# The Cyclical Nature of the Productivity Distribution[*,**]

Matthias Kehrig[†]

23 January 2015

**Abstract**

Using plant-level data, I show that the dispersion of total factor productivity levels in U.S. manufacturing is greater in recessions than in booms. This phenomenon primarily reflects a relatively higher share of unproductive firms in a recession and is particularly pronounced in industries with higher levels of overhead inputs and industries that produce durable goods. The increased dispersion and left-skewness of productivity in a recession is characteristic of both incumbent and entering or exiting establishments, suggesting that the productivity dynamics are not solely driven by compositional changes. I construct a business cycle model where production requires overhead inputs. In a recession, the least productive establishments exit while all surviving establishments produce at a less efficient scale, which results in a more productivity distribution, which is more dispersed at the unproductive tail. The model endogenously delivers procyclical aggregate total factor productivity, entry, employment and a countercyclical relative price of durables.

[*]I am very grateful to my dissertation committee Lawrence Christiano, Martin Eichenbaum, Andrea Eisfeldt and Giorgio Primiceri for helpful discussions and encouragement. Discussions with Klaus Adam, Gadi Barlevy, Susantu Basu, Ariel Burstein, Christian Bayer, Nick Bloom, Jeff Campbell, Olivier Coibion, Nicola Fuchs-Schündeln, Alejandro Justiniano, Giuseppe Moscarini, Ezra Oberfield, Kjetil Storesletten and Nic Ziebarth as well as seminar participants at Northwestern, Wharton, Penn State, UT Austin, Duke, the Universities of Western Ontario, Montréal, Mannheim, Bonn, Cologne and Frankfurt, the Federal Reserve Board of Governors, the Federal Reserve Banks of Atlanta, Chicago, New York and Richmond, the 2010 Census RDC Research Conference, the SED Annual Meeting 2011 and the SITE Workshop 2011 have greatly benefitted this work. I would like to express my gratitude to Randy Becker, Hyowook Chiang, Cheryl Grim, Frank Limehouse and Kirk White for valuable help during my work with Census microdata at the Chicago RDC, Randy Becker for obtaining his detailed structures and equipment price deflators, Kirk White for making industry-level inventory price deflators available to me, Enrico Tan and David Wasshausen for explanations with BEA's industry-level capital stock and investment data. All errors are my own.

[**]Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

[†]University of Texas at Austin, Department of Economics, 2225 Speedway Stop C3100, Austin, Texas 78712-0301, matthias.kehrig@austin.utexas.edu. The latest version of this paper can be downloaded at http://papers.ssrn.com/abstract=1854401

# 1 Introduction

*"In the process of creative destruction [...] many firms may have to perish that nevertheless would be able to live on vigorously and usefully if they could weather a particular storm.*

— J.A. Schumpeter, "Capitalism, Socialism, Democracy," Part II, Ch. VIII.

*"Liquidate labor, liquidate stocks, liquidate farmers, liquidate real estate – it will purge the rottenness out of the system."* — Andrew Mellon, Secretary of the Treasury, 1929.

Are recessions cleansing or sullying? That is, do recessions force relatively unproductive firms to exit (the cleansing view) or do they make more firms relatively unproductive (the sullying view)? This question is theoretically ambiguous and empirical evidence on it has so far been limited. The key contribution of this paper is to shed light on this question by utilising detailed estalishment-level data to characterise both the business cycle dynamics of the cross-firm productivity distribution as well as identifying the key factors of firm survival.

Determining whether recessions are cleansing or sullying is central to understanding the inherent nature of recessions. For example, cleansing and sullying recessions have vastly different welfare implications. Sullying recessions are unambiguously negative for welfare because the economy suffers not only a downturn but also a relatively higher share of unproductive firms. Cleansing recessions, in contrast, have the positive aspect that unproductive firms exit the economy. As suggested by Schumpeter (1942), cleansing recessions can be interpreted as a necessary step in the process of "creative destruction" and thus as a defining characteristic of technological progress. In addition, whether recessions are cleansing or sullying will have very different implications for policy. Stabilisation policy has traditionally been used to minimise both the size and frequency of business cycles, but whether such a policy is desirable hinges on the assumption that business cycles are in fact sullying. As the above quote by Andrew Mellon suggests, cleansing recessions warrant a very different policy response than those that have characterised the post-World War II era.

Underlying these different views about the nature of recessions are conflicting predictions about the business cycle dynamics of productivity dispersion across firms. Schumpeterian models of creative destruction predict that productivity dispersion should be positively correlated with output since unproductive firms exit when demand is low. In contrast, in an environment where firms compete for production factors – as developed in Melitz (2003) –, the increased demand for inputs in a boom raises input prices. As a result, only the more productive firms that can afford to pay higher factor costs will survive, so that productivity dispersion will tend to be higher in recessions than booms. Thus, characterising whether productivity dispersion is procyclical (the cleansing view) or countercyclical (the sullying view) can help address one of the fundamental questions about business cycles.

Knowing the cyclicality of productivity dispersion is important for another reason: It can shed light on the relevance of cross-sectional "risk" or "uncertainty shocks" that have been proposed as a source of business cycles by Bloom (2009) and are considered to have played a significant role during the Great Recession of 2008/09. Christiano et al. (2014) for example show theoretically that, in an environment with financial frictions, cross-sectional risk shocks may explain 60% of output fluctuations. The mechanism proposed by these authors relies on a more spread-out distribution of productivity levels in recessions, but lacks the empirical micro support.[1] In the present paper, I fill this gap and provide empirical evidence for both the presence and magnitude of dispersion cyclicality up to and including the Great Recession.

Using confidential Census data, I estimate establishment-level[2] productivity in the U.S. manufacturing sector from 1972-2009. My empirical work establishes three main results: First, cross-sectional productivity dispersion is countercyclical; the distribution of total factor productivity levels across establishments is about 12% more spread-out in a recession than in a boom.[3] Second, the bottom quantiles of the productivity distribution are more cyclical than the top quantiles. In other words, the countercyclicality of productivity dispersion is mostly due to a higher share of relatively unproductive establishments during downturns. Third, the countercyclical pattern of productivity dispersion is more pronounced in durable goods industries than in non-durable goods industries. These results were obtained by estimating productivity using the methodology proposed by Olley and Pakes (1996), but they continue to hold when one infers total factor productivity from simple Solow residuals or from using alternative structural techniques. The cyclicality results also hold for several dispersion measures such as the cross-sectional variance, inter-quartile or inter-decile range.

The empirical results suggest a cyclical survival cutoff at the bottom tail of the productivity distribution so that recessions are sullying. At face value, then, Schumpeterian models seem at odds with my empirical findings. These models need not be dismissed entirely, however. For cleansing could be a feature of recessions, but other factors that lead to sullying may dominate. Cleansing recessions in basic Schumpeterian models typically arise from demand variations in partial equilibrium, which generates procyclical productivity dispersion.[4] To overcome this inconsistency, I expand them in two ways: First, I introduce a cost channel into an environment that, without the cost channel, would lead to cleansing in a recession. Second, I consider the countervailing effects of the demand and cost channels in general equilibrium. This may overturn the cleansing effect of recessions present in the partial equilibrium Schumpeterian model.

I build a model along the lines of Ghironi and Melitz (2005) in which business cycles are driven by aggregate shocks. Although I do not rule out the possibility of exogenous second moment

[1] Chugh (2011), Chen and Zha (2011) and Arellano et al. (2010) propose similar mechanisms featuring financial frictions. Bloom et al. (2012); Bachmann and Bayer (2013, 2014) propose models with real frictions that equally rely on a more spread-out productivity distribution.

[2] Following the previous literature, I use the terms “establishment” and “firm” interchangably.

[3] Using a different sampling and estimation design Bloom et al. (2012) confirm this finding although they focus on the exogenous innovation in total factor productivity rather than the overall productivity level which is the relevant measure for survival.

[4] Caballero and Hammour (1994) is the canonical model of a Schumpeterian model of cleansing. A procyclical productivity dispersion also results from Schumpeterian models extended by endogenous reallocation between surviving firms as in Barlevy (2002).

shocks as in Bloom et al. (2012); Bachmann and Bayer (2014); Christiano et al. (2014), productivity dispersion across firms in my model is entirely endogenous as in Caballero and Hammour (1994). Firms enter endogenously into the economy, differ in their productivity and are active in a durable and a non-durable goods sector. This setup makes it possible to evaluate whether the mechanism in my model can match the sectoral differences found in the data. Production in the durable goods sector requires a fixed input such as overhead labour or organisational capital. The cost for this fixed overhead input is a crucial determinant of firm profitability. As a result, only firms above a certain productivity cutoff will make non-negative profits and be active in durables. This productivity cutoff, which regulates productivity dispersion, depends positively on the price of fixed inputs and negatively on the level of demand.

I use my model to study the dynamics of productivity dispersion in booms and recessions. I model aggregate fluctuations as shocks to household preferences that raise aggregate demand. Consider a firm with productivity exactly equal to the cutoff productivity. An increase in demand increases profits. At first sight, additional profits benefit the firm at the cutoff as in Caballero and Hammour (1994). However, higher profits also increase entry. A larger number of firms raises aggregate demand for production factors. In particular, the price of overhead factors rises. This hurts the firm at the cutoff, because it may no longer be able to cover overhead cost. If that is the case, this firm will no longer be active in durables and the productivity cutoff will be higher, resulting in a more compressed productivity distribution in a boom. The model is key in showing that, for a calibration targeted to match the U.S. economy, the general equilibrium effects of factor prices dominate the effects of demand so that recessions are sullying.[5] In a partial equilibrium environment, these demand factors alone would cause recessions to be cleansing. The critical factor responsible for procyclical prices of fixed inputs is procyclcial firm entry which leads to a congestion effect in the market for corporate overhead. My mechanism for cleansing in booms is a variant of what Lucas (1978) developed in a growth setting: As the economy grows, fixed entrepreneurial inputs become more expensive, so that the least productive units exit.

My model is hence capable of replicating the three main empirical findings: The changing truncation makes the dispersion countercyclical (Result #1), it affects unproductive firms more than productive ones (Result #2) and is more pronounced in durable goods industries (Result #3). The crucial feature to deliver the above results is the fixed factor and its procyclical price. In reality, there exist a wide array of fixed input factors such as managerial labour, organisational capital, supply chains or technical know-how. In this paper, I model the fixed factor as managerial labour input and present empirical evidence for both its presence and procyclicality. As for the former, higher fixed inputs in durable than in non-durable production will show up in a production function estimation as higher returns to scale. I estimate returns to scale at the firm level and find that they are higher in durable than in non-durable goods industries. This finding lends support

[5] Buera and Moll (2012) and Petrosky-Nadeau (2013) propose alternative theoretical mechanisms where financial recessions lead to sullying rather than cleansing.

to my assumption of fixed factors playing a more important role in durables than in non-durables.

Although the model was constructed with the objective of understanding the business cycle properties of productivity dispersion, other implications of the model have support in the data. For example, the productivity cutoff, and hence aggregate TFP, is procyclical, although the source of fluctuations in my model is not disturbances to aggregate total factor productivity. This happens because, in a boom, the underlying productivity dispersion is truncated at the bottom.[6] This endogenous selection of more productive firms in a boom is stronger in durables because overhead inputs in that sector are more important. As a consequence, the productivity in durables relative to non-durables is procyclcial which results in a countercyclical average relative price of durables. The countercyclicality in the relative price of durables has been widely noted, and my model appears to provide a novel explanation for that phenomenon.[7] The model is also consistent with procyclical employment, profits and firm entry. Lastly, the truncation also implies that the cross-sectional distribution of rates of return for firms active in durables is more compressed in a boom. This conforms well with the finance literature that finds that dispersion in the cross section of stock market returns is countercyclical, see for example Heaton and Lucas (1996); Storesletten et al. (2004); Campbell et al. (2001).

The empirical part of my paper is closely related to the work by Bloom et al. (2012). Using a different sampling and estimation design, they confirm the findings in this paper although they focus on the innovations in total factor productivity than the overall levels which is the relevant measure for firm survival. Bachmann and Bayer (2014) document countercyclical dispersion of productivity *growth rates* for German firms. Davis and Haltiwanger (1990) document cyclical patterns in employment dispersion across establishments. Bartelsmann et al. (2013) use international micro-level data to empirically relate fixed inputs to productivity dispersion. My theoretical model is a general equilibrium version of the seminal work by Caballero and Hammour (1994) extended by a non-convex production technology as in Ghironi and Melitz (2005). Syverson (2004) has examined a similar model in different geographic markets. Moll (2009); Buera and Moll (2012); Petrosky-Nadeau (2013); Midrigan and Xu (2014) are examples of recent research that links a worsening in credit conditions to a wider spread of the productivity distribution. Barlevy (2002, 2003) and Ouyang (2009) propose models in which recessions truncate the productivity distribution at the top because less highly productive firm-worker matches are created due to less labour reallocation, lower availability of credit finance and less learning-by-doing. Bloom et al. (2012); Bachmann and Bayer

[6]This implication appears in a number of models featuring productive heterogeneity on the micro level, see for example Lagos (2006); Hsieh and Klenow (2009); Moll (2009).

[7]There has been a long and heated debate about different explanations for this fact. Probably the most relevant strand of research in this area is the investment-specific technological change literature. An alternative strand of research puts increasing returns to scale (in durables) at the heart of a countercyclical relative price of durables. See for example Murphy et al. (1989); Benhabib and Farmer (1996); Hall (1990); Harrison (2003). Greenwood et al. (2000) have proposed exogenous fluctuations in investment-specific technologies to explain both a countercyclical relative price of durables and fluctuations in macroeconomic aggregates. Fisher (2006) and Justiniano and Primiceri (2008) find that a large share of the volatility reduction of macro aggregates is due to a reduction in volatility of investment-specific disturbances.

(2014); Gilchrist et al. (2010); Christiano et al. (2014); Arellano et al. (2010); Chugh (2011) propose various theoretical models that are driven by exogenous second moment shocks while, in my model, the productivity distribution is endogenously countercyclical. Bachmann and Moscarini (2011); D'Erasmo and Boedo (2012) propose alternative models where productivity dispersion evolves endogenously over the business cycle.

The paper is organised as follows: Section 2 describes the data, the econometric strategy and documents the empirical findings. The empirical patterns define the puzzle that cannot be explained in existing models with aggregate disturbances. This is the theoretical challenge that is to be explained in Section 3 which lays out the model. The goal of the model is to provide a unified theoretical explanation of both the micro-level dispersion results as well as the typical macroeconomic dynamics. Section 4 concludes.

# 2 The empirics of productivity dispersion

## 2.1 Data

I use confidential establishment-level[8] manufacturing data collected by the Census Bureau which comprise the Annual Survey of Manufactures (ASM), the Census of Manufactures (CMF), the Plant Capacity Utilization Survey (PCU), the Longitudinal Business Database (LBD) and the COMPUSTAT-SSEL bridge. In addition to these data, I use industry-level data from several publicly available sources: price deflators from the NBER-CES Manufacturing Industry Database (NBER-CES)[9], various asset data from the the Capital Tables published by the Bureau of Labor Statistics (BLS)[10], the Fixed Asset Tables published by the Bureau of Economic Analysis (BEA)[11] and the Industrial Production and Capacity Utilization published by the Federal Reserve Board of Governors (IPCU)[12]. Unless otherwise noted, all datasets are at annual frequency.

From the Census of Manufactures (CMF) and the Annual Survey of Manufactures (ASM) I construct a large dataset of plants in the U.S. manufacturing sector. This panel spans the years 1972-2009 and to my knowledge it is the longest plant-level data set of a significant sector of the U.S. economy. This period contains six NBER recessions (including the most recent "Great Recession" 2008/09) which allows me to study productivity dispersion over several business cycles. Every year, there are about 60k observations (after truncating the top and bottom 1%) which covers a much larger fraction of plants than other micro-level firm datasets. Also, it covers not just

[8]Census defines an establishment as a business location whose primary activity is production. In manufacturing, this can usually be thought of as a production plant.

[9]The NBER-CES Manufacturing Industry Database is a joint program of the National Bureau of Economic Research and the Census Bureau; http://www.nber.org/nberces/.

[10]1987-2008 Capital Data for Manufacturing Industries http://www.bls.gov/mfp/mprdload.htm.

[11]Tables 3.1S, 3.1E, 3.3S, 3.3E, 3.7S, 3.7E, 3.8S and 3.8E at http://www.bea.gov/national/FA2004/SelectTable.asp.

[12]Industrial Production and Capacity Utilization – G.17; dataset compiled by the Federal Reserve; http://www.federalreserve.gov/datadownload/Build.aspx?rel=G17.

publicly traded firms as in COMPUSTAT which could potentially exhibit different productivity dispersion dynamics than privately held firms. This comparatively broad coverage reduces the risk of misleading conclusions that are based on the specificity of the sample.

I use the price deflators in the NBER-CES manufacturing data to get real quantities of output, materials and energy.[13] Furthermore, I combine this panel with the LBD and the COMPUSTAT-SSEL bridge which helps me identify additional plant characteristics. The LBD contains information about the birth and death year of all plants in the economy while with the COMPUSTAT-SSEL bridge the CMF/ASM panel can be linked to the COMPUSTAT dataset. Using these additional datasets I can differentiate further plant charactersitics such as plant age, whether a plant is an entrant, an incumbent, simply idle or an exiter and whether a plant belongs to a firm that is publicly traded, i.e. has access to equity finance. All these characteristics have been theoretically linked to productivity dispersion[14] and with my panel at hand, I can address these aspects.

Earlier versions of the CMF or ASM have been used before in a number of studies (see for example Baily et al. (1992); Ábrahám and White (2006); Hsieh and Klenow (2009); Petrin et al. (2011)). Previous research has typically focused on estimating returns to scale, the persistence of productivity or aggregate productivity growth in one or repeated cross sections.[15] To my knowledge, the present paper is the first attempt to analyse the empirical productivity distribution in U.S. manufacturing at annual frequency and to document the cyclical properties over the horizon 1972-2009. In addition to this new research interest, the data that are used in the present study span not only a longer period, but are also substantially improved (as described in detail in Appendix A) over the versions used in the above-cited research.

With its wealth of information on the plant level, the ASM/CMF panel is an excellent source to assess the dynamics of cross-sectional productivity dispersion. Still, it is not perfect: Census samples large establishments above a certain employee or asset value threshold with certainty, smaller establishments are selected with probability $p < 1$. Census chooses the sampling probability $p$ such that the inverse reflects the sampling weight, i.e., the number of establishments that the sampled observations are representative for. In my analysis, I weight observations with the inverse of $p$ to roughly replicate the underlying population of all manufacturing plants. Omitting this step would underrepresent small plants which are known to exhibit different dynamics (see for example Gertler and Gilchrist (1994); Moscarini and Postel-Vinay (2012)). I furthermore weigh observations by plant size[16] to avoid small outlier observations unduly driving my results.[17]

[13] Note that these are price deflators on the 6 digit NAICS industry level. Ideally, plant-specific prices are needed, but there is no way to get around this data limitation in the full panel, so my productivity measure contains within-industry price dispersion. This measure is commonly referred to as $TFPR$ (Foster et al. (2008); Hsieh and Klenow (2009)). In the context of this paper, in turn, revenue productivity is actually the relevant measure for firm survival.

[14] See for example Hopenhayn (1992) and Clementi and Hopenhayn (2006).

[15] Petrin et al. (2011) for example use the estimator developed by Levinsohn and Petrin (2003) to decompose aggregate TFP growth into terms reflecting technical efficiency and reallocation.

[16] The baseline specification uses plant's assets, but value added or employment deliver similar results.

[17] Even the dispersion of unweighted productivities is still countercyclcial suggesting that small plants are more volatile than large plants.

A second potential problem is that Census rotates this subsection of small establishments with $p < 1$ in order to maintain a representative sample of the manufacturing sector. This rotation happens in years ending in 4 and 9 and could thus create a quinquennial selection cycle in the productivity dispersion of small firms.[18] I control for this sampling peculiarity in Section 2.4 and results are also displayed in Table 4. Lastly, the treatment of new entrants changed in the ASM in 1989. This change is dealt with in Section 2.4.3.

Lastly, I limit my attention to the "ASM establishments" (identified by $\texttt{ET} = \texttt{0}$) in Census years (years ending in 2 and 7), when, in principle, the entire manufacturing sector is sampled.[19] This step maintains longitudinal consistency and results in a large panel: over 1972-2009, there are about 2.2 million observations in my sample which corresponds to about sixty thousand plants every year.

## 2.2 Productivity estimation

Studying cross-sectional productivity dispersion requires plant-level productivity estimates. A large literature is concerned with the estimation of production functions and productivity. I build on this literature and the methods it developed. This paper follows the vast strand of previous research and assumes a Cobb-Douglas[20] gross output production function on the plant level:

$$y_{ijt} = a_{ijt} + \beta_j^k k_{ijt} + \beta_j^l l_{ijt} + \beta_j^m m_{ijt} + \beta_j^e e_{ijt} \tag{1}$$

where $y$ denotes the log of production, $a$ total factor productivity and $k$, $l$, $m$ and $e$ are logged real inputs of capital, hours worked, material use and energy use, respectively. $\beta^k$, $\beta^l$, $\beta^m$ and $\beta^e$ are production function elasticities. The subscript index $t$ denotes time, $i$ the plant which belongs to industry $j$. Unless otherwise noted, industry denotes one of the 473 6-digit NAICS industries in the manufacturing sector. Note that the production function elasticities are industry-specific reflecting a common technological structure within an industry. Inference on $a_{ijt}$ will be described below.

The preferred specification is gross output rather than valued added. Basu and Fernald (1995) have shown that the value added specification may lead to an upward bias of production elasticity estimates if firms have market power. This upward bias is not present in the gross output specification. The construction of each input and output variable from the raw datasets is described in great detail in Appendix A. At this point, I just want to mention that materials $m$ and output $q$ have been corrected for inventory changes to accurately reflect goods used and produced rather than goods paid for and sold. Otherwise dispersion in inventory adjustment across plants would generate a changing productivity dispersion. Capital inputs are constructed corrected for

[18]This matters a lot if one is interested in the evolution of dispersion of TFP growth rates rather than TFP levels. This is also a problem if one estimates TFP in a way that requires lagged variables.

[19]Note that this procedure also drops all "administrative records." These are establishments with less than three employees, so-called "AR establishments," that are not sampled, but imputed by Census based on administrative records from IRS. This procedure follows Foster et al. (2008).

[20]Like Lee and Nguyen (2002) I have also estimated a translog production function without much differences in the productivity dynamics.

capital-embodied technical change; since this process is very specific to the asset type, I distinguish between structure-specific and equipment-specific technical change when computing the structure and equipment capital stock. Capital $k$ is the logged sum of both.[21] Lastly, note that both capital and energy are included in the production function. As pointed out by Burnside et al. (1995), the capital stock per se is not productive. Rather, it requires energy (fuels or electricity) to be utilised in production. I therefore use some form of energy to proxy for capital services.[22]

**Endogeneity and selection** Estimating production functions to learn about either returns to scale ($\beta$'s) or productivity ($a$) has been the subject of a considerable body of research.[23] My baseline specification employs the estimation technique proposed by Olley and Pakes (1996). The Olley-Pakes method is widely used in the current productivity literature because it addresses two problems: first, an endogeneity problem (contemporaneous inputs in (1) are correlated with productivity $a$) and, second, a selection problem (plants with a very low productivity $a$ will exit the industry). While previous research has developed several ways to overcome the endogeneity problem, Olley and Pakes propose the only method that overcomes the selection bias from plant exit. Exit is in fact cyclical (see Jarmin and Miranda (2002)), so the selection bias is cyclical too. In my context, a cyclical selection bias could then drive measured cyclical productivity dispersion. Since exit matters a fair amount in business cycles (the average exit rate in my data is 7%), I am worried that this selection bias could be quantitatively relevant, so my preferred method is Olley-Pakes.

