GLOBAL ECONOMY & DEVELOPMENT
WORKING PAPER 107 | September 2017

Global Economy and Development at BROOKINGS

THE MIDDLE PRODUCTIVITY TRAP:

DYNAMICS OF PRODUCTIVITY DISPERSION

Dany Bahar

Dany Bahar is David M. Rubenstein Fellow in the Global Economy and Development program at the Brookings Institution.

Abstract:

This paper uses a worldwide firm-level panel dataset to document a U-shaped relationship between growth and initial productivity levels: fast growth is concentrated at both ends of the productivity distribution, a result that implies productivity dispersion. This result uses within country-industry variation only, and could partly explain the documented increasing productivity dispersion within narrowly defined sectors.

Acknowledgements:

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.

Brookings recognizes that the value it provides is in its absolute commitment to quality, independence and impact. Activities supported by its donors reflect this commitment and the analysis and recommendations are not determined or influenced by any donation. A full list of contributors to the Brookings Institution can be found in the Annual Report at https://www.brookings.edu/about-us/annual-report/.

The author would like to thank Martin Baily, Joel Bell, Jose Miguel Benavente, Chiara Criscuolo, Diego Comin, Kemal Dervis, Romain Duval, Karim Foda, Peter Gal, Homi Kharas, Santiago Levy, Zia Qureshi, Martin Rotemberg as well as participants at the KDI-Brookings Productivity workshop and the International Monetary Fund's Jobs and Growth seminar. This paper is background research for the "Great Paradox" project undertaken in partnership with the Chumir Foundation for Ethics in Leadership on technological change, stalled productivity growth, and inequality. Excellent research assistance was provided by Luis Omar Herrera and Sebastian Strauss. All errors are my own. Author's contact information: 1775 Massachusetts Ave NW, Washington D.C. 20036; db21@post.harvard.edu; http://www.danybahar.com.

1 Introduction

A few months before leaving office President Barack Obama wrote about the challenges that his successor would have to tackle.[1] Recent innovations, he claimed, "have not yet substantially boosted measured productivity growth." In fact, since 2004, productivity growth slowed across nearly all advanced economies (Baily and Montalbano, 2016). Productivity being the most important determinant of economic growth, Obama concluded: "Without a faster-growing economy, we will not be able to generate the wage gains people want, regardless of how we divide up the pie." The challenge ahead of us begs a better understanding of productivity growth dynamics, which is what this paper aims to do.

An important set of studies have looked at the possible causes of the slowdown in productivity growth, including mis-measurement (see Syverson, 2016 for discussion), the role of recent innovation in boosting (or not) productivity (e.g., Cowen, 2011; Gordon, 2015), the existence of increasing market frictions or decreasing dynamism (e.g., Decker et al., 2014), and of structural vs. cyclical economic factors (e.g., Fernald, 2014; Adler et al., 2017). Beyond dynamics in average productivity, large productivity dispersion within narrowly defined sectors has been widely documented (e.g., Syverson, 2004; Hsieh and Klenow, 2009). Also, Decker et al. (2016a) have documented increasing productivity dispersion since the 1990s in the U.S.

Yet, there is no consensus on what explains increasing productivity dispersion across time.[2] Increasing productivity dispersion could be a result lack of convergence, which could be interpreted as lack of diffusion of innovations across firms. In fact, Comin and Mestieri (2013) show that penetration rates of technologies have declined using macro data. Consistently, this pa-

[1]http://www.economist.com/news/briefing/21708216-americas-president-writes-us-about-four-crucial-areas-unfinished-business-economic

[2]For Hsieh and Klenow (2009) dispersion is a result of misallocation, which could be interpreted as a static increase in dispersion, unless this misallocation worsens over time.

per documents a robust, yet undocumented thus far, stylized fact linking the structural pattern of firm-level productivity growth to within-industry dispersion: convergence-divergence dynamics, where the fastest firm-level TFP growth rate is concentrated at both extremes of the initial productivity distribution, generating a "U-shaped" convergence curve.

The rest of this paper is organized as follows. Section 2 describes the dataset and sample. Section 3 presents some summary statistics, estimates productivity growth regressions, and discuss the main results. Section 4 concludes.

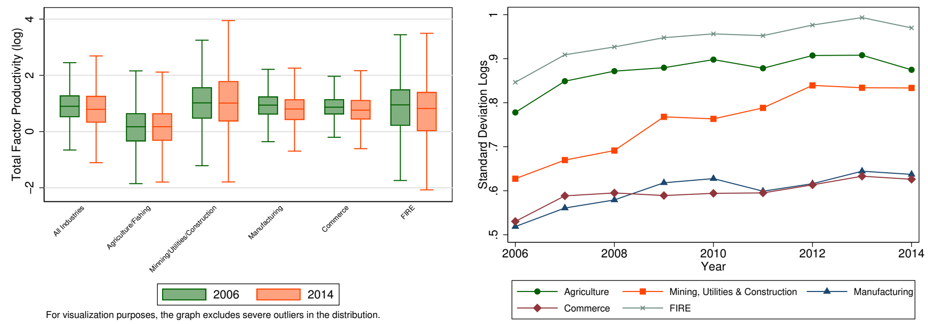
2 Data description and stylized facts

The main data source is the Orbis database from Bureau Van Dijk.[3] The database samples firms worldwide on a yearly basis with their unconsolidated financial information, including operating revenue, cost of workers, value of total assets, and cost of materials. Given that the coverage of US firms in Orbis is poor and scarce, I use COMPUSTAT as an additional source of data to include US firms. I compute total factor productivity (TFP) for every firm (plant) in every period (Online Appendix Section A.1 details the construction of the TFP measures and of the sample and addresses the representativeness of the data).

The final sample is an unbalanced panel of about 4 million firms between 2006 and 2014, totaling over 16 million observations. The firms are active in 654 different six-digit NAICS codes distributed in the following one-digit categories: agriculture and fishing; mining, utilities and construction; manufacturing; commerce (retail and wholesale); and finance, insurance, and real estate (FIRE). The sample has firms active in 127 countries. However, about 95% of the firm-year pairs in the dataset are concentrated in the following

[3]This dataset has been employed also by a number of other researchers in the previous years (e.g., Bloom et al. 2012; Fons-Rosen et al. 2013; Andrews et al. 2016; Gopinath et al. 2017; Duval et al. 2017).

Figure 1: Firm TFP distribution and dispersion over time, 2006 and 2014



The left panel of the figure visualizes the distribution of log TFP for all firms by sectors both in 2006 and 2014. The horizontal line inside the box represents the median value and the edges of the box represents values in the 25th and 75th percentiles. The right panel of the figure visualizes the average TFP dispersion from 2006 to 2014 by broad sector (averaged over six-digit NAICS codes).

countries: Croatia, Czech Republic, Bulgaria, Finland, France, Germany, Italy, Korea, Norway, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Ukraine, and the United Kingdom.

3 Convergence-divergence dynamics

Figure 1 presents two stylized facts using the raw data. First, the left panel illustrates how median TFP—represented by the horizontal line inside of the box— has dropped between 2006 and 2014 across all sectors. Negative overall productivity growth is consistent with the data that has been shown by others looking at the post-2008 crisis period (e.g., OECD, 2015; Andrews et al., 2016). Second, the right panel illustrates that productivity dispersion of the sample (as measured by the standard deviation of log TFP) seems to have increased during the same period, particularly in the manufacturing, mining, and FIRE sectors (a fact that also recently been established by Decker et al. (2016a) for the US since the 1990s. much before the 2008 crisis).

