The Cyclicality of Productivity Dispersion[*,**]

Matthias Kehrig[†]

January 16, 2011

Abstract

Using plant-level data, I show that the dispersion of total factor productivity in U.S. durable manufacturing is greater in recessions than in booms. This cyclical property of productivity dispersion is much less pronounced in non-durable manufacturing. In durables, this phenomenon primarily reflects a relatively higher share of unproductive firms in a recession. In order to interpret these findings, I construct a business cycle model where production in durables requires a fixed input. In a boom, when the market price of this fixed input is high, only more productive firms enter and only more productive incumbents survive, which results in a more compressed productivity distribution. The resulting higher average productivity in durables endogenously translates into a lower average relative price of durables. Additionally, my model is consistent with the following business cycle facts: procyclical entry, procyclical aggregate total factor productivity, more procyclicality in durable than non-durable output, procyclical employment and countercyclicality in the relative price of durables and the cross section of stock returns.

[*] I am very grateful to my dissertation committee Lawrence Christiano, Martin Eichenbaum, Andrea Eisfeldt and Giorgio Primiceri for helpful discussions and encouragement. Discussions with Gadi Barlevy, Nick Bloom, Jeff Campbell, Alejandro Justiniano, Camelia Kuhnen, Ezra Oberfield and Nicolas Ziebarth as well as seminar participants at Northwestern University, the Federal Reserve Bank of Chicago and the 2010 Census RDC Research Conference 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 getting access to 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.

[**] The research in this paper was conducted while the author was a Special Sworn Status researcher of the U.S. Census Bureau at the Chicago Research Data Center. 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.

[†] Northwestern University, Department of Economics, 302 Arthur Andersen Hall, 2001 Sheridan Road, Evanston, IL 60208, m-kehrig@northwestern.edu. The latest version of this paper can be downloaded at www.depot.northwestern.edu/~mke816/indexjm.html.

1 Introduction

This paper investigates the cyclical properties of productivity dispersion across plants in the U.S. manufacturing sector. From a business cycle perspective, changes in cross-sectional productivity dispersion are relevant because they determine the dynamics of aggregate total factor productivity, a central object of interest in business cycles.

Although some theories of aggregate fluctuations address productivity dispersion, it is not obvious how dispersion moves over the business cycle. One set of models emphasises the process of Schumpeterian "creative destruction" and predicts that productivity dispersion is positively correlated with output. In a recession, when demand is low, unproductive plants exit and are eventually replaced by highly productive plants.[1] On the other hand, an environment where plants compete over common resources – as developed in Melitz (2003) for international trade – has the potential to deliver the opposite result. In a boom, increased demand for production factors raises factor prices. Only more productive plants can afford to pay the higher factor costs, while unproductive plants exit. In a recession, higher productivity dispersion persists in the face of weak competition.[2] In the context of the latter mechanism, recessions are "sullying," while they are "cleansing" in the context of Schumpeterian creative destruction.

The existing literature has provided little empirical evidence on cyclical properties of the productivity distribution, and it is so far not clear whether recessions are cleansing or sullying.[3] The goal of this paper is to address this question. In doing so, I assess the relative importance of cost and demand factors that have a countervailing impact on dispersion as highlighted above.

Using confidential Census data, I estimate plant-level productivity in the U.S. manufacturing sector from 1972-2005. My empirical work establishes three main results: First, cross-sectional productivity dispersion is countercyclical; i.e. the distribution of plant-level productivity is more spread-out in a recession than in a boom. 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 changes at the bottom end of the productivity

[1] Caballero and Hammour (1994, 1996) formalised the notion of cleansing in a business cycle context which goes back to Schumpeter (1939) and appears in a large number of models, among others Mortensen and Pissarides (1994); Campbell (1998); Gomes, Greenwood, and Rebelo (2001); Lentz and Mortensen (2008).

[2] Examples of this strand of research are Davis and Haltiwanger (1990, 1992, 1999); Melitz (2003); Ghironi and Melitz (2005); Eisfeldt and Rampini (2006, 2008); Melitz and Ottaviano (2008). The same countercyclical dispersion is consistent with models of reallocation if one accounts for a changing plant size. Production factors are shifted from unproductive to productive plants in a boom, which compresses the productivity dispersion.

[3] Bachmann and Bayer (2009a) have documented a countercyclical dispersion in productivity *growth rates* for German firm-level data. Eisfeldt and Rampini (2006) have established a similar result using data on publicly traded firms covered in Compustat. Contrary to them, I examine the dispersion of productivity levels rather than growth rates.

A number of studies has looked at the productivity dispersion in one or repeated cross sections (see for example Baily, Hulten, and Campbell (1992); Bartelsmann and Doms (2000); Syverson (2004) for TFP dispersion in the U.S. Hsieh and Klenow (2009); Moll (2009); Song, Storesletten, and Zilibotti (forthcoming) compare the productivity dispersion in developing and emerging economies to the one in the U.S.).

distribution. 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 Levinsohn and Petrin (2003), but they are robust to using alternative methods to infer total factor productivity. The cyclicality results also hold for several dispersion measures such as the cross-sectional variance, inter-quartile or inter-decile range.

Schumpeterian models typically consider only variation in demand, which generates pro-cyclical productivity dispersion. In a boom, when demand is high, unproductive plants survive more easily and vice versa in a recession. At face value, such models are at odds with my empirical finding of a countercyclical dispersion. To overcome this inconsistency, I introduce a cost channel into an environment that, without the cost channel, would lead to cleansing in a recession. When plants compete over common resources, such an environment leads to higher cost in a boom, which could make it harder for unproductive firms to survive.

I build a model along the lines of Ghironi and Melitz (2005) in which business cycles are driven by aggregate shocks as in Caballero and Hammour (1994). Plants differ in their productivity and are active in two sectors (durables and non-durables). Production in the durable sector requires a fixed input such as overhead labour or organisational capital. The costs for this fixed input are a crucial determinant of plant profitability. As a result, only plants 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. Since fixed inputs in durables are a key feature of my model, I provide empirical evidence for them. Higher fixed inputs in durables than in non-durables will show up in a production function estimation as higher returns to scale. I estimate returns to scale on the plant level and find that they are increasing in durable goods industries, while in non-durable goods industries they are constant.[4] This finding lends support to my assumption of overall constant returns to scale and fixed factors merely in durables. This approach attempts to provide a unified theoretical explanation of both the micro-level dispersion results as well as the typical macroeconomic dynamics.

I use my model to study the dynamics of productivity dispersion in booms and recessions. I conceptualise aggregate fluctuations as shocks to household preferences that raise aggregate demand. Consider a plant with productivity exactly equal to the cutoff productivity. An increase in demand increases profits. At first sight, additional profits benefit the plant at the cutoff. On the other hand, higher profits increase entry into the economy. A larger number of plants raises aggregate demand for production factors. In particular, the price of the fixed factor will rise. This hurts the plant at the cutoff, because it may no longer be able to cover the

[4]This difference in returns to scale estimated on the plant level confirms the findings of Burnside (1996) and Harrison (2003) who estimated returns to scale on the industry level.

costs for the fixed factor. If that is the case, this plant will no longer be active in durables and the productivity cutoff will be higher, resulting in a more compressed productivity distribution in a boom. 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 operates at the bottom end of the distribution (Result #2) and predominantly in durable goods industries (Result #3). The crucial feature to deliver the above results is the procyclical price of the fixed factor. 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 its procyclicality.

Although the model was constructed with the objective of understanding the business cycle properties of productivity dispersion, it has many other implications which provide additional support for the model. For example, the productivity cutoff, and hence average productivity in durables, is procyclical, which results in a countercyclical average relative price of durables. The countercyclicality in the price of durables has been widely noted, and my model appears to provide a novel explanation for that phenomenon.[5] This explanation rests on endogenous selection of more profitable plants into durables in a boom. In addition, the model endogenously predicts that aggregate TFP is procyclical although the source of fluctuations in my model is not disturbances to aggregate total factor productivity. This happens because the underlying productivity dispersion is truncated at the bottom in a boom.[6] The model is also consistent with procyclical employment 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 cross section of stock market returns to be countercyclical, see for example Heaton and Lucas (1996); Storesletten, Telmer, and Yaron (2004).

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

[5]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 at the heart of a countercyclical relative price of durables. See for example Murphy, Shleifer, and Vishny (1989); Benhabib and Farmer (1994, 1996); Hall (1990); Caballero and Lyons (1992); Bartelsmann, Caballero, and Lyons (1994); Harrison (2003). Greenwood, Hercowitz, and Krusell (1997, 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.

[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).

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[7] manufacturing data from 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. These data are supplanted by industry-level data from several publicly available sources: price deflators from the NBER-CES Manufacturing Industry Database (NBER-CES)[8], the Capital Tables published by the Bureau of Labor Statistics (BLS)[9], the Fixed Asset Tables published by the Bureau of Economic Analysis (BEA)[10] and the Industrial Production and Capacity Utilization published by the Federal Reserve Board of Governors (IPCU)[11]. Unless otherwise noted, all datasets are at annual frequency.

The Census data are mainly the Census of Manufactures (CMF) and the Annual Survey of Manufactures (ASM). Combined, they span the period 1972-2007 (due to data limitations with price deflators only observations up to and including 2005 can be used) and contain information on establishment-level inputs and outputs. These data have been used before in a number of studies (see for example Baily, Hulten, and Campbell (1992); Ábrahám and White (2006); Hsieh and Klenow (2009); Petrin, Reiter, and White (forthcoming)). Previous research has typically focused on estimating returns to scale, the persistence of productivity or aggregate productivity growth in one or repeated cross sections. Petrin, Reiter, and White (forthcoming) use the estimator developed by Levinsohn and Petrin (2003) to decompose aggregate TFP growth into terms reflecting technical efficiency and reallocation as proposed by Levinsohn and Petrin (2010). 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 1972-2005. With this longitudinal perspective, I can

[7] 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.

[8] 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/.

[9] 1987-2008 Capital Data for Manufacturing Industries http://www.bls.gov/mfp/mprdload.htm.

[10] 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.

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

assess the empirical validity of different theories of business cycle mechanisms and models of entry and exit using their implications for higher-order moments of productivity.

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. Several datasets were combined with the ASM/CMF: The LBD covers the entire universe of plants in the U.S. economy since 1976. The primary interest is to combine it with the ASM/CMF in order to get reliable information about plant birth and death. The COMPUSTAT-SSEL bridge carries the information if a firm is also listed in the COMPUSTAT dataset, i.e. if it is publicly traded or not. This will help to test implications where access to financing plays an important role.

The ASM/CMF exhibits unparalleled detail on the plant level and is an excellent source to assess the cross-sectional productivity dispersion and how it evolves over time. This level of detail, however, comes at some cost. Ideally, one would want to analyse a balanced panel that reflects the entire universe or is at least a random sample of the population of manufacturing plants. The CMF/ASM dataset, however, is censored in several ways: In "Census years" (years ending in 2 and 7), the dataset covers the entire universe of manufacturing establishments that are active in that year.[12] In "ASM years" (all other years), the data cover only a subset of the universe of manufacturing firms that is not a random sample. Large establishments above a certain employee or asset value threshold are sampled with certainty, smaller establishments are selected with a certain probabiliy $p < 1$. Census chooses the sampling probability such that the inverse reflects the sampling weight, i.e. the number of establishments that the sampled observations is representative for. Using these weights, one can roughly replicate aggregate output and employment from existing plant-level observations at annual frequency.

The goal is to obtain a panel where results about the productivity dispersion are not driven by changes in the sampling design. In the baseline specification, I therefore drop all plants from the CMF that are not part of the ASM sample.[13] Although this procedure underrepresents small establishments, it maintains longitudinal consistency. Also, the number of observations that remain in my sample is still very large: About 60k annual observations sum to about 1.8m data points that stand for the lion's share in aggregate activity.

Alternatively, one could also analyse the productivity dispersion of the observations weighted by the Census-provided sampling weight rather than restricting the analysis to the consistent ASM sample. Although it may provide a robustness check, this approach makes the some strong assumptions. Namely, it requires that small establishment that were not sampled in the ASM, do in fact not only share similar size of assets, employment and output, but also the same TFP with those establishments that were sampled. In future work, I plan to check this property.

[12] Establishments with less than three employees, so-called "AR establishments," are imputed by Census based on administrative records from IRS; those observations are dropped.

[13] Those observations are identified by `ET = 0`.

Lastly, Census rotates the sample of small establishments that are included in the ASM sample every quinquennially in years ending in 4 and 9. This matters a lot if one is interested in the evolution of dispersion of TFP growth rates rather than TFP levels.[14] The TFP growth rate of many small establishments in years 4 and 9 cannot be computed because they were not in the ASM sample in the preceding year. I will also analyse the dispersion of TFP growth rates below. In order to exclude spurious results that stem from sample rotation, I analysed the dispersion-GDP relationship once containing all years and once dropping all rotation years. This robustness check is conducted in Section 2.3.2

2.2 Productivity dispersion

2.2.1 Constructing the dispersion measure

Studying cross-sectional productivity dispersion requires plant-level productivity estimates. This can be assessed by estimating a production function. This research has a long-standing tradition in the estimation of returns to scale. This paper follows the vast strand of previous research and assumes a Cobb-Douglas[15] gross output production function on the plant level:

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

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. β^k, β^l , β^m and β^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. The estimation of a_{ijt} will be described below.

The preferred specification is gross output rather than valued added. Basu and Fernald (1995, 1997) have shown that the value added specification leads to an upward bias of production elasticity estimates if factor markets are imperfect. This upward bias that is not present in the gross output specification. Lastly, note that both capital and energy are included in the production function. As pointed out by Burnside, Eichenbaum, and Rebelo (1995), the capital stock per se is not productive. Rather, it requires energy (fuels or electricity) to be utilised in production. Hence, I use some form of energy to proxy for capital services. This step will play a key role in the estimation describe below. 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

[14] This is also a problem if one estimates TFP in a way that requires lagged variables.

[15] Lee and Nguyen (2002) have estimated a translog production function for some industries without many differences in estimated return to scale.

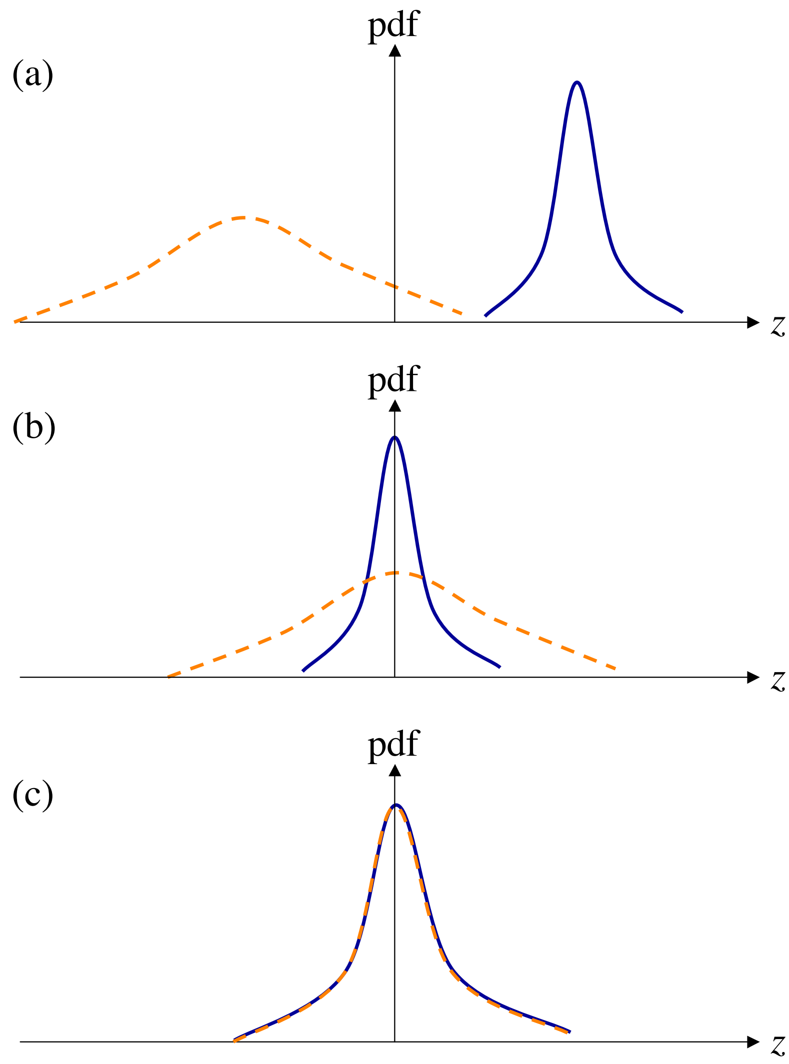

Figure 1: Industry differences in productivity dispersion

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

and more form the average growth trend 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:

$$a_{ijt} = g_j t + z_{ijt} \tag{2}$$

The resulting term z_{ijt} still needs to be corrected in more ways to obtain a proper cross-sectional dispersion measure $Disp_t$:

$$Disp_t \equiv E_t \left[Var_{jt} \left(\frac{z_{ijt} - \overline{z}_j}{\sigma_j} \right) \right] \tag{3}$$

As Figure 1 (a) illustrates, different industries may have a different average level of productivity. Changes in the overall cross-sectional dispersion could hence be driven by changes in "outlier" industries. This can be corrected by subtracting the long-run industry mean, $\overline{z}_j$. The effect of this step is illustrated in Figure 1 (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. As a consequence, I scale each industry by its standard deviation σ_j. This scaling step is illustrated in Figure 1 (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.