This method requires information about exit of individual plants which makes the estimation of the selection step possible. I identify an exiting plant from the Longitudinal Business Database (LBD) which I match to my baseline panel constructed from the ASM/CMF. The LBD covers all establishments in the entire economy and nests the ASM/CMF observations from the manufacturing sector.[24]

**Identification and possible problems** The Olley-Pakes method relies on three main assumptions: First, productivity is a first-order Markov process. Second, capital in period $t$ is a function of last period's capital stock and investment; $k_{ijt}$ is hence predetermined while all other inputs are chosen in time $t$ after productivity $a_{ijt}$ is known. Third, investment is strictly increasing in productivity. While the Olley-Pakes method looks like the best-suited approach in the context of

[21]In a similar spirit, I also tried to distinguish between production and non-production labour, but this distinction does not matter for results.

[22]Alternatively, I used directly observed utilisation rates from the Plant and Capacity Utilisation Survey (PCU) – a small subset of the manufacturing data. This cross check delivers very similar results as using energy to proxy for the used capital stock suggesting that using energy is a good approximation of capital *services*.

[23]See for example Hall (1990); Basu and Fernald (1995, 1997); Benhabib and Farmer (1996); Burnside (1996); Harrison (2003) in the macro literature and Blundell and Bond (2000); Olley and Pakes (1996); Levinsohn and Petrin (2003); Ackerberg et al. (2006) in the Industrial Organisations literature.

[24]A small set of observations in the ASM/CMF appears before (after) it appears (disappears) in the LBD. According to the creators of the LBD, this is a well-known, yet small problem. By default, I stick to the information from the LBD, but if an observation shows up in the manufacturing data but not in the LBD and the variables exist, I adjust the entry/exit year accordingly.

this paper, it is not free of problems. The assumption that investment is strongly increasing in productivity will easily be violated in the presence of non-convex adjustment costs which lead to lumpy investment; those observations (with $i_t = 0$) have to be omitted. Although this is a valid concern, it is fortunately not a very pressing problem in my data (the share of observations with $i_t = 0$ is about 15%).

A second downside of the Olley-Pakes method is the assumption that technical efficiency differences are the only factor underlying profitability differences. But it is plausible to think that, in a response to a (firm-specific) demand shock, the firm raises prices rather than production. Prices on the micro level are not available to me, so my productivity measure will be contaminated by within-industry price dispersion. Mapping my baseline specification to the observables in the data, I can write

$$y_i = a_i + \beta^k k_i + \beta^l l_i + \beta^e e_i + \beta^m m_i \tag{1}$$

$$\Leftrightarrow p_i - \overline{p} + y_i = \underbrace{\overbrace{a_i}^{TFPQ_i} + (p_i - \overline{p})}_{TFPR_i} + \beta^k k_i + \beta^l l_i + \beta^e e_i + \beta^m m_i \tag{1'}$$

The value of production that I observe in the data is the left hand side of equation (1'), so the resulting Olley-Pakes estimate will be $TFPR_i = a_i + p_i - \overline{p}$, an object commonly called "revenue productivity" or profitability (see Foster et al. (2008); Hsieh and Klenow (2009)) which is not to be confused with technical efficiency, $TFPQ_i = a_i$. As Foster et al. (2008) mention, it is $TFPR_i$ which matters for plant survival rather than $TFPQ_i$. To answer whether recession are cleansing or sullying, the central question of the present paper, $TFPR_i$ the right dispersion measure to look at. Although I am unable to identify dispersion in technical efficiency, I am still able to identify the (revenue) productivity measure that is relevant in this context.

Nevertheless, I want to make sure that my results are not specific to the structural setup of the Olley-Pakes method with its assumptions about investment monotonicity and timing. In Appendix C.3, I infer productivity from Solow residuals and find that the dispersion dynamics of Solow residuals are preserved.

**Constructing the dispersion measure** Plant-level total factor productivity is jointly estimated with returns to scale in the above procedure. It corresponds to $a_{ijt}$ in equation (1). The object of interest is the cross sectional dispersion of $a_{ijt}$ within each of the 473 industries. I want to collapse these 473 dispersion measures into one statistic that reflects the average dispersion in the manufacturing sector. In order to avoid certain outlier industries dominating this representative measure, I correct dispersion in every industry by some industry characteristics. These corrections are described in detail in Appendix B.1, here is a brief overview: First, I correct industry dispersion for an industry-specific growth trend.[25] The remaining portion is denoted by $z_{ijt}$; changes in the

[25] Otherwise industries with very strong or very weak long-run productivity growth will essentially drive my results.

dispersion of $z_{ijt}$ will reflect business cycle changes around the long-run dispersion trend. Lastly, I recenter $z_{ijt}$ and scale it by its long-run standard deviation $\sigma_j$[26]. The resulting normalised industry dispersion is defined as:

$$Sd_{jt}\left(\frac{z_{ijt}-\overline{z}_j}{\sigma_j}\right)$$

To organise my results, I report the dispersion of the median of the 473 industries[27] for each period $t$; one measure for non-durable manufacturing ($Disp_t^n$), one for durable industries ($Disp_t^d$). The split into non-durables and durables is motivated by the different time series properties of neutral and investment-specific technological progress (see Greenwood et al. (1997); Fisher (2006)). If productivity behaves differently in non-durables and durables on the aggregate level, productivity dispersion dynamics on the micro level, too, might be different.

$$Disp_t \equiv Median_t\left[Sd_{jt}\left(\frac{z_{ijt}-\overline{z}_j}{\sigma_j}\right)\right] \tag{2}$$

## 2.3 The empirics of cross-sectional productivity dispersion

I now turn to the results of the empirical investigation. Recall that $Disp_t$, the typical within-industry dispersion of productivity levels, will decrease in a recession if strict cleansing would take place, while $Disp_t$ will increase in a recession if recessions are sullying. The two time series $Disp_t^n$ and $Disp_t^d$ as defined in (2) for both non-durables and durables are plotted in Figure 1 and their time series properties are displayed in Table 1. One can easily see that dispersion fluctuates over time. In fact, it is almost as volatile as durable production (i.e. investment goods) and almost three times as much as GDP. Dispersion in durables is slightly more volatile than dispersion in non-durables (about 1.3 times as volatile) which is not surprising given the overall larger fluctuations in the durables than in non-durables. Dispersion in durables is also more more persistent (autocorrelation 0.34) than dispersion in non-durables (autocorrelation 0.07).

### Result 1: Productivity dispersion is countercyclial

Figure 1 displays the HP filtered residuals of logged $Disp_t$ measure defined in (2), that is, one can interpret the values as percentage deviations of dispersion from its non-linear trend. The main objective of this paper is to study if the business cycle has an impact on the productivity distribution as predicted by cleansing or sullying theories described in the introduction. The basic cleansing theories would predict that dispersion should decrease in a recession and expand in a

[26]Otherwise, industries where productivity is very spread-out in general like cement will dominate my results. In contrast to correcting for the long-run growth and the scaling by long-run standard deviation, additionally taking out the long-run skewness hardly makes a difference suggesting that higher-order moments do not matter as much.

[27]Considering the mean dispersion across industries rather than the dispersion of the median industry leads to basically the same results.

Table 1: Productivity dispersion: Summary statstics

| Variable $X$ | Corr$(X_t, X_{t-1})$ | $\sigma$(X) | $\frac{\sigma(X)}{\sigma(Y)}$ | $\frac{\sigma(X)}{\sigma(GDP)}$ |
|---|---|---|---|---|
| $Disp_t^n$ | 0.069 | 0.043 | 1.662 | 2.002 |
| $Disp_t^d$ | 0.341 | 0.053 | 0.889 | 2.487 |
| $Y_t^n$ | 0.409 | 0.026 | 1 | 1.204 |
| $Y_t^d$ | 0.394 | 0.060 | 1 | 2.798 |
| $GDP_t$ | 0.505 | 0.021 | – | 1 |

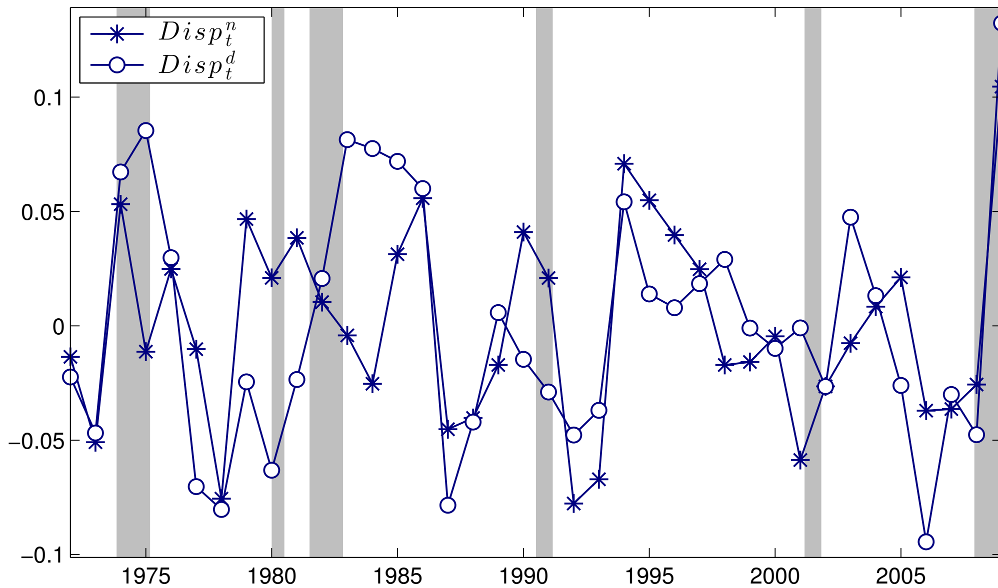
*Note:* $Y_t^n$ denotes the output of non-durable, $Y_t^d$ that of durable manufacturing. They were obtained from the NBER-CES manufacturing database by dividing the value of shipments (`VSHIP`) by the price deflator (`PISHIP`) for each industry and summing up the resulting real output within all non-durable manufacturing industries (NAICS 311111-316999 and NAICS 322111-326999) to obtain $Y_t^n$; analogous procedure for durable output $Y_t^d$ (NAICS 321111-321999 and NAICS 327111-339999). $GDP$ is obtained from the NIPA tables by dividing nominal gross domestic product by the GDP deflator. All variables haven been HP-100 filtered.

boom. But even a casual glance at Figure 1 indicates that most peaks in productivity dispersion coincide with recessions as defined by the NBER (indicated as shaded bars). This looks especially true for the early recessions in the sample: 1974, 1980, 1982 and again – and most strongly – in 2008/09. As Figure 1 shows, the productivity distribution in durables increases by 6.7% (about 1.4% in non-durables) on average in the six NBER recessions in the sample. The increase in dispersion is strong in the 1974 and 1980/82 recessions: From the previous peak to the trough of the recession dispersion in durables increases by 13% and 16% respectively. During the 1990's and 2000's recessions, dispersion does not increase as much and is not as correlated with the cycle as before. In the "Great Recession" of 2008/09, however, the dispersion increase is the largest ever and pervasive through durables and non-durables: at the bottom of the recession in 2009, the distribution is about 20% more spread-out than at the peak in 2007. This shows that risk shocks that have recently been proposed as a business cycle driver (Bloom (2009); Bloom et al. (2012); Christiano et al. (2014); Chen and Zha (2011)) might have a prominent role in explaining the latest recession. To my knowledge, this is the first evidence of a more spread-out distribution in productivity levels, thus providing the empirical foundation for the models proposed by these authors.

To formalise the evidence on cyclicality, I correlate the business cycle components of output and dispersion. The resulting scatter plot in Figure 2 shows the negative relationship between production (in durable and non-durable manufacturing respectively) and productivity dispersion. The negative relationship is pervasive but more pronounced in durables – a fact that is confirmed in the Correlograms (displayed in Figure 4). In durable manufacturing, the correlation between output and productivity dispersion is -0.503 and statistically significant at the 95% level. While the correlation between output and productivity dispersion in non-durable manufacturing is negative (-0.293) it is only borderline significant.

Below, I will display the results of several robustness checks. First, I will check if various

Figure 1: Dispersion in productivity levels




*Note:* Time series plot of the cyclical components (HP-100 filtered annual time series) of productivity dispersion in non-durable goods producing (stars) and durable goods producing (circles) manufacturing. The dispersion measure is as defined in equation (2). Shaded bars denote NBER recessions.

measures of productivity dispersion such as the inter-quartile range, the variance etc. are countercyclical as well. Second, I will check if dispersion is still negatively correlated with various output measures (manufacturing output, GDP) and different filtering techniques.

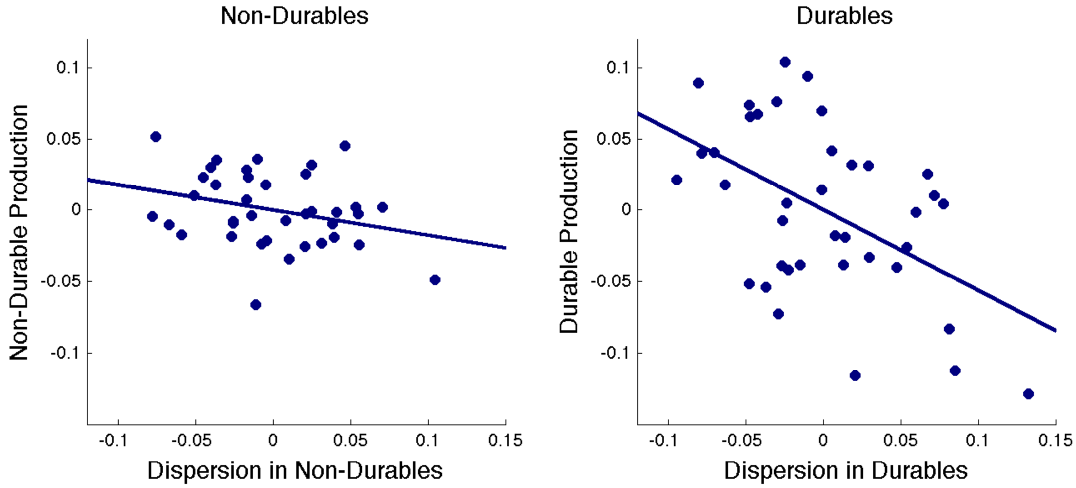
**Result 2: Unproductive units drive dispersion dynamics**

As shown above, the productivity distribution as a whole is negatively correlated with output. This result means that the distribution is *more dispersed* in a recession. This pattern could be explained by the overall distribution fanning out as conjectured by Davis and Haltiwanger (1990). Alternatively, this countercyclicality is consistent with most movements happening at the top of the distribution (as in Gabaix (2011))[28] or at the bottom of the distribution (as in Ghironi and Melitz (2005)). To that end, I look at the behaviour of individual quantiles over recessions.

The cleansing view postulates a tougher truncation at the bottom end of the productivity distribution in a recession, i.e. one should see the bottom quantiles picking up in a recession. Figure 3 paints a different picture, however. First, it is predominantly the bottom end of the productivity distribution that changes in a typical NBER recession. Second, these plants *decrease*,

[28] Gabaix (2011) looks at firm- rather than plant-level data. So it is plausible that some very large high-productivity firms dominate dynamics.

Figure 2: Plot of productivity dispersion and output variations

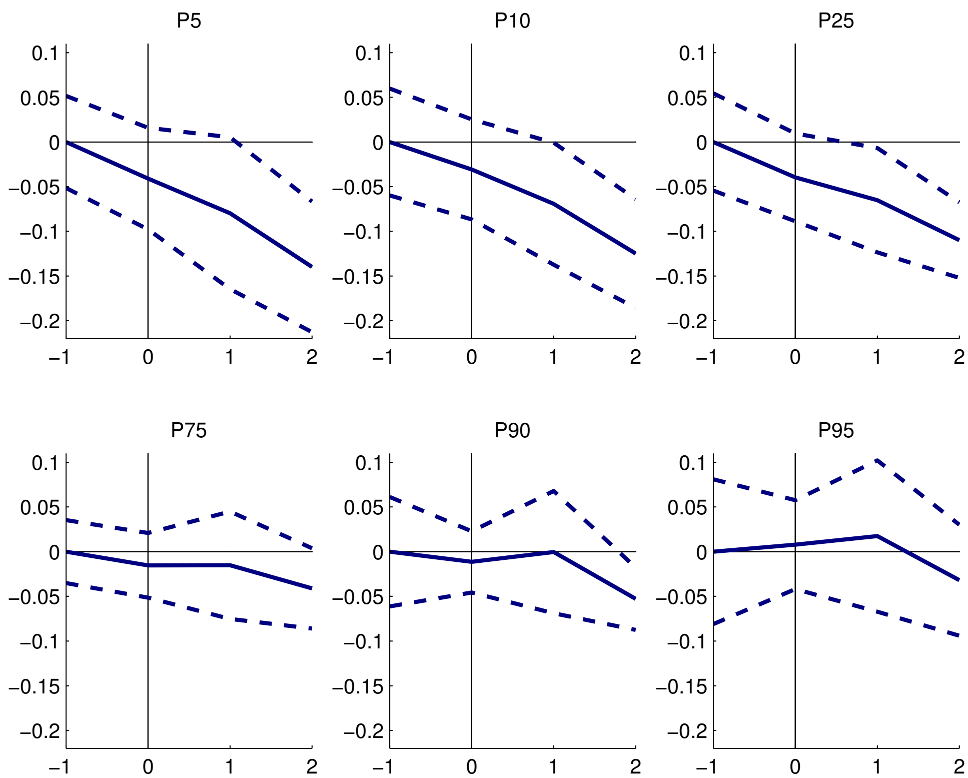



i.e. the unproductive tail of the distribution becomes even less productive. This means that the unproductive surviving plants tend to be less efficient in recessions than in booms. The most productive plants in the cross section, in contrast, barely change their productivity over the business cycle. The strong dynamics at the bottom end of the productivity distribution are consistent with the view that a truncation is changing over the business cycle. Such a truncation has been proposed in a number of models: see Caballero and Hammour (1994); Melitz (2003); Melitz and Ottaviano (2008) just to name a few. The evidence that lower quantiles tend to be procyclical suggests that this truncation is *higher* in a boom than in a recession. This result is inconsistent with the view proposed in Caballero and Hammour (1994) that first, the productivity dispersion is procyclical and that, second, the truncation should be higher in a recession. The latter implies that the lower quantiles are countercyclical rather than procyclical. Result 2 conforms well with findings in the literature on product-level pricing and labour income. Recall that my productivity measure is revenue productivity which contains the ratio of plant-level to industry-level prices. Berger and Vavra (2011) find that price change dispersion is countercyclical which comes predominantly from prices decreases at the lower tail of the distribution rather than price increases at the top end. Guvenen et al. (2012) examine the dynamics of the labour income distribution and find that in a recession the distribution is more skewed towards the bottom.

**Result 3: The dispersion in durables is more countercyclical than in non-durables**

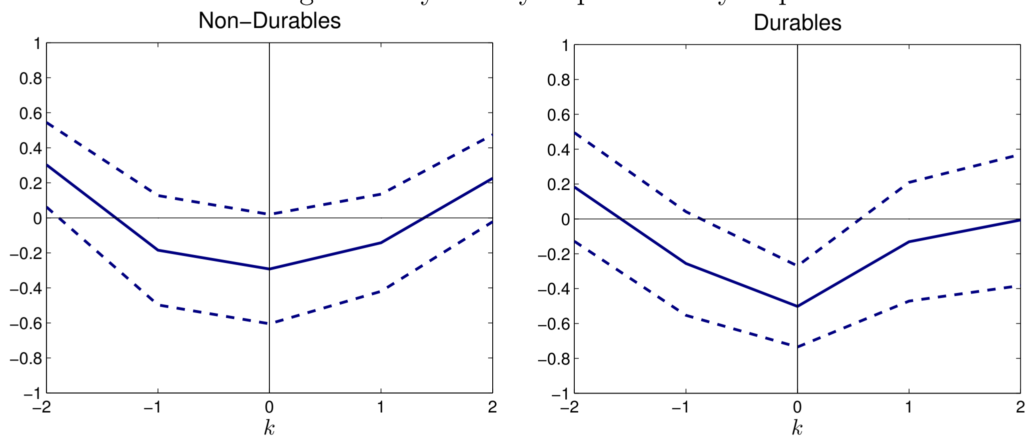
The results of the correlations between output and dispersion in the two sectors is displayed in in Figure 4 and Table 2. As we can easily see from those graphs, the productivity dispersion in more pronouncedly and significantly countercyclial in durables than in non-durables. The estimated contemporaneous correlation is $-0.502$ with a standard error of 0.119. This makes the result

Figure 3: Behaviour of quantiles over NBER recessions




*Note:* Each panel displays the behaviour of the indicated quantile over NBER recessions. The graphs are obtained by constructing a time series for each quantile from an industry's productivity distribution that has been corrected for industry-specific productivity growth and long-run dispersion and skewness. As above, I use the quantile of the median industry. For the time series of each quantile I cut out a subsample (a "worm") that starts one year before the onset of a NBER recession (denoted by $-1$) and lasts until years after the onset of a recession (denoted by 2). The six worms, one for each NBER recession, are normalised to zero in the year before the recession. Solid lines denote the mean of each "recession worm," dashed lines are the standard deviation over the six recessions.

Figure 4: Cyclicality of productivity dispersion




*Note:* Correlograms of the cyclical components (HP-100 filtered annual time series) of output and productivity dispersion in non-durable (left) and durable (right) goods producing manufacturing respectively: $Corr(Y_t, Disp_{t+k})$. The construction of productivity dispersion is detailed in Appendix B.1. Dashed lines denote 95% confidence intervals computed using GMM controling for autocorrelation and heteroscedasticity à la Newey and West (1987).

significant at the 95% level. This negative correlation is also present – albeit weaker – with a one-year lead suggesting that upshots in dispersion predate contractions. The correlation is slightly less negative and the error bands wider, but the negative correlation is still significantly different from zero at the 90% level.

Table 2: Cyclicality of productivity dispersion

| Lead/Lag | Correlation of output and dispersion in ... | |
|---|---|---|
| in years | Non-Durables | Durables |
| -2 | 0.303** | 0.183 |
| | (0.123) | (0.159) |
| -1 | -0.185 | -0.256* |
| | (0.159) | (0.152) |
| 0 | -0.293* | -0.502*** |
| | (0.159) | (0.119) |
| 1 | -0.143 | -0.131 |
| | (0.142) | (0.174) |
| 2 | 0.226* | -0.006 |
| | (0.127) | (0.192) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

In non-durables goods industries, the overall cyclicality pattern is similar, but with a correlation of $-0.293$ weaker and only significant at the 90% level. Also, it is far from being countercyclical at several leads and lags. The standard error bands are so much wider than in the durable goods industries that it is impossible to reject the null hypothesis of no correlation at the 95% level or even the 90% level. This probably weighs more importantly: While the dispersion dynamics in durables are significantly countercyclical, the dynamics of dispersion in non-durables are not. This is the second key result of the empirical work on productivity dispersion.

Results 1-3 taken together have an important implication: The countercyclical dispersion is predominantly driven by changes at the bottom tail. If one were to aggregate up micro-level productivity, aggregate productivity would then be *endogenously* procyclical.[29] Viewed through the lens of a real business cycle model, this is an empirical micro foundation of technology shocks. Since I find the micro-level dispersion dynamics to be more pronounced in durables, Results 1-3 can be combined to an empirical micro foundation of technology shocks that are *specific to the production of investment goods*. Investment-specific technological fluctuations have recently attracted attention in the literature as a potential business cycle driver (see Justiniano et al. (2010, 2011); Fisher (2006)). In Section 3 below, I'll capture this micro foundation theoretically and show how my model can provide a unified explanation of micro-level dispersion dynamics and aggregate technolgoy.