There are two processes that could drive an increase in productivity dispersion in between two periods. First, reallocation: firms entering and/or exiting the market in ways that increases dispersion in the distribution of productivity (e.g., high entry of very productive firms and low exit of very unproductive firms, high entry of both very unproductive and very productive firms, etc.). Second, within-firm productivity dynamics: firms at the top becoming more productive relative to firms at the bottom. Exploring the first process is challenging in this setting, since sampling methods in each country could vary and are unknown, and entry and exit might not be reliable. Therefore, I focus on the second process: dynamics of productivity growth within each country and sector for incumbent firms. To study this, I estimate the following TFP growth convergence regression:

$$TFPGrowth_{i,p,c,t\rightarrow T} = \beta_l lnTFP_{i,p,c,t} + \beta_q lnTFP^2_{i,p,c,t} + \theta_{c,p} + \eta_{p,t} + \omega_{c,t} + \varepsilon_{i,p,c,t} \quad (1)$$

Where $TFPGrowth_{i,p,c,t\rightarrow T}$ is the three-year compound average growth rate (CAGR) for firm i's TFP between period t and T.[4][5] Each firm is active in six-digit NAICS industry denoted by p, in country c. The right hand side includes the initial level (in period t) of log TFP for the same firm, both in its linear and quadratic form, to allow for non-linearities in the estimation. All estimations include a battery of fixed effects: country-industry ($\theta_{c,p}$),

[4]The CAGR is computed as $({}^{TFP_T}/_{TFP_t})^{1/T-t} - 1$. Given that the sample is not a fully balanced panel (firms enter and exit frequently), the construction of the sample used to compute growth rates requires an explanation. The sample is composed by all firms in the 2006-2014 period that are present in any three-year period pooled in one regression, without overlap. For example, if a firm's fist appearance in the sample is in the year 2006 and the same firm appears in the sample in the year 2009, then that firm represents one observation. The same firm would be part of the sample a second time only for the 2009-2012 period, regardless on whether it is present in 2007 and/or 2008. Other firms for which the first appearance in the sample is 2007 are included in the sample if they have data for 2010, too, and included again a second time for 2010-2013 (if data for 2013 is available).

[5]Online Appendix Section B explores results using different period lengths to compute growth. The longer the periods the less pronounced the right tail of the U-shaped relationship.

Table 1: TFP Growth Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Max
TFP CAGR	2,888,055	-0.02	0.25	-0.98	37.76
Baseline log TFP	2,888,055	0.78	0.94	-11.60	9.72
Sales (log)	2,888,055	6.74	1.87	-0.29	19.72
Employment costs (log)	2,888,055	1.66	1.29	0.00	13.24
Capital (log)	2,888,055	6.46	1.95	-0.09	21.52
Materials (log)	2,888,055	5.63	2.42	-0.15	19.53
Firms per Country-NAICS	2,888,055	5,241.26	9,547.92	1.00	56,248.00

industry-year ($\eta_{p,t}$) and country-year ($\omega_{c,t}$). $\varepsilon_{i,p,c,t}$ represents idiosyncratic shocks. Table 1 shows the summary statistics for the sample used with close to 2.9 million observations. Average TFP growth is negative, consistently with Figure 1.

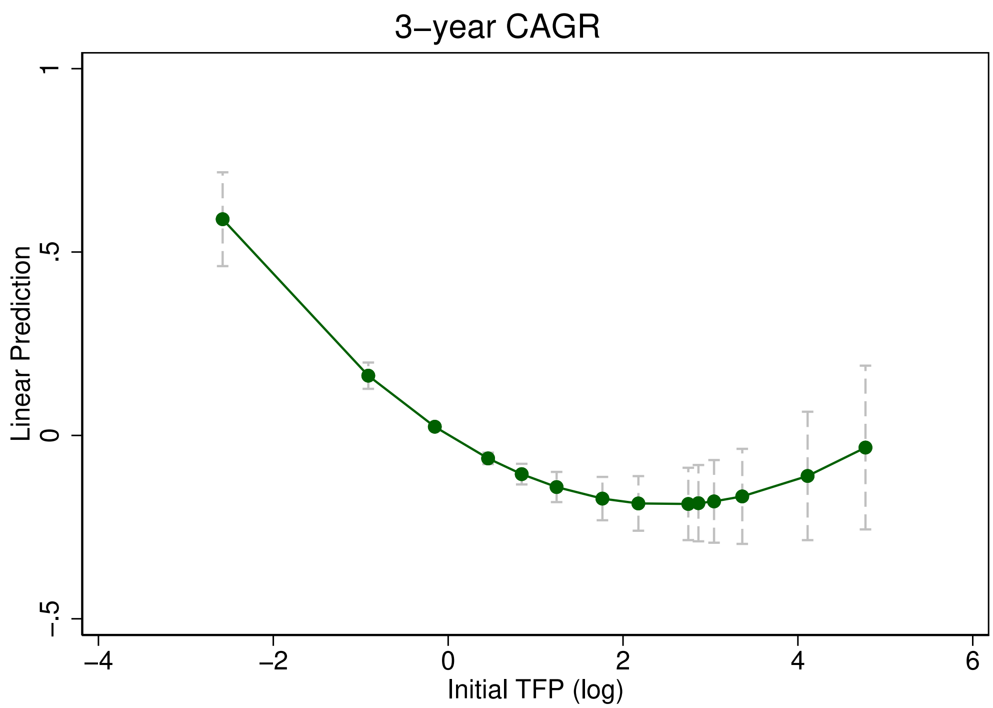
The results of the estimation of the convergence regression are presented in Table 2. Column (1) presents the results of a linear growth regression, while column (2) allows for a quadratic term. The results are consistent with the convergence intuition: firms with higher initial levels of TFP would tend to grow slower (hence, the negative estimator). Yet, in column (2) β_l is estimated to be negative while β_q is estimated to be positive, both statistically different from zero.

This implies that —on average— there is a non-linear, U-shaped, relationship that describes TFP growth dynamics given the baseline level. That is, fast growth occurs for firms both in the lower end and the upper end of the TFP distribution. Note that this result controls for country-industry fixed effects, implying that these dynamics occur within each country-industry. The line titled "IP," which stands for inflection point, shows the percentile after which the growth rate starts increasing, which in this case is the 98th percentile. Figure 2 visualizes this result.[6]

Table 3 replicates the results from above for each one-digit NAICS code.

[6] Results are robust, though somewhat weaker, when correcting for markups. See Online Appendix Sections A.3.3 and C.4 for more details.

Figure 2: Expected TFP Growth based on initial TFP levels



This figure visualizes the expected three-year TFP annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code.

Table 2: TFP Growth Regression

Dependent Variable: TFP three-year Growth Rate (CAGR)		
	(1)	(2)
lnTFP	-0.1533	-0.1508
	(0.014)***	(0.017)***
lnTFP × lnTFP		0.0301
		(0.005)***
N	2887380	2887380
Adj r2	0.21	0.26
IP		.98

All columns include country-industry, country-year, and industry-year fixed effects. Standard errors clustered at the country and industry level are presented in parenthesis.

