turn affect the results of subsequent inference.

In particular:

- They imply different levels for the stationary average of productivity. (+41.16%)
- They find a different magnitude of endogenous disruptions. (-18.71%)
- The baseline model cannot capture the entirety of disruption costs, which extend across time. (24.98% of output)
- The components of Aggregate Productivity Growth can differ substantially. (by as much as -79.49%)

Related Literature

Current standard for production function estimation:

- Olley and Pakes (1996) (henceforth OP)
- Levinsohn and Petrin (2003) (henceforth LP)
- Ackerberg, Caves, and Frazer (2015) (henceforth ACF)

Issues with gross-output:

- Bond and Söderbom (2005): Gross-output not uniquely identifiable in OP/LP/ACF framework

Issues with value-added:

- Bruno (1978), Diewert (1978): Value-added requires additional non-trivial assumptions

Solution:

- Gandhi, Navarro, and Rivers (2020) (henceforth GNR): Extend ACF to accommodate gross-output

Related Literature

Evidence of disruption effects:

- Sakellaris (2004): Investment spikes followed by TFP drops in US Manufacturing plants
- Cooper and Haltiwanger (2006): They fit a series of models with convex and non-convex adjustment costs on US Manufacturing plant data. They find that a mix of convex and non-convex costs, triggered by investment spikes, fit the data best.
- Gradzewicz (2020): A sample matching experiment with Polish firms reveals that investment spikes lead to an average drop of TFP by 6.25%. Then, it gradually recovers.

Related Literature

Other work on endogenous productivity:

- Doraszelski and Jaumandreu (2013): R&D expenditures
- De Loecker (2013): exports
- Khan and Khederlarian (2021): inventory costs

The Baseline Model

We estimate by GNR the gross-output production function of a baseline model, which ignores investment spikes, as per the standard approach. The model is described by

$$Y_{it} = e^{\omega_{it}+\varepsilon_{it}} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it}^{\beta_M} \tag{1}$$

and

$$\omega_{it} = c + \rho\omega_{it-1} + \xi_{it} \tag{2}$$

with $E(\varepsilon_{it}|\mathcal{I}_{it-1}) = E(\xi_{it}|\mathcal{I}_{it-1}) = 0$.

ω_{it} is the observable (by the firm) part of productivity and ε_{it} is unobservable.

Equation (2) is augmented by a selection probability term in the estimation step, following OP and the literature. The probability is found by logit estimation on third degree polynomials of the investment rate, log inputs, and log output (without cross terms), and a complete set of time dummies.

Disruption Evidence

We conduct a sample matching experiment, similar to Gradzewicz (2020), to compare the performance of productivity between a sub-sample of firms that adjust via spike and a sub-sample of statistically similar firms that do not.

We select the sub-sample of matched observations using a set of approximate and exact matching variables, according to Rosenbaum and Rubin (1985).

The approximate matching variables are firm age, real sales and sales growth, and lagged leverage. The exact matching variables are the year of the spike, legal form, the manufacturing sub-sector, and trading status.

Disruption Evidence

Investment spikes are identified by the following dummy variable

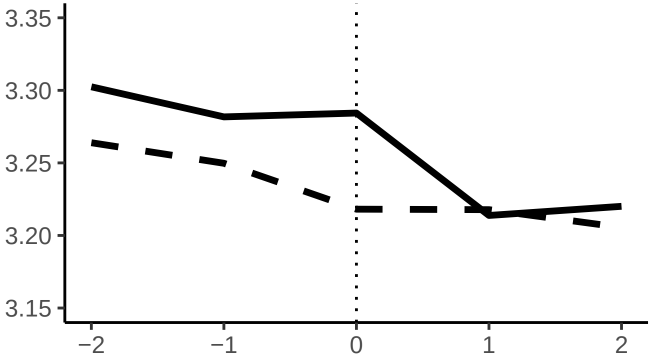
$$d_{it}^{I} = \begin{cases} 1, & \frac{I_{it}}{K_{it}} > \max\left\{2.75 \cdot E\left(\frac{I_{it}}{K_{it}} \middle| K_{it}\right), 0.20\right\} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where the conditional expectation $E\left(\frac{I_{it}}{K_{it}} \middle| K_{it}\right)$ is the fitted value of the regression of $\frac{I_{it}}{K_{it}}$ on a third degree polynomial of log K_{it} and a complete set of time and sector dummies. We also control for the 2008 Greek financial crisis, by fitting separate regressions for before and after 2008.