2.2.2 The Levinsohn-Petrin estimator

How can a_{ijt} be estimated? Naturally, production inputs, k_t, l_t, m_t and e_t, will be positively correlated with plant productivity, a_t. An OLS estimation will then result in biased estimates of both production elasticities $\beta^k, \beta^l, \beta^m$ and β^e and productivity a_t. Previous research that estimated returns to scale on the industry[16] rather than the plant level had to grapple with the same problem. One way this literature overcame the endogeneity problem was to obtain demand instruments that are not correlated with productivity.[17] This approach is feasible if one estimates returns to scale using aggregate or industry level data. Running (1) on the plant level, however, would require plant-level instruments, otherwise a lot of plant-specific information in the data would be lost. This information loss would hurt the endeavour of assessing plant-level productivity. Alternatively, one could assume constant returns to scale and perfectly

[16] See among others Hall (1990); Caballero and Lyons (1992); Bartelsmann, Caballero, and Lyons (1994); Burnside, Eichenbaum, and Rebelo (1995); Basu and Fernald (1995, 1997); Harrison (2003).

[17] Hall (1990) has proposed the oil price, military spending and the political party of the president as instruments. Burnside, Eichenbaum, and Rebelo (1995) have extended this set for a measure of monetary policy shocks.

competitive factor markets. Under these assumptions, the production elasticities are readily approximated by the cost shares of each production factor. This approach is easily implemented and therefore very popular in the literature (see for example Ábrahám and White (2006); Castro, Clementi, and Lee (2009); Lee and Mukoyama (2008)). This approach is implemented as a robustness check in Section B.3. It has the downside to impose constant returns to scale, so it is not informative about fixed factors in production. Fixed factors, however, play an important role in several theoretical models that feature productivity dispersion and a survival cutoff.

For that reason, I turn to the structural estimation technique developed by Levinsohn and Petrin (2003), henceforth LP. A detailed description of the procedure can be found in Appendix B.1. At this point, I will merely give a concise description. The authors assume the same production function as in equation (1) including energy e_t and k_t as the capital *stock* (rather capital services). It is (adopting their notation and omitting industry subscripts):

$$y_{it} = \omega_{it} + \eta_{it} + \beta^l l_{it} + \beta^k k_{it} + \beta^m m_{it} + \beta^e e_{it} \tag{1'}$$

The firm productivity term is decomposed into a portion that the plant observes, ω_{it}. The authors assume that ω_{it} has a first-order Markov structure. η_{it} on the other hand is not observed by the plant and hence does not influence contemporaneous decision rules. The capital stock in period t is assumed to be fixed. The plant's state vector is hence $(k_{it}, \omega_{it}, P_t)$. P_t denotes the vector of firm input prices. The key step in the LP technique is to utilise the information that input demand is generally a function of productivity (and the other state variables). This input demand function can be inverted to obtain a (non-parametric) function of productivity in terms of the state and firm observed inputs. This can be plugged into (1') and estimated semi-parametrically.

2.2.3 Results – 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. Also, there are no significant changes when separating out the capital stock into structures and equipment. As an alternative to electricity, I also used overall energy expenditures (including fuels), but this, too, does not change the results significantly. I use the above-described LP 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 1 displays the results of this regression in durables goods industries (NAICS 321, 327-339), Table 2 the analogous results for non-durables (NAICS 311-316, 322-326).

Table 1: Returns to scale – Durables

Coefficient	Estimate	Std. Err.	95% Confidence Interval	
Hours Worked	0.3415	0.0015	0.3385	0.3444
Materials	0.5055	0.0023	0.5009	0.5101
Capital	0.0321	0.0003	0.0313	0.0328
Electricity	0.1747	0.0033	0.1682	0.1814
Returns to scale	1.055	0.0043		

Wald test of constant returns to scale: $\chi^2 = 200.37$ ($p = 0.0000$).

Panel comprises 781,004 establishments in durable industries 1972-2005. Durable goods industries defined as those with NAICS 321 and 327-339.

Table 2: Returns to scale – Non-durables

Coefficient	Estimate	Std. Err.	95% Confidence Interval	
Hours Worked	0.2702	0.0003	0.2697	0.2707
Materials	0.4825	0.0020	0.4796	0.4874
Capital	0.0479	0.0010	0.0459	0.0499
Electricity	0.2019	0.0094	0.1833	0.2205
Returns to scale	1.0025	0.0097		

Wald test of constant returns to scale: $\chi^2 = 0.09$ ($p = 0.7692$).

Panel comprises 959,688 establishments in non-durable industries 1972-2005. Non-durable goods industries defined as those with NAICS 321 and 327-339.

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.[18] Remarkably, the coefficient on the capital stock is extremely small while the coefficient on energy is 0.17 in durables and 0.2 in non-durables. This shows 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, Eichenbaum, and Rebelo (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 add up to unity. A Wald test of the null hypothesis of constant returns cannot be rejected. This is

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

in line with previous research by Basu and Fernald (1995); Burnside, Eichenbaum, and Rebelo (1995); Harrison (2003) who also find constant returns to scale in non-durable industries.

The production function regression in durable goods industries look mostly similar to the results in non-durables with the exception of returns to scale being increasing. The production function coefficients add up to 1.055. This is only slightly above unity, but a Wald test can reject the null hypothesis of constant returns to scale at the 5% level. This finding will play an important role in explaining differences in the cyclicality of productivity dispersion below. The results of slightly increasing returns to scale are consistent with a fixed factor of production and otherwise constant returns to scale. The estimation procedure above did not explicitly take into account a fixed factor (that may be firm- or industry specific). If the true production function in durable goods industries is constant-returns to scale and there is a fixed production factor, then the estimated returns to scale will be increasing. 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 (2010) 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.

2.2.4 Results – Cross-sectional productivity dispersion

Plant-level total factor productivity is jointly estimated with return to scale in the above procedure. It corresponds to ω_t in equation (1'). I obtain this residual and detrend it as described in the previous section. The resulting object corresponds to the variable called z_{ijt} above. This detrended productivity is now recentered and scaled by the industry mean and standard deviation. Finally, I consider dispersion changes at business cycle frequency by HP-filtering the time series.[19] I now take the cross-sectional variance *within* each industry and average across all non-durable industries. This gives the dispersion measure in non-durables (for durables analogously) as outlined above:

$$Disp_t \equiv E_t \left[Var_{jt} \left(\frac{z_{ijt} - \overline{z}_j}{\sigma_j} \right) \right] \tag{3}$$

The two time series of the dispersion measures $Disp_t$ in durables and non-durables are displayed in Figure 2.

As Figure 2 shows, the cross-sectional dispersion varies over time. Casual observation of the

[19] Appendix B.2 discusses the long-run trend in dispersion.

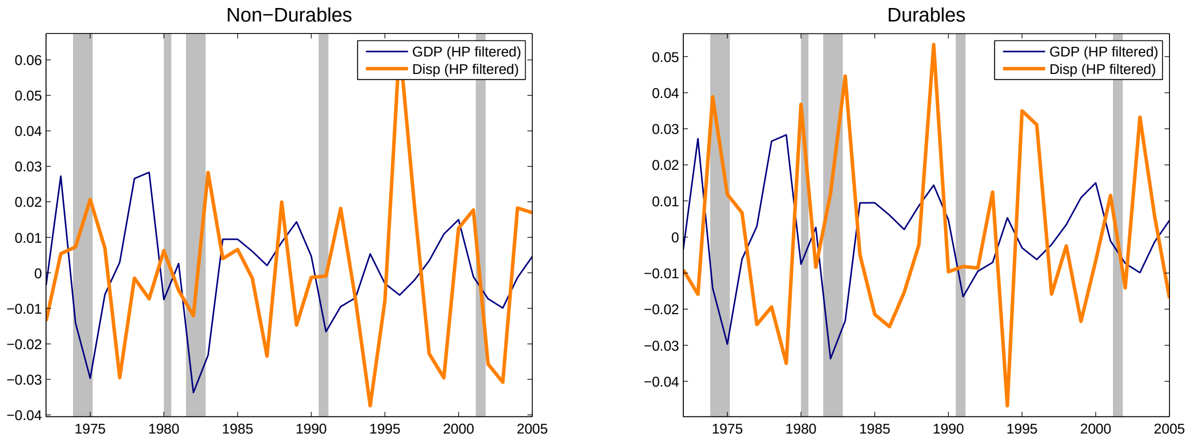

Figure 2: Time series of GDP growth rate and Dispersion

Time series plot of the annual growth rate of real GDP and of *Disp*, the dispersion measure defined in equation (3). Left panel displays non-durables, right panel displays durables, shaded bars denote NBER recessions. The *Disp* measure has been demeaned for better visibility.

graph reveals that dispersion has some peaks that coincide with recessions as defined by the NBER. This seems especially to be true for the early recessions in the sample: 1974, 1980, 1982 and – with some qualifications – 1991. There is an interesting rise in dispersion starting after this recession in both durables and non-durables. This secular rise starts in the early 1990's and the dispersion measure almost doubles by the end of the sample in 2005. This looks like an interesting fact for future research, but in the context of this paper I shall not devote further attention to it. The spikes in dispersion are clearly stronger in durables, displayed in the right panel on Figure 2.

Result 1: The productivity dispersion is countercyclial

The casual observation of the dispersion spiking in recessions warrants more formal evidence. I will correlate the dispersion measure with several output measures: the growth rate of real GDP, HP-filtered residuals and the number of recessionary months/year as determined by the NBER. As mentioned above, the dispersion rises secularly starting in the early 1990's. For that reason, I HP-filter the *Disp* measure. This takes out the (non-linear) trend.[20] As Table 3 shows, all three measures of aggregate fluctuations are negatively correlated with dispersion

[20]The resulting HP-filtered time series looks almost identical for the first part of the sample until the early 1990's. To make sure that my results are not driven by this HP filtering of dispersion, I also correlated the unfiltered time series with the annual growth rate of GDP. The results are similar albeit slightly weaker in durables. The contemporaneous cross correlation is still -0.21 (0.092), which means that this correlation is significantly different from zero at the 95% level.

contemporaneously and at least with a one-year lag. The negative contemporaneous correlation is statistically significant at the 95% level for all measures. I prefer to correlate dispersion with the growth rate of real GDP because it is a fairly theory-free measure that is not subject to the specific assumptions about some filtering technique or the definition of the business cycle dating committee. Unless otherwise noted, I will therefore focus on GDP growth in the following.

Table 3: Cross-correlations of dispersion and output measures

Lead/Lag	Correlation of dispersion in Durables with ...		
	GDP growth	GDP (HP)	No. boom months/year (NBER)
-2	0.139	0.232	0.102
-1	0.179	0.353	-0.034
0	-0.420	-0.528	-0.361
1	-0.327	-0.121	-0.278
2	-0.119	0.179	-0.249

Result 2: The dispersion in durables is stronger then in non-durables

The results of the corrlations between GDP and dispersion in the two sectors is displayed in in Table 4 and Figure 3. As we can easily see from those graphs, the productivity dispersion in more countercyclial in durables than in non-durables. The estimated contemporaneous correlation is -0.42 with a standard error of 0.12. This makes the result significant at the 95% level. This negative correlation continues with a one-year lag suggesting that one year into the recession the dispersion is still above its average level. The correlation is slightly less negative and the error bands wider, but the negative correlation is still significantly different from zero at the 95% level.

Table 4: Cross-correlations of dispersion in Non-Durables and Durables

Lead/Lag	Correlation of GDP growth with dispersion in ...	
	Non-durables	Durables
-2	0.257	0.139
-1	0.137	0.179
0	-0.172	-0.420
1	-0.194	-0.327
2	0.044	-0.119

In non-durables goods industries, the overall cyclicality pattern is similar. It is slightly countercyclical contemporaneously as well as with a one-year lag. It is not as strong, however,

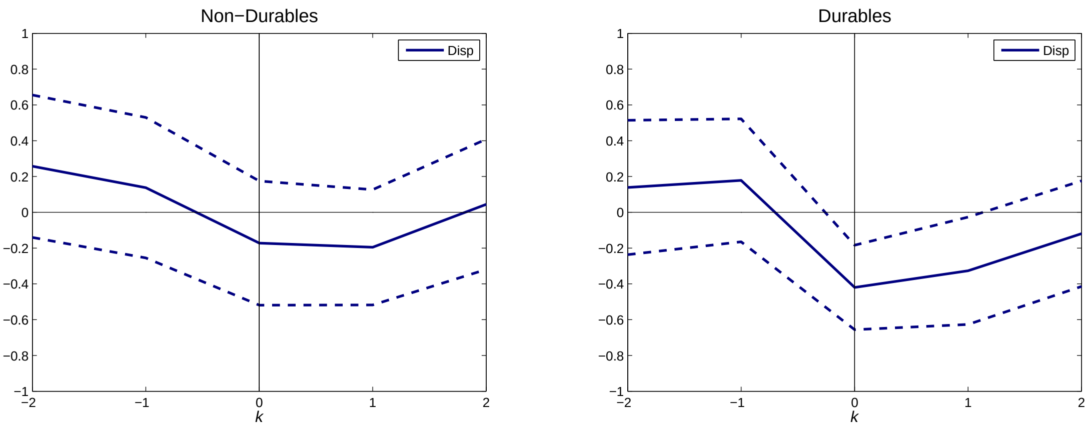

Figure 3: Cyclicality of Dispersion in Non-Durables (left) and Durables (right)

Correlograms display the correlation between the growth rate of real GDP and the dispersion measure defined in equation (3): $Corr(GDP_t, Disp_{t+k})$; dashed lines denote 95% confidence interval; confidence region computed using GMM.

as in durables. The correlation is -0.17 for contemporaneous correlation and -0.19 for a one-year lag. The standard error bands are much wider than in the durable goods industries. In particular, they are so wide 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.

Note that durable and non-durable goods industries were different along another dimension. The estimated returns to scale were higher in durables then in non-durables suggesting unobserved fixed factors are higher in durables than in non-durables. In the model section below, we shall see how to explore this combination of empirical facts. These two empirical findings related to differences between durables and non-durables are interesting complements to the literature on investment-specific fluctuations. Greenwood, Hercowitz, and Krusell (1997, 2000) claim that a large share of aggregate fluctuations are driven by technology shocks specific to the marginal efficiency of new investment, i.e. to goods produced by durables goods manufacturers. In the model section below, we will explore these relationships to the theory of investment-specific fluctuations further.