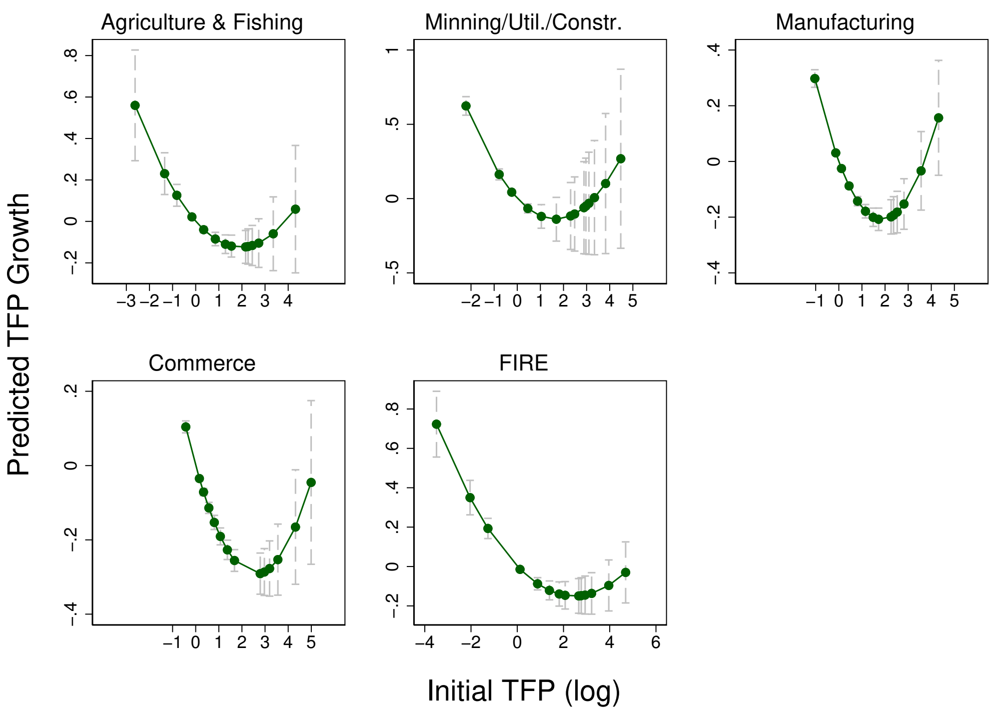
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

The industries are agriculture (AGR); mining, utilities, and construction (MUC); manufacturing (MNFTR); commerce (COM); and financial, insurance, and real estate services (FIRE). The U-shaped relationship is robust across all industries and the inflection point occurs within the distribution (titled IP in the last line of the table). In general, the inflection point is at the very end of the distribution (above the 97th percentile), except for MUC (column 2), for which the inflection point is at the 81st percentile, implying that growth starts increasing for the about the firms at the top 20 percent. Figure 3 plots the relationships in the table.

These results are consistent with Andrews et al. (2016) who show evidence of divergence in productivity levels between global frontier and laggard firms, whereas all my results focus on growth (not levels) and are within-country.[7] The main takeaway from these results is that such dynamics could generate increasing productivity dispersion in the long run (consistently with to what documented by Decker et al. (2016b)), and in turn generate higher market

[7]Note that Andrews et al. (2016) compare levels of laggard and of global frontier firms where the firm composition of the latter changes year-by-year.

Figure 3: Expected TFP Growth, by one-digit NAICS



This figure visualizes the expected three-year TFP annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the nth to the 99.99th percentile, for each one-digit NAICS industry. The estimation controls for country-industry, industry-year and country-year fixed effects, where each industry is a six digit NAICS code.

Table 3: TFP Growth Regression, by one-digit NAICS

Dependent Variable: TFP three-year Growth Rate (CAGR)					
	AGR	MUC	MNFTR	COM	FIRE
lnTFP	-0.1276	-0.1684	-0.2247	-0.2248	-0.1213
	(0.024)***	(0.029)***	(0.011)***	(0.016)***	(0.018)***
lnTFP_sq	0.0327	0.0510	0.0605	0.0432	0.0245
	(0.011)***	(0.008)***	(0.006)***	(0.007)***	(0.002)***
N	99849	420095	595342	1080389	691705
Adj r2	0.31	0.32	0.39	0.30	0.23
IP	.99	.81	.97	.98	.97

The table estimates a TFP growth convergence regression, by one-digit NAICS sectors. All columns include country-industry, industry-year, and country-year fixed effects. Standard errors clustered at the country and industry level are presented in parenthesis.
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

concentration as firms at the top become more productive, similarly to the rise of superstar firms documented by Autor et al. (2017). In fact, the results suggest something else: Firms at the top are innovating and appropriating the returns to those innovations, but those innovations, perhaps given prevalent market frictions, are not trickling down to laggard firms.

4 Concluding Remarks

This paper shows one particular stylized fact on cross-firm productivity convergence: There is no full convergence, but rather divergence driven by fast growth of the firms at the frontier leaving the rest behind, generating a "middle productivity trap." This process could partly explain both the increase of productivity dispersion (Decker et al., 2016a) and the increasing market share of top firms (Autor et al., 2017).

Moreover, if TFP firm-level dispersion indeed reflects an inefficient allocation of resources (e.g., Hsieh and Klenow, 2009), then the increasing dispersion over the long run that comes out of the above documented dynamics implies in turn worse misallocation of resources across firms over time. Thus, the documented convergence-divergence dynamics could partially explain the post-2005 slowdown in aggregate productivity. Even if establishing the causal link between all these processes at both at the firm and aggregate levels represents important empirical challenges, the results presented in this study aim to shed light on these relationships.

References

Adler, Gustavo, Romain Duval, Davide Furceri, Sinem Kili, Ksenia Koloskova, and Marcos Poplawski-Ribeiro. "Gone with the Headwinds : Global Productivity." *IMF Staff Discussion Note* .

Andrews, Dan, Chiara Criscuolo, and Peter N Gal. "The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy." *OECD Productivity Working Papers* , 5.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. "The Fall of the Labor Share and the Rise of Superstar Firms." *MIT Working Paper* .

Baily, Martin, and Nicholas Montalbano. "Why is US productivity growth so slow? Possible explanations and policy responses." *Hutchins Center Working Paper* , September: (2016) 1–28.

Bloom, Nicholas, Raffaella Sadun, and J. Van Reenen. "The Organization of Firms Across Countries." *The Quarterly Journal of Economics* 127, 4: (2012) 1663–1705.

Comin, Diego A, and Marti Mestieri. "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *NBER working paper* , No. 19010: (2013) 66.

Cowen, Tyler. *The Great Stagnation: How America Ate All the Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better*. New York: Dutton, 2011.

De Loecker, Jan, and Frederic Warzynski. "Markups and Firm-Level Export Status." *American Economic Review* 102, 6: (2012) 2437–2471.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28, 3: (2014) 3–24.

———. "Declining business dynamism: Implications for productivity?" *Hutchins Center Working Paper* , 23.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda. "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness.", 2016b.

Duval, Romain, Gee Hee Hong, and Yannick Timmer. "Financial Frictions and the Great Productivity Slowdown." *IMF Working Papers* 17, 129.

Fernald, John G. "Productivity and Potential Output Before, During , and After the Great Recession." *Federal Reserve Bank of San Francisco Working Paper Series* .

Fons-Rosen, Christian, Sebnem Kalemli-Rosen, Bent E. Sorenson, Carolina Villegas-Sanchez, and Vadym Volosovych. "Quantifying Productivity Gains from Foreign Investment." *NBER Working Papers Series* 18920: (2013) p. 69, Cambridge MA.