We identify 3661 investment spikes in 41768 observations. The matched sub-samples cover 591 investment spike events.

Disruption Evidence

Figure: Disruption effect in the baseline model: log TFP around spike events



$N = 591$
The solid line reports for the spike sub-sample and the dashed line reports for the matched sub-sample.

Disruption Evidence

The baseline model finds that log TFP of the spike firms experiences a slight boost in the period of the investment spike, relative to the non-spike firms, and drops on average by 7.07 percentage points in the period after.

A two-sided Wilcoxon test rejects the null of no shift from the spike period to the next with a p-value of zero.

These findings are inline with the existing literature, but at odds with the underlying assumptions used to produce them.

The Disruption Model

We extend the model by introducing investment spike decisions in the productivity process. Thus, equation (1) is the same as in the baseline model

$$Y_{it} = e^{\omega_{it}+\varepsilon_{it}} K_{it}^{\beta_K} L_{it}^{\beta_L} M_{it}^{\beta_M} \tag{1}$$

while equation (2) becomes

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

and $E(\varepsilon_{it}|\mathcal{I}_{it-1}) = E(\xi_{it}|\mathcal{I}_{it-1}) = 0.$

The estimation procedure is identical to that of the baseline model.

Production Function Estimates

Table: Production Function Estimates

	β_K	β_L	β_M
Baseline	0.1029 (0.0098)	0.1449 (0.0225)	0.6777 (0.0016)
Disruption	0.0799 (0.0046)	0.2299 (0.0061)	0.6777 (0.0016)

$N = 41768$

Values in parentheses are standard errors, calculated using 1000 bootstrap samples, following LP. Hahn and Liao (2021) find that bootstrap errors overestimate the true standard errors.

In the disruption model, the estimate for β_K decreases by 2.3302 baseline standard deviations and for β_L it increases by 3.7753 standard deviations. β_M is estimated identically by both models, so the estimates are also identical.

Productivity Process Estimates

Table: Productivity Process Estimates

	c	ρ	δ
Baseline	0.4812	0.8427	
	(0.0147)	(0.0048)	
	(0.0000)	(0.0000)	
Disruption	0.3109	0.9284	-0.0179
	(0.0080)	(0.0025)	(0.0004)
	(0.0000)	(0.0000)	(0.0000)

$N = 41768$

Values in parentheses are heteroskedasticity and autocorrelation robust standard errors and corresponding p-values.

The disruption model finds that there is an omitted disruption coefficient in the productivity process that is negative and statistically significant.

Productivity Process Estimates

Retrieving log TFP according to the estimated parameters of both models shows that the baseline model finds a 3.2056 average log TFP for this sample, whereas the disruption model finds a sample average of 3.2408.

The implied long-run average for the baseline model can be calculated as

$$\frac{\hat{c}}{1-\hat{\rho}}$$

which is 3.0594.

For the disruption model is it

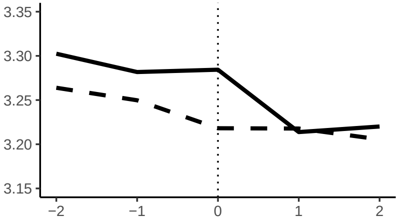
$$\frac{\hat{c}+\hat{\delta}\cdot P(\widehat{d^{I}_{it-1}=1})}{1-\hat{\rho}}$$

Using the in-sample frequency of spikes for $P(\widehat{d^{I}_{it-1}=1})$, we get 4.3187.