Note that durable and non-durable goods industries are commonly found to differ along another dimension: The estimated returns to scale are higher in durables then in non-durables. This has been found in previous research (for example Burnside (1996); Harrison (2003) who sectoral or sub-sectoral data) and I confirm this result in Appendix C.1 for my plant-level data. Higher returns to scale in durables suggest that unobserved fixed factors are higher in durables than in non-durables. I shall use this finding to motivate higher fixed factors in durable production which are an important driver of cross-sectional productivity dispersion.

## 2.4 Robustness checks

### 2.4.1 Alternative dispersion measures

Up to this point I have only studied the dynamics of the cross-sectional standard deviation. This cross-sectional measure could be entirely driven by a small number of outlier observations. I want to make sure that the countercyclcial dispersion is pervasive throughout the distribution. To this end, I also study the cyclical patterns of other dispersion measures such as the variance, the inter-quartile and inter-decile range. The results are displayed in Table 3 and confirm the previous results: dispersion is strongly and significantly countercyclical in durables, but only weakly and partially signifiant in non-durables. In line with Result 2 (most action comes from the bottom tail) on would expect the countercyclicality to become stronger by looking at the inter-quartile and inter-decile range, a result that is borne out in the data.

[29] Keep in mind that I used the Census sampling weights and firm size when computing the cross-sectional standard deviation in (2). This is important when thinking about aggregates.

Table 3: Correlation of output and several measures of productivity dispersion

| Correlation of output in with cross-sectional... | Non-durables | Durables |
|---|---|---|
| *A. Olley-Pakes estimate* | | |
| Standard Deviation | -0.293* | -0.502*** |
| | (0.159) | (0.119) |
| Variance | -0.286* | -0.455*** |
| | (0.166) | (0.141) |
| Inter-quartile range | -0.311* | -0.509*** |
| | (0.179) | (0.130) |
| Inter-decile range | -0.058 | -0.540*** |
| | (0.166) | (0.120) |
| *B. Solow Residuals (see Appendix C.3)* | | |
| Standard Deviation | -0.202 | -0.490*** |
| | (0.181) | (0.095) |
| *C. Levinsohn-Petrin estimate* | | |
| Standard Deviation | -0.172 | -0.420*** |
| | (0.321) | (0.115) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

#### 2.4.2 Alternative output measures

In addition to checking the robustness across different dispersion measures, I now check the robustness through different output measures. All of these results are displayed in Table 4. So far, output was defined as the production of the respective subsector of manufacturing. While this is the most plausible measure to matter for the productivity dispersion of the surviving plants, manufacturing – though very cyclical – is only one sector in the economy. To check if dispersion dynamics are correlated with the overall business cycle, I correlate dispersion with HP-filtered GDP. Following Ravn and Uhlig (2002), I choose the HP coefficient to be 6.25 rather than the more conventional 100. To avoid the implicit assumptions of any filtering technique, I also look at the correlation of productivity dispersion and aggregate output growth. All different output measures are negatively correlated with dispersion and the durable-non-durable split is still present.

Lastly, I am concerned about years ending in 4 and 9 when Census rotates the sample of small plants. I am worried that the firms entering the sample in one of those sample rotations could be systematically different in their productivity. In fact, one of the criteria for updating the sample is

Table 4: Correlation of Productivity Dispersion and Several Output Measures

| Correlation of dispersion in with... | Non-durables | Durables |
|---|---|---|
| Production – HP filtered ($\lambda = 100$) | -0.293* | -0.502*** |
| | (0.159) | (0.119) |
| Production – HP filtered ($\lambda = 6.25$) | -0.273* | -0.511*** |
| | (0.146) | (0.121) |
| Production growth rate | -0.160 | -0.442*** |
| | (0.186) | (0.163) |
| GDP – HP filtered ($\lambda = 100$) | -0.336* | -0.537*** |
| | (0.172) | (0.122) |
| GDP – HP filtered ($\lambda = 6.25$) | -0.183 | -0.439*** |
| | (0.152) | (0.127) |
| GDP growth rate | -0.250 | -0.413** |
| | (0.190) | (0.168) |
| No. of NBER boom quarters/year | -0.209 | -0.408** |
| | (0.171) | (0.163) |
| Census rotation years dropped | -0.306** | -0.496*** |
| | (0.143) | (0.124) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Description of standard errors see Figure 4.

to maintain a representative size distribution.[30] If large and small plants have different productivity distributions, then this could artificially change the productivity distribution in years 4 and 9 – all of which (except 1974) are boom years when I find the distribution to be more compressed. When I drop these years from my time series, dispersion is still significantly countercyclial, thus suggesting that the plant selection induced by Census does not distort the underlying dispersion dynamics. More robustness checks are conducted in Appendix B.2.

### 2.4.3 Extensive versus intensive margin

The universe of plants is made up of different types of plants each of which has a distinct productivity distribution. The most important types of plants in that respect are entering, incumbent and exiting plants. Most theoretical models predict that exiting plants have lower productivity while entering plants – that typically embody newer vintages – have higher productivity.[31] Previous empirical

[30] In the non-rotation years, Census tries to merely adjust the sample in a way to account for entry and exit. Census attempts to keep the age distribution of plants representative.

[31] This is mostly true for "successful entrants," plants that enter and are productive enough to remain in a competitive environment. "Unsuccessful entrants," in contrast, are plants that enter and are so unproductive that they have to leave immediately.

work has established that entry and exit vary systematically with the business cycle (see Lee and Mukoyama (2011) for plant entry and exit and Davis and Haltiwanger (1990) for job creation and destruction) although the magnitude of entry and exit is debated.

Is the productivity distribution more spread-out in a recession because productivity changes across all plants (intensive margin) or because the composition of plant types systematically changes over the business cycle (extensive margin)? The countercyclical dispersion could result from procyclical entry if the productivity distribution of entrants has similar support to that of incumbent plants but is more concentrated. An analogous argument holds for exit.

Table 5: Summary Statistics: Share of plant types

| Share of... | Non-Durables | Durables |
|---|---|---|
| Entrants | 5.5% | 5.4% |
| Exiters | 6.2% | 6.0% |
| Idle | 0.7% | 0.7% |
| Active | 87.6% | 87.9% |
| No. annual obs. | 26,510 | 28,395 |

One class of models stresses the importance of the intensive margin where the productivity of all plants – entering, exiting and incumbent – is more spread-out in a recession. The interpretation of cross-sectional productivity dispersion as risk shocks lies at the heart of the this view. Another class of models emphasise the extensive margin where cyclical entry and exit are the main driver of productivity dispersion while the productivity of individual surviving plants is fixed (e.g., Caballero and Hammour (1994)). In my sample, entrants and exiters make up about 12% of all observations; this might be strong enough to impact the dispersion dynamics of the entire sample.

To check if compositional bias is driving my results, I split the sample into incumbent plants on the one hand and entering/exiting[32] plants on the other hand. If dispersion dynamics were solely driven by the extensive margin, the sample of incumbent plants would not exhibit dispersion dynamics. If, on the other hand, dispersion dynamics are pervasive throughout all plants, both subsamples should exhibit the countercyclical dispersion.

As Table 6 shows, both groups exhibit similar dispersion dynamics. In the group of incumbent plants, productivity dispersion is countercyclical in durables and (again weakly so) in non-durables. Both volatility and cyclicality among incumbent plants resemble that of the overall sample which one would expect given that incumbents make up 87% of the overall data. The subsample of entrants and exiters, in contrast, is much more volatile than the overall sample. The cross sectional standard deviation fluctuates about four to five times as much as that among incumbent plants. Dispersion among entrants/exiters is also more pronouncedly countercyclical than that among incumbent plants, especially in non-durables which is in contrast to the weak cyclicality

[32] An exiting plant is defined as a plant that is active, but will exit at the end of this period.

Table 6: Productivity dispersion – extensive and intensive margin

| | Non-durables | Durables |
|---|---|---|
| *A. Volatility* | | |
| Full sample | 0.043 | 0.053 |
| Incuments | 0.045 | 0.049 |
| Entrants/Exiters | 0.163 | 0.196 |
| *B. Cyclicality* | | |
| Full sample | -0.293* | -0.502*** |
| | (0.159) | (0.119) |
| Incumbents | -0.299* | -0.482*** |
| | (0.168) | (0.132) |
| Entrants/Exiters | -0.577*** | -0.220* |
| | (0.078) | (0.122) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Volatility is the standard deviation of the dispersion time series. Cyclicality denotes the correlation between manufacturing output and dispersion (non-durable and durable, respectively). Description of standard errors see Figure 4.

of productivity dispersion among non-durables incumbents.[33] Because the countercyclical dispersion is pervasive throughout incumbents, entrants and exiters, we can rule out that productivity dispersion is *solely* driven by compositional changes. Productivity changes along the intensive margin must take place. The much stronger volatility across entrants/exiters, however, suggests that changes along the extensive margin amplify the countercyclical dispersion.

# 3 The model

The empirical work presented in the previous section has established three main results: the measured cross-sectional productivity dispersion is countercyclical, unproductive establishments are the main driver of the dynamics of the productivity distribution, the countercyclicality is stronger in durables than in non-durables and it is stronger in industries with greater importance of overheads. The standard cleansing-of-recessions view has trouble addressing the first two facts. The goal of this theoretical section is to develop a business cycle model that is consistent with the empirical findings on the micro level without ruling out cleansing a priori. In contrast to previous research,

[33]Lee and Mukoyama (2011) find that entrants in recessions tend to be more productive entrants in booms. To be consistent with my results about the entire universe of firms, dispersion of exiters and incumbents must be strongly countercyclical. Note that the difference may also stem from the shorter time horizon they are using.

standard shocks to aggregate total factor productivity (TFP) drive business cycles in my model.[34] Given the importance of dynamics at the bottom of the distribution, my model will feature a fixed overhead factor of production as in Melitz (2003); Ghironi and Melitz (2005) which will give rise to a truncation of the productivity distribution from below. The presence of a fixed factor is supported by the evidence that managerial labour is an important driver of the productivity distribution and my estimation of returns to scale (see Appendix C.1). All evidence points into the direction that overhead inputs are more prevalent in durables; this is precisely that manufacturing subsector whose productivity dispersion is most cyclical.

The model features endogenous entry in the economy (exit is random) and firm-level heterogeneity in productivity. In the model, I adopt the assumption from Olley and Pakes (1996) that all cross-sectional profitability differences are driven by underlying technological differences ($a_i$ in equation (1')) rather than mark-up differences ($p_i - \overline{p}$ in equation (1')). I am aware that this interpretation of my empirical results is specific in that it ignores possible changes in mark-ups and a varying degree of competition, both of which probably are relevant in the real world. These elements could be added to the model [35], but the more simple model version chosen here is sufficient to explain the empirically observed results.

## 3.1 Final producers

A continuum of perfectly competitive final goods firms combines intermediate inputs to one homogeneous final good that can be sold to final consumers:

$$Y_t = \int_{i=0}^{N_t} y_{it} di$$

where $N_t$ is the measure of varieties available at time $t$ and $y_{it}$ is the input of intermediate variety $i$. Since final producers can perfectly substitute one variety for another – the implied elasticity of substitution is infinity – the price of every variety must be the same: $p_{it} = p_{jt} \quad \forall i, j \in [0, N_t], i \neq j$. I further assume that final output is the numéra ire and normalize its price to unity.

## 3.2 Intermediate producers

### 3.2.1 Production

A continuum of perfectly competitive intermediate goods firms each produces output that is sold to the final goods producers. In period $t$, there is a measure $N_t$ of intermediate producers active. Each intermediate goods firm hires labour $l_{it}$ and rents capital $k_{it}$ in order to produce intermediate

[34] Models that are driven by idiosyncratic productivity shocks are Gabaix (2011); Christiano et al. (2014, 2010); Bloom et al. (2012); Bachmann and Bayer (2013, 2014).

[35] I could introduce imperfections – "wedges" or deviations in the firms' first-order condition as in Hsieh and Klenow (2009) or varying substitutability between products as in Melitz and Ottaviano (2008); Foster et al. (2008) – that lead to a changing mark-up dispersion across firms when aggregate conditions change.

output $y_{it}$:

$$y_{it} = A_t z_{it} k_{it}^{\alpha} l_{it}^{\gamma}$$

where $A_t$ is an aggregate productivity level, $z_{it}$ is the firm's idiosyncratic productivity and $\alpha$ and $\gamma$ are the production elasticities of all inputs.[36] In line with the empirical results, I assume overall decreasing returns to scale $\xi \equiv \alpha + \gamma < 1$.

Business cycles will be driven by fluctuations in the common productivity component, $A_t$. Again in line with the empirical results, I assume that each of the two productivity components follows an AR(1) stochastic process:

$$\log A_t = (1 - \rho^A) \log \overline{A} + \rho^A \log A_{t-1} + \eta_t \tag{3}$$

$$\log z_{it} = (1 - \rho^z) \log \overline{z}_i + \rho^z \log z_{it-1} + \varepsilon_{it} \tag{4}$$

where $\eta_t$ and $\varepsilon_{it}$ are the aggregate and idiosyncratic components of a technology shock, respectively. Both $\eta$ and $\varepsilon$ are drawn from a time-invariant distribution, $F_\eta$ and $F_\varepsilon$ respectively. The distribution of idiosyncratic technology shocks $\varepsilon_{it}$ is time invariant and, importantly, independent of the level of aggregate technology $A_t$.[37] $\overline{A}$ and $\overline{z}_i$ are the long-run averages of each productivity component.

Key to the production of intermediates is a fixed overhead input, denoted $c_f$, which must be hired every period by all firms regardless of their size, productivity or other characteristics. Since overheads are required to produce at all but are not related to the production amount, it is hard to identify their impact in the data.[38] Part of why that is the case is that overheads might be required in terms of any production input – capital, labour, materials, blueprints, reputation, legal licenses – or any combination of those. Examples for capital overheads could be the financing and maintenance cost of capital structures; examples of overhead labour are managerial workers, human resources, finance, organization, advertising or management etc. Regardless of their origin, these costs are most appropriately modeled as independent of the firm's production volume decision. Though overheads are ubiquitous, I will focus on overhead labour, e.g. managers, in the model. In the data, I found a way to attribute some of the overheads on non-production labour, so this choice seems supported empirically.

Conditional on being active and hiring overhead inputs, every firm $i$ optimally chooses production labour, $l_{it}$, and capital, $k_{it}$, in perfectly competitive markets in order to maximize static

[36] For simplicity, I omit materials as in input in the model. For an example of a model with materials and an explicit network structure see Gabaix (2011).

[37] Several recent authors (see for example Bloom (2009); Bloom et al. (2012); Arellano et al. (2010); Schaal (2012); Vavra (2014); Christiano et al. (2014)) assume that the variance of $\varepsilon$ is larger when the level of $A$ is low. Though not necessary to generate the countercyclical dispersion in TFP levels documented earlier, this feature could be added and would reinforce the countercyclical nature of TFP levels.

[38] Overhead inputs are similar in nature to "intangible inputs" in that the econometrician observes more their indirect than their direct effects. Yet, researchers attribute important firm decisions to them: Atalay et al. (2014), for example, propose intangibles as a main driver of decisions of vertical integration, and Eisfeldt and Papanikolaou (2013) propose "organizational capital" – a concept similar to overhead inputs – as an important driver of a firm's investment and financing decisions.

profits

$$\pi_{it} = A_t z_{it} k_{it}^{\alpha} l_{it}^{\gamma} - r_t k_{it} - w_t l_{it} - \omega_t c_f$$

where $r_t, w_t, \omega_t$ are the prices of capital, labour and the overhead input in period $t$. Standard profit maximization leads to the firm's factor demand:

$$r_t = \alpha A_t z_{it} k_{it}^{\alpha-1} l_{it}^{\gamma} \quad \Leftrightarrow \quad k_{it} = \left[ A_t z_{it} \left( \frac{\alpha}{r_t} \right)^{1-\gamma} \left( \frac{\gamma}{w_t} \right)^{\gamma} \right]^{\frac{1}{1-\xi}} \tag{5}$$

$$w_t = \gamma A_t z_{it} k_{it}^{\alpha} l_{it}^{\gamma-1} \quad \Leftrightarrow \quad l_{it} = \left[ A_t z_{it} \left( \frac{\alpha}{r_t} \right)^{\alpha} \left( \frac{\gamma}{w_t} \right)^{1-\alpha} \right]^{\frac{1}{1-\xi}}. \tag{6}$$

Note that firms will not optimize over overhead inputs as they are required as long as the firm is active at all.

### 3.2.2 Profitability, exit and entry

While the previous section described the optimal behavior of active firms, I also model the life-cycle pattern of firms. This seems crucial because entry and exit are at the heart of the Schumpeterian models of business cycles. They rest on the idea that recessions are times when unproductive firms become unprofitable and are "cleansed out" of the economy. Obviously, productivity-driven exit has direct consequences for the cyclical nature of the productivity distribution. A key objective of my model is hence to feature endogenous exit along profitability. Although they are related in the present model, profitability, rather than productivity, determines whether or not a firm survives. This is in line with related empirical findings by Foster et al. (2008) who document that even unproductive firms may survive as long as they are able to extract a sufficient mark-up. To understand exit along the profitability margin, it is instructive to look at maximized firm profits:

$$\pi_{it} = (1-\xi) \left[ A_t z_{it} \left( \frac{\alpha}{r_t} \right)^{\alpha} \left( \frac{\gamma}{w_t} \right)^{\gamma} \right]^{\frac{1}{1-\xi}} - \omega_t c_f. \tag{7}$$

Equation (7) shows that more productive firms (higher $z$) make larger profits and that the profit function is convex in productivity $z$ as long as returns to scale are decreasing ($0 < \xi < 1$) – which is supported by the empirics. The most productive firms then post the highest profits and the least productive firms ($z$ close to zero) post gross profits close to zero.

The overhead costs are responsible for why some unproductive firms become unprofitable at all. If there were no overhead costs, even the least productive firm would generate profits amounting to a $(1-\xi)$-share of their sales. When there are overheads, however, only the more productive firms break even after paying the overheads and post positive net profits. The least productive firms, which are also the smallest producers, do not generate enough gross profits to also cover the fixed overhead costs. They become unprofitable and exit. This immediate exit can be motivated by credit constraints: firms that temporarily make negative profits will be unable to pay their production factors without credit and have to shut down operations. Alternatively, one may think that firms may become idle and "wait out" the recession until the macro-economic environment becomes profitable again. Such a strategy makes sense if there are no sunk costs of re-entry. Whatsoever scenario one picks, firms that are idle or temporarily shut down operations due to lack of finance do not appear in the data and hence they would be counted as exiting firms exactly like in the present model where firms permanently exit once they become unprofitable.

The convex gross profits and constant fixed overhead costs mean that there is a well-defined productivity threshold, denoted $z_t^*$, which yield zero profits and below which firms would make negative profits:

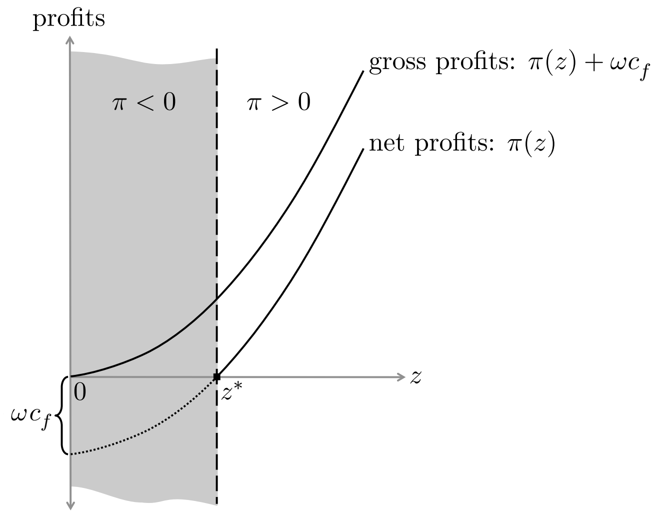
$$\begin{aligned}\pi(z_t^*) = 0 &= (1-\xi)\left[A_t z_t^* \left(\frac{\alpha}{r_t}\right)^{\alpha} \left(\frac{\gamma}{w_t}\right)^{\gamma}\right]^{\frac{1}{1-\xi}} - \omega_t c_f \\ \Leftrightarrow z_t^* &= \frac{1}{A_t}\left(\frac{\omega_t c_f}{1-\xi}\right)^{1-\xi} \left(\frac{r_t}{\alpha}\right)^{\alpha} \left(\frac{w_t}{\gamma}\right)^{\gamma} \qquad (8)\end{aligned}$$

Equation (8) is a key relationship as it regulates the measure of active firms in the economy, $1 - F(z_t^*)$, and the productivity distribution of active firms. Unproductive firms (firms with a low $z_{it}$) survive more easily when aggregate technology rises, i.e. when $A_t$ is large. Then, the economy is in a boom and the cutoff level $z_t^t$ is low. Conversely, a downturn poses harsher conditions for unproductive firms that may have to exit. This aspect reflects the view that recessions are weeding out unproductive firms – the "cleansing effect of recessions" (Caballero and Hammour (1994)). Note that endogenous factor prices $r_t$, $w_t$ and $\omega_t$ will dampen the effect of $A_t$, but they will not overturn it (for details see Section 3.5.2). Figure 5 illustrates gross and net profits as a function of $z$ and the productivity cutoff $z^*$ for positive profits.

Above, I have described how firms endogenously exit according to their productivity level. Key to the exit decision are negative profits which is consistent with the notion that re-entry is costless. Such an interpretation is consistent with temporarily unprofitable firms that merely become idle and do not have to pay entry costs. If there are no entry costs, however, I have to assume an exogenously fixed measure of entrants every period; this keeps the equilibrium measure of active firms stationary.[39] Upon entry, each firms receives its long-run productivity draw $\overline{z}_i$. The idiosyncratic productivity will be reverting to that long-run firm-specific mean.

[39] As an alternative, one may assume that firms do have to pay an entry cost $c_e$, denominated in units of labour for example. While the conclusions about the cyclicality of the productivity distribution would remain unchanged, the exit behavior of firms would be more complicated: Firms with a long-run productivity draw $\overline{z}_i$ below the steady state-level of $z^*$ would immediately exit after entry, those with a long-run productivity draw above the steady state-level of $z^*$ would only exit if their discounted future profits cover the temporary negative profits. The movements of the survival cutoff $z^*$ over the business cycle would be weaker quantitatively, but unchanged qualitatively. A weaker movement in the cutoff, however, would strengthen the counter cyclicality of productivity dispersion.

Figure 5: Firm productivity and firm profitability




*Note:* Firm profits gross and net of the overhead costs, $\omega c_f$, as a function of the firm-specific productivity $z$. Any firm with productivity level below $z^*$ will make losses (shaded area) because they are too small to generate enough revenue to cover both production and overhead costs. The productivity cutoff $z^*$ is computed as the productivity level generating zero net profits.