Foster, Lucia, John Haltiwanger, and Chad Syverson. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98, 1: (2008) 394–425.

Gal, Peter N. “Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS.” *OECD Economics Department Working Papers* , 1049.

Gopinath, Gita, Sebnem Kalemli-Ozcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. “Capital Allocation and Productivity in South Europe.” *Quarterly Journal of Economics* .

Gordon, Robert. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton, NJ: Princeton University Press., 2015.

Hsieh, CT, and P.J. Klenow. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics* 124, 4: (2009) 1403–1448.

Levinsohn, James, and Amil Petrin. “Estimating Production Inputs to Functions Using Control for Unobservables.” *The Review of Economic Studies* 70, 2: (2003) 317–341.

OECD. *The Future of Productivity*. OECD Publishing, 2015. `http://www.oecd-ilibrary.org/economics/the-future-of-productivity{_}9789264248533-en`.

Syverson, Chad. “Product Substitutability and Productivity Dispersion.” *Review of Economics and Statistics* 86, 2: (2004) 534–550.

———. “Challenges To Mismeasurement Explanations for the U.S. Productivity Slowdown.” In *NBER Working Paper*, 2016, Working Paper 21974.

Wooldridge, Jeffrey M. “On estimating firm-level production functions using proxy variables to control for unobservables.” *Economics Letters* 104, 3: (2009) 112–114.

Online Appendix for
The middle productivity trap: dynamics of productivity dispersion

Dany Bahar (Brookings Institution and Harvard CID)

September 18, 2017

A Data Construction

A.1 Data source and representativeness of the sample

The main data source for this paper is the Orbis dataset, compiled by Bureau van Dijk. It contains relevant accounting and performance indicators over time for millions of firms across the globe. I use COMPUSTAT as an additional source of data to include U.S. (publicly listed) firms, given that Orbis selection of U.S. firms is poor and sparse. I restrict the dataset to those firm-year observations with relevant variables needed to estimate Total Factor Productivity, namely operating revenue, number of workers, tangible fixed assets, and cost of materials. The variables are unconsolidated, implying they represent the financials of each plant in the sample separately. I also have information on the location of the plant (city and country) and its economy activity categorized in six-digit NAICS. The sample is an unbalanced panel of plants between 2005 and 2014, though most of the information is concentrated in post-2008 years. We follow Gal (2013) suggestions when it comes to the imputation of some missing values for few firms. In particular, we imputed the value of material costs (e.g., intermediate goods) by computing the difference between operating revenue and added value.

The sample is not a random sample of firms, and therefore there could be serious concerns of the representativeness of such sample. In order to reduce those concerns, I use additional data from the Structural and Demographic

Business Statistics database (SBDS) from the Organization for Economic Cooperation and Development (OECD). In particular, we construct weights for each cell of country, industry (using two-digit NAICS codes) and five different groups of firm sizes: less than 10 employees, between 10 and 19 employees, between 20 and 49 employees, between 50 and 2490 employees, and 250 employees or more. The weights in each cell are computed as the ratio of total employment according to the Orbis sample to the total employment in SBDS, in each year of the sample. About 25 percent of the observations were given a weight of 1 (i.e., no weight) for lack of data on the SBDS database.

A.2 Deflating values

We convert all the monetary variables to be in constant U.S. dollar values based on year 2010. We use different deflating indexes for different values. In particular, we use the following indexes:

1. We use the Producers Price Index (PPI) for all commodities from the Federal Reserve, to deflate operating revenue values by industry.

2. We compute a index to deflate cost of materials by computing a weighted PPI by industry, using the Bureau of Economic Analysis' Input-Output table.

3. We use the Employment Cost Index from the Bureau of Labor Statistics, to deflate cost of employees by industry.

4. We use the PPI for investment capital goods from the Federal Reserve to deflate fixed tangible assets, computed as the ratio of current to real value of the stock of private equipment.

Creating deflators at the industry level often relied on the availability of industry concordance tables and merging at different levels of disaggregation. Typically, the first merge attempt occurs at four-digit NAICS. For those observations for which the merge was unsuccessfully, a new merge attempt occurs at three-digits, and then a new iteration using two-digits. The attempt uses the mean deflator value for all industries in a given year.

First, in order to alleviate concerns regarding the representativity of the dataset, I weight firms based on their country, industry, year, and size class using data from Structural Analysis Database (STAN) from the Organization for Economic Cooperation and Development website, following the methodology by Gal (2013). This corrects for the total number of workers in a country, industry, year, and size class in the sample using as benchmark the homologue number in the economy.[8] All the results that follow use analytical weights to improve representativeness of the sample.

A.3 Estimation of the production function

A.3.1 Estimation of elasticities

I use different methodologies widely accepted in the literature to estimate the parameters of the production function that serves to estimate Total Factor Productivity. In particular, I assume a Cobb-Douglas production function of the following form:

$$Y_i = K^{\beta_K} wL^{\beta_L} M^{\beta_M} \tag{2}$$

In the data, for a given year, Y is represented by operating revenue, K is represented by fixed tangible assets, wL is represented by the cost of employees, and M is represented by cost of materials. I estimate the values of β_K, β_L and β_M using:

1. Cost share values, for every country and industry (three-digit NAICS code) levels. This assumes constant return to scale. It uses the weights to estimate total costs.

[8] Appendix Section A.1 details the construction of the analytical weights.

2. Cost share values at the plant level. This assumes constant return to scale.

3. Ordinary Least Squares, estimating a log-transformation of equation (2), for every country and three-digit NAICS. I exclude country-industry cells with less than 30 observations. I drop country-industries for which any elasticity is estimated to be negative or larger than 1.

4. Levinsohn and Petrin (2003) based on the implementation suggested by Wooldridge (2009). I assume materials is a flexible input, and compute it for every country and three-digit NAICS, using two lagged values for the estimation. I exclude from the estimation country-industry cells with less than 30 observations. I drop country-industries for which any elasticity is estimated to be negative or larger than one.

A.3.2 Estimation of Total Factor Productivity estimation

With the elasticities in hand, I estimate TFP for each observation (firm-year) in the sample by computing:

$$lnTFP = lnY - \hat{\beta}_K ln(K) - \hat{\beta}_L ln(wL) - \hat{\beta}_M ln(M)$$

Note that each firm-year observation uses the corresponding elasticity based on its country and industry.

I turn to estimate the parameters of the production function for each country and three-digit NAICS industry, through a number of different methods established in the literature: (i) cost shares at the country and three-digit NAICS industry level, (ii) cost shares at the plant level, (iii) ordinary least squares at the country and three-digit NAICS industry, and (iv) using the methodology devised by Levinsohn and Petrin (2003) following Wooldridge (2009), at the country and industry three-digit NAICS level. With the estimated parameters, namely β_K, β_L and β_M, I compute four different measures of TFP for each firm-year by computing the residual:

Table A1: Correlation between TFP using different methods

Variables	CS	CSPlant	OLS	LP
CS	1.000			
CSPlant	0.243	1.000		
OLS	0.586	0.229	1.000	
LP	0.528	0.380	0.680	1.000

This table presents the correlation matrix of log TFP for all observations in the sample using different estimation methodologies.