Result 3: Bottom quantiles drive the dispersion dynamics

As shown above, the productivity distribution as a whole is negatively correlated with GDP. 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

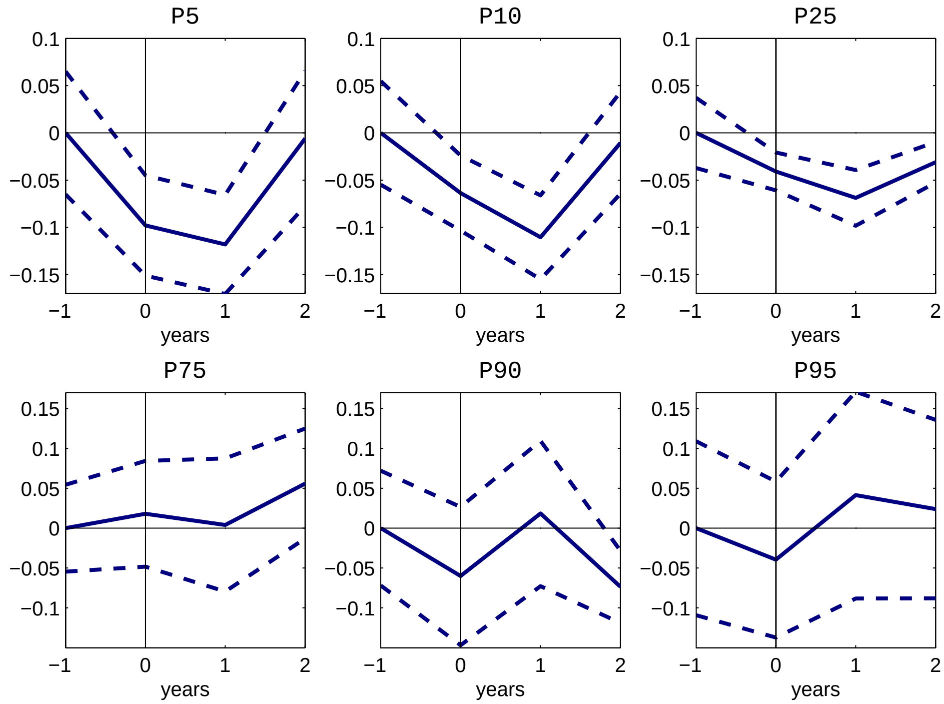

Figure 4: Behaviour of Quantiles over NBER recessions

Each panel displays the average behaviour of the indicated quantile over NBER recessions. The graphs are obtained by cutting out a subsample (a "worm") of the time series that starts one year before the onset of a NBER recession and lasts until years after the onset of a recession. Solid lines denote the mean of each recession worm, dashed lines are the standard deviation over the five recessions.

(1990). Alternatively, this countercyclicality is consistent with most movements happening at the top of the distribution (as in Gabaix (forthcoming)) or at the bottom of the distribution (as in Ghironi and Melitz (2005)). To that end, I look at the correlation of individual quantiles with the business cycle. Figure 4 paints a fairly clear picture. The bottom quantiles in the productivity distribution are more cyclical than the top quantiles. In particular, they go down in a recession. This means that in a recession 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 active here. 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 are 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.

Above, I have mentioned the relationship of my empirical results to the literature on investment-specific fluctuations. This research does not only lay the source for a large share of fluctuations in the durable goods sector. It also attributes a significant share in the volatility reduction in macroeconomic aggregates to fluctuations in that sector. This has been proposed in a number of papers: Stock and Watson (2002); Fisher (2006); Justiniano and Primiceri (2008); Sargent, Williams, and Zha (2006) to cite a few. This motivates me to look for a structural breaks in the cyclicality of dispersion in durables. There is some debate in the above-cited papers about the timing of this structural break. Some estimates revolve around 1980, other around 1984. I follow the latter approach and split my sample in 1984. Figure 5 displays the cyclicality in the full sample and the subsamples before and after 1985. Clearly, there is a

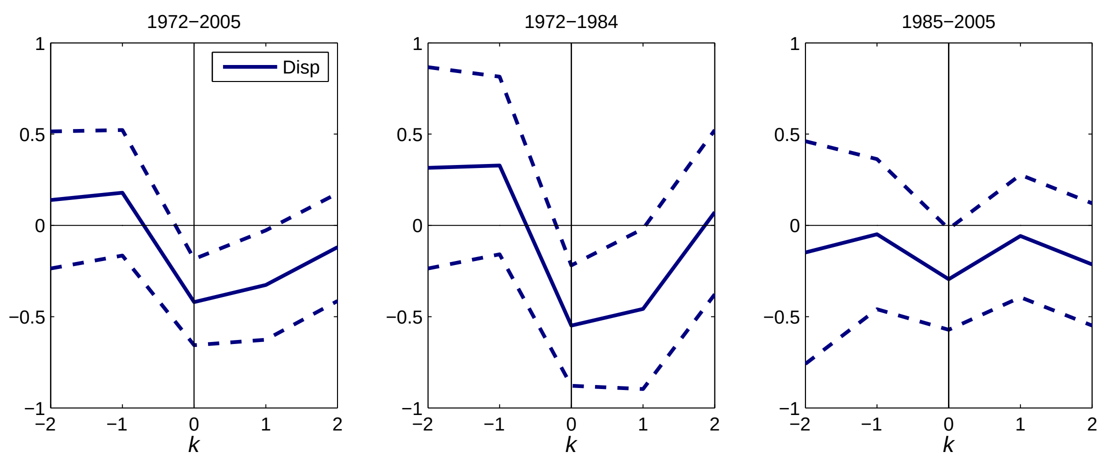

Figure 5: Cyclicality of Dispersion in Durables over Time

structural break in the cyclicality of dispersion. It is markedly countercyclial before 1984 and slightly countercyclical after 1984. This result is interesting in its own right. In addition to that, it provides a nice micro foundation to previous research that attribute a large share in the volatility reduction on fluctuations in aggregate TFP in the durable goods sector. Standard business cycle models feature only aggregate TFP, which – in reality – depends on the underlying TFP distribution. A changing distribution will impact aggregate TFP. If the underlying distribution become acyclical in the mid 1980's, this would imply in most standard models that (measured) aggregate TFP becomes acyclical. This implication makes it plausible to think about the volatility reduction in TFP in durables as being caused by the underlying productivity dispersion rather than exogenous aggregate TFP fluctuations. This result is certainly an interesting curiosity, which I plan to pick up again in future research.

2.2.5 Results: Dispersion in TFP growth rates

The main focus of this project is the dispersion of productivity *levels*. A related strand of research has worked on the cross-sectional dispersion of the productivity *growth rate*. Bachmann and Bayer (2009a,b) have recently established that in German data the cross sectional dispersion of growth rates of firm-level Solow residuals is countercyclical. A similar result was established by Eisfeldt and Rampini (2006) for Compustat firms. It would be interesting to see if similar results are present in the ASM dataset. This dataset has the advantage over Eisfeldt and Rampini (2006) that it is not confined to publicly traded firms of the Compustat sample. Publicly traded firms make up only a very small fraction of the ASM sample. On the other hand, I can only look at plants in the manufacturing sector. Therefore, my ASM data are also somewhat more limited than those by Bachmann and Bayer (2009a,b) who look at units across the economy. Following these authors, I construct the dispersion measure as[21]

$$Disp = Var_t(z_{it} - z_{it-1}).$$

The resulting time series and correlogram are displayed in Figure 6.

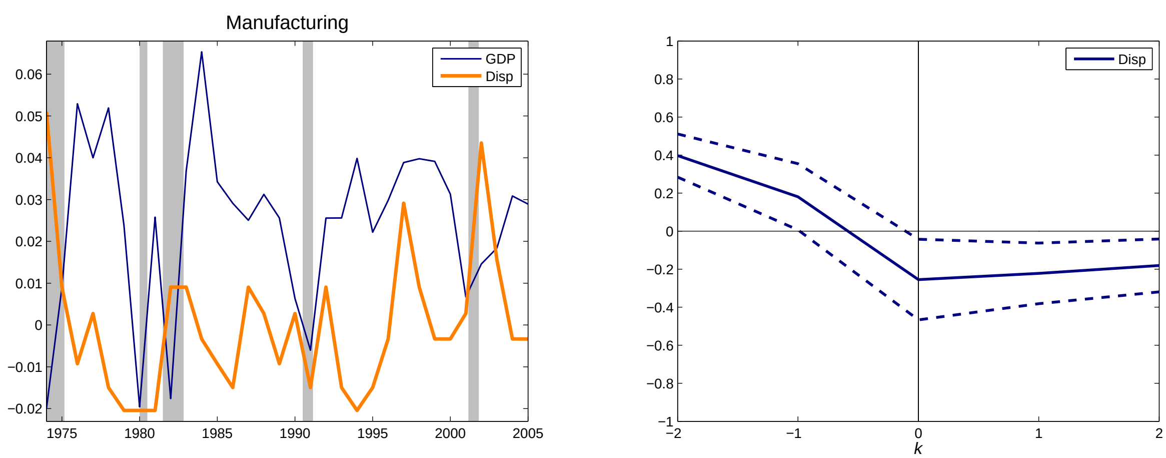

Figure 6: Dispersion of TFP Growth Rates

[21]I still detrended plant productivity before computing the growth rates and the variance. Note that I omitted the recentering and scaling exercises across industries. Looking the dispersion of growth rates should not be affected at all by differences in long-run industry location and spread.

2.3 Robustnes checks

2.3.1 Price dispersion

The analysis in this paper is hampered by some data limitations that have to be addressed. Most pressingly is a common measurement problem of inputs and outputs. Production y, capital k, materials m and fuels (but not electricity) are measured in nominal values that are then deflated by some price deflator. These deflators are different for each of the variables and available on the 6-digit NAICS industry level for output, materials and fuels, on the 3-digit NAICS level for capital. Although this is already a fairly fine level, price dispersion among firms *within* each industry could cause a measurement problem that confounds the estimates of plant-level productivity. The productivity measure I consider is hence "revenue productivity" as in Foster, Haltiwanger, and Syverson (2008); Hsieh and Klenow (2009). For the analysis of plant survival over the business cycle, revenue productivity looks almost like the more relevant measure to look at because whether a plant survives depends more on its profits rather than physical productivity.[22]

In the following, I will abstract from price dispersion in inputs. This is equivalent to assuming that these inputs are homogeneous across firms. If on thinks about standardisation of material inputs or energy, then this assumption does not seem too unrealistic. Output price dispersion might pose a more serious problem. Let p_i be the firm specific price and $\overline{p}$ the (industry-wide) price deflator used to transform nominal into real production. Omitting time subscripts and the η component of firm productivity I rewrite the production function as

$$\begin{aligned} y_i &= \omega_i + \beta^k k_i + \beta^l l_i + \beta^e e_i + \beta^m m_i \\ \Leftrightarrow p_i - \overline{p} + y_i &= \underbrace{\overbrace{\omega_i}^{\text{TFPQ}_i} + (p_i - \overline{p})}_{\text{TFPR}_i} + \beta^k k_i + \beta^l l_i + \beta^e e_i + \beta^m m_i \end{aligned}$$

The object of interest is ω_i, but as the above transformation illustrates within-industry price dispersion, $p_i - \overline{p}$, confounds our inference on ω_i. What we measure in reality, is "revenue productivity," $TFPR_i$, rather than physical productivity $TFPQ_i$ (terms are borrowed from Foster, Haltiwanger, and Syverson (2008)). As a consequence, variance in measured $TFPR_i$ if composed of variance in physical productivity and prices:

$$V(TFPR_i) = V(TFPQ_i) + V(p_i) + 2Cov(p_i, \omega_i)$$

What is known about these terms? Evidence on a large scale is not available, but Foster,

[22] Note that within industry price dispersion only becomes a problem when price dispersion changes over the business cycle. The mere existence of unaccounted price dispersion does not alter the *cyclical* properties of measured productivity dispersion.

Haltiwanger, and Syverson (2008) have used a small subsample of industries in a repeated cross section in the Census of Manufactures where firm-level prices are available. They find that in a cross section physical productivity is more dispersed than revenue productivity, i.e. $V(TFPR_i) < V(TFPQ_i)$[23]. This implies that the term $2Cov(p_i, \omega_i)$ must be very negative and larger in absolute value than $V(p_i)$. If this covariance is negative, unproductive firms tend to charge higher prices than productive firms. They apparently have some market power to price differentiate. When is this pricing power particularly high? The literature on mark-ups (Jaimovich (2007)) and market power (Balasubramanian and Sivadasan (2009)) suggest that recessions are times of more pricing power. This means that in a recession the term $2Cov(\omega_i, p_i)$ should be even more negative than in a boom. This should bias $TFPR$ downward in a recession. On the other hand, $V(p_i)$ will be higher if pricing power increases and prices become more dispersed. So I cannot make a definitive statement about which of the two terms dominates. I showed above, however, that the covariance term must be larger than than the variance term. Taking this as suggestive evidence implies that my estimates of the cyclicality of productivity dispersion should be a lower bound on the true productivity dispersion.

2.3.2 Rotation of ASM sample

The data for this work consist of the ASM panel of the Census of Manufactures. While this panel is longitudinally consistent and has a large number of observations, it is subject to the sampling rules developed by Census. In particular, Census rotates the ASM panel of plants every five years in years ending in 4 and 9.[24] I am worried that the firms entering the sample in one of those sample rotations could be systematically different in their productivity. If that is the case, then we should see systematics changes in those rotation years. The previous results suggest, this is not the case. To be safe, I will drop these rotation years and check if the negative correlation between GDP growth and dispersion persists. Dropping the rotation years leads to a correlation coefficient of $-0.403(0.151)$. This coefficient is still significantly negative in the same ballpark as the main result ($Corr(GDP, Disp_t) = -0.420$). Therefore, I conclude that the main result is not biased by the sample of rotation in the ASM.

[23]A similar point was made in Hsieh and Klenow (2009). They build a model where firms everywhere have identical technology (production function coefficients that sum up to 1), but inputs are distorted by firm-specific "wedges." In their structural model, physical technology can be identified from the distribution of these wedges. They compare their measure of physical productivity and revenue productivity to find that the former is about twice as dispersed as the latter.

[24]In the years between, Census tries to merely adjust the same in a way to account for entry and exit. Census attempts to keep the age distribution of plants constant.

3 The Model

We saw in the previous section that the cross-sectional productivity dispersion is countercyclical, that most dynamics result from the bottom end of the distribution and that the countercyclicality is stronger in durables. 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. In contrast to previous research, I will not assume idiosyncratic shocks, but merely rely on aggregate shocks to drive my business cycle model.[25] Given the importance of dynamics at the bottom of the distribution, my model will feature a fixed 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 finding of higher returns to scale in durables. This suggests the presence of a fixed factor in precisely that subsector of manufacturing whose productivity dispersion is most cyclical.

The model features endogenous entry in the economy (exit is random) and firm-level heterogeneity in productivity. There are two sectors – durables and non-durables – which is motivated by the empirical differences between those sectors in terms of returns to scale and dispersion cyclicality.

3.1 Final goods producers

There is a non-durable and a durable goods sector, denoted by n and d respectively, that each produce a homogeneous final good by assembling heterogeneous varieties:

$$Y_t^n = \left[\int_{i\in\Omega} y_{it}^{\frac{\sigma-1}{\sigma}}\, di\right]^{\frac{\sigma}{\sigma-1}} \qquad \sigma > 1$$
$$Y_t^d = \left[\int_{i\in\Omega^*} y_{it}^{\frac{\varrho-1}{\varrho}}\, di\right]^{\frac{\varrho}{\varrho-1}} \qquad \varrho > 1$$

where Ω is the set of varieties used in Sector n and Ω^* that of varieties in Sector d. We will later see that $\Omega^* \subset \Omega$. The elasticity of substitution between varieties in the two sectors is allowed to differ. Relaxing the assumption of an identical elasticity of substitution in both sectors allows for the possibility that intermediate firms can price discriminate between final customers. The residual demand curve for each variety in both sectors is identical except for

[25] Models that are driven by idiosyncratic productivity shocks are Gabaix (forthcoming); Christiano, Motto, and Rostagno (2009); Christiano, Trabandt, and Walentin (2010); Bloom, Floetotto, and Jaimovich (2009); Bachmann and Bayer (2009a,b).

the different elasticity of substitution

$$p_{it} = P_t^n \left(\frac{Y_t^n}{y_{it}}\right)^{\frac{1}{\sigma}}$$
$$p_{it} = P_t^d \left(\frac{Y_t^d}{y_{it}}\right)^{\frac{1}{\varrho}}.$$

Therefore, the price indices in both sectors can be expressed as

$$P_t^n = \left[\int (p_{it}^n)^{1-\sigma} di\right]^{\frac{1}{1-\sigma}}$$
$$\tilde{P}_t^d = \left[\int (p_{it}^d)^{1-\varrho} di\right]^{\frac{1}{1-\varrho}}.$$

It is important to keep track of the two different prices because the relative price of durables, $\frac{\tilde{P}_t^d}{P_t^n}$, will change when economic conditions change.