The cutoff also determines the measure of active firms in the economy. The total measure of firms in the economy, $N_t$, is composed of surviving incumbents active last period , $N^*_{t-1}$, and new entrants, $N^E_t$: $N_t = N^*_{t-1} + N^E_t$. Since all incumbents receive productivity draws from the same distribution as the incumbents, their productivity distribution looks the same. Whether a firm is an incumbent or an entrant, it will only survive if its productivity is above this period's survival cutoff $z^*_t$, so the measure of firms that are active this period is:

$$N^*_t = [1 - F_z(z^*_t)]N_t = [1 - F_z(z^*_t)](N^*_{t-1} + N^E_t) \tag{9}$$

To keep the measure of active firms, $N^*_t$, stationary, I assume that the measure of entrants, $N^E$, equals the measure of exiting firms in steady state, i.e. that measure of firms below the steady-state cutoff level.

### 3.2.3 Characterising the productivity distribution

Before studying the dynamics of the productivity distribution over the business cycle, it is necessary to understand the characteristics of the productivity distribution in the (steady state) equilibrium. The productivity dispersion measure in the empirical section was constructed using production inputs that correspond to $k, l$ and gross output which corresponds to $y$ in the model. Overhead inputs may be contained in any of these two inputs and it is almost impossible to distinguish overhead from marginal inputs. Empirically, I established that proxies for overhead labour showed an effect on the measured productivity distribution.

While each firm observes and decides on inputs, outputs and exit after knowing its physical productivity $z$, it is the measured productivity $\tilde{z}$ that is observed by the econometrician. Keeping the model analysis congruent with the empirical work, measured productivity can be expressed as a function of physical productivity:

$$\begin{aligned}\tilde{z}_{it} &= \frac{y_{it}}{A_t k_{it}^{\alpha} \left(l_{it} + c_f\right)^{\gamma}} < \frac{y_{it}}{A_t k_{it}^{\alpha} l_{it}^{\gamma}} = z_{it} \\ &= z_{it} \left(\frac{l_{it}}{l_{it} + c_f}\right)^{\gamma}\end{aligned}$$

In the empirical analysis, I measured the dispersion of the logarithm of productivity, i.e.

$$\log \tilde{z}_{it} = \log z_{it} + \gamma \log \left(\frac{l_{it}}{l_{it} + c_f}\right). \tag{10}$$

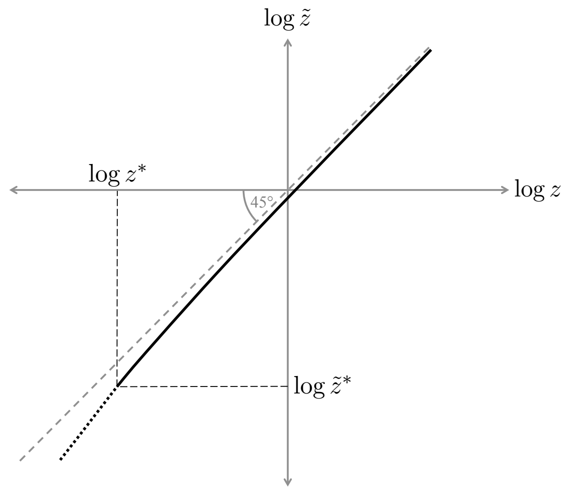
Equation (10) shows that measured productivity, $\log \tilde{z}$, is always lower than physical productivity, $\log z$. Crucially, this difference is greatest for the least productive surviving firms and monotonically declines with productivity. This is because unproductive firms will only hire little production labour $l$. For those firms, the share of overhead inputs in total labour is greatest (their term $\log \left(\frac{l}{l+c_f}\right)$ is very negative) and they will hence look very unproductive. The more productive a firm, the more labour it will hire and the negative term $\log \left(\frac{l}{l+c_f}\right)$ will vanish. For the most productive firms overheads do not confound the proper measurement of physical productivity. Figure 6 illustrates the relationship between physical and measured productivity.

The overall distribution of productivity as measured in the empirical portion of this paper will depend on the distribution of physical productivity up to the truncation, $\log z^*$, and distribution of firm size in terms of production employment, $l(z)$. Compared to the distribution of physical productivity, the distribution of measured productivity will be more dispersed and this dispersion increases when the size distribution shifts left as it does in a recession. I will explain this in Section 3.5.2.

### 3.2.4 The firm size distribution and aggregation

Profit-maximising firm behavior determines some features of the cross section of active firms. I will start by showing that more productive firms will be larger in general, hiring more inputs and producing more output. Since all firms operate the same technology (summarized by $\alpha$ and $\gamma$) and since all firms face the same prices $r_t$ and $w_t$, their capital intensities, $\frac{k_{it}}{l_{it}}$, must be the same. This can be simplified to writing the relative input choices of two distinct firms, $i$ and $j$, as a function

Figure 6: Physical and measured productivity




*Note:* Measured productivity (in logs), $\log \tilde{z}$, as a function of physical productivity $\log z$, as defined in equation (10). Overheads in any input will act as a concave transformation of physical to measured productivity. This makes the distribution of measured productivity more dispersed than that of physical productivity. The measured productivity distribution is truncated at $\log \tilde{z}^* = \log z^* + \gamma \left( \frac{l(z^*)}{l(z^*)+c_f} \right)$.

of their relative productivity only:

$$\begin{aligned} \alpha A_t z_{it} k_{it}^{\alpha-1} l_{it}^{\gamma} &= r_t = \alpha A_t z_{jt} k_{jt}^{\alpha-1} l_{jt}^{\gamma} \\ \Leftrightarrow \frac{k_{it}}{k_{jt}} &= \left( \frac{z_{it}}{z_{jt}} \right)^{\frac{1}{1-\xi}} \end{aligned}$$

and similarly for labour. Thus, all relative factor inputs and hence the relative size of firm output are a function of the relative productivities:

$$\frac{k_{it}}{k_{jt}} = \frac{l_{it}}{l_{jt}} = \frac{y_{it}}{y_{jt}} = \left( \frac{z_{it}}{z_{jt}} \right)^{\frac{1}{1-\xi}} \qquad \forall i, j \in [0, N_t],\ i \neq j. \tag{11}$$

It will be convenient to compute firm size at the cutoff, $y(A_t, z_t^*)$ using the firm first-order conditions and the expression for $z_t^*$:

$$\begin{aligned} y(A_t, z_t^*) &= A_t z_t^* k(A_t, z_{it})^{\alpha} l(A_t, z_{it})^{\gamma} \\ &= \frac{\omega_t c_f}{1-\xi} \end{aligned} \tag{12}$$

**Aggregation** In order to close the model, I need to aggregate firm-level production and employment to the economy-wide level. I can use the above-listed expressions for optimal firm production and employment and use the production function to obtain aggregates inputs and aggregate output. Despite the firm heterogeneity and the non-convex production structure, production and inputs aggregate nicely. In general, aggregates will be determined by the underlying productivity distribution of firms, the survival cutoff $z_t^*$ and the measure of firms in the economy. Once, we know aggregate production and factor demand, will posit aggregate factor supply curves to compute factor prices and close the model. Aggregate output is

$$Y_t = \int_0^{N_t} y_{it} di.$$

I now rewrite aggregate output in various ways: First, I commute the integration to $z$ rather than $i$; second, I take into account that only firms above the productivity threshold $z_t^*$ are active in the economy; third, I make use of the expression for relative fir size in equation (11); fourth, I use the expression for output at the cutoff (12).

$$\begin{aligned} Y_t &= N_t \int_0^{\infty} y(A_t, z_{it}) dF_z(z) \\ &= N_t \int_{z_t^*}^{\infty} y(A_t, z_{it}) dF_z(z) \\ &= N_t^* \frac{1}{1 - F_z(z_t^*)} \int_{z_t^*}^{\infty} y(A_t, z_{it}) dF_z(z) \\ &= N_t^* y(A_t, z_t^*) \frac{1}{1 - F_z(z_t^*)} \int_{z_t^*}^{\infty} \frac{y(A_t, z_{it})}{y(A_t, z_t^*)} dF_z(z) \\ &= N_t^* y(A_t, z_t^*) \frac{1}{1 - F_z(z_t^*)} \int_{z_t^*}^{\infty} \left(\frac{z_{it}}{z_t^*}\right)^{\frac{1}{1-\xi}} dF_z(z) \\ Y_t &= N_t^* \frac{\omega_t c_f}{1 - \xi} \left[\frac{\zeta(z_t^*)}{z_t^*}\right]^{\frac{1}{1-\xi}} \end{aligned} \tag{13}$$

where $N_t$ is the total number of firms (active and exiting) and $N_t^*$ is the measure of firms above the cutoff and who are producing output; $N_t^*$ is what we see in the data. $\zeta(z_t^*) \equiv \left(\frac{1}{1-F_z(z_t^*)} \int_{z_t^*}^{\infty} z^{\frac{1}{1-\xi}} dF_z(z)\right)^{1-\xi} = \left(E\left[z^{\frac{1}{1-\xi}} \middle| z > z_t^*\right]\right)^{1-\xi}$ is an aggregate productivity index that is useful in expressing aggregate quantities. Aggregate output increases with the number of active firms, $N_t^*$, the firm size at the survival cutoff, $y(A, z^*)$, and declines with the cutoff $z^*$ (Leibniz rule).

In a similar vein, one can compute aggregate demand for capital, labour and materials as

$$K_t = \int_{i=0}^{N_t} k_{it} di = \frac{\alpha}{r_t} Y_t \tag{14}$$

$$L_t = \int_{i=0}^{N_t} l_{it} di = \frac{\gamma}{w_t} Y_t. \tag{15}$$

Aggregate demand for overheads is given by the measure of active firms, $N_t^*$, because every active firm needs overhead inputs $c_f$ in order to produce:

$$M_t = N_t^* c_f. \tag{16}$$

Aggregate profits are

$$\begin{aligned}\Pi_t &= \int_{i=0}^{N_t} \pi_{it} di \\ &= \int_{i=0}^{N_t} (1-\xi) y_{it} - \omega_t c_f di \\ &= (1-\xi) Y_t - N_t^* \omega_t c_f\end{aligned}$$

I assume that net profits $\Pi$ are distributed lump-sum to consumers who also provide all production factors to firms. That way, final consumers are paid income $r_t K_t + w_t L_t + N_t^* \omega_t c_f + \Pi_t$ which equals the value of aggregate output in the economy $Y_t$.

### 3.3 Households

The remaining part of the model is just for closure: A representative household has standard preferences over consumption, $C_t$, and disutility from labour. His supply of production labour, $L_t$, and overhead labour, $M_t$, generates labour income $w_t L_t$ and $\omega_t M_t$, respectively. Since households own all assets in the economy – the physical capital stock and firms, they also earn rental income $r_t K_t$ and profits $\Pi_t$. Income is spent on consumption or invested in the capital stock next period. Any new firms immediately belong to the household's endowment so he is entitled to their profits, too. The household problem is hence to maximize

$$\begin{aligned}&\max_{C_t, K_{t+1}, L_t, M_t} U = E_0 \sum_t \beta^t \frac{\left[C_t - \frac{L_t^{1+\psi}}{1+\psi} - \chi \frac{(M_t + \overline{M})^{1+\phi}}{1+\phi}\right]^{1-\theta}}{1-\theta} \\ &s.t. \quad C_t + K_{t+1} - (1-\delta) K_t \leq r_t K_t + w_t L_t + \omega_t M_t + \Pi_t\end{aligned}$$

The household is endowed with a unit measure of time and $L_t$ isis hours worked, $\beta$ is his discount factor and $\theta$, the inverse of intertemporal elasticity of substitution. $\psi$ and $\phi$ denote the supply elasticities of production and overhead labour, respectively, and $\chi$ denotes the additional disutility from supplying overhead labour. Imagine $\chi$ captures the extra effort a manger must exert compared to a production worker. $\overline{M}$ is a fixed utility cost the household incurs from supplying overhead labour at all. Importantly, the supply of either type of labour is independent of the consumption/savings decision as in Greenwood et al. (1988). The household's inter- and intratemporal first-order conditions are

$$C_t^{-\theta} = \beta E_t C_{t+1}^{-\theta}(1 - \delta + r_{t+1}) \tag{17}$$
$$w_t = L_t^{\psi} \tag{18}$$
$$\omega_t = \chi(M_t + \overline{M})^{\phi} \tag{19}$$

while the household rents out the entire capital stock in period $t$ there is.

The resource constraint is

$$Y_t = C_t + K_{t+1} - (1 - \delta)K_t. \tag{20}$$

## 3.4 Equilibrium

An equilibrium is a set of prices $\{r_t, w_t, \omega_t\}$, aggregate quantities $\{C_t, K_{t+1}, L_t, M_t, Y_t\}$, firm-specific quantities $\{k_{it}, l_{it}\}$, a productivity cutoff level $z_t^*$, a measure of active firms $N_t^*$ such that

- given prices, the aggregate quantities maxmise household utility, i.e. equations (17)-(19) hold;
- given prices, the firm-specific quantities maximize profits, i.e. equations (5) and (6) hold;
- firms optimally exit if profits are negative and the productivity cutoff $z_t^*$ is defined by the zero-profit equation (8);
- given the productivity cutoff, firm-specific quantities integrate to aggregate quantities, i.e. equations (14), (15) and (16) hold;
- aggregate demand for capital equals the capital stock;
- aggregate demand for production and overhead labour ((15) and (16)) equals aggregate supply (18)) and (19));
- aggregate supply for final goods (equation (13)) equals aggregate demand; i.e. the resource constraint (equation (20)) holds;
- the measure of active firms follows the law of motion defined in equation (9).

## 3.5 Business cycles

I will now study the dynamic properties of the model. The main goal of the model is to interpret the empirical patterns of the productivity distribution which I documented in Section 2. A particular emphasis will be on the effects of entry and exit as well as the overhead inputs in production. Business cycles in the model are driven by standard shocks to aggregate TFP, $\eta_t$. Since this is a real model, it seems natural to consider this standard shock though other shocks such as discount rate shocks etc. could be added.

### 3.5.1 Calibration

The model shares most common features of the real business cycle model. Most standard parameters will be calibrated to values that the empirical macro literature has estimated: The household discount rate $\beta$ is calibrated to 0.98 which corresponds to a long-run real interest rate of about 2% p.a. The inter temporal elasticity of substitution is calibrated at 1 meaning that the utility function is logarithmic. Other estimates, especially in the finance literature, put $\theta$ around 2 which implies a lower inter temporal elasticity. The parameters $\psi$ and $\phi$ regulate the slope of the supply curves of production and non-production labour and thus the volatility of wages and quantities. I use CPS data to calibrate these parameters such that the model generates wage volatility for both types of labour that matches those in the CPS data. Evidently, the much larger value for $\phi$ implies that, consistent with empirics, overhead wages are more volatile than those of production workers.

Most parameters on the production side of the economy can be directly calibrated to estimates from Section 2. The production elasticities for capital and production labour, $\alpha$ and $\gamma$, are estimated at the same time as productivity and displayed in Table 9; note that I transform them into a value-added specification and combine the estimates for capital and energy to obtain a production elasticity for capital services. The long-run technology level $\overline{A}$ is set to unity which centers the cross-sectional distribution of log productivity at zero. There is a stark difference in the persistence and volatility of aggregate versus idiosyncratic productivity: The common component is very persistent and smooth while the idiosyncratic productivity is almost one order of magnitude larger but almost not autocorrelated at all. In line with the data, I set $\rho^A = 0.95$ and $\sigma^A = 0.041$ and $\sigma^z = 0.174$. $\rho^z$ is set to zero which simplifies the entry and exit dynamics as I explained above. Given the empirical estimate of $\rho^z \approx 0.15$, this is not far from the empirical picture.

The level of overheads $c_f$ is not standard and overhead inputs are hard to measure in the data. But the level of $c_f$ regulates the level of the survival cutoff $z^*$ and thus the measure of exiting firms. Exit rates are observed in the data to be around 6.6% which puts $c_f$ at ???. The measure of newly entering firms, $N^E$, is set such that it equals the measure of exiting firms in steady state: $N^E = \frac{F(\overline{z}^*)}{1-F(\overline{z}^*)}\overline{N}^*$. The calibration, their targeted moments and the corresponding data source are listed in Table 7.

### 3.5.2 The dynamics of the productivity distribution

The main goal of the model was to find a way to interpret the cyclical changes of the cross-sectional productivity distribution documented in the data. I have shown in the previous model sections that the measured productivity distribution depends on the survival cutoff $z_t^*$ and the size distribution $l_{it}$. The former matters for the productivity composition of firms in the economy while the latter matters for how efficiently active firms utilize their overhead inputs. First, I am going to show that the productivity cutoff $z_t^*$ increases in a recession. Then I show that all surviving firms will utilize their inputs less efficiently. While there is some cleansing going on in a recession, the measured

Table 7: Calibration

| Label | Parameter | Value | Targeted Moment | Data Source |
|---|---|---|---|---|
| $\alpha$ | Production elasticity $k$ | 0.29 | estimated in Section 2 | ASM/CMF |
| $\gamma$ | Production elasticity $l$ | 0.65 | estimated in Section 2 | ASM/CMF |
| $\overline{A}$ | Long-run aggregate TFP average | 1 | | ASM/CMF |
| $\rho^A$ | Persistence of aggregate TFP | 0.90 | | ASM/CMF |
| $\rho^z$ | Persistence of idiosyncratic TFP | 0 | | ASM/CMF |
| $\sigma^A$ | St. Dev. of aggregate TFP shocks | 0.041 | estimated in Section 2 | ASM/CMF |
| $\sigma^z$ | St. Dev. of idiosyncratic TFP shocks | 0.174 | estimated in Section 2 | ASM/CMF |
| $c_f$ | Overhead inputs | ??? | share non-production workers | ASM/CMF |
| $N^E$ | Measure of entrants | ??? | steady-state share of exiting firms | ASM/CMF |
| $\beta$ | Discount Rate | 0.98 | standard | |
| $\theta$ | Intertemporal elasticity of substitution | 1 | standard | |
| $\psi$ | Elasticity of production labour supply | ??? | St. Dev. of production worker wages | NBER Mfg DB |
| $\phi$ | Elasticity of overhead labour supply | ??? | St. Dev. of non-production worker wages | CPS |
| $\overline{M}$ | autonomous disutility from overhead labour supply | 1 | ??? | |

productivity distribution need not look more compressed due to a truncation of its left tail because all firms, especially the surviving unproductive firms make much less efficient use of their overhead inputs.

The productivity cutoff, equation (8), is determined by firm profits, equation (7), which is plotted in Figure 5. Recessions are conceptualized by standard aggregate technology shocks, i.e. a fall in $A_t$. Though this will also lead to a fall in interest rates $r_t$ and wages $w_t$, but their fall will be small enough, so that the term $\left[A_t\left(\frac{\alpha}{r_t}\right)^{\alpha}\left(\frac{\gamma}{w_t}\right)^{\gamma}\right]$ decreases in a recession:

**Proposition 1** *If the supply curves of capital, production and overhead labour are non-decreasing, then the term* $\left[A_t\left(\frac{\alpha}{r_t}\right)^{\alpha}\left(\frac{\gamma}{w_t}\right)^{\gamma}\right]$ *comoves positively with* $A_t$*, i.e.* $\hat{A}_t > \alpha\hat{r}_t + \gamma\hat{w}_t$.

**Proof** See Appendix D.

**Lemma 1** *If the supply curves of capital, production and overhead labour are non-decreasing and if* $\hat{A}_t < \alpha\hat{r}_t + \gamma\hat{w}_t$*, then the cutoff* $z_t^*$ *rises after a rise in* $A_t$.

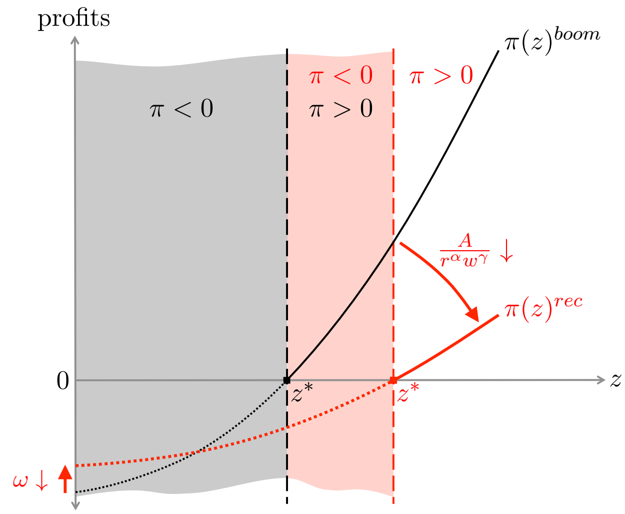
**Proof** See Appendix D.

Proposition 1 and Lemma 1 show that factor price changes in general equilibrium may dampen the effects of aggregate productivity on profits, but they do not reverse them. This means that lower aggregate productivity $A_t$ will make all firms less profitable and erode the profitability advantage of high-productivity (high $z$) firms. Graphically, this corresponds to a downward-rotation of the profit function in Figure 7. But the price for overheads will adjust in a recession. Graphically, this means that the intercept of the profit function, $-\omega_t c_f$, will change as well. If the cutoff ends up increasing in a recession, then the measure of active firms will be lower. Lower demand for overheads will reduce their price $\omega_t$, which means that the profit function shifts up in a parallel fashion. This shift dampens the rise of the survival cutoff in recessions, but does not reverse it.

**Proposition 2** *If* $A_t$ *declines, then the falling price for overheads* $\omega_t$ *dampens the rise of the survival cutoff* $z_t^*$*, but it does not reverse it.*

The results of Propositions 1 and 2 correspond to the right-rotation and the vertical shift of the profit function in Figure 7. Both changes show that, in a recession, the survival cutoff increases and that all surviving firms make lower profits. The rising survival cutoff implies says that the composition of firms is more productive in general, but the econometrician may measure lower productivity as I will demonstrate below. The profits of every surviving firm, however, decline. Because the low-productivity producers got cleansed out, because there are fewer firms in the economy and because the profits of every surviving firm are lower, aggregate profits decline in a recession. The changes to the profits function and the survival cutoff are schematically displayed in Figure 7.

Figure 7: Firm profitability and survival in recessions



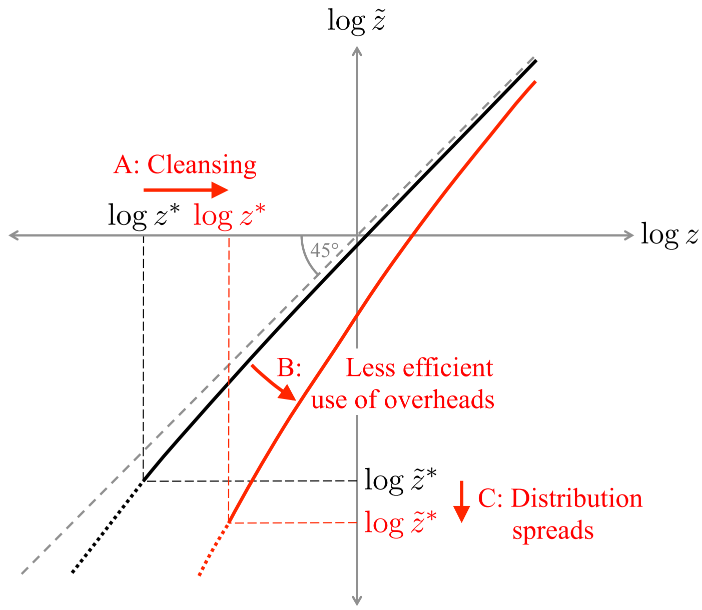

*Note:* Profits as a function of firm-specific productivity. Compared to booms, recessions are times when the profit function is flatter (Proposition 1) and the survival condition tougher (Proposition 2): $z^*$ in recessions is higher than in booms. This means more firms are exiting in a recession – the shaded red area is larger than the shaded grey area – and the surviving firms are making lower profits – $\pi(z)^{rec}$ is strictly beneath $\pi(z)^{boom}$.