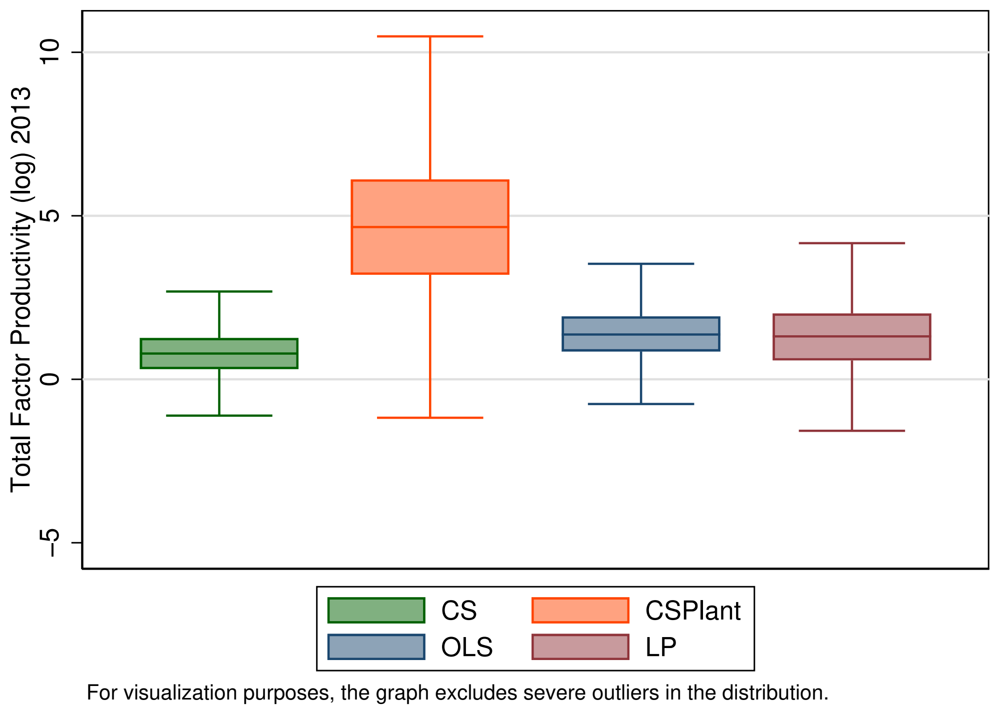
$$lnTFP = lnY - \hat{\beta}_K ln(K) - \hat{\beta}_L ln(wL) - \hat{\beta}_M ln(M)$$

There is a positive and often strong correlation between the (log) TFP estimates in the sample using all different methods, as shown in Table A1. Figure A1 shows the distribution of log TFP for 2014 (the last year of our sample) side by side, using different methodologies (the TFP estimated using cost share values at the plant level is, naturally, much more dispersed than the other measures). In the main body of the paper I present results using revenue-based TFP based on cost shares, unless otherwise specified. This choice allows me to have the largest sample in terms of firms and countries.

A.3.3 Markups

Note that the TFP values are revenue-based, since operating revenue includes pricing. This is less of a concern for this exercise for two main reasons. First, the analysis compares firms within a six-digit NAICS code, which is a very narrow definition of an industry, in which we would expect less pricing differences. In addition, the work by Foster et al. (2008) shows that firms with larger TFP tend to have lower prices, implying that markups are higher among less productive firms. This fact would work against the results I present in the analysis. However, in order to relieve any remaining concerns, I follow the methodology put forward by De Loecker and Warzynski (2012) and follow the implementation suggested by Andrews et al. (2016),

Figure A1: TFP comparison using different methods



This figure visualizes the distribution of log TFP for all observations in the year 2014 using different estimation methodologies. The horizontal line inside the box represents the median value. The edges of the box represents the values in between the 25th and 75th percentile, while the whiskers' edges represent adjacent values.

to compute markups for each firm, assuming materials is a flexible output.[9] Results are mostly robust to using a transformation of TFP that corrects for markups using this methodology.

The computed values for TFP correspond to revenue-based TFP, because operating revenues includes both price and output. I follow the work by De Loecker and Warzynski (2012) and its implementation by Andrews et al. (2016), and compute a firm's markup in a given year in the following manner:

$$markup_i = \frac{\beta_M}{M_i/(K_i + wL_i + M_i)}$$

I then compute a measure of TFP net of markups by computing:

$$lnTFP_i^{nomarkup} = lnTFP_i - log(markup_i)$$

All results are robust to using this measure of TFP.

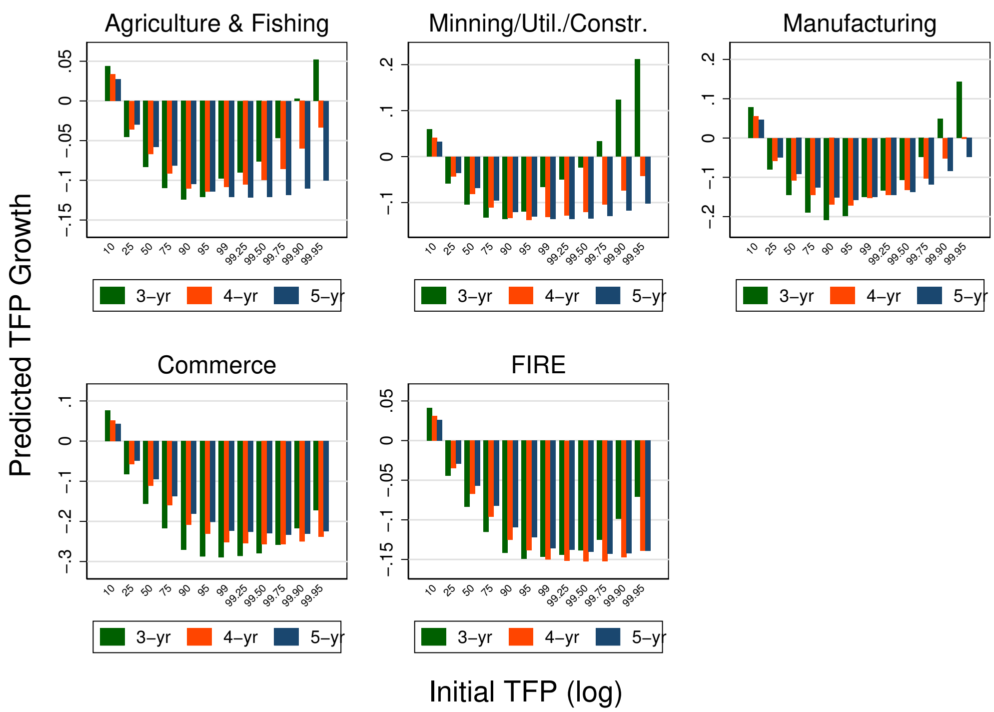
B Using longer periods to compute growth

When redoing the exercise above for different lengths when computing growth rates, I find that the longer the period the less pronounced the right part of the U-shaped convergence curve. This is shown in Figure A2, which plots the expected growth rate for different percentile values of the distribution, from the 10th percentile to the 99.95th percentile. Across all industries, the three-year growth rate is expected to increase for values of TFP above the 99th percentile much more than four-year and five-year growth rates. Yet, the U-shaped curve is still there, particularly in firms in both the manufacturing and the mining, utility, and construction sectors.

Naturally, it is important to consider that even when the growth at the top is relatively slower when using longer time periods of growth as compared to shorter periods, those smaller differences in growth rates between the frontier

[9] See Appendix A.3.3 for details.