3.2 Intermediate goods producers

There is a continuum of monopolistic intermediate goods producers, each of which produces a particular variety $i, i \in [0,1]$. For reasons of simplicity they just employ labour.[26] Apart from producing different varieties, each firm is endowed with an idiosyncratic productivity draw, denoted by z_i, upon birth which it keeps for the rest of its existence. Every firm hires labour l_{it} to produce output which it can sell either in Sector n or in both Sectors n and d. The specific characteristic of the durable goods sector is a fixed factor in production. A firm has to pay a cost c_f if it wants to sell its goods there as well. This fixed costs can be interpreted as overhead expenditures for advertising, management or R&D. The fixed factor in durables is consistent with my empirical finding that returns to scale are higher in durables. This finding is consistent with the results of previous research (see for example Burnside (1996); Harrison (2003)). Behold this result should be interpreted as a *difference* in fixed factors rather than fixed factors being absent from non-durables. *Higher* returns to scale in durables then suggest *higher* fixed factors in durables. Since I am only interested in differences in returns to scale between the two sectors, I keep returns to scale constant in Sector n and add the fixed factor in Sector d which make Sector d look like increasing returns to scale. The production functions

[26] This simplifying assumption could be relaxed at the expense of expositional clarity. It wouldn't change the qualitative results, however.

for either sector are

$$y^n_{it} = z_{it} l_{it}$$
$$y^d_{it} = z_{it}(l_{it} - c_f)$$

Firms take nominal wages, W_t, and their demand curve as given. Profit-maximising behaviour leads to the familiar pricing rule

$$p^n_{it} = \frac{\sigma}{\sigma - 1} \frac{W_t}{z_{it}}$$
$$p^d_{it} = \frac{\varrho}{\varrho - 1} \frac{W_t}{z_{it}}.$$

I define the real wage as the nominal wage in terms of the non-durable output good: $w_t \equiv W_t / P^n_t$ and let the relative price of durables be $P^d_t \equiv \tilde{P}^d_t / P^n_t$. From now on we express every variable in terms of output n, for example profits are $\pi^d_t = \pi^{d\ \mathrm{nom}}_t / P^n_t$. Using the profit-maximising pricing rule we can derive firm output, labour demand and profits in both sectors in terms of wages and aggregate sector sales. Firm profits in the non-durable sector are always positive because firms charge a mark-up over marginal cost and there are no fixed cost. In durables, in contrast the presence of the fixed factor implies fixed cost and thus potentially negative profits. The profit function in durables is:

$$\pi^d_{it}(z_{it}) = \frac{1}{\varrho} \left(\frac{\varrho - 1}{\varrho} \frac{z_{it}}{w_t} \right)^{\varrho - 1} \left(P^d_t \right)^{\varrho} Y^d_t - w_t c_f$$

The productivity cutoff The fixed cost preclude very inefficient firms to produce for Sector d, i.e. there is a productivity level z^*_t such that profits from producing for Sector d with productivity z^*_t are zero. Solving the above profit function for this cutoff yields

$$z^*_t = \frac{1}{\varrho - 1} \left[\left(\varrho \frac{w_t}{P^d_t} \right)^{\varrho} \frac{c_f}{Y^d_t} \right]^{\frac{1}{\varrho - 1}} \tag{4}$$

Equation (4) is a key relationship as it regulates the equilibrium distribution of productivity and therefore deserves some discussion. Although there are still endogenous variables contained in this expression, it is instructive to take a close look at it. Unproductive firms with a low z can more easily survive in this environment if the value of their sales is high (high Y^d_t or high P^d_t). 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)). Because the cleansing effect emerges due to changes in aggregate demand (to be precise, the value of aggregate demand), I label this channel the “demand channel.”

There is a second channel that I label the “cost channel” which has the opposite effect on the productivity cutoff. If the cost for the fixed factor, w_t, are high, it is fairly costly for a firm to operate profitably regardless of its production level. It must hence be productive enough to have a relatively low price and thusly attract enough demand to make profits at all. Since real wages are procyclical, a downturn may hence be associated with a *decrease* in the productivity cutoff. In the same way as the cleansing effect operates through the demand channel, there is a “permissive effect of recessions” which operates thourhg the cost channel. A decrease in the real wage is permissive in that it allows inefficient firms to stay in the economy. Note that while the cleansing effects of recession literature emphasised the importance of demand, I also consider cost factors in firm survival. It is unclear ex ante, whether the cleansing effect via the demand channel or the permissive effect of recessions via the cost channel would dominate. This depends on the response of real wages to aggregate output and the elasticity of substitution ϱ. The strength of the cost channel is increasing in the elasticity of substitution ϱ. The higher ϱ, the more easily single varieties are substituted against other ones and the lower the mark-up of single firms. In that case, small changes in the real wage have a strong impact on the survival of unproductive firms and the cost channel dominates the demand channel. Higher wages in a boom will then lead to a higher cutoff and a more compressed productivity dispersion.

If single varieties are bad substitutes, however, unproductive firms can more easily survive an increase in real wages. This is because bad substitutability leads to higher mark-ups which in turn allows unproductive firms to pass through the higher wages to final producers. Then the demand channel will dominate the cost channel. Higher aggregate demand in a boom will then lead to a lower cutoff and a more spread-out productivity dispersion.

How cyclical are the cost of fixed factors? One criticism against the importance of the permissive effect could be that wages are generally believed to be very sticky. This would greatly dampen the permissive effect. One needs to take a close look at the wage in expression (4): 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. I focus on the interpretation 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 get empirical evidence for this claim, I turn to data on managerial compensation and analyse the cyclicality of their income. The data come from ExecuComp, a database that overs the top executive pay in a large cross section of firms. Their real income growth rate is computed and correlated with the business cycle in Figure 7.

If they behaved like normal production worker wages, one would expect the managerial wages to be mostly acyclical. Figures 7 paints a different picture. All components of manage-

Table 5: Components of executive compensation

Compensation Component	Average Share	Volatility (over time)
Base Pay	15.28%	0.094
Stock Options – exercised	19.39%	0.328
Stock Options – unexercised	65.32%	0.282
Total Earnings	100.00%	0.105

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`).

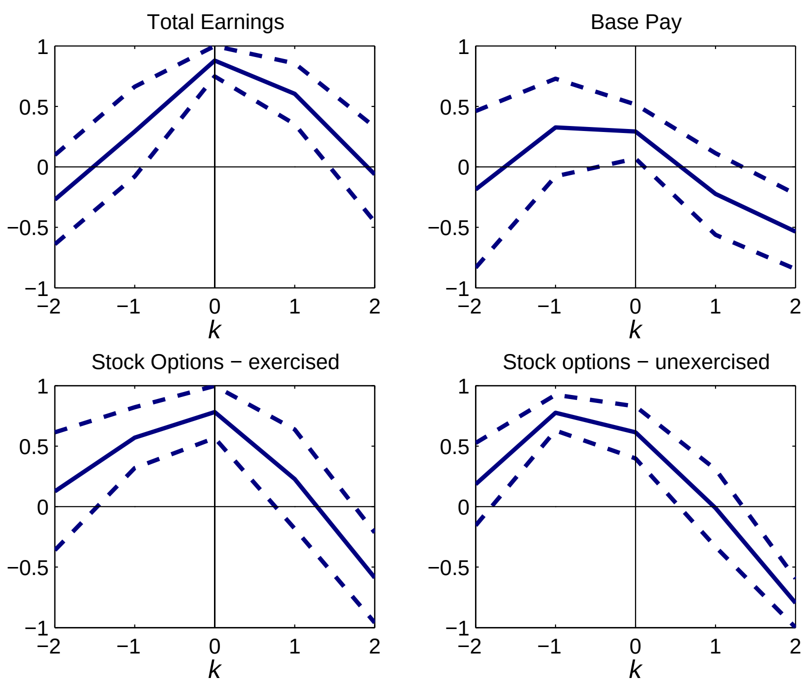

Figure 7: Cyclicality of Executive Compensation

Correlogram of different portions of executive compensation: $Corr(GDP_t, w_{t+k})$, dashed lines denote 95% confidence intervals constructed as describe in Figure 3. Correlated data are HP filtered residuals of GDP and the aggregate real earnings.

rial compensation – be it base salary, payment in shares or stock options – are pronouncedly procyclical.

3.3 Aggregation and averages

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 aggregate production function to obtain aggregate output in both sectors. Keep in mind that I can integrate over all varieties i or over all productivity levels z. If I do the latter, I have to take into account that there is a certain measure of firms active in the economy which is defined as N_t. This measure varies with entry and exit (described below), but in period t it is a state variable. Let the average productivity levels be

$$\begin{aligned}
\overline{z}^n &= \left[\int_{z_L}^{z_H} z^{\sigma-1} dF(z)\right]^{\frac{1}{\sigma-1}} \\
\overline{z}_t^d &= \left[\frac{1}{1-F(z_t^*)}\int_{z_t^*}^{z_H} z^{\varrho-1} dF(z)\right]^{\frac{1}{\varrho-1}}
\end{aligned}$$

Then we can aggregate intermediate output

$$\begin{aligned}
Y_t^n &= \left[\int y_{it}^{\frac{\sigma-1}{\sigma}} di\right]^{\frac{\sigma}{\sigma-1}} \\
&= \left(\frac{\sigma-1}{\sigma}\frac{1}{w_t}\right)^{\sigma} Y_t^n \left[\int z_{it}^{\sigma-1} di\right]^{\frac{\sigma}{\sigma-1}} \\
&= \left(\frac{\sigma-1}{\sigma}\frac{1}{w_t}\right)^{\sigma} Y_t^n \left[\int_{z_L}^{\infty} z^{\sigma-1} N_t dF(z)\right]^{\frac{\sigma}{\sigma-1}} \\
1 &= \left(\frac{\sigma-1}{\sigma}\frac{1}{w_t}\right)^{\sigma} N_t^{\frac{\sigma}{\sigma-1}} \left[\int_{z_L}^{\infty} z^{\sigma-1} dF(z)\right]^{\frac{\sigma}{\sigma-1}} \\
1 &= \left(\frac{\sigma-1}{\sigma}\frac{1}{w_t}\right)^{\sigma} N_t^{\frac{\sigma}{\sigma-1}} \left(\overline{z}^n\right)^{\sigma}
\end{aligned}$$

Note that the last expression shows that the wage is a state variable because $\overline{z}^n$ is fixed and N_t is a state variable. The wage will change as more firms enter the economy. The fact that wages are completely determined by labour demand only follows from the firm production technology which has constant returns to scale and uses labour only.

As in Melitz (2003) the present economy with heterogeneous firms is isomorphic to one with a measure of N_t representative firms producing with exactly that "average" productivity level $\overline{z}^n$. This average productivity does not change because the underlying productivity distribution does not change over time and any firm will be producing in equilibrium. Accordingly, there is a similar expression of aggregate productivity in Sector d with the only difference that the relevant cutoff is z_t^* rather than z_L. Up to truncation the productivity distribution of firms active in sector d is the same as in Sector n, a feature which shall be explained in greater detail below.

If I parametrise the productivity distribution as bounded Pareto with shape parameter k on support $[z_L, \infty)$, then these expressions have a convenient closed-form solution

$$\overline{z}^n = \left(\frac{k}{k+1-\sigma}\right)^{\frac{1}{\sigma-1}} z_L \tag{5}$$

$$\overline{z}_t^d = \left(\frac{k}{k+1-\varrho}\right)^{\frac{1}{\varrho-1}} z_t^* \tag{6}$$

$$1 - F(z_t^*) = \frac{N_t^d}{N_t} = \left(\frac{z_L}{z_t^*}\right)^k = \left(\frac{z_L}{\overline{z}_t^d}\right)^k \left(\frac{k}{k+1-\varrho}\right)^{\frac{k}{\varrho-1}} \tag{7}$$

Note that $\overline{z}_t^d$ depends on t via z_t^* which may fluctuate over time with economy-wide wages and durable output and prices as equation (4) shows. I defined N_t as the measure of all active firms in the economy (which is also the measure of firms active in Sector n). Let N_t^d be the measure of the subset of all firms also active in Sector d. I can now express all equilibrium quantities as averages of the "representative firm" in the economy (that has productivity $\overline{z}^n$ or $\overline{z}_t^d$ respectively) times the measure of active firms, N_t and N_t^d, respectively. These average prices, quantities and profits are

$$P_t^n \equiv 1 = N^{\frac{1}{1-\sigma}} p_t^n(\overline{z}^n) = \frac{\sigma}{\sigma-1}\frac{w_t}{\overline{z}^n} N^{\frac{1}{1-\sigma}} \tag{8}$$

$$\begin{aligned} P_t^d &= (N_t^d)^{\frac{1}{1-\varrho}} p_t^d(\overline{z}^d) = \frac{\varrho}{\varrho-1}\frac{w_t}{\overline{z}_t^d}(N_t^d)^{\frac{1}{1-\varrho}} \\ &= \frac{\varrho}{\varrho-1}\frac{\sigma-1}{\sigma}\left(\frac{z_t^*}{z_L}\right)^{\frac{k+1-\varrho}{\varrho-1}}\left[\frac{k}{k+1-\varrho}N_t\right]^{\frac{\varrho-\sigma}{(\varrho-1)(\sigma-1)}} \end{aligned} \tag{9}$$

$$y_t^n(\overline{z}^n) = \left(\frac{\sigma-1}{\sigma}\frac{\overline{z}^n}{w_t}\right)^{\sigma} Y_t^n \tag{10}$$

$$y_t^d(\overline{z}_t^d) = \left(\frac{\varrho-1}{\varrho}\frac{\overline{z}_t^d P_t^d}{w_t}\right)^{\varrho} Y_t^d \tag{11}$$

$$l_t^n(\overline{z}^n) = \frac{y_{it}^n}{\overline{z}^n} = \left(\frac{\sigma-1}{\sigma}\frac{1}{w_t}\right)^{\sigma} (\overline{z}^n)^{\sigma-1} Y_t^n \tag{12}$$

$$l_t^d(\overline{z}_t^d, c_f) = \frac{y_{it}^d}{\overline{z}_t^d} + c_f = \left(\frac{\varrho-1}{\varrho}\frac{P_t^d}{w_t}\right)^{\varrho}\left(\overline{z}_t^d\right)^{\varrho-1} Y_t^d + c_f \tag{13}$$

$$\pi_t^n(\overline{z}^n) = \frac{1}{\sigma}\left(\frac{\sigma-1}{\sigma}\frac{\overline{z}^n}{w_t}\right)^{\sigma-1} Y_t^n \tag{14}$$

$$\pi_t^d(\overline{z}_t^d) = \frac{1}{\varrho}\left(\frac{\varrho-1}{\varrho}\frac{\overline{z}_t^d}{w_t}\right)^{\varrho-1}\left(P_t^d\right)^{\varrho} Y_t^d - w_t c_f \tag{15}$$

The relative average price of durables deserves some more attention:

$$\frac{p_t^d}{p_t^n} = \frac{\varrho}{\varrho-1}\frac{\sigma-1}{\sigma}\frac{\overline{z}^n}{\overline{z}_t^d}$$

As I can see, the relative price of durables, $p_t^d(\overline{z}_t^d)/p_t^n(\overline{z}^n)$, varies over time with entry/exit and with the productivity cutoff z_t^*. The cutoff moving up leads to more efficient firms in durables on average. Equivalently, when the cutoff is high, the marginal cost in durables are lower. Because firms charge a price which is constant markup over marginal cost, and their marginal cost are lower on average, their price will be lower as well.

3.4 Firm entry and exit

For tractability reasons, firms in the durables goods sector are also active in non-durables. One may think of Sector n to also include services firms typically provide along their products. A firm that sells cars typically also generates revenue from servicing these cars and selling spare parts. It is plausible to assume that service-related revenues are not as cyclical as car sales themselves. Keep in mind that a firm can always serve both markets if it elects so.

As in Melitz (2003), active firms die randomly at rate ζ (death shock). This is of course a strong assumption that might be related to my statements about productivity. I maintain this assumption, however, because it keeps the analysis tractable. Recent empirical research on firm entry and exit by Lee and Mukoyama (2008) have established that firm exit rates are fairly acyclical (contrary to firm entry rates which are strongly procyclical). This lends support to the assumption of an acyclical exit rate. Given this death rate, the probability of survival until period T from today is $(1-\zeta)^T$.