The firm size distribution is the second main determinant of the measured productivity distribution. The firm first-order condition, equation (6), shows that labour demand will fall in a recession for every active firm. This means that measured productivity, equation (10) will decline for all firms, but especially the small firms which employ a high share of overhead labour. Although a rising cutoff will tend to compress the productivity distribution, the measured productivity distribution may look more dispersed and left-skewed because the least productive surviving firms will look very onefficient in utilizing their overhead inputs. The calibration shows that this second effect can be quite large and lead to a sizable change in the productivity distribution observed in the data. The changing productivity distribution is displayed in Figure 8.

It is possible to show how the size of the least productive surviving firm changes over the business cycle. That is, one can compute $l(z^*)$ in booms and recessions. This is important for the behavior of the productivity distribution as the size of the least productive firm determines how the unproductive tail of the distribution looks like in booms and recessions. By the firm's first-order condition (6) and using the expression for the survival cutoff (8) we know:

$$\begin{aligned}\frac{l(z^*)^{Rec}}{l(z^*)^{Boom}} &= \left[\frac{A^{Rec}(z^*)^{Rec}\left(\frac{\alpha}{r^{Rec}}\right)^{\alpha}\left(\frac{\gamma}{w^{Rec}}\right)^{1-\alpha}}{A^{Boom}(z^*)^{Boom}\left(\frac{\alpha}{r^{Boom}}\right)^{\alpha}\left(\frac{\gamma}{w^{Boom}}\right)^{1-\alpha}}\right]^{\frac{1}{1-\xi}} \\ &= \frac{w^{Boom}}{w^{Rec}}\frac{\omega^{Rec}}{\omega^{Boom}}. \qquad (21)\end{aligned}$$

Figure 8: Physical and measured productivity in a recession




*Note:* The distribution of measured productivity, $\tilde{z}$, in booms and recessions. Although the truncation at the left tail becomes slightly more pronounced (A: Cleansing), overhead inputs are used less efficiently (B: Less efficient ...). This results in a measured productivity distribution that is more spread-out at the left tail (C: Dispersion spreads).

Equation (21) shows that the size of the marginal firm at the cutoff depends on the relative volatility of the wages for the two types of labour. In the data, wages for overhead inputs are much more volatile over the business cycle than those for production workers. This means that the marginal firm shrinks in size in a recession. This feature can be represented graphically in Figure 8: it means that the left end of the solid red line depicting $\tilde{z}$ in a recession is further from the 45-degree line than the left end of the solid black line depicting $\tilde{z}$ in a boom.

If the size difference of the marginal firm is large enough over the business cycle, then it will dominate the increase in the survival cutoff and the actually measured productivity of the least productive firm observed in the data *declines* in a recession as seen in the data:

$$
\begin{aligned}
&(\tilde{z}^*)^{Rec} < (\tilde{z}^*)^{Boom} \\
\Leftrightarrow\ & (z^*)^{Rec} - (z^*)^{Boom} < \gamma \log\left(\frac{l(z^*)^{Boom}}{l(z^*)^{Rec}} \times \frac{l(z^*)^{Rec} + c_f}{l(z^*)^{Boom} + c_f}\right)
\end{aligned}
$$

# 4 Conclusion

This paper established the dynamics of the empirical productivity distribution over the business cycle. Among U.S. manufacturing plants, this dispersion is higher in a recession than in a boom. These cyclical dynamics are predominantly driven by changes in the unproductive tail of the distribution. The empirical evidence points at non-production overhead labour as systematically related to the dynamics of the productivity distribution.

This countercyclical productivity dispersion is at odds with conventional cleansing models of the business cycle which posit a more compressed productivity dispersion in recessions. In order to reconcile these models with the empirical findings, I build a business cycle model featuring overheads in production and endogenous exit. In my model, a negative aggregate shock does lead to some cleansing of unproductive firms, but the surviving firms now operate at a more inefficient level, so they look more unproductive. This inefficient use of overheads is concentrated in the unproductive tail of the productivity distribution where firms are smallest. Since the size distribution shifts left so much, the measured productivity dispersion is more spread-out in a recession. In addition to that, this shock is also consistent with the typical macroeconomic business cycle acts such as procyclical consumption, investment, wages, interest rates and employment while firm exit is countercyclical.

The University of Texas at Austin.

# Appendix

## A Census manufacturing data

### A.1 General description

The data used in this project are compiled by the U.S. Census Bureau and comprise the Census of Manufactures (CMF), the Annual Survey of Manufactures (ASM) and the Plant Capacity Utilization Survey (PCU). Additional data come from the NBER-CES productivity database, the Federal Reserve Board of Governors (data on capacity utilization), the Bureau of Economic Analysis (BEA; data on capital stocks and investment prices), the Bureau of Labor Statistics (BLS; data on depreciation rates and inventory price deflators). The Compustat-SSEL bridge (CPST-SSEL) is used to determine which establishments are publicly traded (are covered in Compustat).

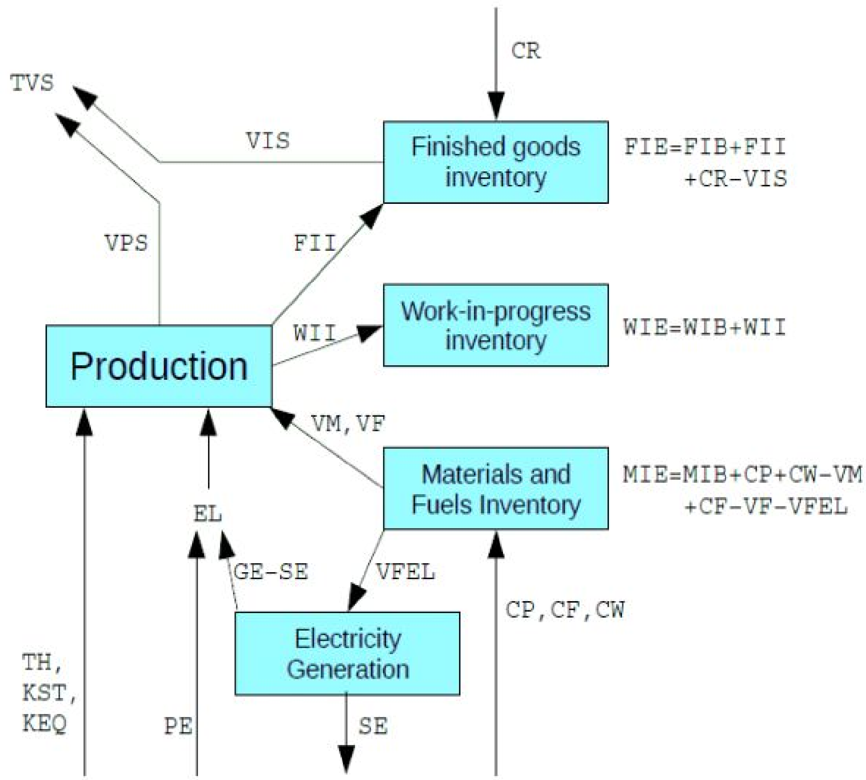
The main data sources are the CMF/ASM. They are both mail-back surveys and cover the U.S. manufacturing sector (NAICS 31-33) on the establishment level where establishment is defined as any distinct unit of a manufacturing firm where the predominant activity is production. Purely administrative establishments are hence excluded. Each establishment carries the Permanent Plant Number (`PPN`), a unique establishment identifier that does not change in case of ownership change or temporary plant shutdown. If an establishment dies permanently, the `PPN` is not reassigned to a new-born establishment. Since 2002, the `PPN` is superseded by the Survey Unit ID (`SURVU_ID`). This more recent identifier was carefully mapped to the `PPN` using `LEGPPN` and `LBDNUM` or assigned a new `PPN` if an establishment was born after 2002. Establishments that belong to the same legal firm carry the same firm identifier `FIRMID`. Firms are called multi-unit firms (MUF) if they operate more than one establishment, single-unit firm (SUF) if they operate merely one.

The Census of Manufactures is conducted at quinquennial frequency (years ending in 2 and 7) and covers all existing 300-350k establishments in the manufacturing sector. The ASM is conducted in non-Census years for about 50-60k establishments taken from the "mail stratum" of the manufacturing sector. The "non-mail stratum" generally consists of small establishments that together make up a very small fraction of activity; their chance to be selected in to the ASM panel is zero. I drop all observations from the non-mail stratum (denotedby `ET` = 0) because this is the only way to obtain a consistent panel over time where the number of (weighted) observations is not driven by the sampling constraints of Census. Of the mail stratum, the ASM covers all "large" establishments with certainty and a selection of "small" establishments. The criteria for an establishment to qualify as large are cutoffs changed over time. In principle, these are cutoffs in terms of asset size, employment or industry share and. For all establishments in the ASM, Census provides frequency weights which are the inverse of the sampling probability and can be used to replicate the underlying population where the sampled small establishments are representative of the establishments not sampled in the ASM. Every five years (years ending in 4 and 9) Census updates its small establishment sample according to the preceding Census to accurately reflect the underlying age and size population. Census attempts to sample the same small establishments in consecutive years until the next sample update.

The data carry a wide array of variables only some of which are of interest for this project. These are data on sales, inventories, employment and hours, capital stocks and investment, intermediates and energy. The following sections describe how observed variables are used to construct measures needed for the estimation.

## A.2 Measurement of real production

Figure 9: ASM flowchart




*Note:* Flow of inputs, outputs and inventories as they are measured in the ASM. All variables are nominal except hours worked (TH) and the real value of the capital stocks of structures (KST) and equipment (KEQ).

The object of interest is a real measure of goods produced (Q). It consists of goods that are produced and sold in the same year (PS) and produced goods that are stored in either of two inventories: finished-goods inventory investment ($\mathtt{FII}^{\mathtt{real}}$) and work-in-progress inventory investment ($\mathtt{WII}^{\mathtt{real}}$).

$$\mathtt{Q} = \mathtt{PS} + \mathtt{FII}^{\mathtt{real}} + \mathtt{WII}^{\mathtt{real}}$$

The first term (PS) comprises receipts from goods produced and sold in the same period. Census collects information about some components of this term (such as product value of shipments, receipts for contract work), but their quality is not consistently reliable throughout the entire sample. Fortunately, total value of shipments (TVS) is considered by Census to be of superior quality. We can use this variable to infer PS as shown in Figure 9.

$$\mathtt{PS} = \frac{\mathtt{VPS}}{\mathtt{PISHIP}} = \frac{\mathtt{TVS} - \mathtt{VIS}}{\mathtt{PISHIP}}$$

where VPS is the nominal value of product shipments, PISHIP is a price deflator on the 4-digit (SIC) industry level from the NBER-CES Manufacturing Productivity Database and VIS is value of inventory sales. The last variable is not directly observed but will conveniently cancel out as explained below.

The second term, $\mathtt{FII}^{\mathtt{real}}$, can be constructed from nominal finished goods inventory investment which in turn can be constructed from the accounting identity: $\mathtt{FIE} = \mathtt{FIB} + \mathtt{FII} + (\mathtt{CR} - \mathtt{VIS})\frac{\mathtt{PIFI}}{\mathtt{PISHIP}}$. This expression contrasts with previous work and deserves more explanation. FIB and FIE denote the nominal value of finished goods inventory at the beginning and end of the period. FII is the value of produced goods that go into finished-goods inventories rather than being sold on the market in the same period. Note that FII is non-negative, because finished goods never flow back from the inventory to production. The last inflow into the finished goods inventory are resales (CR), finished goods purchased from other establishments that are resold without further changes or additions. Inventories that are sold in the current period are denoted by VIS. We do not observe VIS directly (though this shall not be a problem); we only know the portion of VIS that are resales (VR).[40]

Resales (CR) and inventory sales (VIS) are traded in the goods market at the market price (PISHIP), while inventory stocks (FIB and FIE) and inventory investment (FII) are valued with a price index for finished-goods inventories (PIFI). This is why the former three variables have to be adjusted for that. Empirically, PIFI is much more volatile than PISHIP and also exhibits a slightly different trend growth rate[41], so this difference might matter when one computes finished inventory investment:

$$\mathtt{FII}^{\mathtt{real}} = \frac{\mathtt{FIE} - \mathtt{FIB}}{\mathtt{PIFI}} - \frac{\mathtt{CR} - \mathtt{VIS}}{\mathtt{PISHIP}}$$

I assume that both FIB and FIE are nominal stocks of inventories that are valued with the inventory price deflator from period $t$, which is supported by the fact that in many cases $\mathtt{FIE}_{\mathtt{t-1}} \neq \mathtt{FIB}_{\mathtt{t}}$. Census sends establishments the ASM/CMF forms at the beginning of the period with end-of-year inventory stock pre-printed in the FIB cell. Establishments are allowed, however, to make changes; this is how last year's end-of-year inventories may differ from this year's beginning-of-year inventories.

The third term, $\mathtt{WII}^{\mathtt{real}}$, can be constructed from the accounting identity: $\mathtt{WIE} = \mathtt{WIB} + \mathtt{WII}$ where, contrary to above, $\mathtt{WII} \gtreqless \mathtt{0}$. No work-in-progress inventories are traded in markets, so terms merely have to be deflated by the price index for work-in-progress inventories (PIWI):

$$\mathtt{WII}^{\mathtt{real}} = \frac{\mathtt{WIE} - \mathtt{WIB}}{\mathtt{PIWI}}$$

Putting all three terms together yields:

$$\begin{aligned}\mathtt{Q} &= \mathtt{PS} + \mathtt{FII}^{\mathtt{real}} + \mathtt{WII}^{\mathtt{real}} \\ &= \frac{\mathtt{TVS} - \mathtt{VIS}}{\mathtt{PISHIP}} + \frac{\mathtt{FIE} - \mathtt{FIB}}{\mathtt{PIFI}} - \frac{\mathtt{CR} - \mathtt{VIS}}{\mathtt{PISHIP}} + \frac{\mathtt{WIE} - \mathtt{WIB}}{\mathtt{PIWI}} \\ \mathtt{Q} &= \frac{\mathtt{TVS} - \mathtt{CR}}{\mathtt{PISHIP}} + \frac{\mathtt{FIE} - \mathtt{FIB}}{\mathtt{PIFI}} + \frac{\mathtt{WIE} - \mathtt{WIB}}{\mathtt{PIWI}} \qquad (22)\end{aligned}$$

[40]Note that resales (CR and VR) are already finished goods, so they will not enter the materials inventory and eventually put through the production process, as was assumed by other researchers. In fact, counting them as material inputs would lead to biased results of production elasticities and productivity (more details below).

[41]This is because inventories are typically older goods of lower quality than those produced in the current period. Quality-adjusted price indices for inventories exhibit hence a higher growth rate than shipment price indices of the same product.

All of these variables are directly observed in the ASM/CMF except for the price deflators, which are obviously not available on the establishment level. I approximate PISHIP by the 4 digit-level industry price index for shipments from the NBER-CES Manufacturing Productivity Database; PIFI and PIWI are ideally industry-level price index for inventory investment (finished goods and work-in-progress goods respectively). BEA does produce inventory price deflators adjusted for quality on the industry level and separately for both finished and unfinished goods, but unfortunately, these are not publicly available, only to BEA sworn status researchers.[42] BLS published an inventory price deflator on the industry level, but this one contains a mix of finished goods, unfinished goods and materials inventories, so it merely looks like a crude measure. For that reason, I have to fall back to use shipment price deflators instead of inventory price deflators. Future researchers that have access to industry-wide deflators for inventories by type can easily combine them with the existing data and produce more accurate measures of output. While the present procedure is as good as one can possibly do to correct for prices, this can lead to inefficient estimates and possibly to further problems estimating total factor productivity, which will be discussed below.[43]

The construction of the ouptut variable improves on previous research in two ways: First, some work has ignored the role of inventories when constructing output variables (exceptions are Hyowook Chiang's measure or Petrin et al. (2011)). This seems problematic since inventory investment is known to fluctuate a lot; for example, it has a much higher volatility than investment in new capital (see Christiano (1988)). Second, in contrast to previous researchers, I classify resales (CR) as finished goods rather than a materials. Classifying cost of resales as a material input used in production and not correcting the output measure by the value of resales (VR) seems misplaced: By definition, resales are products that are bought and then resold without any change to the product. They are therefore not going through the production process and provide no information about the firm's productivity as a producer of goods. Even worse, a researcher running a production function regression to study productivity will obtain biased estimates of production elasticities and as a consequence also biased estimates of productivity. Counting CR as material input and not correcting the output measure will bias the coefficient estimate of materials towards 1 (i.e upwards) and it will also bias all other coefficient estimates (downward). Even small values of resales (CR is on average 5% of overall materials purchases) bias the estimates significantly.[44]

## A.3 Measurement of labour input

The ideal measure is hours worked of all workers. The ASM/CMF only carries information on plant hours worked (PH), which covers only production workers, so hours of non-production workers have to be imputed. In addition to the number of total employees (TE), production workers (PW) and production worker hours (PH), the ASM/CMF carries information about wage payments for all employees (SW) and production workers (WW), which contain some information about the hours worked if one has an idea about the level of wages. Wages and salaries can be exploited to construct a more accurate measure of total hours worked. Let WP and WNP denote the average wages for production and non-production workers, respectively. Then, total hours (TH) can be expressed as

[42] Census researchers that have special sworn status are not entitled to obtain the data either.

[43] At this point, I am following the large productivity literature and estimate revenue factor productivity (TFPR in Foster et al. (2008) or Hsieh and Klenow (2009)).

[44] As a check on the strength of this bias I simulated 1000 observations of the following technology: $Y = K^a M^b$ with $a = 0.1$ and $b = 0.45$. Estimating $a$ and $b$ using $Y = Y + CR$ and $M = M + CR$ instead yields the following estimates $\hat{a} \approx 0.05$ and $\hat{b} \approx 0.52$ even when $\texttt{CR} = 0.05M$. This bias obviously becomes stronger the larger CR.

the sum of production worker hours ($\mathtt{PH}$) and non-production worker hours ($\mathtt{NPH}$):

$$\mathtt{TH} = \mathtt{PH} + \mathtt{NPH} = \mathtt{PH} + \frac{\mathtt{SW} - \mathtt{WW}}{\mathtt{WNP}}.$$

Wages for production workers can be computed as $\mathtt{WP} = \frac{\mathtt{WW}}{\mathtt{PH}}$. Unfortunately, wages of non-production workers are not observed in the ASM/CMF. I assume that the wages for non-production workers ($\mathtt{WNP}$) are 150% of those for production workers ($\mathtt{WP}$): $\mathtt{WNP} = 1.5 \times \frac{\mathtt{WW}}{\mathtt{PH}}$.[45]

Total hours under this assumption can be calculated as:

$$\begin{aligned} \mathtt{TH} &= \mathtt{PH} + \mathtt{NPH} \\ &= \mathtt{PH} + \frac{\mathtt{SW} - \mathtt{WW}}{\mathtt{WNP}} \\ &= \mathtt{PH} + \frac{\mathtt{SW} - \mathtt{WW}}{1.5 \times \mathtt{WP}} \\ \mathtt{TH} &= \mathtt{PH}\, \frac{\mathtt{SW} + 0.5 \times \mathtt{WW}}{1.5 \times \mathtt{WW}} \end{aligned} \tag{23}$$

Total hours worked can be constructed in this way for about 97.6% of all observations. The remaining observations do not have information on either of $\mathtt{PH}$, $\mathtt{SW}$ or $\mathtt{WW}$. In that case, I set $\mathtt{TH} = 2 \times \mathtt{TE}$ (50 weeks of 40 hrs/week each).

There is not a major improvement in the construction of the hours worked variable over previous research. If I do get the CPS data in then the imputation of non-production worker hours would be a substantial improvement.

## A.4 Measurement of capital input

Capital input (or capital services) in production, $\tilde{K}_t$, are determined by both the existing productive capital *stock* available to the firm, $K_t$, and the *utilisation* at which this stock is run, $u_t$. The latter is a percentage, so the object of interest, capital services are defined as the product of stock and utilisation:

$$\tilde{K}_t \equiv u_t K_t. \tag{24}$$

First, I shall describe how I measure the capital stock that is available to the firm for production, then the utilisation of the capital stock.

[45] A very proper way would be to utilise external information from the Current Population Survey to construct annual industry-region-specific average wages for both production workers and non-production workers, which gives an industry-region-year-specific ratio of the two average wages: $\mathtt{a} = \frac{\mathtt{WNP}}{\mathtt{WP}}$. Then, total hours can be computed on the establishment level as:

$$\mathtt{TH} = \mathtt{PH} + \frac{\mathtt{SW} - \mathtt{WW}}{\mathtt{WNP}} = \mathtt{PH} + \frac{\mathtt{SW} - \mathtt{WW}}{\mathtt{a} \times \mathtt{WP}} = \mathtt{PH}\, \frac{\mathtt{SW} + (\mathtt{a} - \mathtt{1}) \times \mathtt{WW}}{\mathtt{a} \times \mathtt{WW}}.$$

ALTERNATIVE:

One could get data on hours worked per employee in both production and non-production: $\mathtt{HRS}_{\mathtt{PW}}$ and $\mathtt{HRS}_{\mathtt{NPW}}$. These data should be available on an industry-region level in the CPS. Then, total hours can be computed as $\mathtt{TH} = \mathtt{HRS}_{\mathtt{PW}} \times \mathtt{PW} + \mathtt{HRS}_{\mathtt{NPW}} \times \mathtt{NPW}$ The disadvantage of this approach is that it implicitly assumes that all workers within an industry work the same amount of hours. Overtime work is not accounted for. As outlined above, wage payments on the other hand, do contain information about establishment-level overtime (and possibly part-time). Therefore, this approach based on industry-wide hours worked per employee would forgo all the information about hours worked contained in wage payments.

### A.4.1 Capital stocks

The capital stock is – ideally – the replacement value of fixed assets in constant dollars. In the absence of frictions, this is the value another firm would be willing to pay to acquire and operate this capital stock itself. In this sense, the replacement value should be an accurate measure of the productivity of the capital stock. Below, I will describe how I infer the closest approximation possible to this constant-dollars replacement value.

The ASM/CMF contains the following information related to capital:

- beginning-of-year and end-of-year total assets (`TAB` and `TAE`)
  - annually 1972-1988; in those years total assets are also separated into buildings (structures) and machinery (equipment): `BAB`, `BAE`, `MAB` and `MAE`
  - quinquennially 1992-2007,
- nominal investment expenditures for buildings (`NB`) and machinery (`NM`) for all years; in 1977-1996 investment expenditures are separated into investment in new and used capital: `NB`, `UB`, `NM` and `UM`,
- nominal building and machinery retirements (1977-1988, 1992): `BRT` and `MRT`,
- nominal building and machinery depreciation (1977-1988, 1992): `BD` and `MD`,
- nominal cost of rented building and machinery (1977-1988, 1992): `BR` and `MR`.

Investment, retirement and depreciation of assets are measured in period-$t$ dollars. Assets stocks (`TAB`, `TAE`, `BAB`, `BAE`, `MAB`, `MAE`), however, are somehow resembling book values rather than resale values. To obtain constant-dollar market values I perform three steps:

1. Transformation of reported values into book values,
2. Transformation of book values into period-$t$ market values,
3. Transformation of period-$t$ market values into constant-dollar values.

**Transformation into book values** The questionnaire of the ASM/CMF asks to list as asset stock values "the original cost of today's assets when they were purchased" in the past. It is not clear from the information given in the documentation whether or not this value takes (physical[46]) depreciation into account or not. If respondents answered the question literally, then it does not include depreciation and is not exactly a book value. If it does, then the reported data really are book values. I tried imputing the capital stock both ways. When I aggregate my capital stock measure and compare it to BEA's industry-wide capital stock, the level of my capital stock is slightly too high while the trend compares well to the BEA capital stock, so my level is off by a constant factor. This level gap is much smaller when I correct the initial values for depreciation.[47]

[46] It is important to consider physical depreciation rather than depreciation on the books. The latter is an accounting measure and does not necessarily reflect the accurate loss of productive capability of structures or equipment. An establishment might use a machine in production that is already entirely written off on the books.