Figure A2: Expected TFP growth varying period length, by one-digit NAICS



This figure visualizes the expected TFP annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile, using different lengths for computing the annual growth (three-, four- and five-year periods), for each one-digit NAICS industry. The estimation controls for country-industry, industry-year and country-year fixed effects, where each industry is a six digit NAICS code.

and laggard firms could be crucial in explaining difference in levels the longer the period under consideration.

C Results using different TFP measures

C.1 Cost shares plant level

Figures A3 and A4 replicates Figures 2 and 3 in the main body of the paper, using a measure of TFP based on a production function estimated through ordinary least squares (see Online Appendix Section A.3.1 for more details).

C.2 Ordinary Least Squares

Figures A5 and A6 replicates Figures 2 and 3 in the main body of the paper, using a measure of TFP based on a production function estimated through ordinary least squares (see Online Appendix Section A.3.1 for more details).

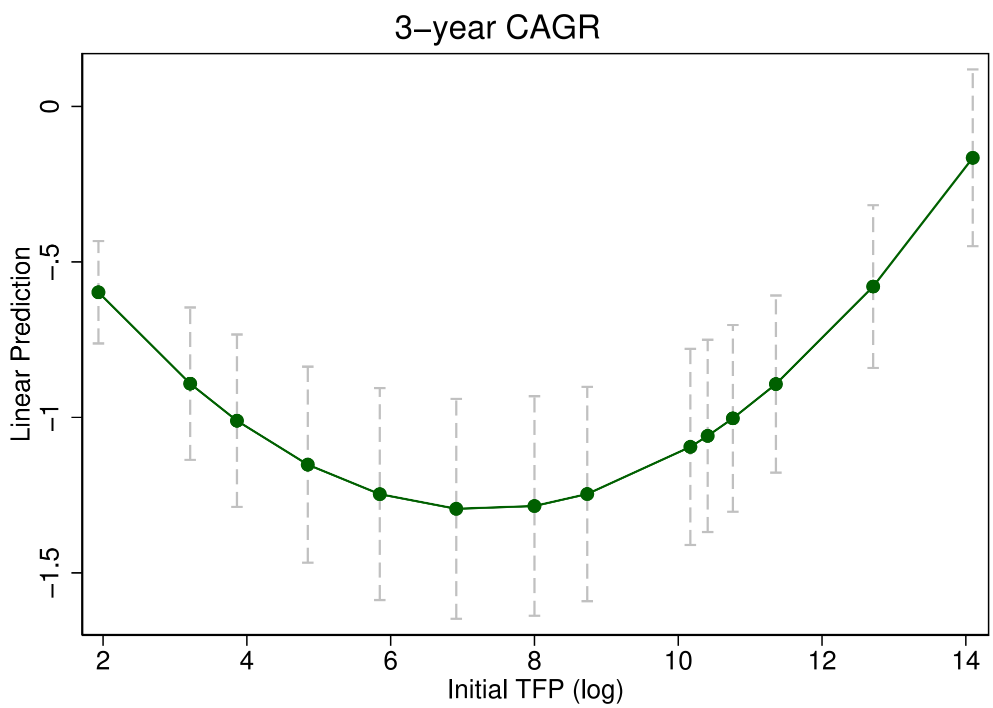
C.3 Wooldridge (2009) TFP

Figures A7 and A8 replicates Figures 2 and 3 in the main body of the paper, using a measure of TFP based on a production function estimated through the methodology in Wooldridge (2009) (see Online Appendix Section A.3.1 for more details).

C.4 Wooldridge (2009) TFP correcting for markups

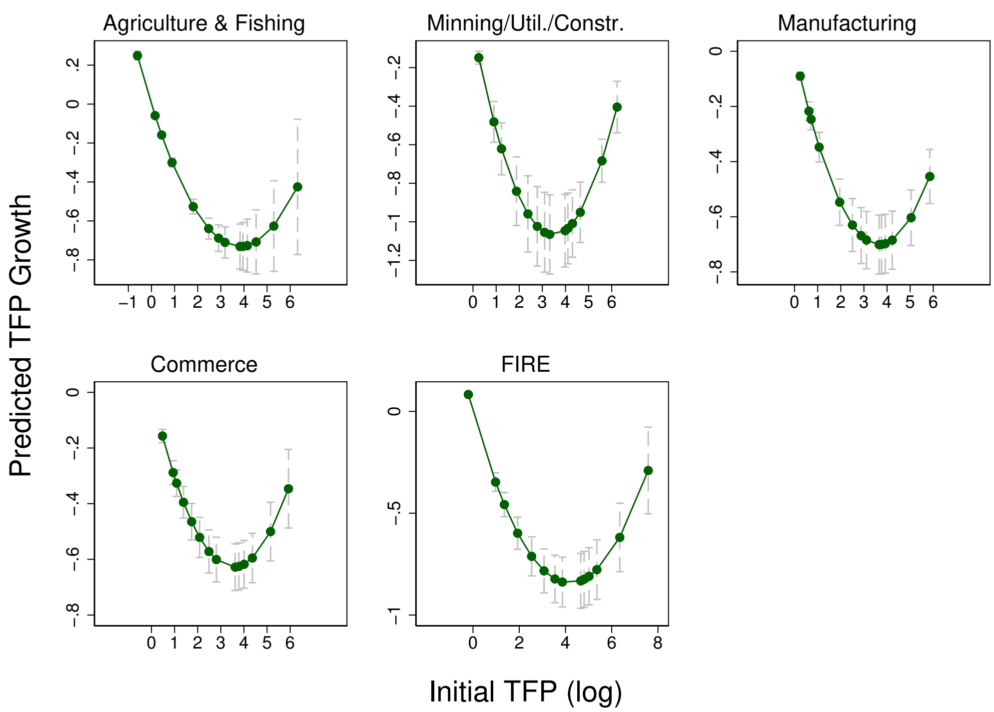
Figures A9 and A10 replicates Figures 2 and 3 in the main body of the paper, using a measure of TFP based on a production function estimated through the methodology in Wooldridge (2009), correcting for markups following the guidelines by De Loecker and Warzynski (2012) and implementation by Andrews et al. (2016) (see Online Appendix Sections A.3.1 and A.3.3 for more details).

Figure A3: Expected TFP Growth based on initial TFP levels



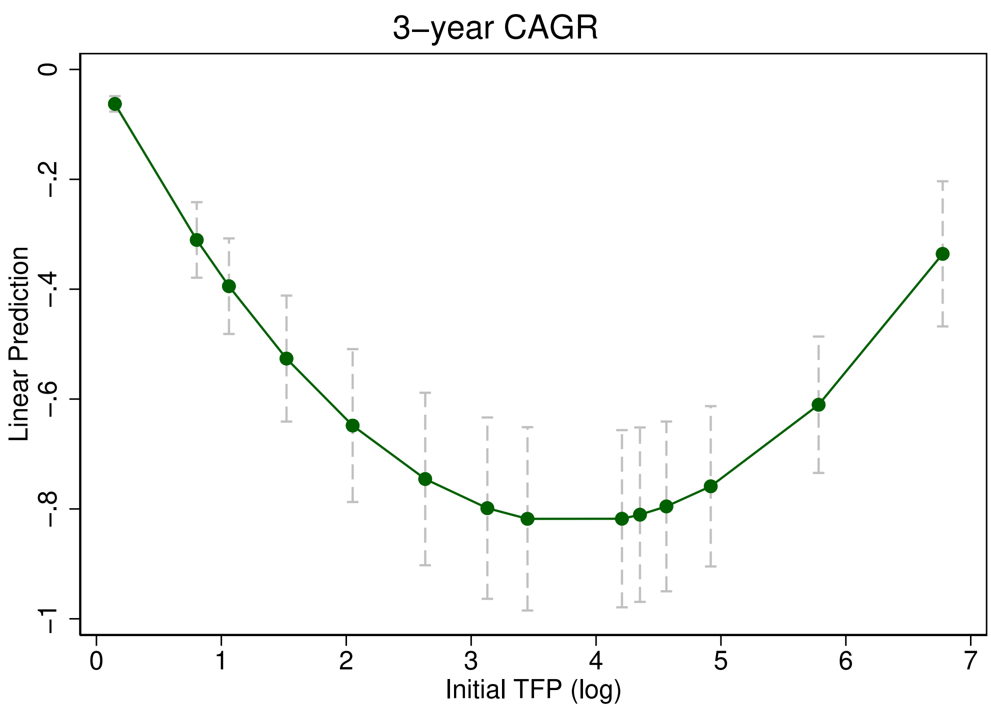
This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using cost shares at the plant level as parameters of the production function (see Online Appendix Section A.3.1 for details).