New firms enter if the expected net present value of profits are larger than a sunk entry cost. Sunk entry cost are denoted in units of labour c_e and receive the same wage as labour employed in production or as a fixed input. A firm makes profits every period after entry until it dies exogenously. Let v_t denote the expected pre-entry net present value (in utils – because firms are owned by households) of the profit stream. This can be written as

$$v_t = E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} (1-\zeta)^{\tau} \pi_{t+\tau} \right]$$

where β is the household discount factor, λ_t the Lagrange multiplier on the household budget constraint and π_t average firm profits. $\beta^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t}$ is the stochastic discount factor. It denominates the period t marginal utility a household hasÁ from obtaining one unit good in period $t+\tau$. $(1-\zeta)^{\tau}$ denotes the probability of firm survival until period $t+\tau$. In any period, entry will occur until the expected profits are low enough to make the free entry condition hold.

$$v_t = c_e w_t \tag{16}$$

3.5 Households

The representative household has preferences over two goods, consumption, C_t, and services from durable goods, and leisure. I assume that utility from durable goods is simply a linear function in the stock of durables: ηD_t. It is assumed that durables evolve according to the following law of motion

$$D_{t+1} = (1-\delta)D_t + I_t.$$

He offers his labour in a competitive labour market and earns wage rate w_t. In addition to his labour income, he receives repayment and interest from bond holdings he invested in last period, $(1+r_t)B_t$ and holds shares in intermediate firms, $s_t N_t$ that entitle him to current-period profits π_t. Note that profits that go into the household budget constraint are the flow profits per share equity:

$$\pi_t = \pi_t^n + \frac{N_t^d}{N_t} \pi_t^d \tag{17}$$

The household problem is hence to maximise

$$\max_{C_t, D_{t+1}, L_t, s_{t+1}} U = \sum_t \beta^t \left[\frac{\left[C_t^{\alpha} (\eta D_t)^{\gamma} (1 - \phi_t L_t)^{\psi} \right]^{1-\theta}}{1-\theta} \right]$$

$$s.t. \quad (1+\tau_t^c) C_t + P_t^d (1+\tau_t^I) \left[D_{t+1} - (1-\delta) D_t \right] + v_t s_{t+1} \left(N_t + N_t^E \right) \leq w_t L_t + (v_t + \pi_t) s_t N_t + T_t$$

The household has to pay consumption and capital goods taxes, τ_t^c and τ_t^I, respectively. The government uses tax revenues to redistribute them back to hosueholds as lump-sum.

$$T_t = \tau_t^c C_t + \tau_t^i P_t^d I_t \tag{18}$$

These distortionary taxes are redistributed by the government lump-sum: T_t. How the tax rates are set will be explained momentarily.

The household is endowed with a unit measure of time and L_t is hours worked, β is his discount factor and θ, the inverse of intertemporal elasticity of substitution. ϕ_t denotes an intratemporal preference shock regulating the trade off between consumption and leisure. I assume that ϕ_t is small enough so that the optimal labor supply is still interior on the unit interval. B_{t+1} and s_{t+1} are bond and equity holdings, respectively, at the beginning of period $t+1$. v_t is the value of equity and π_t is the per-period flow profit the households receives from holding equity (dividends). It is assumed that the household portfolio is perfectly diversified across firms, so all idiosyncratic risk (death shock, productivity draw of entrants) washes out. Let λ_t be the Lagrange multiplier on on the household's budget constraint and let $X_t \equiv \left[C_t^{\alpha} (\eta D_t)^{\gamma} (1 - \phi_t L_t)^{\psi} \right]$, then the first-order conditions are

$$\frac{\alpha}{C_t} X_t^{1-\theta} = \lambda_t (1+\tau_t^c) \tag{19}$$

$$\lambda_t P_t^d (1+\tau_t^I) = \beta \lambda_{t+1} (1+\tau_{t+1}^I) P_{t+1}^d (1-\delta) + \beta X_{t+1}^{1-\theta} \frac{\gamma}{D_{t+1}} \tag{20}$$

$$X_t^{1-\theta} \frac{\psi \phi_t}{(1-\phi_t L_t)} = \lambda_t w_t \tag{21}$$

$$\lambda_t v_t (N_t + N_t^E) = \beta \lambda_{t+1} (v_{t+1} + \pi_{t+1}) N_{t+1}$$

$$\lambda_t v_t = \beta (1-\zeta) \lambda_{t+1} (v_{t+1} + \pi_{t+1}) \tag{22}$$

Keep in mind that the timing assumptions about firm entry and death imply that

$$N_{t+1} = (1-\zeta)(N_t + N_t^E). \tag{23}$$

The death shock hits at the end of period t; it hits all incumbent firms, N_t, and all firms that just entered in t and planned to take up production in period $t+1$.

3.6 Equilibrium

Determinacy The model may have multiple equilibria. This feature arises from the endogenous selection of firms into durable goods production. To illustrate the possibility of multiplicity, consider the following scenario near a deterministic steady state: In the absence of shocks, households think about increasing their demand for durables today, I_t increases, and reducing it tomorrow, I_{t+1} falls. If the equilibrium was unique, such a strategy would lead to an increase in marginal utility of consumption today (λ_t) and decrease of marginal utility from both consumption and durables tomorrow ($\lambda_{t+1}, u'(D_{t+1})$) and equation (20) would not hold. In particular, increasing demand for durables would increase the price of them today (higher P_t^d) and lower demand tomorrow would decrease the relative price of durables (lower P_{t+1}^d), thus violating equation (20) even more. In the present context, the scenario just described is feasible without violating any equilibrium conditions. Additional demand for durables today triggers firm entry. This in turn increases labour demand and hence the wage, which is the main component of the productivity cutoff (4). If the cutoff increases, average productivity increases and the relative price of durables *falls* thus justifying the initial beliefs of households. The key feature of this model is that an *increase* in demand for durables can lead to a *lower* relative price of durables.

This little example shows, how beliefs can become self-fulfilling. This is an interesting feature of the model as it allows for endogenous fluctuations. Two factors can mitigate this decrease in the relative price and keep the equilibrium in the model economy determinate: a low intertemporal elasticity of substitution (high θ) and taxes on investment goods that increase when demand is high (high τ_t^I).[27] The multiplicity feature of this model context is subject of ongoing research. In the present paper, I focus on a unique equilibrium in the model. θ is calibrated to economically plausible values consistent with other research. The taxes are allowed to be a function of economic activity:

$$\tau_t^c = \overline{\tau}^c + x_c^c C_t + x_c^L L_t$$
$$\tau_t^i = \overline{\tau}^i + x_i^i I_t + x_i^L L_t$$

Tax rates in the current period, τ_t^c and τ_t^i are a function of the long-run average tax rate, $\overline{\tau}^c$ and $\overline{\tau}^I$, and the levels of consumption and investment. We can think of the fact that taxes are dependent on current economic activity as as implicit policy maker that attempts to design tax policy according to economic conditions. If $x_\bullet^\bullet > 0$ this corresponds to countercyclical fiscal policy. This "countercyclical fiscal policy" does not seem implausible. Relaxing this assumption and analysing endogenous fluctuations in the context of this model is subject of future research.

[27] This approach is chosen in Christiano and Harrison (1999).

Labour market Aggregate labour demand is key as shocks to aggregate demand drive labour demand which in turn drives the wage and therefore fixed cost. Aggregate labour demand is

$$L_t = N_t l(\overline{z}^n) + N_t^d l(\overline{z}_t^d) + N_t^E c_e \tag{24}$$

Equation (24) denotes aggregate labour demand while equation (21) determines aggregate labour supply. Both taken together determine equilibrium.

Goods market There are two goods, durables and non-durables that are each produced by the final goods producer in each sector and purchasd by households:

$$Y_t^n = C_t \tag{25}$$

$$Y_t^d = I_t = D_{t+1} - (1-\delta)D_t \tag{26}$$

Resource Constraint This is basically the budget constraint reformulated

$$w_t L_t + N_t \pi_t + N_t v_t = C_t + P_t^d \left[D_{t+1} - (1-\delta)D_t\right] + (N_t + N_t^E)v_t \tag{27}$$

General equilibrium The equilibrium consists of a set of endogenous variables

$$z_t^*, \overline{z}_t^d, P_t^d, Y_t^n, Y_t^d, y_t^n, y_t^d, l_t^n, l_t^d, \pi_t^n, \pi_t^d, \pi_t, N_t, N_t^E, N_t^d, v_t, C_t, I_t, \lambda_t, L_t, w_t$$

that satisfies firm and household optimality as well as feasibility as prescribed by equations (4), (6)-(27).

Shock and Calibration At present, the model is driven by one shock in the household sector that evolves over time as follows (in logs)

$$\phi_t = (1-\rho_\phi)\overline{\phi} + \rho_\phi \phi_{t-1} + \varepsilon_t^\phi$$

where $\overline{\phi}$ is the steady state value, ρ_ϕ the autocorrelation and the disturbances is independent and distributed normally as

$$\varepsilon_t^\phi \sim \mathcal{N}(0, \sigma_\phi).$$

Table 6 displays the parameters chosen for calibration. The productivity distribution is parametrised as a Pareto with lower bound $z_L = 1$ and shape parameter k. The shape parameter is the smallest admissible value that satisfies the condition $k > \sigma - 1$ and $k > \varrho - 1$. Recall that these conditions are necessary if we want to restrict attention to positive equilibrium quantities and prices. In accordance with the literature on mark-ups σ and ϱ are chosen to equal 5.4 and 3.5 respectively which gives the smallest possible value for k as 4.5. The rate of death shock, ζ, is chosen to match the exit rate in my sample. The fixed input factor c_f is chosen to match the observed steady state returns to scale in the durable goods sector. The depreciation rate in consumer durables, δ, is set to 0.085 to match an annual depreciation rate of 30% as common for consumer durables. The discount rate $\beta = 0.9925$ reflects an annual real interest rate of about 3%. The intertemporal elasticity of substitution $\theta = 1.5$ is chosen to accord with empirical studies. The utility weight on consumption, α, is set equal to 0.66 in order to match expenditure shares for consumption goods.

TABLE 6: CALIBRATION

Parameter	Symbol	Value
Pareto Distribution: Shape	k	4.5
Pareto Distribution: Lower bound	z_L	1
Rate of death shock	ζ	0.05
EOS non-durables	σ	5.4
EOS durables	ϱ	3.5
Sunk entry cost	c_e	5
Fixed Cost in Durables	c_f	0.1
Deprecitation rate durables	δ	0.085
Discount Factor	β	0.9925
Inv. intertemporal EOS	θ	1.5
Utility weight non-durables	α	0.66
Utility weight durables	γ	$1-\alpha$
Utility weight leisure	ψ	0.7
Mean consumption tax rate	$\overline{\tau}^c$	0.1
Mean investment tax rate	$\overline{\tau}^i$	0.2
Inv. Tax dependence on I	x_i^i	2.5
Inv. Tax dependence on L	x_i^L	0.5
Mean demand shock	$\overline{\phi}$	1
Autocorrelation demand shock	ρ_ϕ	0.8
Std. Dev. demand shock	σ_ϕ	1

3.7 Dynamics of the economy

How does the economy respond to a shock to ϕ? This shock originally shifts the labour supply schedule and also change the consumption plan. In that way, we can think about it as a shock that alters aggregate demand. I shall first describe the initial adjustment on impact and the dynamic response. The outward shift in labour supply increases employment while leaving the wage initially unaffected. This is rather unusual and it is a consequence of firm technology. It is linear homogeneous in labour, so that the marginal product of labour is independent of the level of labour. The increased employment leads to higher labour income which immediately raises demand for consumption and investment goods. With a constant wage this raises profits. In light of higher profits, new firms enter up to the point that the free entry condition equation (16) holds. The new equilibrium right after the impact of the shock involves higher employment, output, consumption, investment and more entrants (that will not be active until the next period).

In the periods after the shock, new firm entry continues, but it gradually converges back to its steady state level. This is illustrated in panel (3,2) of Figure 8. The additional entry leads to a hump-shaped rise in the mass of incumbent firms panel (3,1) which in turn increases labour demand in a similar fashion – see panel (2,1). Because aggregate labour demand increases, so do real wages, displayed in panel (2,2).[28] Real wages are a key factor in the productivity cutoff. As the cutoff rises, see panel (1,1), the productivity distribution in durables becomes more compressed, thus leading to negative co-movement between output and productivity dispersion. This is the key result of the theoretical part of the model: It is possible to reconcile a boom with a more compressed productivity dispersion. This happens although the model in principle does allow for cleansing. A second consequence of a rising productivity cutoff is that the average productivity in durables increases. This lowers the average marginal cost and thus the average relative price of durables which is displayed in panel (1,2). This feature is a nice complement to the literature on investemnt-specific technological change which has to rely on exogenous technology shocks in the durable goods sector. In this model, the relative price of durables fluctuates endogenously due to the selection along a rising productivity cutoff in a boom. Lastly, both consumption and durables purchases increase, see panels (4,1) and (4,2). Households dispose over a higher income, so they increase their demand for goods overall. As panel (4,2) illustrates, this rise is stronger in durables purchases which is partly due to the relative price of durables declining.

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. The dispersion appears to be more negatively correlated with the business cycle in durable goods industries relative ot non-durable goods industries. Lastly, I found the dynamics to be driven predominantly by changes in the lower quantiles. This evidence can be interpreted as a changing truncation at the bottom end of the productivity distribution in durables.

[28] Note that wages also increase because of additional entry into the economy and a higher mass of firms paying the fixed cost to produce in durables. This is merely reinforcing the effect of additional labour demand in production.

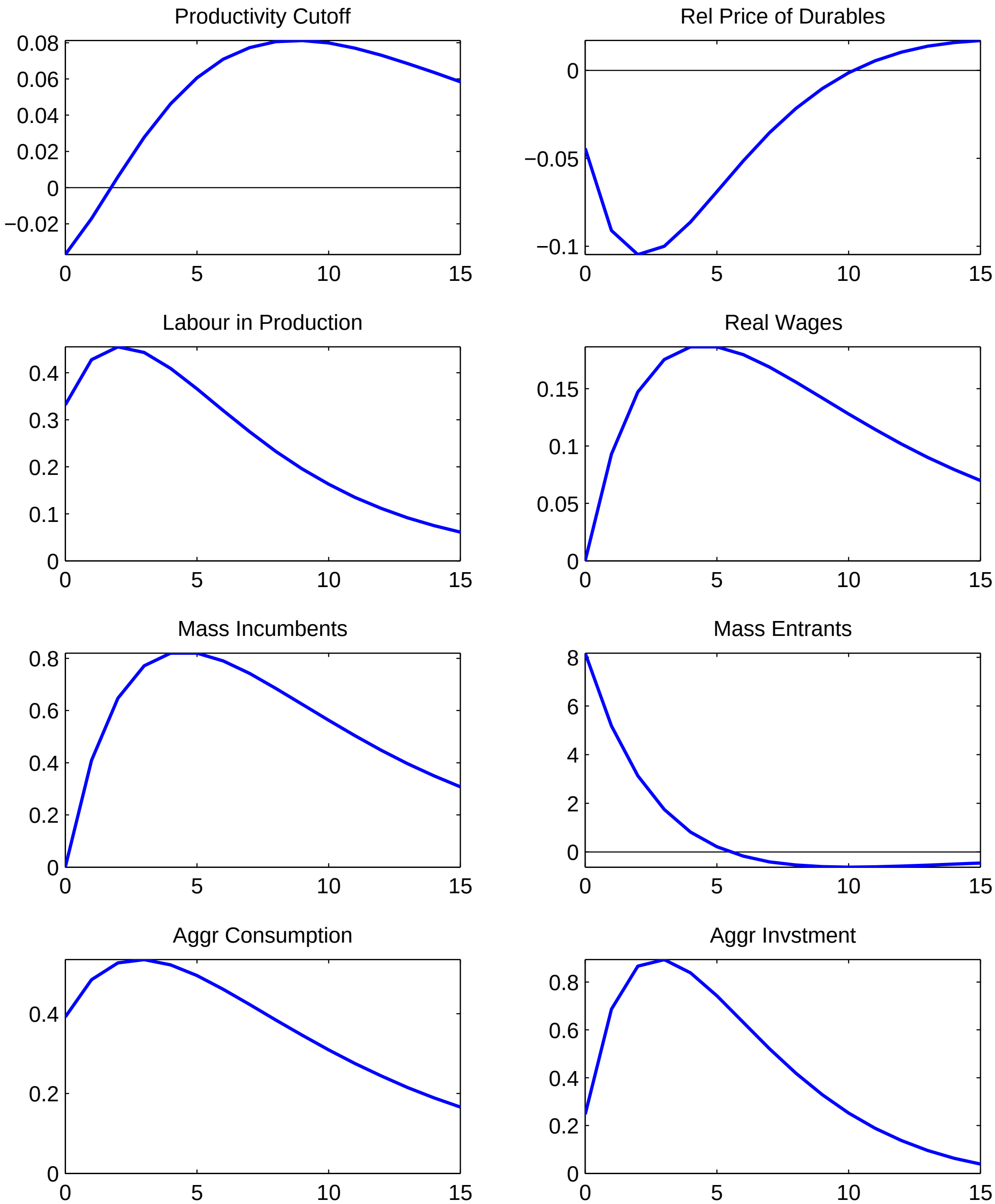

Figure 8: Impulse Response Function of an aggregate demand shock to ϕ

This countercyclical productivity dispersion is at odds with conventional cleansing models of the business cycle which posit a procyclical productivity dispersion. In order to reconcile these models with the empirical findings, I built a business cycle model along the lines of Ghironi and Melitz (2005). In my model, a shock that originates in the household sector and changes aggregate demand is consistent with a countercyclical productivity dispersion. In addition to that, this shock is also consistent with the typical macroeconomic business cycle acts such as procyclical consumption, investment, wages and employment.