[47] Multiplying the reported initial measure by $(1-\delta)$ implicitly assumes that capital stocks are one year old. An alternative, more refined method would be to construct the average age, $\tilde{T}$, of an establishment's capital stock from past investment expenditures and multiply the reported capital stock value by $(1-\delta)^{\tilde{T}}$. The former way (assuming average age of one year) yields aggregate capital stocks that are slightly too high, an approximation of the latter way (assuming average age as reported by BEA, which is about 22 years for structures and 6 years for equipment capital) yields aggregate capital stocks that are distinctly too low.

This suggests, that some respondents took the question literally and reported as asset values the initial expenditures unaccounted for by depreciation, others did take depreciation into account. Multiplying the reported capital stock values by $(1-\delta)$ transforms the observations into book values that will yield an aggregate time series that precisely matches the trend growth and roughly matches the absolute level.

**Transformation into market values** Transforming book values into market values requires (a) knowledge about the vintage structure of each establishment and (b) knowledge about the productivity of each vintage. This cannot be determined on the establishment level because we just know the dollar amount of investment but hardly the quality of the purchased capital.[48] The quality of the vintage, however, is crucial to determine the replacement value.

Due to the paucity of information on the establishment level, I turn to industry-level capital stock data published by the Bureau of Economic Analysis (BEA).[49] BEA publishes historical-cost, current-cost and real-cost estimates of capital stocks of 3-digit NAICS (2-digit SIC) industries that can help turn the ASM book values into real market values. For a single asset type, these end-of-year estimates[50] are defined as follows:

$$
\begin{aligned}
HC_t &= \sum_{\tau=0} \left(1-\frac{\delta}{2}\right)(1-\delta)^{\tau} I_{t-\tau} \\
CC_t &= P_t \sum_{\tau=0} \left(1-\frac{\delta}{2}\right)(1-\delta)^{\tau} \frac{I_{t-\tau}}{P_{t-\tau}} \\
RC_t &= \sum_{\tau=0} \left(1-\frac{\delta}{2}\right)(1-\delta)^{\tau} \frac{I_{t-\tau}}{P_{t-\tau}}
\end{aligned}
$$

where $\tau$ is the vintage (purchased $\tau$ periods before period $t$), $\delta$ is the depreciation rate, and $I_t$ are nominal investment expenditures in period $t$. The term $\left(1-\frac{\delta}{2}\right)$ appears because BEA assumes that new capital is put into place in the middle of the period. Note how the historical-cost capital stock is the industry analogue to the establishment book value. The current-dollar value, in contrast, is the nominal value of the capital stock in year-$t$ dollars where expenditures for every vintage have been deflated by the corresponding period price index and then reinflated by the current-period price index (hence the name). In this way, the $CC_t$ measure denotes the value of the capital stock as if it had been purchased at the end of the previous period. I assume that all establishments within an industry have a similar ratio of current-dollar market values to book values. Then I can use the ratio of $\frac{CC_t}{HC_t}$ to determine the period-$t$ market value of an establishment's capital stock.

**Transformation into constant dollars** This is then easily expressed in constant dollars by deflating the resulting measure by an investment price deflator.[51] Investment price deflators are published by the Bureau of Labor Statistics (BLS) and the Bureau of Economic Analysis (BEA) on the 3-digit NAICS industry level and on the 4-digit industry level as underlying table to the

[48] As mentioned above, investment in new and used capital goods are reported in the data only for a subsample.

[49] Tables 3.1E, 3.1S, 3.2E, 3.2S, 3.3E and 3.3S of BEA's Fixed Asset Tables; downloaded from http://www.bea.gov/national/FA2004/SelectTable.asp.

[50] Because I use beginning-of-year capital stocks BEA's data are rolled forward one year.

[51] Note that this is an investment price deflator rather than a capital price deflator because the capital stock is now expressed as if it had been an investment at the end of last period.

NBER Manufacturing Database.[52] I choose the BEA deflators because they were revised recently (in 2009), which matters a lot for capital goods (esp. equipment).[53] All three transformation steps (reported to book value, book value to market value and period-$t$ to constant-dollars transformation) combined give us the replacement value of an establishment's capital stock in constant dollars as:

$$K_t^{st} = \texttt{BAB}_t(1 - \delta^{st})\frac{CC_t^{st}}{HC_t^{st}}\frac{1}{P_t^{st}} \tag{25}$$

and analoguously for equipment capital. This procedure is accurate if all establishments in an industry exhibit the same profile across asset types and have the same vintage structure over time. This is obviously a strong assumption which is likely to be violated and lead to establishment-level measurement error.

For the years 1972-1988, I observe the capital stock annually and could compute the capital stocks in the above-mentioned way. Alternatively, I could iterate the capital stock every period using the perpetual inventory method:

$$K_{t+1}^{st} = (1 - \delta_t^{st})K_t^{st} + I_t^{st} \tag{26}$$

where $K_t^{st}$ is the stock of structure capital observed at the beginning of the period, $I_t^{st} = \frac{\texttt{NB}_t}{P_t^{st}}$ is real structure investment (nominal new and used[54] investment expenditures divided by an investment price index) and $\delta_t$ is a depreciation rate, published by BLS on the 3-digit NAICS industry level in period $t$. The former way of directly deflating the capital stock every period has the advantage of following the establishment-level information very closely. The latter perpetual inventory method shows exactly how the existing capital stock came about and follows a common procedure (see for example Becker et al. (2004)). I tried both alternatives and for equipment capital there is hardly any difference which supports the consistency of our above deflation technique. The procedure to directly deflate the capital stock every period underestimates structure capital compared to aggregate data on structures from BEA. Over the course of 20 years (1972-1992) the aggregate structure capital stock grows only at an annual rate of 0.15% which translates into a share of structures in total assets of 33.5% (while it should be about 45%). This implies that my interpretation of the structures measure in the CMF/ASM is flawed, which casts some doubt on the initialisation procedure as shown in equation. Therefore, I am sceptical of resetting the capital stock back to the value implied by equation (25) every time I observe it for continuing establishments. The perpetual inventory method, in contrast, does a good job at generating data that – aggregated to the industry level – resemble outside sources in terms of long-run growth. For this reason, I choose the perpetual inventory method and use the asset stock data observed every year to merely adjust the level of

[52] The investment price deflator could also be obtained from BEA by dividing CC/RC, but BEA warns researchers that the latter measure is not very reliable for years reaching far back. For that reason, I make use of the price indices published by BLS.

[53] I also tried the NBER and BLS deflators; the former do almost as good a job as the BEA deflators when one aggregates the establishment-level data and compares them to publicly available industry aggregates of capital stocks by type. BLS deflators cannot generate aggregates that resemble publicly available aggregates as well, which is mostly due to their price indices being only revised for the last 20 years. Once NBER deflators are updated in the future, they might be a superior measure as they go down to the 4-digit NAICS industry level.

[54] Census collected investment expenditures separately for new and used investment 1977-1996; in those years I sum the two groups and that in others years reported investment comprises both expenditures for new and used investment.

the implied total capital stock (keeping the asset split implied by the perpetual inventory method). I only use equation (25) to impute structures and equipment stocks directly when I observe an establishment for the first time.

From 1988 on, asset stock values, retirement and depreciation data are no longer observed. So I have to iterate and face the question of resetting or continuing the perpetual inventory method every five years. For the same reason as above, I proceed with the perpetual inventory method and merely adjust the implied book value of the imputed capital stocks by the book value (accounted for by depreciation) that is observed in the Census years. This procedure can be applied to both buildings (structures) and machinery (equipment) separately as the ASM/CMF contains investment data about both types.

**Improvements in the measurement of the capital stock** The capital stock measures differ from previous work about imputing capital stocks in the ASM. This is different because previous work omitted the second deflation step (period-$t$ market values to constant-dollars market values) and because deflators used in that work have been revised repeatedly. As a consequence, the old capital stock measures were too small at the beginning and too large at the end of the sample (as Figure 10 shows). Because the second deflation was omitted, the capital stock is a nominal value rather than a real one. It is not surprising in this light that the capital stock in the old sample is growing at an annual rate of 4.9% for structures (!) and 5.6% for equipment respectively. This barely squares with industry-wide aggregates where the capital stock grows at 1.1% and 2.6% (structures and equipment resp.). My measures end up at 1.2% and 2.7% which looks pretty close to the data published by BEA. This will have some important implication for researchers that used/are using his data. In my assessment of long- and short-run productivity the nominal trend picks up a lot of the upward trend in production. Second, because the investment price deflators are industry-year-specific, this essentially introduces industry-year dummies into a regression analysis. The former will put an upward bias to the coefficient on capital while the former will pick up industry-year-specifics that are not necessarily rooted in the capital stock.

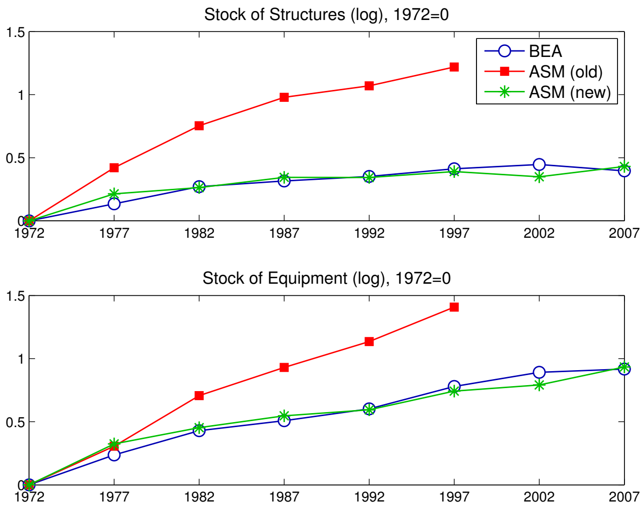
In a similar vein, I find that the old investment measure for equipment is off the benchmark as well, while structure investment comes fairly close. For this I have no other explanation than that the price indices for investment were revised very often.

In many years, (esp. 1972-1976) beginning-of-year capital stocks are not or only partially measured. I can use the end-of-year capital stock from the previous year as far as that is available. This usually leads to some hundreds replacements per year, but to many more in 1973 (24k) and 1982 (9k). In all years subsequent to a Census year and after 1988 (when annual measurement of `BAB` and `MAB` stops), we can naturally impute the beginning-of-year capital stock in this way for almost all observations (50-60k observations).

### A.4.2 Accounting for capacity utilisation

In addition to the capital stock available to the establishment, we need to know the utilisation of this capital stock to determine the capital services going into production. As pointed out in previous research (see for example Jorgenson and Griliches (1967); Basu (1996); Burnside et al. (1995)), failure to control for capacity utilisation will bias TFP to be more procyclical than it actually is because measured TFP merely reflects unmeasured (procyclical) capacity utilisation. If capacity utilisation rates become more heterogeneous in a downturn, then measured dispersion in TFP would just be a figment of specification error. Burnside et al. (1995) have suggested to use electricity or energy instead of capital. The idea in that paper is that energy/electricity is

Figure 10: Aggregate capital stock in U.S. manufacturing




*Note:* Capital stocks of the U.S. manufacturing sector published by BEA (blue circles), aggregating the old (red squares) and my refined (green stars) capital stock data in the CMF/ASM. Normalised to 0 in 1972. Clearly visible that the trend growth is off in the old ASM capital measures.

predominantly used to power capital, so variations in energy used will be a good approximation for capacity utilisation. Using energy instead of capital stocks may have the advantage that one measures actual capital services, but, on the other hand, this approach assumes constant energy efficiency. Without further knowledge of the capital stock's energy efficiency one cannot distinguish highly energy efficienct machines running at high capacity of low energy efficiency machines running at low capacity. The capital services supplied by the former are higher for two reasons: more energy-efficienct machines are presumable newer, so their productivity is likely to be much higher. Second, these more productive machines are running at full capacity. For those reasons, I gladly make use of the directly observed capacity utilistion measure from the Plant Capacity Utilisation Survey (PCU). The PCU is a subset of the ASM/CMF and collects explicit information on the utilisation of an establishment's existing capacities. This allows me to construct an explicit capital services measure and omit the energy/electricity inputs that are implicit in the utilisation rates.

Utilisation rates in the PCU are only observed for a small subsample of the data in the ASM/CMF (for about 280k of 4m observations total). I therefore use the data in the PCU to compute industry-wide utilisation rates and use them as a proxy for the other establishments. The idea is that increased demand for a certain good makes most establishments in this industry run at higher capacity.[55] This works for all years after 1974 when the PCU started. For 1972/73, I use utilization rates computed by the Federal Reserve Board[56].

So far, I have outlined how to compute the utilised replacement value in constant dollars of the capital stock the firm *owns*. In addition to its on capital stock, an establishment may rent capital

[55] Of course, this is not true if even within an industry products are imperfect substitutes due to transportation or branding.

[56] Industrial Production and Capacity Utilization – G.17; compiled by the Federal Reserve; downloaded at http://www.federalreserve.gov/datadownload/Build.aspx?rel=G17.

to produce. It would be ideal to deduce the real amount of rented capital and include it into the capital measure. Due to data limitations, I have to omit this step: Rented capital is only reported in the years until 1988 and rental payments are hard to transfer into units that correspond to the constant-dollar measure used for the establishment's own capital stock.

## A.5 Measurement of materials input

Materials are purchased on the market (materials & parts, $\texttt{CP}$, contract work, $\texttt{CW}$) or come from the materials inventory and are then used in production. Measurement is complicated by the fact that materials inventories ($\texttt{MIB}$ and $\texttt{MIE}$) comprise both materials and fuels. Therefore, I have to make an assumption about how much of changes in material inventories are driven by changes in fuel inventory. I assume that all changes in materials inventory are due to changes in materials, while the stock of fuels stays constant. Given the fact that several fuels are storable only at high cost (e.g., natural gas) this seems like a reasonable assumption. Then, I can express the value of materials used in the production process ($\texttt{VM}$) through the inventory identity

$$\begin{aligned} \texttt{MIE} &= \texttt{MIB} + (\texttt{CP} + \texttt{CW} - \texttt{VM}) \times \frac{\texttt{PIMI}}{\texttt{PIMAT}} \\ \Leftrightarrow \texttt{M} \equiv \frac{\texttt{VM}}{\texttt{PIMI}} &= \frac{\texttt{MIB} - \texttt{MIE}}{\texttt{PIMI}} + \frac{\texttt{CP} + \texttt{CW}}{\texttt{PIMAT}}. \end{aligned} \tag{27}$$

As with goods inventories above, inflows into materials inventories have to be deflated by market prices ($\texttt{PIMAT}$), while materials stocks have to be deflated by inventory prices ($\texttt{PIMI}$). The former comes from the NBER productivity database, the latter could in part be obtained on the industry level from BLS's multifactor productivity tables. As with goods inventory deflators above, these are only available since 1987 and for the same reasons as above I approximate $\texttt{PIMI}$ with $\texttt{PIMAT}$.

## A.6 Measurement of energy input

I use several measures of energy inputs: electricity, fuels and a combination of them.

### A.6.1 Electricity

Electricity used in the production process ($\texttt{EL}$) is easily measured. It consists of the quantity of purchased electricity ($\texttt{PE}$) and the difference between generated and sold electricity ($\texttt{GE} - \texttt{SE}$). Since electricity is hardly storable, we do not have to worry about something like an electricity inventory:

$$\texttt{EL} = \texttt{PE} + \texttt{GE} - \texttt{SE}. \tag{28}$$

For later purposes, it makes sense to impute a price for electricity the establishment pays: $\texttt{PIEL} = \frac{\texttt{EE}}{\texttt{PE}}$. N.B.: If $\texttt{GE} = 0$, then the fuel used for electricity generation ($\texttt{VFEL}$) is zero as well.

### A.6.2 Fuels

Fuels used in production (nominally expressed as $\texttt{VF}$) can come from fuel purchases ($\texttt{CF}$) or from the materials/fuels inventory. As outlined above, I assume that any change in materials inventory ($\texttt{MIE} - \texttt{MIB}$) is due to materials only and that the fuel stock in the inventory stays constant. Then, fuel purchases can be used in the production process (oil used to produce plastics) or for electricity generation (oil burned in an electricity generator). The latter quantity is not observed, but must be zero for the vast majority of observations that do not produce any electricity; for those observations $\texttt{VF} = \texttt{CF}$. If this is not the case, then I assume that generated electricity is produced with a linear technology. In particular, I assume that 1$ of fuel expenditures can be converted into electricity that could be sold for 1$ (taking into account overhead etc). The idea is that a firm will only find it profitable to produce its own electricity rather than purchasing it when the price of fuel (contained in VFEL) relative that of electricity is not too high and that it can relatively easily substitute among different fuels. $\texttt{GE} = \frac{\texttt{VFEL}}{\texttt{PIEL}} \Leftrightarrow \texttt{VFEL} = \texttt{GE} \times \frac{\texttt{EE}}{\texttt{PE}}$. Fuel used in production (F) equals the value of fuels (VF) deflated by the energy price index PIEN,

$$\begin{aligned}
\texttt{VF} &= \texttt{CF} - \texttt{VFEL} \\
\texttt{VF} &= \texttt{CF} - \texttt{GE} \times \frac{\texttt{EE}}{\texttt{PE}} \\
\texttt{F} &\equiv \frac{\texttt{VF}}{\texttt{PIEN}} = \frac{\texttt{CF} - \texttt{GE} \times \frac{\texttt{EE}}{\texttt{PE}}}{\texttt{PIEN}} \qquad (29)
\end{aligned}$$

where I assume that the price for fuels equals PIEN, the price deflator for overall energy from the NBER-CES database.

### A.6.3 Total energy

Again, I assume that fuels inventory (recorded as part of materials inventory in the ASM/CMF) is unchanged. This means that all fuel purchases are immediately consumed in production or in electricity generation. Total energy expenditures (VE) comprise those for fuels (CF) and electricity (EE); the nominal value is:

$$\begin{aligned}
\texttt{VE} &= \texttt{CF} + \texttt{EE} \\
\texttt{E} &\equiv \frac{\texttt{VEN}}{\texttt{PIEN}} = \frac{\texttt{CF} + \texttt{EE}}{\texttt{PIEN}} \qquad (30)
\end{aligned}$$

where I use PIEN, the industry-specific energy price deflator from the NBER-CES productivity database, to obtain real energy input, E.

## A.7 Construction of cost shares

The baseline estimation described in Appendix C.3 requires knowing the cost shares of factor inputs. They are constructed as follows

$$\begin{aligned}
c_L &= \frac{\texttt{SW}}{TC} \\
c_K &= \frac{rK}{TC} \\
c_M &= \frac{\texttt{VM}}{TC} = \frac{\texttt{CP} + \texttt{CW} + \texttt{MIB} - \texttt{MIE}}{TC} \\
c_E &= \frac{\texttt{VE}}{TC} = \frac{\texttt{CF} + \texttt{EE}}{TC} \\
TC &= \texttt{SW} + rK + \texttt{VM} + \texttt{VE}.
\end{aligned}$$

All variables are expressed in period-$t$ nominal costs and except $r$ and $K$ are observed in the original dataset. $K$ is the real capital stock (in year-2005 dollars) constructed as described in Appendix A.4, $r_t$ denotes the *nominal* rental rate (year-$t$ dollars rent paid per one year-2005 dollar worth of capital). Multiplying this rental rate, $r_t$, by the real capital stock, $K_t$, gives the nominal period-$t$ capital cost of financing the stock in period $t$. This makes it accord with the other nominal values. The rental rate is constructed from the BLS Capital Tables[57] by dividing corporate capital income (Table 3a) by the real capital stock (Table 4a). The latter variable is expressed in constant (year-2005 dollar), while the former is expressed in current-period dollars, so $rK$ are the capital cost expressed in period-$t$ dollars. Note that capital cost merely includes rent and depreciation, not physical utilisation cost which is captured in the energy cost share.[58] Table 8 displays industry summary statistics on the average plant variables.

# B Empirics – details

## B.1 Correcting the productivity estimate

In contrast to looking at the dispersion of TFP growth rates, analysing TFP levels is not as straightforward. The distribution of TFP levels is likely to differ substantially across industries in terms of the central and higher moments which reflects the industry's technological and competitive environment. Changes in the industry-level TFP dispersion might hence not be observed by just looking at the entire cross section, a problem that did not appear in the cross-sectional dispersion of growth rates. For that reason, I will have to look at cross-sectional dispersion *within* industries. The definition of industry should be narrow enough to overcome this between industry heterogeneity as far as possible. 6-digit NAICS level industries are feasible in my data and should be reasonably narrow. Still, some of the remaining within-industry level heterogeneity might be driven by differences in the type of products rather than productivity. I am aware of this limitation, but the limitation of data do not permit a more in-depth analysis.

The plant-level productivity has to be corrected in several ways. First, it needs to be detrended, second, recentered at zero, thirdly scaled by the long-run variance. These normalisation steps warrant more explanation. First, industries may well have different long-run productivity growth. As a result, industries that diverge more and more form the average growth trend (think semiconductors) are more and more important in determining the cross-sectional dispersion. To correct for that, I fit a simple econometric model with a linear trend. This trend is allowed to differ across industries: $a_{ijt} = g_j t + z_{ijt}$ The resulting term $z_{ijt}$ reflects productivity dispersion around the industry's long-run growth trend. It still needs to be corrected in more ways to obtain a proper cross-sectional dispersion measure $Disp_t$:

$$Disp_t \equiv Median_t \left[ Var_{jt} \left( \frac{z_{ijt} - \overline{z}_j}{\sigma_j} \right) \right] \tag{2}$$

As Figure 11 (a) illustrates, different industries may have a different average *level*, $\overline{z}_j$, of productivity even after correcting for the growth *trend*. Changes in the overall cross-sectional dispersion could hence be driven by changes in "outlier" industries. I hence normalise each industry by its long-run

[57] "Capital by Asset Type for NIPA-level Manufacturing Industries" downloaded from http://www.bls.gov/mfp/mprdload.htm.

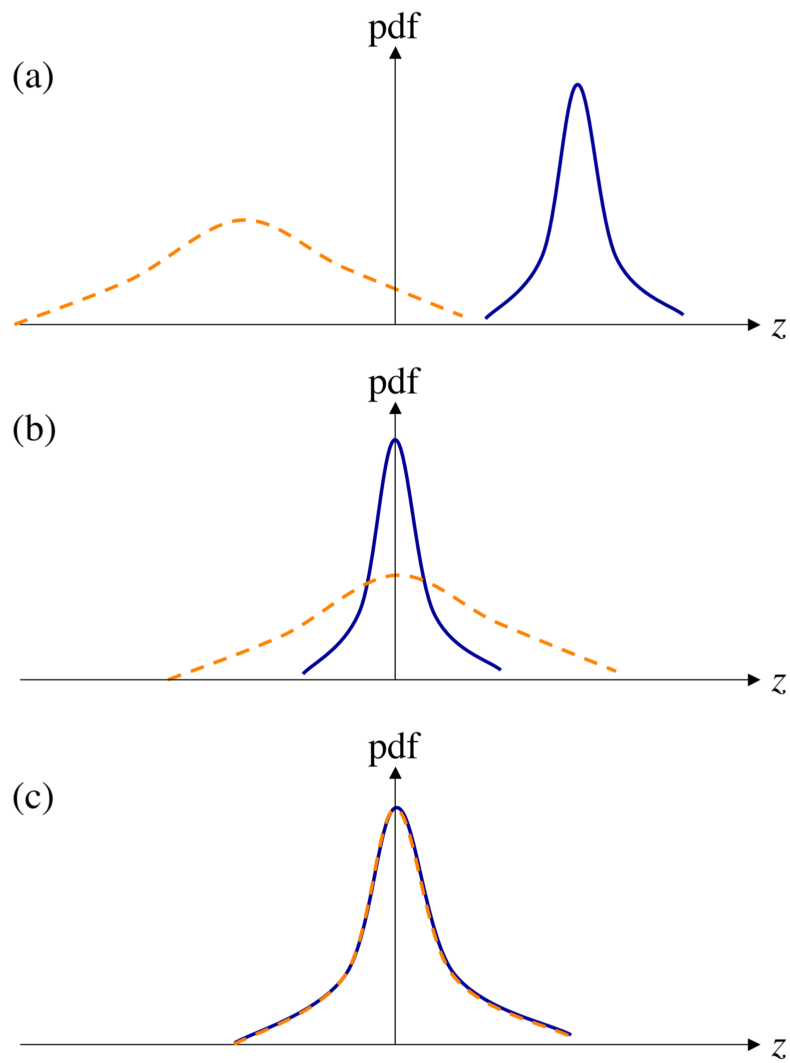
[58] This obviously assumes that depreciation is not influenced by utilisation.