Figure A4: Expected TFP Growth, by one-digit NAICS



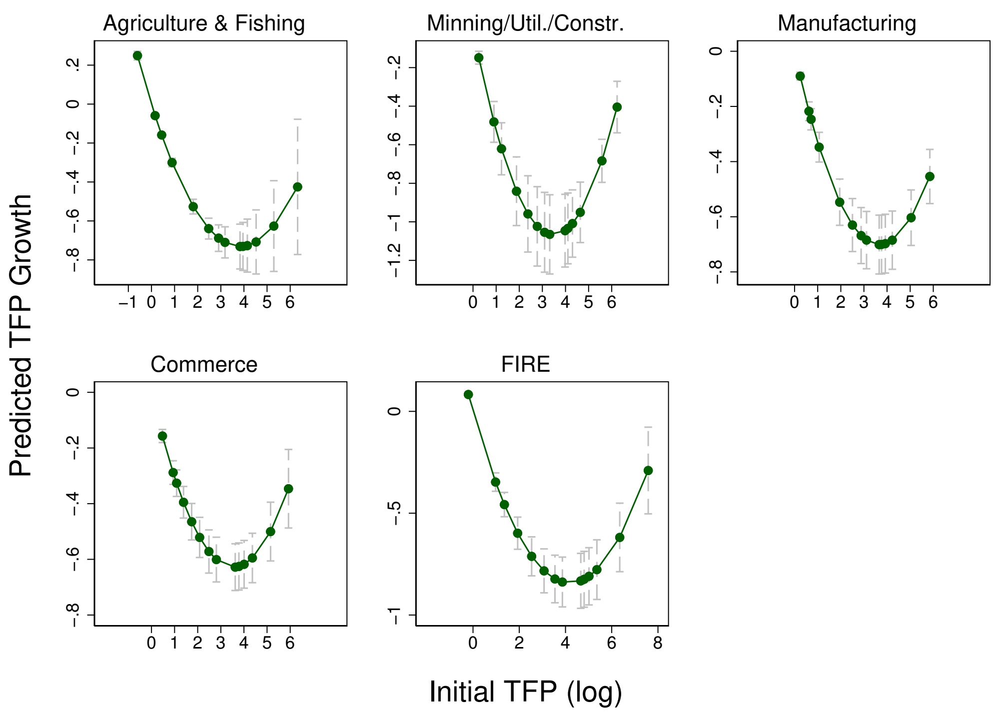
This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. It plots the three-year TFP annual growth rate for each one-digit NAICS industry. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using cost shares at the plant level as parameters of the production function (see Online Appendix Section A.3.1 for details).

Figure A5: Expected TFP Growth based on initial TFP levels



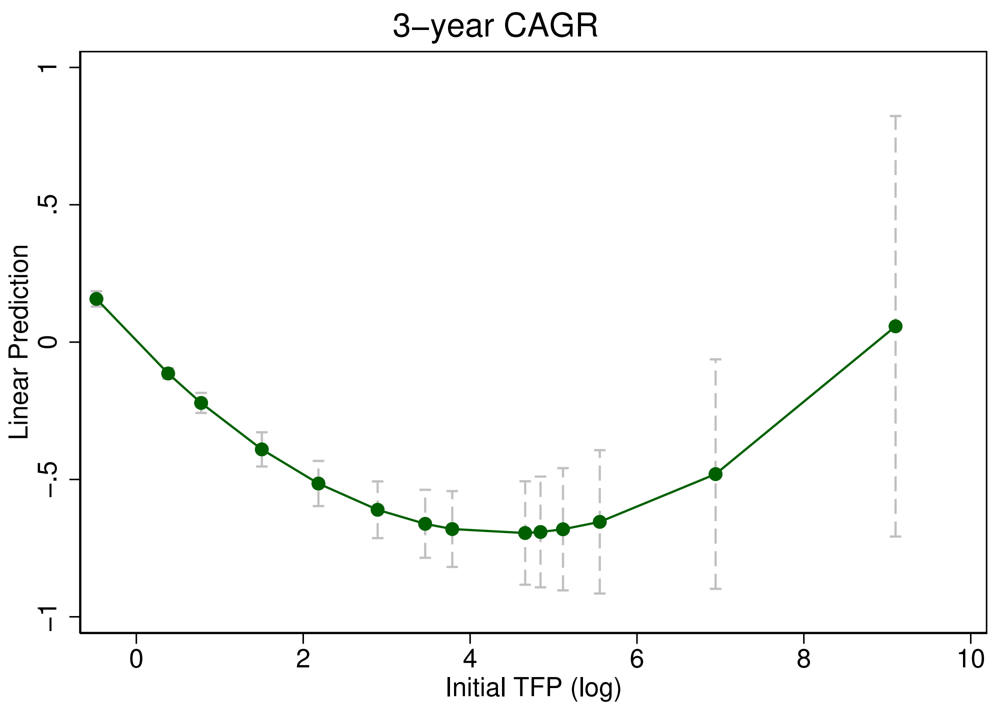
This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using ordinary least squares to estimate the parameters of production function (see Online Appendix Section A.3.1 for details).