I will direct future research at introducing more shocks into the model. It is plausible to assume that some shocks deliver the empirically observed outcome and others don't. Which are these and what is their distinct criterion? A second possible extension is to focus on a micro-founded theory of investment-specific fluctuations. The endogenous selection of firms into durables on a productivity threshold is an interesting mechanism that can contribute to the research programme on investment-specific technical change. Furthermore, this avenue will also open the possibility to look into endogenous fluctuations that arise from indeterminacy.

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.

Rüdiger Bachmann and Christian Bayer. Firm-specific productivity risk over the business cycle: Facts and aggregate implications. *Working Paper*, September 2009a.

Rüdiger Bachmann and Christian Bayer. The cross-section of firms over the business cycle: New facts and a DSGE exploration. *Working Paper*, September 2009b.

Yan Bai, José-Victor Ríos-Rull, and Kjetil Storesletten. Demand shocks that look like productivity shocks. *Working Paper*, November 2010.

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

Natarajan Balasubramanian and Jagadeesh Sivadasan. Capital resalability, productivity dispersion, and market structure. *Review of Economics and Statistics*, 91(3):547–557, August 2009.

Eric J. Bartelsmann and Mark Doms. Understanding productivity: Lessons from longitudinal data. *Journal of Economic Literature*, 38(3):569–594, September 2000.

Eric J. Bartelsmann, Ricardo J. Caballero, and Richard K. Lyons. Customer- and supplier-driven externalities. *American Economic Review*, 84(4):1075–1084, September 1994.

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 increasing returns. *Journal of Economic Theory*, 63(1):19–41, June 1994.

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

Nicholas Bloom, Max Floetotto, and Nir Jaimovich. Really uncertain business cycles. *Working Paper*, 2009.

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. *NBER Macroeconomics Annual*, 10(1):67–110, 1995.

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

Ricardo J. Caballero and Mohamad L. Hammour. On the timing and efficiency of creative destruction. *Quarterly Journal of Economics*, 111(3):805–852, August 1996.

Ricardo J. Caballero and Richard K. Lyons. External effects in U.S. procyclical productivity. *Journal of Monetary Economics*, 29(2):209–225, April 1992.

Jeffrey R. Campbell. Entry, exit, embodied technology, and business cycles. *Review of Economic Dynamics*, 1(2):371–408, April 1998.

Rui Castro, Gian Luca Clementi, and Yoonsoo Lee. Cross-sectoral variation in firm-level idiosyncratic risk. *FRB Cleveland Working Paper 08-12*, January 2009.

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 and Sharon G. Harrison. Chaos, sunspots and automatic stabilizers. *Journal of Monetary Economics*, 44(1):3–31, August 1999.

Lawrence J. Christiano, Roberto Motto, and Massimo Rostagno. Financial factors in economic fluctuations. *Working Paper*, May 2009.

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.

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.

Gian Luca Clementi and Dino Palazzo. Entry, exit, firm dynamics, and aggregate fluctuations. *Working Paper*, September 2010.

Russell Cooper and John C. Haltiwanger. Inventories and the propagation of sectoral shocks. *American Economic Review*, 80(1):170–190, March 1990.

Russell Cooper and John C. Haltiwanger. The aggregate implications of machine replacement: Theory and evidence. *American Economic Review*, 83(3):181–186, June 1993.

Russell Cooper, John C. Haltiwanger, and Laura Power. Machine replacement and the business cycle: Lumps and bumps. *American Economic Review*, 89(4):921–946, September 1999.

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

Steve J. Davis and John C. Haltiwanger. Gross job creation, gross job destruction, and employment reallocation. *Quarterly Journal of Economics*, 107(3):819–863, August 1992.

Steven J. Davis and John C. Haltiwanger. Gross job creation and destruction: Microeconomic evidence and macroeconomic implications. *NBER Macroeconomics Annual*, 5:123–168, 1990.

Steven J. Davis and John C. Haltiwanger. On the driving forces behind cyclical movements in employment and job reallocation. *American Economic Review*, 89(5):1234–1258, December 1999.

Andrea L. Eisfeldt and Dimitris Papanikolaou. Organization capital and the cross-section of expected returns. *Working Paper*, June 2010.

Andrea L. Eisfeldt and Adriano A. Rampini. Capital reallocation and liquidity. *Journal of Monetary Economics*, 53(3):369–399, April 2006.

Andrea L. Eisfeldt and Adriano A. Rampini. Managerial incentives, capital reallocation, and the business cycle. *Journal of Financial Economics*, 87(1):177–199, January 2008.

Jonas D. 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.

Jeremy T. Fox and Valérie Smeets. Does input quality drive measured differences in firm productivity? *International Economic Review*, forthcoming.

Xavier Gabaix. The granular originals of aggregate fluctuations. *Econometrica*, forthcoming.

Amit Gandhi, Salvador Navarro, and David Rivers. Identifying production functions using restrictions from economic theory. *Working Paper*, June 2009.

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

João Gomes, Jeremy Greenwood, and Sergio Rebelo. Equilibrium unemployment. *Journal of Monetary Economics*, 48(1):109–152, August 2001.

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.

Robert E. Hall. Invariance properties of Solow's residual. In Peter A. Diamond, editor, *Growth, Productivity, Employment: Essays to Celebrate Bob Solow's Birthday*. MIT Press, 1990.

Robert E. Hall. Labor demand, labor supply, and employment volatility. *NBER Macroeconomics Annual*, 6:17–47, 1991.

Robert E. Hall. Reorganization. *Carnegie-Rochester Conference Series on Public Policy*, 52: 1–22, June 2000.

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.

Nir Jaimovich. Firm dynamics and markup variations: Implications for sunspot equilibria and endogenous economic fluctuations. *Journal of Economic Theory*, 137(1):300–325, November 2007.

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.

Aubhik Khan and Julia K. Thomas. Nonconvex factor adjustments in equilibrium business cycle models: Do nonlinearities matter? *Journal of Monetary Economics*, 50(2):331–360, March 2003.

Aubhik Khan and Julia K. Thomas. Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436, March 2008.

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. *Census Discussion Paper CES 08-17*, June 2008.

Rasmus Lentz and Dale T. Mortensen. An empirical model of growth through product innovation. *Econometrica*, 76(6):1317–1373, November 2008.

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

James Levinsohn and Amil Petrin. Measuring aggregate productivity growth using plant-level data. *Working Paper*, May 2010.

Robert E. Lucas. On the size distribution of business firms. *The 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.

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

Dale T. Mortensen and Christopher A. Pissarides. Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, 61(3):397–415, July 1994.

Kevin M. Murphy, Andrei Shleifer, and Robert W. Vishny. Building blocks of market clearing business cycle models. *NBER Macroeconomics Annual*, 4:247–287, 1989.

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

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*, forthcoming.

Thomas J. Sargent, Noah Williams, and Tao Zha. Shocks and government beliefs: The rise and fall of American inflation. *American Economic Review*, 96(4):1193–1224, September 2006.

Joseph Alois Schumpeter. *Business Cycles*. McGraw-Hill, New York, 1939.

Zheng Song, Kjetil Storesletten, and Fabrizio Zilibotti. Growing like China. *American Economic Review*, forthcoming.

James H. Stock and Mark W. Watson. Has the business cycle changed and why? *NBER Macroeconomics Annual*, 17:159–218, 2002.

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.

Julia K. Thomas. Is lumpy investment relevant for the business cycle? *Journal of Political Economy*, 110(3):508–534, June 2002.

Marcelo L. Veracierto. Plant-level irreversible investment and equilibrium business cycles. *American Economic Review*, 92(1):181–197, March 2002.

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

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 Manufacturing Survey (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 (year 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 that is 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

The object of interest is a real measure of goods produced ($\mathtt{Q}$). It consists of goods that are produced and sold in the same year ($\mathtt{PS}$) 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 ($\mathtt{PS}$) 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 ($\mathtt{TVS}$) is considered by Census to be of superior quality. We can use this variable to infer $\mathtt{PS}$ as shown in Figure 9.

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

where $\mathtt{VPS}$ is the nominal value of product shipments, $\mathtt{PISHIP}$ is a price deflator on the 4-digit (SIC) industry level from the NBER-CES Manufacturing Productivity Database and $\mathtt{VIS}$ 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. $\mathtt{FIB}$ and $\mathtt{FIE}$ denote the nominal value of finished goods inventory at the beginning and end of the period. $\mathtt{FII}$ 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 $\mathtt{FII}$ is non-negative, because finished goods never flow back from the inventory to production. The last

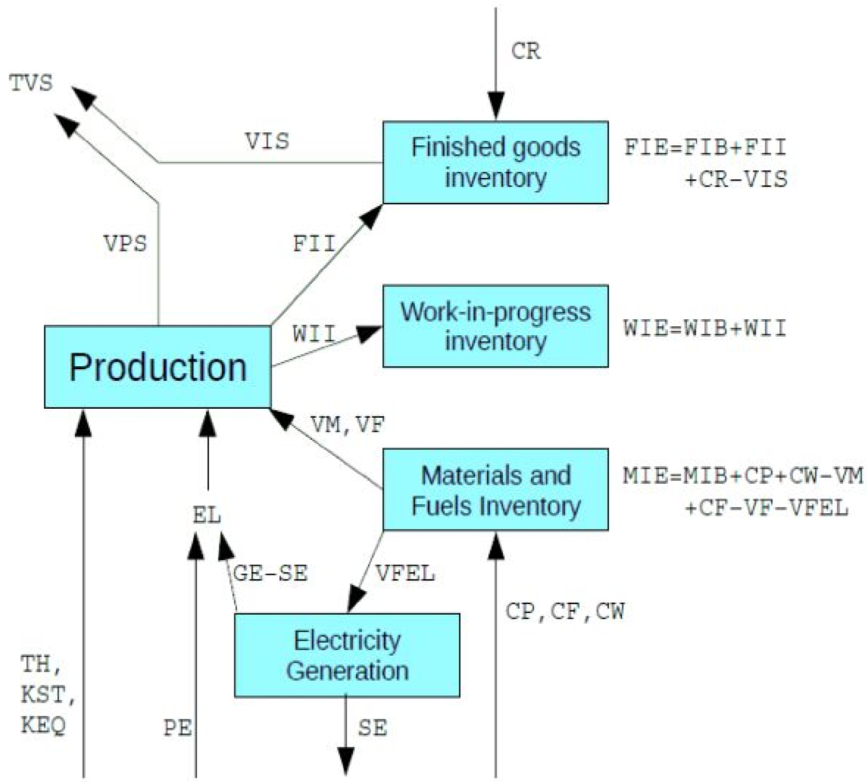

Figure 9: ASM flowchart; all variables are nominal except hours worked (TH) and the capital stocks of structures (KST) and equipment (KEQ).

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).[29]

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[30], so this difference might matter when one computes finished inventory investment:

[29]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.

[30]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.

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

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:

$$\begin{aligned}
\mathtt{Q} &= \mathtt{PS} + \mathtt{FII}^{\mathtt{real}} + \mathtt{WII}^{\mathtt{real}} \\
&= \frac{\mathtt{TVS} - \mathtt{VIS}}{\mathtt{PISHIP}} + \frac{\mathtt{FIE} - \mathtt{FIB}}{\mathtt{PIFI}} - \frac{\mathtt{CR} - \mathtt{VIS}}{\mathtt{PISHIP}} + \frac{\mathtt{WIE} - \mathtt{WIB}}{\mathtt{PIWI}} \\
\mathtt{Q} &= \frac{\mathtt{TVS} - \mathtt{CR}}{\mathtt{PISHIP}} + \frac{\mathtt{FIE} - \mathtt{FIB}}{\mathtt{PIFI}} + \frac{\mathtt{WIE} - \mathtt{WIB}}{\mathtt{PIWI}} \qquad (28)
\end{aligned}$$

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 are 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.[31] 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.[32]

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

[32] At this point, I am following the large productivity literature and estimate revenue factor productivity (TFPR

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, Reiter, and White (forthcoming)). 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.[33]

A.3 Measurement of labor 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 the sum of production worker hours (`PH`) and non-production worker hours (`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 (`WNP`) are 150% of those for production workers (`WP`): $\mathtt{WNP} = 1.5 \times \frac{\mathtt{WW}}{\mathtt{PH}}$.[34]

in Foster, Haltiwanger, and Syverson (2008) or Hsieh and Klenow (2009)).

[33] 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 $\mathtt{CR} = 0.05M$. This bias obviously becomes stronger the larger `CR`.

[34] A very proper way would be to utilise external information from the Current Population Survey to construct

Total hours under this assumption can be calculated as:

$$\begin{aligned}
\texttt{TH} &= \texttt{PH} + \texttt{NPH} \\
&= \texttt{PH} + \frac{\texttt{SW} - \texttt{WW}}{\texttt{WNP}} \\
&= \texttt{PH} + \frac{\texttt{SW} - \texttt{WW}}{1.5 \times \texttt{WP}} \\
\texttt{TH} &= \texttt{PH}\,\frac{\texttt{SW} + 0.5 \times \texttt{WW}}{1.5 \times \texttt{WW}}
\end{aligned} \tag{29}$$

Total hours worked can be constructed in this way for about 97.6% of all observations. The remaining observations don't have information on either of PH, SW or WW. In that case, I set $\texttt{TH} = 2 \times \texttt{TE}$ (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:

$$\tilde{K}_t \equiv u_t K_t. \tag{30}$$

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.

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: $\texttt{a} = \frac{\texttt{WNP}}{\texttt{WP}}$. Then, total hours can be computed on the establishment level as:

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

ALTERNATIVE:

One could get data on hours worked per employee in both production and non-production: $\texttt{HRS}_{\texttt{PW}}$ and $\texttt{HRS}_{\texttt{NPW}}$. These data should be available on an industry-region level in the CPS. Then, total hours can be computed as $\texttt{TH} = \texttt{HRS}_{\texttt{PW}} \times \texttt{PW} + \texttt{HRS}_{\texttt{NPW}} \times \texttt{NPW}$ 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[35]) 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

[35] 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.

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.[36] 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 yiel 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.[37] 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).[38] 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[39] 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 τ is the vintage (purchased τ periods before period t), δ 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-

[36]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.

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

[38]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.

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

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.[40] 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 NBER Manufacturing Database.[41] I choose the BEA deflators because they were revised recently (in 2009), which matters a lot for capital goods (esp. equipment).[42] 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:

$$K_t^{st} = \mathtt{BAB}_t(1-\delta^{st})\frac{CC_t^{st}}{HC_t^{st}}\frac{1}{P_t^{st}} \tag{31}$$

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

[40] 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.

[41] 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.

[42] 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.

using the perpetual inventory method:

$$K_{t+1}^{st} = (1 - \delta_t^{st}) K_t^{st} + I_t^{st} \tag{32}$$

where K_t^{st} is the stock of structure capital observed at the beginning of the period, $I_t^{st} = \frac{\mathtt{NB}_t}{P_t^{st}}$ is real structure investment (nominal new and used[43] investment expenditures divided by an investment price index) and δ_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, Haltiwanger, Jarmin, Klimek, and Wilson (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 I 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'm sceptic of resetting the capital stock back to the value implied by equation (31) 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 the implied total capital stock (keeping the asset split implied by the perpetual inventory method). I only use equation (31) 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.

[43] 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.

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% (structures - !) and 5.6% (equipment) respectively. This barely squares with industry-wide aggregates where the capital stock grows at 1.3% and 2.5% (structures and equipment resp.). My measures end up at 1.3% 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-specific, this essentially introduces industry dummies into a regression analysis. The former will put an upward bias to the coefficient on capital while the former will pick up industry-specifics that are not necessarily rooted in the capital stock.

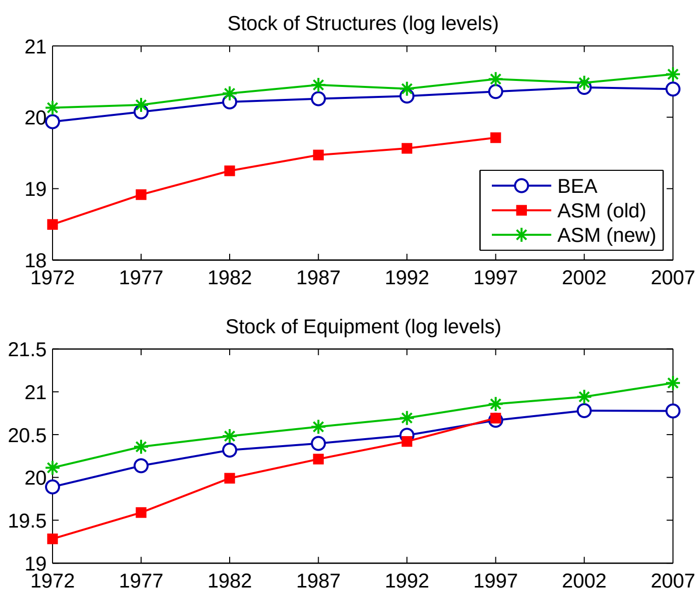

Figure 10: Capital stocks of the U.S. manufacturing sector published by BEA (blue circles), aggregating the old (red squares) and my refined (green stars) capital stock data in the CMF/ASM. CMF data for 2007 are still preliminary.