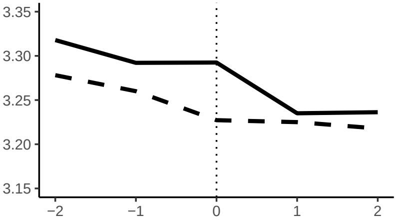
Disruption Effect Between Models

Figure: Average log TFP in the baseline and the disruption model

(a) Baseline $log(TFP)$



(b) Disruption $log(TFP)$



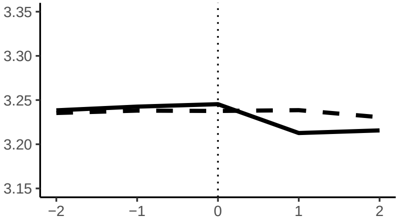
$N = 591$

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

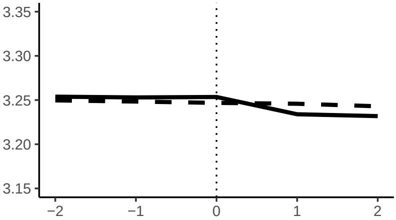
Disruption Effect Between Models

Figure: Average ω in the baseline and the disruption model

(a) Baseline ω



(b) Disruption ω



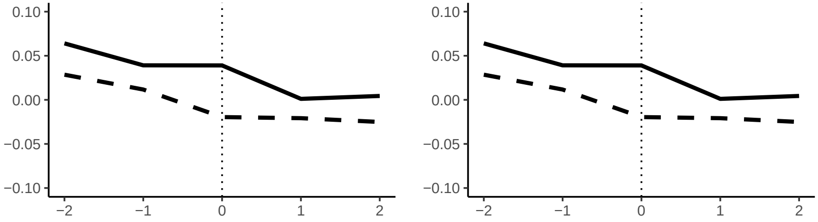
$N = 591$

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

Disruption Effect Between Models

Figure: Average ε in the baseline and the disruption model

(a) Baseline ε (b) Disruption ε



$N = 591$

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

Overall Disruption Effect

The size of the δ coefficient estimate implies that investment spikes results in a drop in ω_{it} by 1.79 percentage points in the immediate next period.

The persistent nature of the process means that investment spikes continue to have an effect in later periods. We can trace the overall effect as the sum of a geometric series

$$\delta + \rho\delta + \rho^2\delta + \rho^3\delta + \cdots = \frac{\delta}{1-\rho} \tag{5}$$

Ceteris paribus, it amounts to an overall decrease in output by 24.98%, which is spread across time in the form of forgone output.

This notable amount is an extra implicit adjustment cost the firm has to bear, which is completely missed by the baseline model.

Of course, after an investment spike, things are not ceteris paribus, since the amount of physical capital has substantially increased.

Aggregate Productivity Growth

We also calculate Aggregate Productivity Growth (APG) and its components, technical efficiency (TE) and reallocation efficiency (RE), as given by both models, following Petrin and Levinsohn (2012).

Table: APG breakdown by each model. Averages per subperiod.

Period	Model	APG	TE	RE
2001-2007	Baseline	0.0016	-0.0020	0.0036
	Disruption	0.0016	-0.0016	-0.0032
2008-2015	Baseline	-0.0034	-0.0136	0.0102
	Disruption	-0.0034	-0.0055	0.0021
2016-2017	Baseline	0.0088	-0.0109	-0.0021
	Disruption	0.0088	-0.0092	-0.0004

Aggregate Productivity Growth

Table: APG breakdown by each model. Averages per subperiod. Disruption over Baseline ratios.

Period	TE	RE
2001-2007	0.7957	0.8861
2008-2015	0.4071	0.2072
2016-2017	0.8464	0.2051

The two models assign significantly different levels of contribution of each component to total APG.

Conclusions

The literature has associated spiky adjustment with productivity drops for many years. Ignoring this finding when estimating production functions can introduce bias at various levels of inference.

We illustrate such a case, where the underlying assumptions on productivity lead to different production function estimates, productivity processes, disruption cost sizes, and APG component magnitudes.

We argue that studies based on production function estimates should take endogenous productivity disruptions into account.

References I

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

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

Bruno, M. (1978). Duality, Intermediate Inputs and Value-Added. In M. Fuss & D. McFadden (Eds.), *Production economics: A dual approach to theory and applications* (Chap. 1, Vol. 23, pp. 3–16). McMaster University Archive for the History of Economic Thought. Retrieved from https://eml.berkeley.edu/prodecon/apps/ch8.pdf

Collard-Wexler, A., & De Loecker, J. (2016). *Production Function Estimation and Capital Measurement Error*. (Working Paper No. 22437). National Bureau of Economic Research. doi:10.3386/w22437

References II

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. (2013). Detecting Learning by Exporting. *American Economic Journal: Microeconomics*, *5*(3), 1–21. doi:10.1257/mic.5.3.1

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* (Chap. 2, Vol. 23). McMaster University Archive for the History of Economic Thought. Retrieved from https://eml.berkeley.edu/prodecon/apps/ch9.pdf

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

References III

Gandhi, A., Navarro, S., & Rivers, D. A. (2020). On the Identification of Gross Output Production Functions. *Journal of Political Economy*, *128*(8), 2973–3016. doi: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. doi:10.1080/07350015.2019.1708369

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

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

References IV

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

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

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

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. doi:10.2307/2683903

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