Figure A6: Expected TFP Growth, by one-digit NAICS



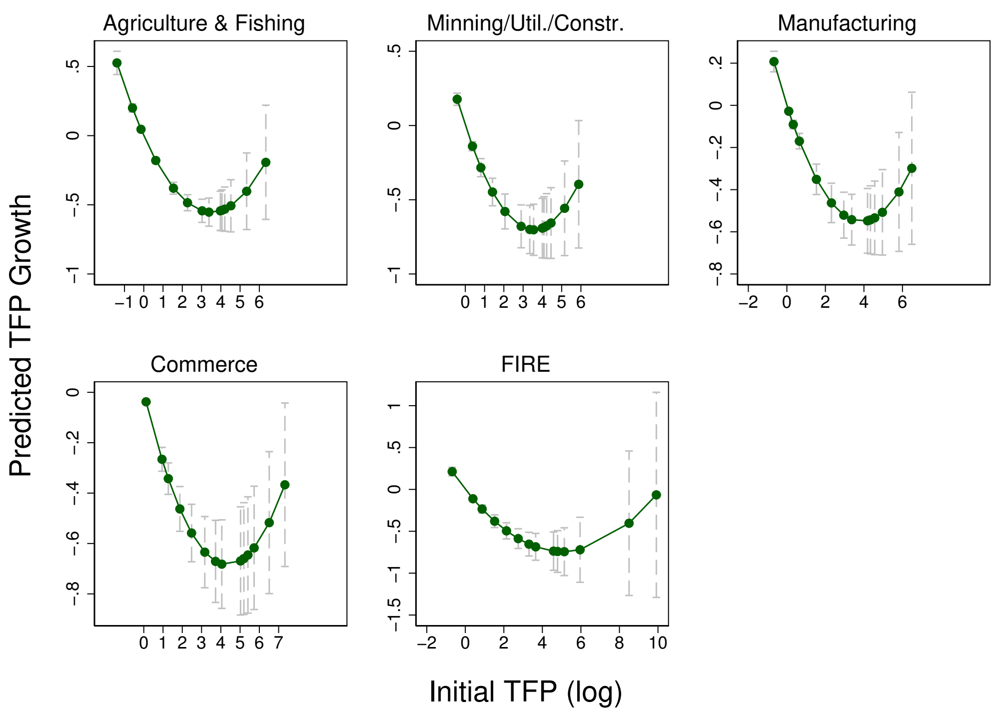
This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. It plots the three-year TFP annual growth rate for each one-digit NAICS industry. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using ordinary least squares to estimate the parameters of production function (see Online Appendix Section A.3.1 for details).

Figure A7: Expected TFP Growth based on initial TFP levels



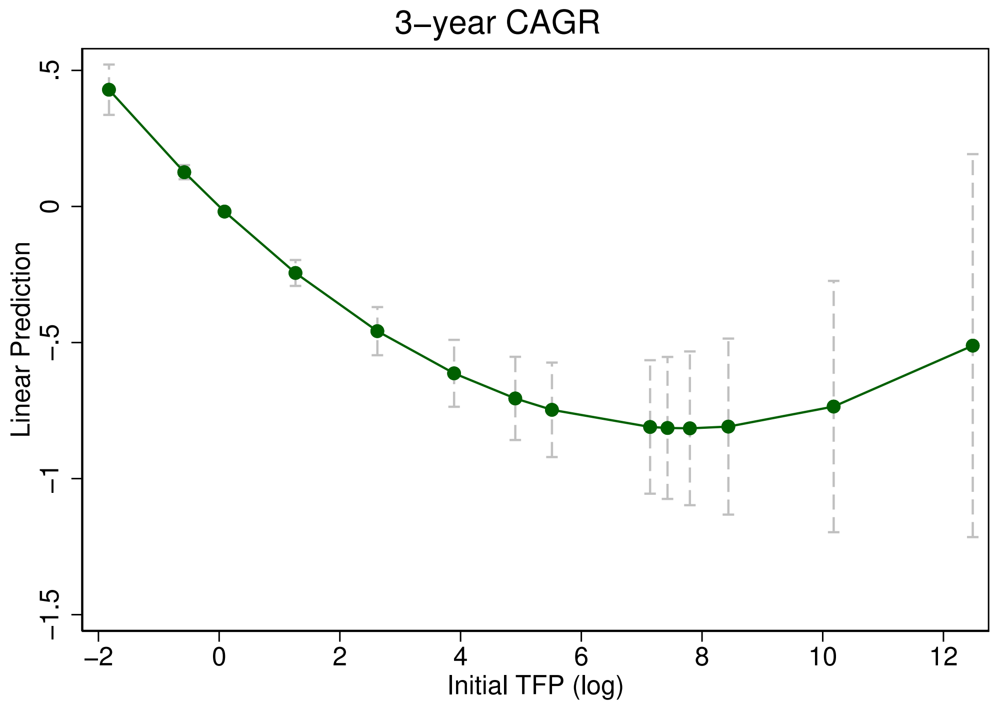
This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using the methodology suggested by Wooldridge (2009) to estimate the parameters of production function (see Online Appendix Section A.3.1 for details).

Figure A8: Expected TFP Growth, by one-digit NAICS



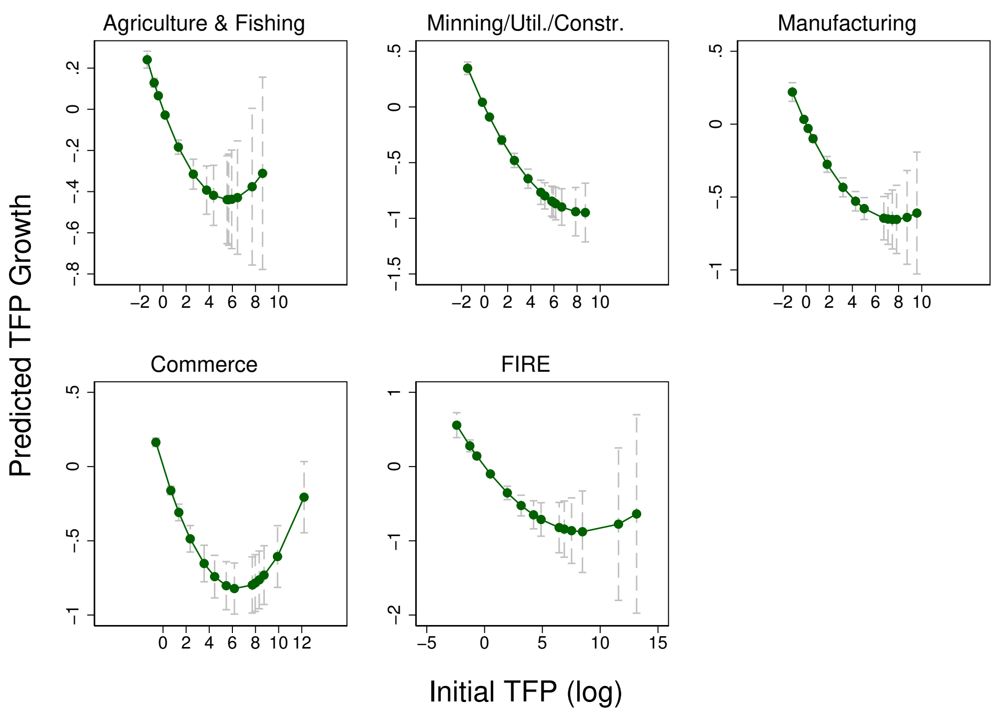
This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. It plots the three-year TFP annual growth rate for each one-digit NAICS industry. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using the methodology suggested by Wooldridge (2009) to estimate the parameters of production function (see Online Appendix Section A.3.1 for details).

Figure A9: Expected TFP Growth based on initial TFP levels



This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. The estimation controls for country-industry, industry-year, and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using the methodology suggested by Wooldridge (2009) to estimate the parameters of production function, correcting for markups (see Online Appendix Sections A.3.1 and A.3.3 for details)

Figure A10: Expected TFP Growth, by one-digit NAICS



This figure visualizes the expected TFP three-year annual growth (CAGR) as a function of baseline log TFP level using the range in the sample that goes from the 1st to the 99.99th percentile. It plots the three-year TFP annual growth rate for each one-digit NAICS industry. The estimation controls for country-industry, industry-year and country-year fixed effects, where each industry is a six digit NAICS code. TFP is estimated using the methodology suggested by Wooldridge (2009) to estimate the parameters of production function, correcting for markups (see Online Appendix Sections A.3.1 and A.3.3 for details)

BROOKINGS

1775 Massachusetts Avenue, NW
Washington, DC 20036
202-797-6000
www.brookings.edu/global