In a similar vein, I find that the old investment measure for equipment is off the benchmark

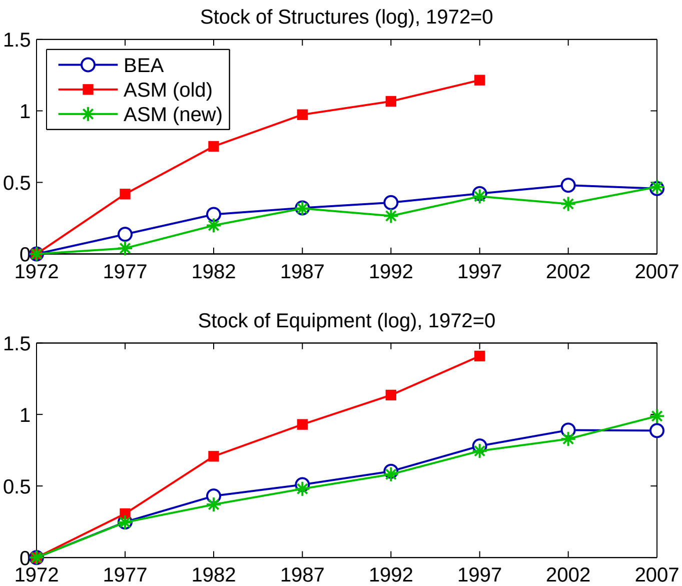

Figure 11: Capital stocks of the U.S. manufacturing sector published by BEA (blue circles), aggregating the old (red squares) and my refined (green stars) capital stock data in the CMF/ASM. CMF data for 2007 are still preliminary. Normalised to 0 in 1972. Clearly visible that the trend growth is off in the old ASM capital measures.

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. All this is documented in the figures in the appendix.

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 it is very prevalent 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, Eichenbaum, and Rebelo (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, Eichenbaum, and Rebelo (1995) have suggested to use electricity or energy instead of capital. The idea in that paper is that energy/electricity is 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.[44] This works for all years after 1974 when the PCU started. For 1972/73, I use utilization rates computed by the Federal Reserve Board[45].

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 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, `CP`, contract work, `CW`) or come from the materials inventory and are then used in production. Measurement is complicated by the fact that materials inventories (`MIB` and `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 storable only

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

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

at high cost (natural gas e.g.) this seems like a reasonable assumption. Then, I can express the value of materials used in the production process ($\mathtt{VM}$) through the inventory identity

$$\begin{aligned} \mathtt{MIE} &= \mathtt{MIB} + (\mathtt{CP} + \mathtt{CW} - \mathtt{VM}) \times \frac{\mathtt{PIMI}}{\mathtt{PIMAT}} \\ \Leftrightarrow \mathtt{M} \equiv \frac{\mathtt{VM}}{\mathtt{PIMI}} &= \frac{\mathtt{MIB} - \mathtt{MIE}}{\mathtt{PIMI}} + \frac{\mathtt{CP} + \mathtt{CW}}{\mathtt{PIMAT}}. \end{aligned} \tag{33}$$

As with goods inventories above, inflows into materials inventories have to be deflated by market prices ($\mathtt{PIMAT}$), while materials stocks have to be deflated by inventory prices ($\mathtt{PIMI}$). 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 $\mathtt{PIMI}$ with $\mathtt{PIMAT}$.

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

$$\mathtt{EL} = \mathtt{PE} + \mathtt{GE} - \mathtt{SE}. \tag{34}$$

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

A.6.2 Fuels

Fuels used in production (nominally expressed as $\mathtt{VF}$) can come from fuel purchases ($\mathtt{CF}$) or from the materials/fuels inventory. As outlined above, I assume that any change in materials inventory ($\mathtt{MIE} - \mathtt{MIB}$) 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 $\mathtt{VF} = \mathtt{CF}$. 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,

$$\begin{aligned}
\texttt{VF} &= \texttt{CF} - \texttt{VFEL} \\
\texttt{VF} &= \texttt{CF} - \texttt{GE} \times \frac{\texttt{EE}}{\texttt{PE}} \\
\texttt{F} &\equiv \frac{\texttt{VF}}{\texttt{PIEN}} = \frac{\texttt{CF} - \texttt{GE} \times \frac{\texttt{EE}}{\texttt{PE}}}{\texttt{PIEN}} \qquad (35)
\end{aligned}$$

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 the production process or in the generation of electricity. Total energy expenditures (VE) comprise those for fuels (CF) and electricity (EE); the nominal value is:

$$\begin{aligned}
\texttt{VE} &= \texttt{CF} + \texttt{EE} \\
\texttt{E} &\equiv \frac{\texttt{VEN}}{\texttt{PIEN}} = \frac{\texttt{CF} + \texttt{EE}}{\texttt{PIEN}} \qquad (36)
\end{aligned}$$

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 B.3 requires knowing the cost shares of factor inputs. They are constructed as follows

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

Note that all variables are expressed in period-t nominal costs. All of these variables except r and K are observed in the original dataset. K is the real capital stock (in year-2000 dollars) constructed as described in Appendix A.4, r_t denotes the *nominal* rental rate (year-t dollars rent paid per one year-2000 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 by dividing corporate capital income (Table 3a) by the real capital stock (Table 4a). The latter variable is expressed in constant (year-2000 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 include rent and depreciation, not utilisation cost which is captured in the energy cost share.[46]

B Empirics

B.1 The Levinsohn-Petrin estimator

This section describes the estimator by Levinsohn and Petrin (2003) in more detail. Recall the assumed production function from equation (1):

$$y_{it} = \omega_{it} + \eta_{it} + \beta^l l_{it} + \beta^k k_{it} + \beta^m m_{it} + \beta^e e_{it} \tag{1'}$$

The firm productivity term is decomposed into a portion that the plant observes, ω_{it}. The authors assume that ω_{it} has a first-order Markov structure. η_{it} on the other hand is not observed by the plant and hence does not influence contemporaneous decision rules. The capital stock in period t is assumed to be fixed. The plant's state vector is hence $(k_{it}, \omega_{it}, P_t)$. P_t denotes

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

the vector of firm input prices. I assume they are the same across firms and firms take them as given. The other production inputs, labour l_{it}, materials m_{it} and energy, denoted by e_{it}, are chosen period t after the productivity is observed. This means that those inputs are a function in the state vector, in particular:

$$e_{it} = f_t(\omega_{it}, k_{it}) \tag{37}$$

Note that P_t appears implicitly in that the functions are time-dependent. LP assume that the function $f_t(\cdot)$ it is strongly monotone. This assumption allows us to invert the function and express plant-level productivity as

$$\omega_{it} = g_t(k_{it}, e_{it}) \tag{38}$$

which can be plugged into (1'). I combine the capital, material and energy inputs with the inverted expression to define a new expression

$$\phi_t = \beta^k k_{it} + \beta^m m_{it} + \beta^e e_{it} + g_t(k_{it}, e_{it}) \tag{39}$$

Of course, I have to make an assumption about $\phi_t(\cdot)$ and LP propose to estimate it non-parametrically as a third-order polynomial. Note that introducing the information contained in energy consumption implicitly accounts for capacity utilsation. It is similar in spirit to Burnside, Eichenbaum, and Rebelo (1995) who propose to replace capital in production function regressions with electricity.

The estimation procedure happens in two steps. First, regress output on labour and the third-order polynomial expansion of k_{it}, m_{it} and e_{it}:

$$y_{it} = \beta^l l_{it} + \phi_t(k_{it}, m_{it}, e_{it}) + \eta_{it} \tag{40}$$

and obtain $\hat{\beta}^l, \hat{\phi}_t$. The latter is an complicated expression in its arguments

$$\hat{\phi}(t) = \hat{\gamma}_0(t) + \sum_{p=0}^{3} \sum_{q=0}^{3-p} \sum_{t=0}^{3-p-q} \hat{\gamma}_{pqr}(t)\ k_i^p(t)\ m_i^q(t)\ e_i^r(t)$$

The coefficients on the capital stock, β^k, materials, β^m, and energy β^e can be teased out of $\hat{\phi}$ using the assumption that ω is first-order Markov. First, I express ω as function of $\hat{\phi}_t$ and arbitrary production coefficients on capital, materials and energy:

$$\omega_{it}(b^k, b^m, b^e) = \hat{\phi}_t - b^k k_{it} - b^m m_{it} - b^e e_{it} \tag{41}$$

Given the first-order Markov assumption of productivity this expression can be used to estimate a forecast $E[\omega_t|\omega_{t-1}]$

$$\hat{\omega}_{it}(b^k, b^m, b^e) = \delta_0 + \delta_1\omega_{it-1}(b^k, b^m, b^e) + \delta_2 \left[\omega_{it-1}(b^k, b^m, b^e)\right]^2 + \delta_3 \left[\omega_{it-1}(b^k, b^m, b^e)\right]^3 \quad (42)$$

Keep in mind that this is still a function in (b^k, b^m, b^e). The productivity forecast error is defined as

$$\xi_{it} \equiv \omega_{it} - E_{t-1}[\omega_{it}|\omega_{it-1}] = \omega_{it} - \hat{\omega}_{it}$$

The last step is to exploit the fact that the sum of the forecast error ξ_{it} and the unobserved part of productivity, η_{it}, are orthogonal to yesterday's variable inputs $l_{it-1}, m_{it-1}, e_{it-1}$ and today's capital stock k_{it}. Let

$$\hat{\zeta}_{it}(b^k, b^m, b^e) \equiv \eta_{it} + \xi_{it} \quad (43)$$

$$= y_{it} - \hat{\beta}^l l_{it} - b^k k_{it} - b^m m_{it} - b^e e_{it} - \hat{\omega}_{it}(b^k, b^m, b^e) \quad (44)$$

Then I can express the moment conditions as

$$E\left[\zeta_{it}, k_{it}\right] = 0$$
$$E\left[\zeta_{it}, m_{it-1}\right] = 0$$
$$E\left[\zeta_{it}, e_{it-1}\right] = 0$$
$$E\left[\zeta_{it}, l_{it-1}\right] = 0$$

The first three conditions are sufficient to identify β^k, β^m and β^e, the last conditions (or even more lags) can be added for efficiency. The parameters can then be estimated using GMM

$$[\hat{\beta}^k, \hat{\beta}^m, \hat{\beta}^e] = \underset{b^k, b^m, b^e}{\operatorname{argmin}} \sum_h \sum_t \left[\hat{\zeta}_t(b^k, b^m, b^e) \times Z_{th}\right]^2 \quad (45)$$

where $Z_{ith} = [k_{it}, m_{it-1}, e_{it-1}, l_{it-1}]$ and possibly more lags if we wish.

B.2 Long-run changes in productivity dispersion

Figure 12 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. It is constant until the late 1980's, from then on dispersion exhibits a secular upward trend. Within the last 15 years of the sample dispersion more than doubles. Similar long-run changes have been noted in Beaudry, Caglayan, and Schiantarelli

(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.

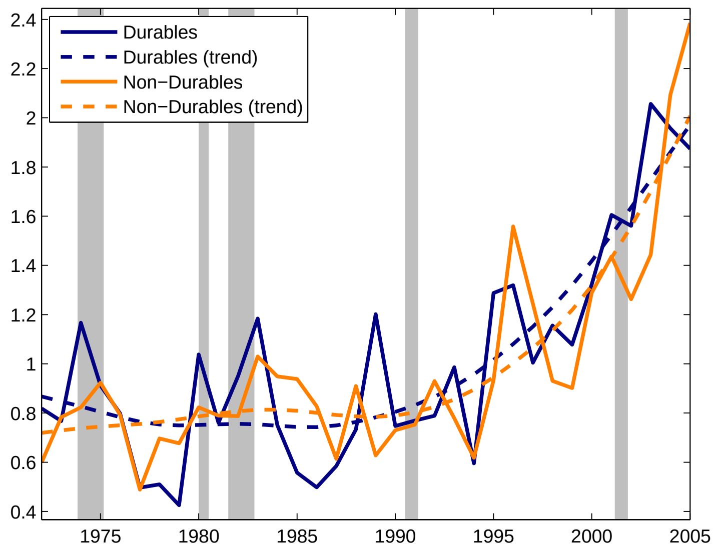

Figure 12: Trends and Cycles in Dispersion

B.3 Robustness: Cost Shares

The structural estimator proposed by Levinsohn and Petrin (2003) 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 inputs in productivity ω. These assumptions look fairly weak, but I want to make sure that they are not responsible for driving my results. This estimator has also been criticised by Ackerberg, Caves, and Frazer (2006). They claim that LP (and to a lesser extent a similar method proposed by Olley and Pakes (1996)) suffers from a collinearity problem that casts doubt on identification. An easy alternative is to assume constant returns to scale and competitive factor markets. These assumptions look reasonable given that I found close to constant returns to scale in Section 2. They are also in line hat many other people have found empirically, see for example Burnside, Eichenbaum, and Rebelo (1995); Burnside (1996); Basu and Fernald (1997) for industry-level evidence and Lee and Nguyen (2002) for plant-level evidence in the clothing industry.

Under these assumptions, production elasticities can be approximated with cost shares.[47] 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 I assume that cost shares are industry specific and constant over time. As above, we still need to detrend a_{ijt}, recenter and scale the residula z_{ijt}.

B.3.1 Solow residuals – cross-sectional dispersion

In the following, we shall take a look at the dispersion in TFP levels. This is more interesting as a higher variance in TFP levels means not only that firms simply become more risky, but has implications for aggregate TFP as well. Changes in the dispersion of TFP levels also has implications in models of frictions and lending as in Christiano, Motto, and Rostagno (2009); Christiano, Trabandt, and Walentin (2010). Firms will have more trouble getting credit if they are more risky in the sense that their TFP will be drawn from a more spread-out distribution.

In contrast to looking at the dispersion of TFP growth rates, analysing TFP levels is not as straight-forward. The distribution of TFP levels is likely to differ substantially across industries in terms of the central and higher moments. 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 don't permit a more in-depth analysis.

To make dispersion comparable across industries and over time, I will recenter the logged TFP distribution in each industry at the industry-specific median, denoted $\overline{z}_j$, and normalise it by the industry-specific standard deviation, σ_j. This will make the TFP distribution of each industry look similar on average but will bear out the time variation in the dispersion I am looking for. The statistic of interest is hence

$$
Disp_{jt} = \frac{z_{ijt} - \overline{z}_j}{\sigma_j}
$$

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

To make statements about the dynamics of the economy-wide distribution, I take the averages across industries at each point in time:

$$Disp_t = E_t \left[Disp_{jt}\right]$$

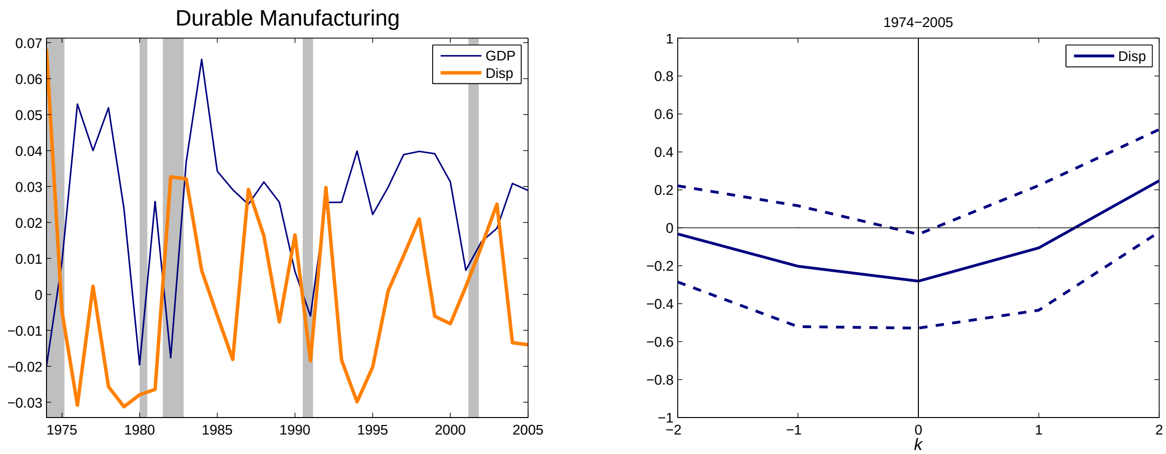

Figure 13: Productivity Dispersion of TFP levels

Plant-level productivity inferred by subtracting cost-share weighted inputs from output. Cross-sectional dispersion of plant-level productivity constructed as described in Section 2.2.1. Left panel displays time series, right panel correlogram with GDP growth: $Corr(\hat{Y}_t, Var(z_{it+k}))$

B.3.2 Solow residuals – the behviour of cross-sectional quantiles

As shown above, the productivity distribution as a whole is negatively correlated with GDP. Like in the main empirical section, I want to analyse the dynamics of different quantiles in the productivity distribution. Figure 14 displays the correlations of these quantiles with GDP. Again, it appears the lower quantiles in the distribution are positively correlated with the business cycle, while the upper quantiles are mostly acyclical. The 75th percentile looks moderately procyclical, but the estimated correlation coefficient is no significantly different from 0. The overall picture that emerges is that not only the dispersion of technology growth rates is countercyclical, but also the dispersion of levels. This dispersion in TFP levels, in turn, appears to be driven mostly by the lower quantiles in the distribution. In a recession, the productivity of some firms maybe fall, but they nevertheless manage to survive in a boom. This is a surprising result and is at variance with the cleansing effect of recessions literature. Therefore I conclude that, that some other driving force must be active in reality that counteracts the driving forces present in the cleansing effect models.