Table 8: Summary statistics: Average plant size in industries

| NAICS | Industry | Gross Output | Value Added | Capital | Hours worked | Materials | Energy |
|---|---|---|---|---|---|---|---|
| 311 | Food | 60.1 | 23.6 | 25.4 | 372.2 | 35.3 | 1.1 |
| 312 | Beverage and tobacco | 139.6 | 97.2 | 61.6 | 389.9 | 41.3 | 1.1 |
| 313 | Textiles and fabrics | 26.3 | 10.3 | 28.4 | 540.5 | 14.8 | 1.1 |
| 314 | Textile mill products | 23.8 | 10.3 | 10.5 | 306.5 | 13.0 | 0.4 |
| 315 | Apparel and accessories | 10.9 | 5.6 | 2.9 | 272.7 | 5.2 | 0.1 |
| 316 | Leather and allied products | 18.3 | 9.0 | 5.8 | 406.5 | 9.1 | 0.2 |
| 321 | Wood products | 18.0 | 7.3 | 10.3 | 201.6 | 10.2 | 0.4 |
| 322 | Paper | 56.1 | 26.9 | 60.2 | 407.0 | 25.9 | 3.2 |
| 323 | Printed matter | 14.3 | 8.8 | 8.9 | 195.1 | 5.3 | 0.2 |
| 324 | Petroleum and coal | 351.9 | 47.1 | 198.0 | 552.2 | 292.4 | 12.4 |
| 325 | Chemical products | 97.6 | 54.2 | 71.2 | 383.3 | 39.1 | 4.2 |
| 326 | Plastics and rubber | 29.8 | 14.3 | 20.2 | 340.2 | 14.7 | 0.8 |
| 327 | Nonmetallic minerals | 17.4 | 9.8 | 19.7 | 191.8 | 6.0 | 1.5 |
| 331 | Primary metals | 71.8 | 24.7 | 63.3 | 477.8 | 43.2 | 3.9 |
| 332 | Fabricated metals | 23.9 | 13.1 | 12.1 | 278.2 | 10.3 | 0.4 |
| 333 | Machinery | 38.5 | 20.6 | 17.6 | 383.7 | 17.5 | 0.4 |
| 334 | Computer and electronics | 52.4 | 32.6 | 38.6 | 889.5 | 19.1 | 0.6 |
| 335 | Electrical equipment | 45.2 | 23.5 | 22.2 | 550.8 | 21.0 | 0.7 |
| 336.1-3 | Motor vehicles | 113.3 | 40.3 | 47.5 | 689.5 | 72.0 | 1.0 |
| 336.4-9 | Other transportation eqpmt. | 159.1 | 94.3 | 61.4 | 1,348.5 | 63.4 | 1.3 |
| 337 | Furniture | 18.2 | 10.3 | 7.7 | 306.6 | 7.7 | 0.2 |
| 339 | Misc. manufacturing | 19.0 | 12.2 | 8.0 | 250.1 | 6.6 | 0.2 |
| Non-Durable | | 60.3 | 25.1 | 37.7 | 357.8 | 33.3 | 1.9 |
| Durable | | 41.8 | 20.7 | 23.1 | 426.3 | 20.2 | 0.8 |
| Total Manufacturing | | 50.0 | 22.6 | 29.6 | 0.4 | 26.1 | 1.3 |

*Note:* Panel descrition see Table **??**. Average annual value of inputs and outputs of an ASM establishment per industry. Gross output, Value Added, Capital, Materials and Energy are expressed in thousand year 2005-dollars, Hours Worked in thousand hours.

Figure 11: Industry differences in productivity dispersion




*Note:* Recentering and scaling of industry productivity dispersion makes dispersion industries comparable among each other and over time.

industry mean, $\overline{z}_j$. The effect of this step is illustrated in Figure 11 (b).

Finally, it is well-known that the within industry productivity dispersion varies greatly. Syverson (2004) reports that, in a cross section, the within-industry dispersion varies greatly across industries. Some industries like cement are very spread-out in general while highly competitive industries are more compressed. As a consequence, I scale each industry by its long-run standard deviation $\sigma_j$. This scaling step is illustrated in Figure 11 (c). Note that the industry mean and standard deviation used in steps 2 and 3 are not dependent on time $t$. Otherwise, any time variation in dispersion would be lost. As an additional step, I also consider correcting for the long-run skewness. While the typical industry has some positive skew, the results from correcting for skewness and not are very similar.

### B.2 Measurement error

The inputs and outputs in equation (1) could be measured imperfectly and bias my empirical results about productivity dispersion. Measurement error is only a problem if it is cyclical in a way that could generate the observed cyclicality in cross-sectional productivity dispersion. With that in mind, I will address possible sources of measurement error. Inputs on the right hand side of equation (1) may be mismeasured (mismeasurement or misreporting on the plant level or mistabulations at the statistical agency). Similarly, establishment-specific inputs prices in capital, materials and energy will have the same effect as mismeasurement. Census claims that of the variables in (1) the output and labour variables are measured fairly well. Along with Foster et al. (2008) I follow that view and focus on measurement error in materials, energy and capital. As an example, I highlight the effect so measurement error in capital, though the same reasoning holds for the other inputs. Suppose I measure capital with a relative error, so the true capital stock $K$ is mismeasured as $\tilde{K} = K(1 + \varepsilon^k)$. Mismeasurement in capital can arise form within-industry differences in capital-embodied technical change and $\varepsilon^k$ is a percentage deviation how much the quality-adjusted capital stock is misreported or by how much it differs across plants within an industry. Omitting other input factors for simplicity (the same logic will hold) and expressing measured TFP in logs, we have (omitting indices for readability):

$$\begin{aligned}\tilde{a} &= y - \beta^k \left(k + \varepsilon^k\right) \\ &= a - \beta^k \varepsilon^k\end{aligned}$$

where $\tilde{a}$ is the measured productivity and $a$ actual productivity which we do not observe. I assume that the measurement error in one input is orthogonal to output and other inputs: $E[\varepsilon^r \varepsilon^s] = 0 \ \forall r, s = y, k, l, m, e, \ r \neq s$. Then, my dispersion measure, the cross-sectional standard deviation of TFP, is

$$SD(\tilde{a}) = \sqrt{V(\tilde{a})} = \sqrt{V(a) + (\beta^k)^2 V(\varepsilon^k)}$$

On average, the cross-sectional dispersion is 10% higher in a recession than in a boom.

$$SD(\tilde{a}^{rec}) = 1.1 SD(\tilde{a}^{boom})$$

Could this cyclicality be driven by cyclical measurement error? In the worst case, the cross sectional variation in true productivity, $V(a)$, remains constant and all the measured increase in $SD(\tilde{a})$ is due to an increase in measurement error $V(\varepsilon^k)$. If so, by what factor $\xi > 1$ would $SD(\varepsilon^k)$ have to be more spread-out in a recession than in a boom: $\xi SD(\varepsilon^{k\text{ boom}}) = SD(\varepsilon^{k\text{ rec}})$?

$$\begin{aligned}
SD(\tilde{a}^{\text{rec}}) &= 1.1 SD(\tilde{a}^{\text{boom}}) \\
\Leftrightarrow \sqrt{V(a) + \xi^2(\beta^k)^2 V(\varepsilon^{k\text{ boom}})} &= 1.1\sqrt{V(a) + (\beta^k)^2 V(\varepsilon^{k\text{ boom}})} \\
\Leftrightarrow \xi^2(\beta^k)^2 V(\varepsilon^{k\text{ boom}}) &= 0.21 V(a) + 1.21(\beta^k)^2 V(\varepsilon^{k\text{ boom}}) \\
\Leftrightarrow \xi &= \sqrt{\frac{0.21}{(\beta^k)^2}\frac{V(a)}{V(\varepsilon^{k\text{ boom}})} + 1.21} \qquad (31)
\end{aligned}$$

As we can see from equation (31), the smaller $\beta^k$ and the larger $V(a)/V(\varepsilon)$, the larger are the upswings in measurement error that are required in order to explain the measured increase in productivity dispersion.

Figure 12: The effect of measurement error

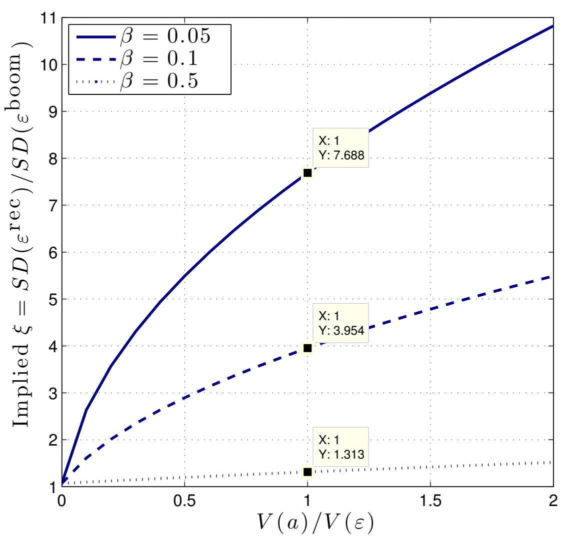



Figure 12 displays quantitative examples for this relationship. In the case of capital, for example, $\beta^k \approx 0.05$ (see Table 9). If one assumes $V(a)/V(\varepsilon) = 1$, i.e. in a boom, mismeasurement is as strong as the true productivity dispersion, then the measurement error in capital has to increase 7.5-fold in order to explain the 10% rise in the measured cross sectional productivity dispersion! Note that these quantitative experiments are a lower bound because I made two very negative assumptions: All of observed dispersion cyclicality is exclusively due to cyclical measurement error and $V(a)/V(\varepsilon) = 1$ looks like a strong assumption too. A similar argument holds for measurement error in materials. Because the coefficient estimate on materials is larger ($\beta^m \approx 0.5$), upswings of "only" 30% in mismeasurement would be sufficient to overturn my results. While I do not deny the possibility of cyclical measurement error driving some of my results, it is hard to believe that the observed dispersion dynamics are entirely driven by mismeasurement of the indicated magnitude.

# C Further results

## C.1 Returns to scale

The dispersion estimates about productivity are the main interest of this paper. They result from a production function estimation and I will display these results. Be reminded that I measure labour as hours worked, materials as the real value of materials used, capital as a quality-adjusted constant-dollar valued assets and energy as electricity. I do not find systematic differences separating hours worked into white collar vs. blue collar workers. As an alternative to energy used, I also used electricity, but this, too, does not change the results significantly. I use the above-described OP procedure to estimate equation (1') separately for durable and non-durable industries. There has been a large body of research suggesting differences in the technology between those two sectors. Therefore, it makes sense to estimate returns to scale and plant-level productivity separately for durables and non-durables. Table 9 displays the results of this regression in durables goods industries (NAICS 321, 327-339) and the analogous results for non-durables (NAICS 311-316, 322-326).

Table 9: Returns to scale in non-durables and durables

| Input | Coefficient estimates ... | |
|---|---|---|
| | Non-durables | Durables |
| Capital | 0.101*** | 0.053*** |
| | (0.002) | (0.010) |
| Hours Worked | 0.235*** | 0.292*** |
| | (0.002) | (0.007) |
| Materials | 0.471*** | 0.520*** |
| | (0.001) | (0.006) |
| Energy | 0.104*** | 0.077*** |
| | (0.001) | (0.001) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

The coefficients on the input factors are consistent with previous estimates on the industry-level. They are also close to the empirically observed cost shares.[59] Compared to the cost shares, the coefficient on the capital stock is smaller than the capital cost share and the coefficient on energy is larger than the energy cost share. This suggests that the capital stock per se is not productive and energy picks up the effects of the *utilised* capital stock. This result confirms the findings of Burnside et al. (1995) who claim that electricity accurately measures capital services. Taken together, the coefficients of capital and electricity add up to a value that is close to the empirically observed cost share of capital.

Returns to scale in non-durables are constant. The production function coefficients sum up to slightly less than unity. In non-durables returns to scale are decreasing (around 0.9) while durable manufacturing is closer to constant returns to scale (around 0.95). These findings are obtained using plant-level data and they conform well with previous research by Basu and Fernald

[59] As an alternative an robustness check, I infer productivity by subtracting cost-weighted inputs from output. These results are displayed in Appendix C.3.

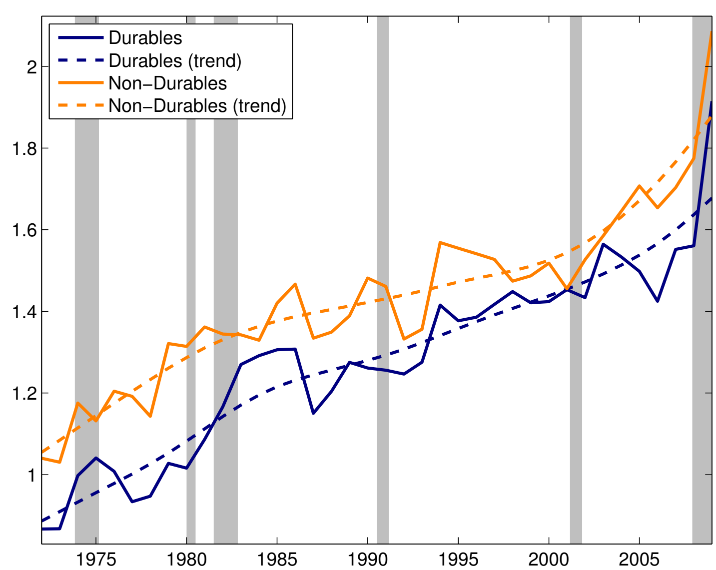
(1995); Burnside et al. (1995); Harrison (2003) who also find higher returns to scale in durable manufacturing. This finding will play an important role in explaining differences in the cyclicality of productivity dispersion below. The results of near constant returns to scale in durables and slightly decreasing returns to scale in non-durables is consistent with a fixed factor of production in durables and otherwise constant slightly decreasing returns to scale. Example of fixed factors of production can be overhead labour or capital such as managers or capital structures.

Note that my results do not claim that fixed factors of production are absent in non-durable goods industries. Higher estimated returns to scale in durables are consistent with the view that fixed factors of production are more prevalent in durables than in non-durable goods industries. Related research has found evidence supporting this result. Eisfeldt and Papanikolaou (2013) for example find that levels of organisational capital or intangible assets such as know-how or supply chain networks tend to be more prevalent in durable goods industries.

## C.2 Long-run changes in productivity dispersion

Figure 13 displays the time series of productivity dispersion before HP-filtering. Quite remarkably, the long-run pattern of productivity dispersion in both durables and non-durables goods industries changes over time. Over the entire sample, it doubles in both durables and non-durables. The growth in long-run dispersion is a bit stronger in the 1970's, 1980's and 2000's then in the rather sluggish 1990's. Similar long-run changes have been noted in Beaudry et al. (2001). They examine the dispersion of firm-level profit *rates* in the UK and find a U-shaped pattern. They explain the long-run changes with learning in the face of macro uncertainty due to policy changes. The same macro-level uncertainty and learning could drive productivity dispersion in the U.S, but my sample is missing the downward-sloping part of the U shape. This is not too surprising because my sample starts later than the UK panel these authors are using.

Figure 13: Trends and cycles in dispersion



## C.3 The dispersion of Solow residuals

The approach to identify production elasticities with cost shares is attractive because it merely requires the assumptions of constant returns to scale and competitive factor markets.

The structural estimator proposed by Olley and Pakes (1996) is attractive, for it permits estimation of plant-level productivity while solving the endogeneity problem. It is, however, subject to some assumptions about timing of input choices and strong monotonicity of investment in productivity. Especially the latter assumption seems strong and non-convex adjustment costs are a simple (and plausible – see Cooper and Haltiwanger (2006)) example that would lead to lumpy investment and violate the strong monotonicity. The estimation then has to discard all observations with zero investment. In my sample, this is only the case for 12.4% of all observations, so by using the Olley-Pakes estimator I do not have to discard a lot of observations. Nevertheless, I want to make sure that my results are not specific to this estimation technique which relies on specific structural assumptions.

A long-standing tradition in the productivity literature consists of identifying $a_{ijt}$ from Solow residuals.This approach requires assuming constant returns to scale and competitive factor markets. These assumptions look reasonable given that I found close to constant returns to scale in Section C.1. They are also in line with what many other people have found empirically, see for example Burnside et al. (1995); Burnside (1996); Basu and Fernald (1997) for industry-level evidence and Lee and Nguyen (2002) for plant-level evidence in the clothing industry. The most important upside is that there are no further assumptions about the timing of input choice and the dependency of investment on productivity. The downside in the context of this paper is that constant returns to scale are inconsistent with the fixed overhead factor in my model. Still, I present the dynamic properties of productivity dispersion when inferred from Solow residuals.

Under the above two weak assumptions, production elasticities can be identified from cost shares.[60] The production function can be rewritten as follows

$$\begin{aligned} y_{ijt} &= a_{ijt} + \beta^k k_{ijt} + \beta^l l_{ijt} + \beta^m m_{ijt} + \beta^e e_{ijt} \\ &= a_{ijt} + c_j^k k_{ijt} + c_j^l l_{ijt} + c_j^m m_{ijt} + c_j^e e_{ijt} \end{aligned}$$

where the cost shares are defined in Section A.7. I assume that these cost shares are constant within 6-digit NAICS industries. This again reflects the idea that the production technology of an industry is constant. As above, we still need to detrend $a_{ijt}$, recenter and scale the residual $z_{ijt}$ as described in B.1. Note that the capital cost share is composed of the cost for both structure and equipment capital (which have very different rates of return $r_t$): $c_j^k = \frac{r_{jt}^{ks} K_t^s + r_{jt}^{ke} K_t^e}{Cost_{jt}}$. I use cost shares on the 6-digit NAICS level to account for fine differences in the production structure of indstries. Table 10 gives an overview of cost shares on the 3-digit NAICS level so that the reader may get an impression of the heterogeneity.

**Results** Like in the main body of the paper, I correlate the HP-100 filtered deviations of the $Disp_t^n$ and $Disp_t^d$ measures with industrial output. Recall that the cleansing view posits dispersion going down in a recession. What we see in Figure 14, however, looks again like the opposite. Rather, the overall time series look similar to the one constructed from Olley-Pakes residuals. Although

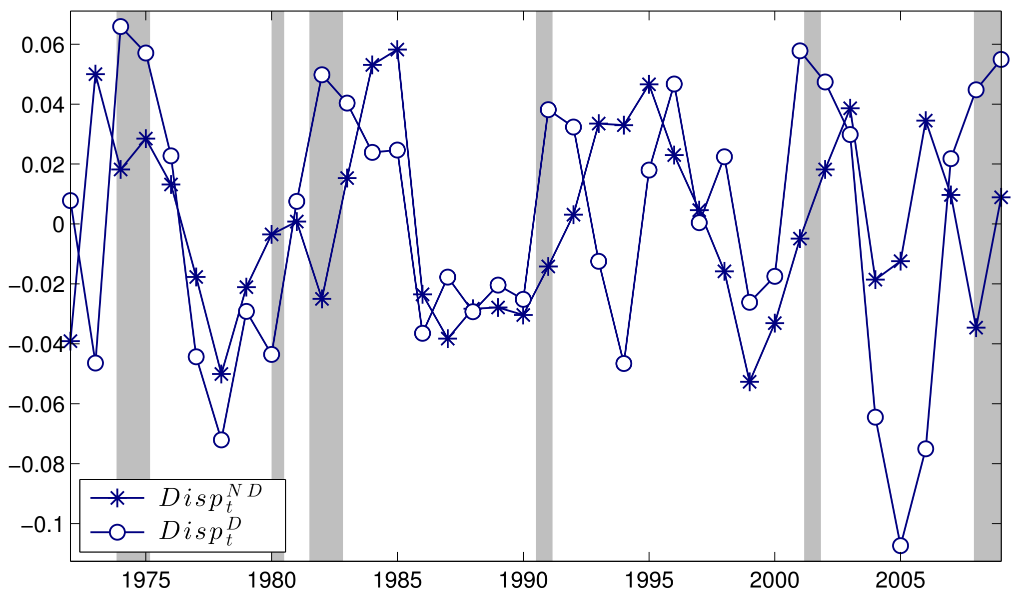
[60] Firms FOC dictate that factor prices are equal to marginal products $w = \partial Y/\partial L = \beta^l Y/L$. Labour costs are then $\alpha Y$ and similarly for the other inputs. If all production elasticities sum to unity, then the cost share is $c_L = wL/Cost = \beta^l Y/[(\beta^k + \beta^l + \beta^m + \beta^e)Y] = \beta^l$.

Table 10: Summary statistics of ASM panel: Cost shares

| NAICS | Industry | Labour | Capital | | Mat. | Energy |
|---|---|---|---|---|---|---|
| | | | Str. | Eqp. | | |
| 311 | Food | 0.13 | 0.06 | 0.07 | 0.72 | 0.02 |
| 312 | Beverage and tobacco | 0.13 | 0.10 | 0.10 | 0.66 | 0.01 |
| 313 | Textiles and fabrics | 0.20 | 0.04 | 0.05 | 0.65 | 0.04 |
| 314 | Textile mill products | 0.28 | 0.02 | 0.02 | 0.66 | 0.02 |
| 315 | Apparel and accessories | 0.27 | 0.02 | 0.03 | 0.68 | 0.01 |
| 316 | Leather and allied products | 0.34 | 0.03 | 0.03 | 0.57 | 0.01 |
| 321 | Wood products | 0.20 | 0.03 | 0.04 | 0.70 | 0.02 |
| 322 | Paper | 0.18 | 0.02 | 0.07 | 0.67 | 0.02 |
| 323 | Printed matter | 0.37 | 0.03 | 0.05 | 0.52 | 0.02 |
| 324 | Petroleum and coal | 0.11 | 0.05 | 0.07 | 0.72 | 0.03 |
| 325 | Chemical products | 0.14 | 0.10 | 0.13 | 0.57 | 0.04 |
| 326 | Plastics and rubber | 0.22 | 0.04 | 0.10 | 0.60 | 0.03 |
| 327 | Nonmetallic minerals | 0.28 | 0.06 | 0.10 | 0.45 | 0.07 |
| 331 | Primary metals | 0.16 | 0.03 | 0.04 | 0.68 | 0.04 |
| 332 | Fabricated metals | 0.30 | 0.04 | 0.08 | 0.55 | 0.02 |
| 333 | Machinery | 0.31 | 0.03 | 0.06 | 0.59 | 0.01 |
| 334 | Computer and electronics | 0.35 | 0.02 | 0.04 | 0.56 | 0.01 |
| 335 | Electrical equipment | 0.23 | 0.04 | 0.07 | 0.62 | 0.01 |
| 336.1-3 | Motor vehicles | 0.21 | 0.02 | 0.06 | 0.70 | 0.01 |
| 336.4-9 | Other transportation eqpmt. | 0.30 | 0.03 | 0.04 | 0.61 | 0.01 |
| 337 | Furniture | 0.31 | 0.04 | 0.05 | 0.57 | 0.02 |
| 339 | Misc. manufacturing | 0.32 | 0.06 | 0.07 | 0.52 | 0.02 |
| Total Manufacturing | | 0.25 | 0.04 | 0.06 | 0.61 | 0.02 |
| Average Non-durable Manufacturing | | 0.17 | 0.05 | 0.12 | 0.60 | 0.02 |
| Average Durable Manufacturing | | 0.27 | 0.04 | 0.10 | 0.55 | 0.02 |

*Note:* Cost shares on the 3-digit NAICS industry level in the ASM panel, averaged over the years 1972-2009. Construction of cost shares described in A.7.