As the correlograms show, the lower quantiles are procyclical, i.e. in a recession the *decrease*.

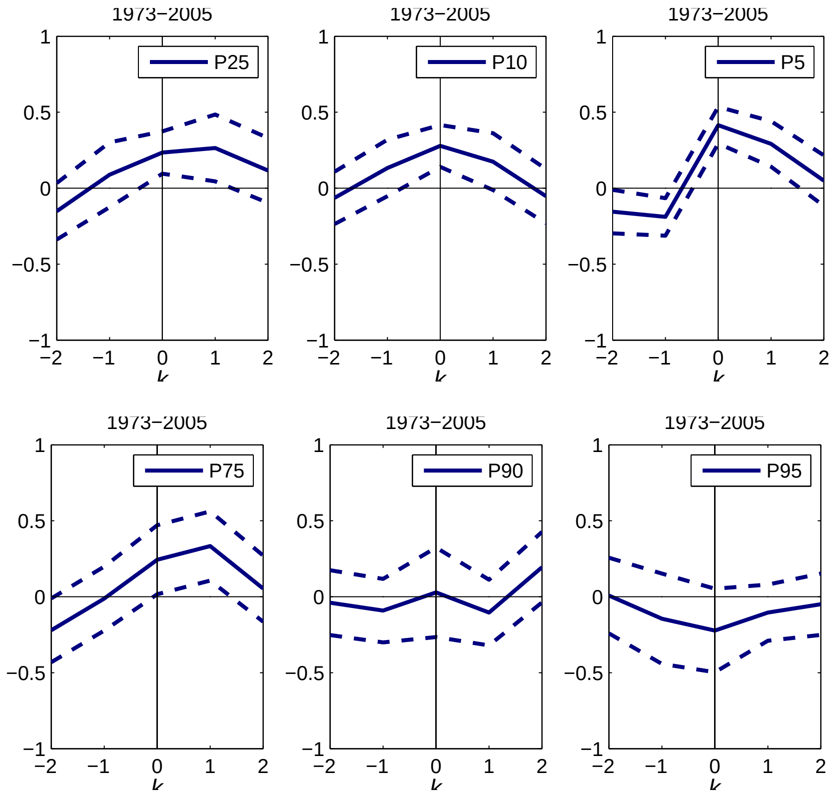

Figure 14: $Corr(GDP_t, P_X(z_{it+k}))$

This seems evidence for truncation at the lower end of the productivity distribution. It is surprising, that this cutoff appears to be higher in a boom.

B.3.3 The cross-sectional dispersion of Solow residual growth rates

Figure 15 displays the cross-sectional variance of TFP growth rates in the baseline specification. As is clear from the visual inspection of Figure 15, the dispersion looks distinctly countercyclial. The estimated correlation coefficient between the cross-sectional variance and HP-filtered GDP is -0.38 with a standard error[48] of 0.11, which means the correlation is significantly different from 0.

Do outliers drive the countercyclicality? Could this countercyclicality be driven by a few extreme outliers while the rest of the TFP distribution remains unchanged over the business cycle? Figure 16 displays the correlation of GDP and several dispersion measures: the variance, the inter-quartile range and the inter-decile range. All three measures are countercyclical which reject the notion that jut some outliers in the tail of the distribution are driving the result. This also suggests that the entire distribution in fanning out as GDP decreases.

[48]For a construction of error bands see comments of Figure 16.

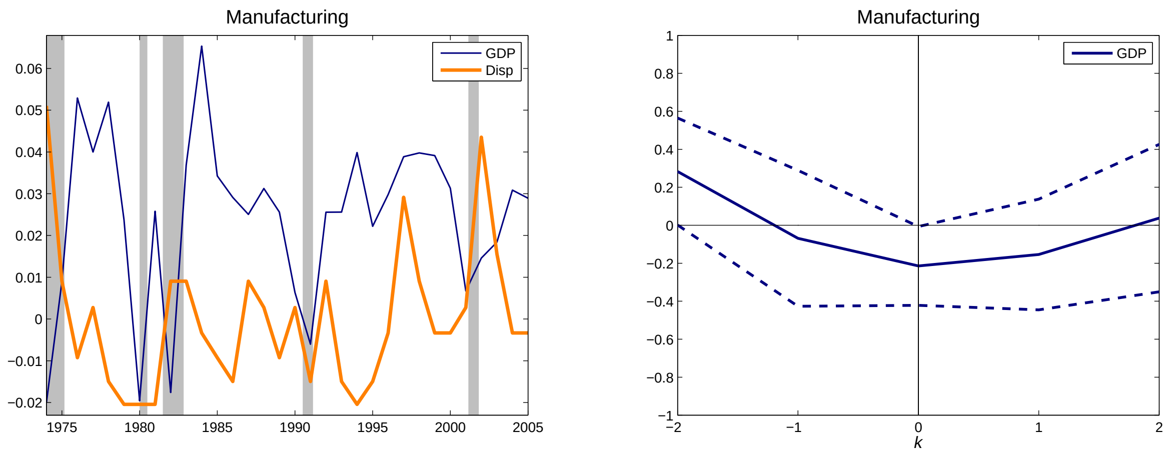

Figure 15: Productivity Dispersion of TFP growth

Plant-level productivity inferred by subtracting cost-share weighted inputs from output. Cross-sectional dispersion of plant-level productivity growth is an variance of unweighted observations. Left panel displays time series, right panel correlogram with GDP growth: $Corr(\hat{Y}_t, Var(\hat{z}_{it+k}))$

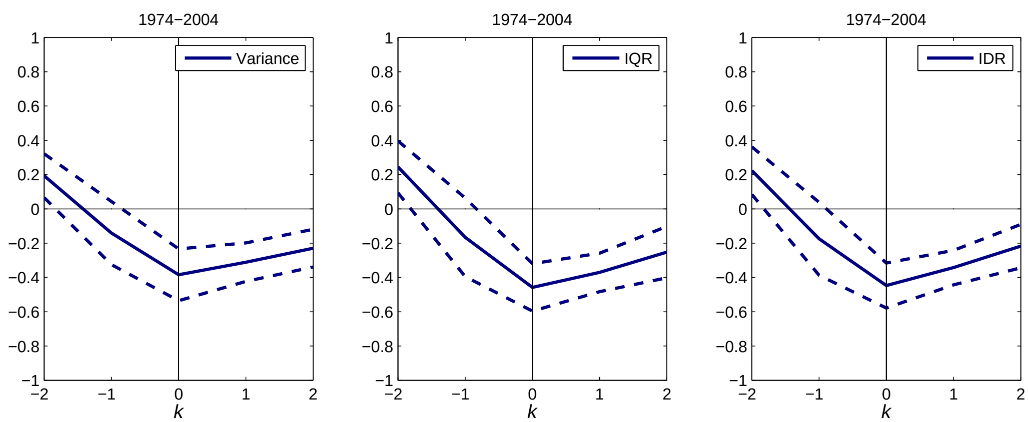

Figure 16: $Corr(GDP_t, Disp_{t+k}(\widehat{z}_{it}))$ – dashed lines 95% error bands

C Detailed solution to the model

First-order conditions

$$
\begin{aligned}
P_t^n \equiv 1 &= \frac{\sigma}{\sigma-1}\frac{w_t}{\overline{z}^n}N_t^{\frac{1}{1-\sigma}} && \text{(Pricing ND Sector)}\\
P_t^d &= \frac{\varrho}{\varrho-1}\frac{w_t}{\overline{z}_t^d}(N_t^d)^{\frac{1}{1-\varrho}} && \text{(Pricing D Sector)}\\
\pi_t &= \pi_t^n + \frac{N_t^d}{N_t}\pi_t^d && \text{(Avge profits)}\\
\pi_t^n(\overline{z}^n) &= \frac{1}{\sigma}\left(\frac{\sigma-1}{\sigma}\frac{\overline{z}^n}{w_t}\right)^{\sigma-1} C_t && \text{(Avge profits in } n\text{)}\\
\pi_t^d(\overline{z}_t^d) &= w_t c_f \frac{\varrho-1}{k+1-\varrho} && \text{(Avge profits in } d\text{)}\\
L_t &= N_t l_t^n + N_t^d l_t^d + N_t^E c_e && \text{(Aggr labour demand)}\\
l_t^n(\overline{z}^n) &= \left(\frac{\sigma-1}{\sigma}\frac{1}{w_t}\right)^{\sigma}(\overline{z}^n)^{\sigma-1} C_t && \text{(labour demand } n\text{)}\\
l_t^d(\overline{z}_t^d, c_f) &= c_f\left[\frac{(\varrho-1)k}{k+1-\varrho}+1\right] && \text{(labour demand } d\text{)}\\
\frac{N_t^d}{N_t} &= \left(\frac{z_L}{\overline{z}_t^d}\right)^k\left(\frac{k}{k+1-\varrho}\right)^{\frac{k}{\varrho-1}} && \text{(Share of firms in D sector)}\\
N_{t+1} &= (1-\zeta)(N_t+N_t^E). && \text{(Dynamics of no. firms)}\\
v_t &= c_e w_t && \text{(Free entry)}\\
\frac{\alpha}{C_t}X_t^{1-\theta} &= \lambda_t(1+\tau_t^c) && \text{(HH FOC: Cons ND)}\\
\lambda_t P_t^d(1+\tau_t^I) &= \beta\left[\frac{\gamma}{D_{t+1}}X_t^{1-\theta}+\lambda_{t+1}P_{t+1}^d(1+\tau_{t+1}^I)(1-\delta)\right] && \text{(HH FOC: Inv durables)}\\
\frac{\psi\phi_t}{1-\phi_t L_t}X_t^{1-\theta} &= \lambda_t w_t && \text{(HH FOC: Intratemp Euler)}\\
\lambda_t v_t &= \beta\mathbb{E}_t\left[(1-\zeta)\lambda_{t+1}(v_{t+1}+\pi_{t+1})\right] && \text{(HH FOC: Intertemp Euler Equity)}\\
w_t L_t + \pi_t N_t &= C_t + P_t^d\left[D_{t+1}-(1-\delta)D_t\right]+v_t N_t^E && \text{(Budget constraint)}
\end{aligned}
$$

Endogenous variables:

$$\overline{z}_t^d, P_t^d, \pi_t, \pi_t^n, \pi_t^d, N_{t+1}, N_t^E, N_t^d, v_t, C_t, D_{t+1}, \lambda_t, L_t, l_t^n, l_t^d, w_t$$

Log-linearised

$$\hat{w}_t = \frac{1}{\sigma - 1}\hat{N}_t \tag{46}$$

$$\hat{P}^d_t - \hat{w}_t + \hat{\bar{z}}^d_t + \frac{1}{\varrho - 1}\hat{N}^d_t = 0 \tag{47}$$

$$-\hat{\pi}_t + \hat{\pi}^n_t\left(1 - \frac{\overline{\pi}^d\overline{N}^d}{\overline{N}\overline{\pi}}\right) + \left[\hat{N}^d_t + \hat{\pi}^d_t\right]\frac{\overline{\pi}^d\overline{N}^d}{\overline{N}\overline{\pi}} = \hat{N}_t\frac{\overline{\pi}^d\overline{N}^d}{\overline{N}\overline{\pi}} \tag{48}$$

$$\hat{\pi}^n_t - (1-\sigma)\hat{w}_t - \hat{C}_t = 0 \tag{49}$$

$$\hat{\pi}^d_t - \hat{w}_t = 0 \tag{50}$$

$$\hat{l}^n_t\frac{\overline{Nl}^n}{\overline{L}} + \hat{N}^d_t\left(1 - \frac{\overline{Nl}^n}{\overline{L}}\right) + \hat{N}^E_t\frac{c_e\overline{N}^E}{\overline{L}} - \hat{L}_t = -\hat{N}_t\frac{\overline{Nl}^n}{\overline{L}} \tag{51}$$

$$\hat{l}^n_t + \sigma\hat{w}_t - \hat{C}_t = 0 \tag{52}$$

$$\hat{l}^d_t = 0 \tag{53}$$

$$\hat{N}^d_t + k\hat{\bar{z}}^d_t = \hat{N}_t \tag{54}$$

$$\hat{N}_{t+1} - \delta\hat{N}^E_t = (1-\delta)\hat{N}_t \tag{55}$$

$$\hat{v}_t - \hat{w}_t = 0 \tag{56}$$

$$\hat{\xi}_t + (1-\theta)\hat{X}_t - \hat{C}_t - \hat{\lambda}_t - \hat{\tau}^c_t\frac{\tau^c}{1+\tau^c} = 0 \tag{57}$$

$$\begin{aligned}&\left[\rho^\xi\hat{\xi}_t + (1-\theta)\hat{X}_t - \hat{D}_{t+1}\right][1-\beta(1-\delta)] + ...\\ &...\left[\hat{\lambda}_{t+1} + \hat{P}^d_{t+1} + \hat{\tau}_{t+1}\frac{\tau^i}{1+\tau^i}\right]\beta(1-\delta) - \hat{\lambda}_t - (1-\rho^b)\hat{b}_t - \hat{P}^d_t - \hat{\tau}^i_t\frac{\tau^i}{1+\tau^i} = 0\end{aligned} \tag{58}$$

$$\hat{\phi}_t\frac{1}{1-\overline{\phi L}} + (1-\theta)\hat{X}_t + \hat{\xi}_t + \hat{L}_t\frac{\phi\overline{L}}{1-\overline{\phi L}} - \hat{\lambda}_t - \hat{w}_t = 0 \tag{59}$$

$$(1-\rho^b)\hat{b}_t + \hat{\lambda}_t + \hat{v}_t - \hat{\lambda}_{t+1} - \hat{v}_{t+1}\frac{\overline{v}}{\overline{v}+\overline{\pi}} - \hat{\pi}_{t+1}\frac{\overline{\pi}}{\overline{v}+\overline{\pi}} = 0 \tag{60}$$

$$\begin{aligned}&-\hat{w}_t\frac{\overline{wL}}{\overline{wL}+\overline{N\pi}} - \hat{L}_t\frac{\overline{wL}}{\overline{wL}+\overline{N\pi}} - \hat{\pi}_t\frac{\overline{N\pi}}{\overline{wl}+\overline{N\pi}} + \hat{C}_t\frac{\overline{C}}{\overline{wl}+\overline{N\pi}} + \hat{P}^d_t\frac{\delta\overline{DP}^d}{\overline{wl}+\overline{N\pi}}\\ &+\hat{D}_{t+1}\frac{\overline{DP}^d}{\overline{wl}+\overline{N\pi}} + \hat{N}^E_t\frac{\overline{\delta vN}}{(1-\delta)(\overline{wl}+\overline{N\pi})} + \hat{v}_t\frac{\overline{\delta vN}}{(1-\delta)(\overline{wl}+\overline{N\pi})} = \hat{N}_t\frac{\overline{N\pi}}{\overline{wl}+\overline{N\pi}} + \hat{D}_t\frac{(1-\delta)\overline{DP}^d}{\overline{wl}+\overline{N\pi}}\end{aligned}$$