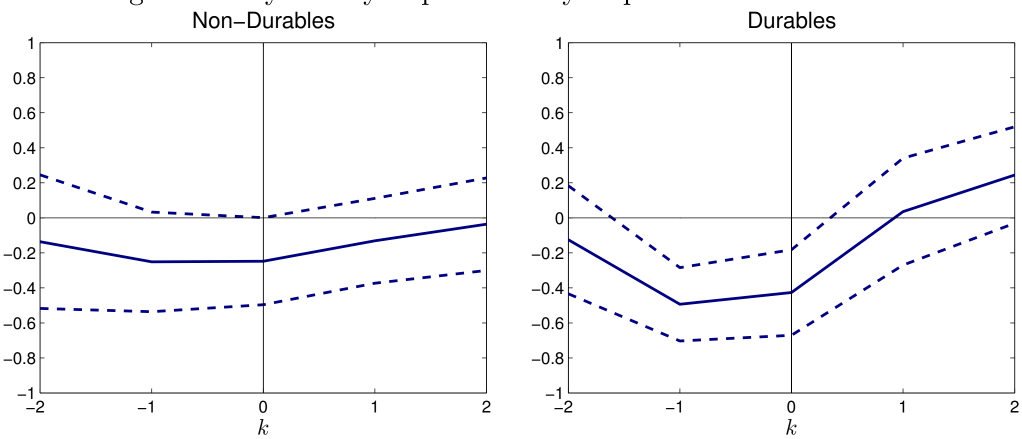
Figure 14: Productivity dispersion – Solow residuals




*Note:* Time series plot of the cyclical components of output and productivity dispersion in durable goods producing manufacturing. The dispersion measure is as defined in equation (2). Left panel displays non-durables, right panel displays durables, shaded bars denote NBER recessions.

dispersion is not as volatile, the peaks of the time series still coincide with NBER recessions. Interestingly, dispersion in non-durables does not spike in 2008/09 as it did with the Olley-Pakes productivity estimates.

Figure 15: Cyclicality of productivity dispersion – Solow residuals




*Note:* Correlograms display the correlation between the cyclical component (HP-100 filtered time series) of output and productivity dispersion in non-durables and durables respectively: $Corr(Y_t, Disp_{t+k})$. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

The overall results of productivity dispersion constructed from Solow residuals conform well with those constructed from Olley-Pakes estimates. As Figure 15 and Table 11 show, dispersion is still strongly countercyclical in durables and weakly countercyclical in non-durables. As Table 12 shows, the countercyclicality is pervasive in various dispersion measures throughout the productivity distribution. Again, Table 13 shows that productivity dispersion is negatively correlated with a variety of the business cycle. The striking similarity between the dispersion measures, once constructed from Olley-Pakes estimates, once from Solow residuals, suggests that the overall negative correlation of dispersion with the business cycle is robust to the specific method used.

## C.4 How volatile are the cost of fixed factors?

One criticism against the importance of the sullying effect could be that wages are generally believed to be very sticky. This would greatly dampen the sullying effect. One needs to take a close look at the wage in expression (**??**): it is the cost of the *fixed* labour input. This fixed labour input can be thought of a place holder for several fixed factors such as managers and rents for structures among others. Eisfeldt and Papanikolaou (2013) provide evidence of organisational capital that closely resembles a fixed input. This organisational capital is more valuable when demand for the firm's output is high, therefore its the compensation of organisational capital is higher in a boom as well. I focus on the interpretation of fixed inputs as a manager. The cyclical behaviour of the cost for this fixed factor is very different than the cost for normal production labour input. To support this claim with empirical evidence, I turn to data on managerial compensation and analyse the cyclicality of their income. The data come from ExecuComp, a database that covers the top executive pay in a large cross section of firms.[61] Their real income growth rate is computed and correlated with the business cycle in Figure 16.

[61] A more comprehensive analysis could be carried out by studying the compensation of supervisory versus non-supervisory workers in the Current Population Survey.

Table 11: Cross-correlations of dispersion – Solow residuals

| Lead/Lag | Correlation of output and dispersion in ... Non-durables | Durables |
|---|---|---|
| -2 | -0.136 | -0.125 |
| | (0.194) | (0.157) |
| -1 | -0.252* | -0.494*** |
| | (0.145) | (0.107) |
| 0 | -0.248* | -0.426*** |
| | (0.127) | (0.125) |
| 1 | -0.131 | 0.036 |
| | (0.124) | (0.156) |
| 2 | -0.036 | 0.245* |
| | (0.135) | (0.140) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

One would expect the managerial wages to be mostly acyclical if they behaved like normal production worker wages. But Figure 16 paints a different picture. All components of managerial compensation – be it base salary, payment in shares or stock options – are pronouncedly procyclical.

# D Proofs

**Proof of Proposition 1** Suppose not, then $\hat{A}_t < \alpha\hat{r}_t + \beta\hat{w}_t$. Given decreasing returns to scale, i.e. $\beta < (1-\alpha)$, this implies that $\hat{A}_t < \alpha\hat{r}_t + (1-\alpha)\hat{w}_t$. By the firm first-order condition, equation (5), demand for capital $k_{it}$ of every active firm would *fall* after a rise in $A_t$. By Lemma 1 (see below), the survival cutoff, $z_t^*$, would then also rise after a rise in $A_t$. This means that aggregate demand for capital, $K_t = N_t \int_{z_t^*}^{\infty} k(z_{it})dF(z)$, would fall because both the integrand, $k(z_{it})$, and the domain of the integration $[z_t^*, \infty)$ decline. Note that the measure of firms, $N_t = N_{t-1}^* + N^E$, is unchanged because the measure of entrants is fixed and the measure of incumbents is given at $t$. Lower aggregate demand for capital is only consistent with a declining interest rate $r_t$ if the capital supply curve is a declining function thus contradicting the original assumption. The same logic applies to labour and the wage rate. ∎

**Proof of Lemma 1** Suppose not, i.e. the cutoff $z_t^*$ declines. Then, the measure of active firms, $N_t^*$, increases because survival conditions are easier. Since every firm is required to hire $c_f$ units of overheads, overall demand for overheads rises with the measure of active firms. Higher demand would then increase the price of overheads, $\omega_t$. But when $\omega_t$ rises after a rise in $A_t$, the cutoff $z_t^*$ can only decline if $\hat{A}_t > \alpha\hat{r}_t + \beta\hat{w}_t$, thus contradicting the original assumption. ∎

**Proof of Proposition 2** Suppose the decline in $\omega_t$ were so strong that it would dominate the decline in $\frac{1}{A_t}\left(\frac{r_t}{\alpha}\right)^{\alpha}\left(\frac{w_t}{\beta}\right)^{\beta}$ such that the survival cutoff $z_t^*$ falls in a recession. Then, the measure of

Table 12: Correlation of dispersion measures and output – Solow residuals

| Correlation of output in with cross-sectional... | Non-durables | Durables |
|---|---|---|
| Variance | -0.242 | -0.49*** |
| | (0.168) | (0.098) |
| Standard Deviation | -0.248* | -0.426*** |
| | (0.127) | (0.125) |
| Inter-quartile range | -0.136 | -0.447*** |
| | (0.155) | (0.098) |
| Inter-decile range | -0.187 | -0.473*** |
| | (0.188) | (0.094) |

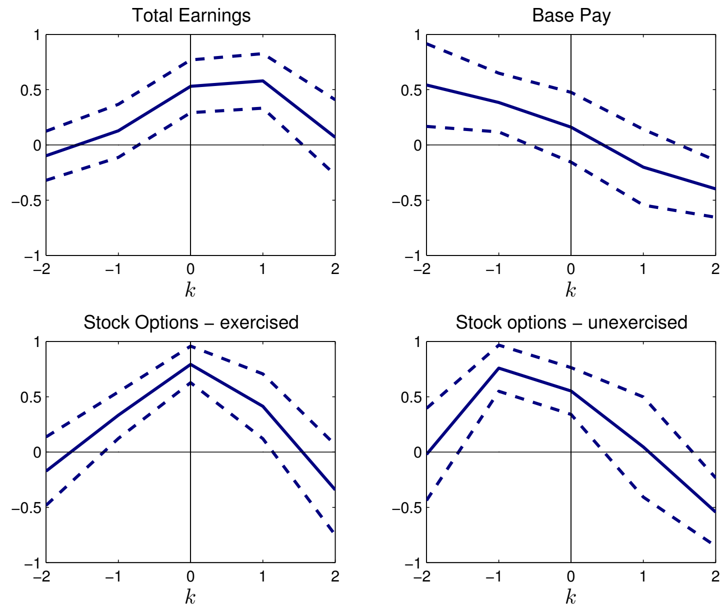
*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

Table 13: Correlation of output measures and dispersion – Solow residuals

| Correlation of dispersion in with... | Non-durables | Durables |
|---|---|---|
| Production – HP filtered ($\lambda = 100$) | -0.248* | -0.426*** |
| | (0.127) | (0.125) |
| Production – HP filtered ($\lambda = 6.25$) | -0.029 | -0.499*** |
| | (0.164) | (0.114) |
| Production growth rate | 0.085 | -0.406*** |
| | (0.16) | (0.148) |
| GDP – HP filtered ($\lambda = 100$) | -0.213 | -0.380*** |
| | (0.171) | (0.128) |
| GDP – HP filtered ($\lambda = 6.25$) | 0.06 | -0.537*** |
| | (0.174) | (0.088) |
| GDP growth rate | 0.127 | -0.442*** |
| | (0.158) | (0.129) |
| No. of NBER boom months/year | 0.053 | -0.441*** |
| | (0.115) | (0.128) |
| Census rotation years dropped | -0.213 | -0.518*** |
| | (0.187) | (0.11) |

*Note:* *, **, *** significantly different from 0 at the 10%, 5%, 1% level, respectively. Dispersion measures defined analogously to equation (2); description of standard errors see Figure 4.

Figure 16: Cyclicality of executive compensation




*Note:* Correlogram of different portions of executive compensation: $Corr(GDP_t, w_{t+k})$, dashed lines denote 95% confidence intervals constructed as describe in Figure 4. Correlated data are HP filtered residuals of GDP and the aggregate real earnings.

Table 14: Components of executive compensation

| Compensation Component | Share | Volatility (over time) | $Corr(GDP_t, w_{t+k})$ |
|---|---|---|---|
| Base Pay | 15.28% | 0.102 | 0.161 |
| Stock Options – exercised | 19.39% | 0.430 | 0.792 |
| Stock Options – unexercised | 65.32% | 0.361 | 0.552 |
| Total Earnings | 100.00% | 0.166 | 0.531 |

*Note:* Components of executive compensation are the average share of each component in aggregate earnings. Volatility of each component is the standard deviation over time of HP filtered residuals of aggregate earnings in each category.
Data come from the Compustat – Execucomp (Annual Compensation) database, a panel of top executives in 3,200 firms 1992-2009. Each component of nominal earnings was deflated using the consumer price index to obtain real earnings. Total Earnings (`TDC1`) comprises Base Pay (`SALARY`), the value of exercised stock options (`OPT_EXER_VAL`) and the value of exercisable stock options that were not exercised (`OPT_UNEX_EXER_EST_VAL`).

active firms would increase along with demand for overheads; this means their price $\omega_t$ could not have fallen in the first place. For the same reason, $\omega_t$ cannot rise in a recession. $\blacksquare$

## References


Árpád Ábrahám and T. Kirk White. The dynamics of plant-level productivity in U.S. manufacturing. *Census Discussion Paper CES 06-20*, July 2006.

Daniel A. Ackerberg, Kevin Caves, and Garth Frazer. Structural identification of production functions. *Working Paper*, 2006.

Cristina Arellano, Yan Bai, and Patrick J. Kehoe. Financial markets and fluctuations in uncertainty. *Federal Reserve Bank of Minneapolis Staff Report*, April 2010.

Enghin Atalay, Ali Hortaçsu, and Chad Syverson. Vertical integration and input flows. *American Economic Review*, 104(4):1120–1148, April 2014.

Rüdiger Bachmann and Christian Bayer. 'wait-and-see' business cycles? *Journal of Monetary Economics*, 60(6):704–719, September 2013.

Rüdiger Bachmann and Christian Bayer. Investment dispersion and the business cycle. *American Economic Review*, 104(4):1392–1416, April 2014.

Rüdiger Bachmann and Guiseppe Moscarini. Business cycles and endogenous uncertainty. *Working Paper*, 2011.

Martin N. Baily, Charles Hulten, and David Campbell. Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity*, 1992:187–267, 1992.

Gadi Barlevy. The sullying effect of recessions. *Review of Economic Studies*, 69(1):65–96, January 2002.

Gadi Barlevy. Credit market frictions and the allocation of resources over the business cycle. *Journal of Monetary Economics*, 50(8):1795–1818, November 2003.

Eric J. Bartelsmann, John C. Haltiwanger, and Stefano Scarpetta. Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1):305–334, February 2013.

Susanto Basu. Procyclical productivity: Increasing returns of cyclical utilization? *Quarterly Journal of Economics*, 111(3):719–751, August 1996.

Susanto Basu and John G. Fernald. Are apparent productive spillovers a figment of specification error? *Journal of Monetary Economics*, 36(1):165–188, August 1995.

Susanto Basu and John G. Fernald. Returns to scale in U.S. production: Estimates and implications. *Journal of Political Economy*, 105(2):249–283, April 1997.

Paul Beaudry, Mustafa Caglayan, and Fabio Schiantarelli. Monetary instability, the predictability of prices, and the allocation of investment: An empirical investigation using U.K. panel data. *American Economic Review*, 91(3):648–662, June 2001.

Randy Becker, John C. Haltiwanger, Ron Jarmin, Shawn Klimek, and Dan Wilson. Micro and macro data integration: The case of capital. *Mimeo*, July 2004.

Jess Benhabib and Roger E. A. Farmer. Indeterminacy and sector-specific externalities. *Journal of Monetary Economics*, 37(3):421–443, June 1996.

David Berger and Joseph Vavra. The dynamics of the U.S. price distribution. *Working Paper*, 2011.

Nicholas Bloom. The impact of uncertainty shocks. *Econometrica*, 77(3):623–685, May 2009.

Nicholas Bloom, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry. Really uncertain business cycles. *NBER Working Paper No. 18245*, July 2012.

Richard Blundell and Stephen Bond. GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3):321–340, 2000.

Francisco J. Buera and Benjamin Moll. Aggregate implications of a credit crunch. *NBER Working Paper No. 17755*, 2012.

Craig Burnside. Production function regressions, returns to scale and externalities. *Journal of Monetary Economics*, 37(2):177–201, April 1996.

Craig Burnside, Martin Eichenbaum, and Sergio Rebelo. Capital utilization and returns to scale. In Ben S. Bernanke and Julio J. Rotemberg, editors, *NBER Macroeconomics Annual*, volume 10, chapter 2, pages 67–110. MIT Press, January 1995.

Ricardo J. Caballero and Mohamad L. Hammour. The cleansing effect of recessions. *American Economic Review*, 84(5):1350–68, December 1994.

John Campbell, Martin Lettau, Burton Malkiel, and Yexiao Xu. Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *Journal of Finance*, 56(1):1–43, February 2001.

Kaiji Chen and Tao Zha. Identifying sources of the cyclicality of capital productivity dispersions through financial frictions. *Working Paper*, 2011.

Lawrence J. Christiano. Why does inventory investment fluctuate so much? *Journal of Monetary Economics*, 21(2-3):247–280, March-May 1988.

Lawrence J. Christiano, Mathias Trabandt, and Karl Walentin. Introducing financial frictions and unemployment into a small open economy model. *Sveriges Riksbank Working Paper Series 214*, April 2010.

Lawrence J. Christiano, Roberto Motto, and Massimo Rostagno. Risk shocks. *American Economic Review*, 104(1):27–65, January 2014.

Sanjay K. Chugh. Firm risk and leverage-based business cycles. *Working Paper*, 2011.

Gian Luca Clementi and Hugo A. Hopenhayn. A theory of financing constraints and firm dynamics. *Quarterly Journal of Economics*, 121(1):229–265, February 2006.

Russell W. Cooper and John C. Haltiwanger. On the nature of capital adjustment costs. *Review of Economic Studies*, 73(3):611–633, July 2006.

Steven J. Davis and John C. Haltiwanger. Gross job creation and destruction: Microeconomic evidence and macroeconomic implications. In Olivier Jean Blanchard and Stanley Fischer, editors, *NBER Macroeconomics Annual*, volume 5, chapter 3, pages 123–168. MIT Press, January 1990.

Pablo N. D'Erasmo and Hernan J. Moscoso Boedo. Intangibles and endogenous firm volatility over the business cycle. *Working Paper*, 2012.

Andrea L. Eisfeldt and Dimitris Papanikolaou. Organization capital and the cross-section of expected returns. *Journal of Finance*, 68(4):1365–1406, August 2013.

Jonas D. M. Fisher. The dynamic effects of neutral and investment-specific technology shocks. *Journal of Political Economy*, 114(3):413–451, June 2006.

Lucia Foster, John C. Haltiwanger, and Chad Syverson. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425, March 2008.

Lucia Foster, Cheryl Grim, and John Haltiwanger. Reallocation in the Great Recession: Cleansing or not? *Center for Economic Studies Working Papers 13-42*, 2013.

Xavier Gabaix. The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772, May 2011.

Mark Gertler and Simon Gilchrist. Monetary policy, business cycles, and the behavior of small manufacturing firms. *Quarterly Journal of Economics*, 109(2):309–340, May 1994.

Fabio Ghironi and Marc J. Melitz. International trade and macroeconomic dynamics with heterogeneous firms. *Quarterly Journal of Economics*, 120(3):865–915, August 2005.

Simon Gilchrist, Jae W. Sim, and Egon Zakrajšek. Uncertainty, credit spreads, and investment dynamics. *Working Paper*, 2010.

Jeremy Greenwood, Zvi Hercowitz, and Gregory W. Huffman. Investment, capacity utilization, and the real business cycle. *American Economic Review*, 78(3):402–417, June 1988.

Jeremy Greenwood, Zvi Hercowitz, and Per Krusell. Long-run implications of investment-specific technological change. *American Economic Review*, 87(3):342–362, June 1997.

Jeremy Greenwood, Zvi Hercowitz, and Per Krusell. The role of investment-specific technological change in the business cycle. *European Economic Review*, 44(1):91–115, January 2000.

Fatih Guvenen, Serdar Ozkan, and Jae Song. The nature of countercyclical income risk. *NBER Working Paper No. 18035*, 2012.

Robert E. Hall. Invariance properties of Solow's residual. In Peter A. Diamond, editor, *Growth, Productivity, Unemployment: Essays to Celebrate Bob Solow's Birthday*, chapter 5, pages 71–111. MIT Press, 1990.

Sharon G. Harrison. Returns to scale and externalities in the consumption and investment sectors. *Review of Economic Dynamics*, 6(4):963–976, October 2003.

John Heaton and Deborah Lucas. Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy*, 104(3):443–487, June 1996.

Hugo A. Hopenhayn. Entry, exit, and firm dynamics in the long run equilibrium. *Econometrica*, 60(5):1127–1150, September 1992.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4), November 2009.

Ron Jarmin and Javier Miranda. The longitudinal business database. Technical report, Center of Economic Studies, U.S. Census Bureau, 2002.

Dale W. Jorgenson and Zvi Griliches. The explanation of productivity change. *Review of Economic Studies*, 34(3):249–283, July 1967.

Alejandro Justiniano and Giorgio E. Primiceri. The time varying volatility of macroeconomic fluctuations. *American Economic Review*, 98(3):604–641, June 2008.

Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti. Investment shocks and business cycles. *Journal of Monetary Economics*, 57(2):132–145, March 2010.

Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti. Investment shocks and the relative price of investment. *Review of Economic Dynamics*, 14(1):102–121, January 2011.

Ricardo Lagos. A model of TFP. *Review of Economic Studies*, 73(4):983–1007, October 2006.

Seong-Hoon Lee and Sang V. Nguyen. Returns to scale in small and large U.S. manufacturing establishments: Further evidence. *Small Business Economics*, 19(1):41–50, August 2002.

Yoonsoo Lee and Toshihiko Mukoyama. Entry, exit, and plant-level dynamics over the business cycle. *Working Paper*, July 2011.

James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2):317–341, April 2003.

Robert E. Lucas. On the size distribution of business firms. *Bell Journal of Economics*, 9(2): 508–523, Autumn 1978.

Marc J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, November 2003.

Marc J. Melitz and Gianmarco I. P. Ottaviano. Market size, trade, and productivity. *Review of Economic Studies*, 75(1):295–316, January 2008.

Virgiliu Midrigan and Daniel Y. Xu. Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–458, February 2014.

Benjamin Moll. Productivity losses from financial frictions: Can self-financing undo capital misallocation? *Working Paper*, December 2009.

Giuseppe Moscarini and Fabien Postel-Vinay. The contribution of large and small employers to job creation in times of high and low unemployment. *American Economic Review*, 102(6):2509–2539, October 2012.

Kevin M. Murphy, Andrei Shleifer, and Robert W. Vishny. Building blocks of market clearing business cycle models. In Olivier Jean Blanchard and Stanley Fischer, editors, *NBER Macroeconomics Annual*, volume 4, chapter 5, pages 247–287. MIT Press, January 1989.

Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, May 1987.

G. Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297, November 1996.

Min Ouyang. The scarring effect of recessions. *Journal of Monetary Economics*, 56(2):184–199, March 2009.

Amil Petrin, Jerome P. Reiter, and T. Kirk White. The impact of plant-level resource reallocation and technical progress on U.S. macroeconomic growth. *Review of Economic Dynamics*, 14(1): 3–26, January 2011.

Nicolas Petrosky-Nadeau. TFP during a credit crunch. *Journal of Economic Theory*, 148(3): 1150–1178, May 2013.

Morten O. Ravn and Harald Uhlig. On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–376, May 2002.

Edouard Schaal. Uncertainty, productivity and unemployment during the great recession. *Working Paper*, 2012.

Joseph A. Schumpeter. *Capitalism, Socialism, Democracy.* Harper Collins Publishers, New York, third edition, 1942.

Kjetil Storesletten, Chris I. Telmer, and Amir Yaron. Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717, June 2004.

Chad Syverson. Product substitutability and productivity dispersion. *Review of Economics and Statistics*, 86(2):534–550, May 2004.

Joseph Vavra. Inflation dynamics and time-varying uncertainty: New evidence and an Ss interpretation. *Quarterly Journal of Economics*, 129(1), February 2014.