Online Appendix for:

"Investment Dispersion and the Business Cycle"

Rüdiger Bachmann and Christian Bayer

A Data

A.1 Description of the Database

Our German firm-level data source is USTAN (*Unternehmensbilanzstatistik*) of Deutsche Bundesbank. It provides annual firm-level data from 1971 to 1998 from the balance sheets and the profit and loss accounts of over 60,000 firms per year. USTAN captures all major balance-sheet items, the major items of the profit and loss statements, and employment. Importantly, USTAN provides separate investment data for structures and equipment. As we will show below, the USTAN sample covers a large fraction of the nonfinancial private business sector (NFPBS).

It originated as a by-product of the Bundesbank's rediscounting, i.e. (overnight-)lending activities. By law, the Bundesbank was required to assess the creditworthiness of all parties backing a *Wechsel*, a promissory note or commercial bill of exchange, put up for discounting. It implemented this regulation by requiring balance-sheet data of all parties involved. These balance-sheet data were then collected and archived into a database.

Promissory notes were a form of trade credit with widespread use throughout the sample period. From the volume of a 0.15% stamp tax on promissory notes and bills of exchange, one can infer that a volume of these titles of roughly 10% of German GDP was issued each year. Moreover, rediscounting promissory notes was a commonly used instrument of monetary policy in Germany. Thus, unlike the Federal Reserve, the Bundesbank did not use T-bills as the major form of collateral but rather private debt. As far as potential cyclical sample selection is concerned, it is important to note that it had to happen only once in a given year that a promissory note from a given firm was used as collateral by someone in order for that firm to appear in USTAN, i.e. it is irrelevant how often that firm issued trade credit and in what volumes.

The quality of the data is particularly high. All mandatory data collected for USTAN have been double-checked by Bundesbank staff. The Bundesbank itself frequently uses the USTAN data for its macroeconomic analyses and for cross-checking national accounting data. We take this as an indication that the bank considers the data as sufficiently representative and of high quality.

One drawback of USTAN is that with the introduction of the euro, the Bundesbank stopped buying commercial bills and collected firm balance-sheet data only irregularly and only from publicly available sources. For this reason, the data set stops being useful in 1999. Therefore, we only use data from 1971 to 1998, which leaves us, after lagging and first-differencing, with 26 years of observations from 1973 to 1998.

The coverage of the sample is broad, although it is technically not a representative sample due to the nonrandom sample design. It was also more common to use promissory notes as trade credit in manufacturing and for incorporated companies, which biases our data somewhat toward these kinds of firms. And, of course, the Bundesbank would only rediscount notes to which it gave a good rating, so that the set of firms in USTAN is also somewhat biased toward financially healthy and larger firms. Nevertheless, USTAN covers a wide range of firms, in fact a wider range in some dimensions (size, ownership, industry) than comparable U.S. data sets (ASM, COMPUSTAT), since short-term financing through promissory notes was a common practice for many German companies across most business sectors. Due to the Bundesbank's rediscounting policy, bills of exchange were very liquid for the creditor.

A.2 Sample Selection

We start with the universe of observations in the USTAN data, merging the files for 1971-1986 and 1987-1998. In a first pass, we then drop all balance sheets that are irregular, e.g., bankruptcy or closing balance sheets, or that stem from a group/holding (*Konzernbilanz*). This leaves us with only regular balance sheets (*Handelsbilanz* or *Steuerbilanz*). We also drop all firms with missing payroll data or missing or negative sales data, which are basically nonoperating firms. A small amount of duplicate balance sheets is removed as well. Finally, we drop the following sectors: hospitality (hotels and restaurants), which only has a small number of firms in the database, financial and insurance institutions, the mostly public health and education sectors, as well as other public companies like museums, etc., and some other small service industries, such as hair cutters, dry cleaners and funeral homes;[28] or when sectoral information was missing. The sectoral aggregate we are studying can be roughly characterized as the nonfinancial private business sector in Germany. This sample selection leaves us with an initial data set of 1,764,846 firm-year observations and 259,614 different firms. The average number of firms per year is 63,030.

[28]The number of firms from the public sector and these small industries is tiny to begin with, as they did not regularly use bills of exchange as a financing instrument.

From this initial sample we remove step-by-step observations, in order to get an economically meaningful data set. We first drop observations from likely East German firms to avoid a break in the series in 1990. We identify a West German firm as a firm that has a West German address or has no address information but enters the sample before 1990. Then we recompute capital stocks with a perpetual inventory method (PIM). In the PIM we drop a small amount of outliers. We remove observations that do not have a log value-added and a log capital stock after PIM.

Another part of the data is removed when firms did not have changes in log firm-level employment (N), capital (K) and real value-added (VA), which obviously requires us to observe firms for two consecutive years. Then we remove outliers in factor changes and real value-added changes. Specifically, we identify as outliers in our sample a firm-year in which the firm-level investment rate or log changes in firm-level real value-added, employment and capital stock fall outside a three-standard-deviations band around the firm and sectoral-year mean. Then we compute firm-level Solow residuals and similarly remove observations with missing log changes in Solow residuals as well as outliers therein. We finally remove – before and after each step of the outlier removal – firms that have less than five observations in firm-level Solow residual changes. We conduct extensive robustness checks of our results to the choices for the outlier and observation thresholds. Table 14 summarizes how many observations are dropped in each step.

Table 14: Sample Creation

Criterion	Firm-Year Observations
Initial Sample	1,764,846
East Germany	-104,299
Outliers in PIM	-7,539
Missing log value-added	-1,349
Missing log capital	-31,819
Missing log-changes in N, K, VA	-161,668
Outliers in N, K and VA log-changes	-41,453
Missing log-changes in Solow residual	-126,086
Outliers in Solow residual log-changes	-18,978
Not enough observations	-417,550
Final Sample	854,105

A.3 Sample Composition

The final sample then consists of 854,105 firm-year observations, which amounts to observations on 72,853 different firms. The average observation length of a firm in the sample is 11.7 years. The average number of firms per year is 32,850. The following Tables 15, 16 and 17 show the average industry,[29] the legal form and the size distributions in our final sample.

USTAN's industry coverage, while somewhat biased toward manufacturing firms, includes the construction, service and the primary sectors. While a bias toward larger firms remains, the size coverage is still fairly broad: 31% of all firm-year observations in our final baseline sample have fewer than 20 employees and 57% have fewer than 50 employees. In terms of ownership structure, only 2% of firm-year observations are from publicly traded firms, just under 60% from limited liability companies and just under 40% from private firms with fully liable partners.

Table 15: Two-Digit Industry Distribution

ID	Sector	Observations	Frequency	WZ 2003
10	Agriculture	12,291	1.44%	A, B
20	Energy & Mining	4,165	0.49%	C, E
31	Chemical Industry, Oil	14,721	1.72%	DF, DG
32	Plastics, Rubber	23,892	2.80%	DH
33	Glass, Ceramics	28,623	3.35%	DI
34	Metals	30,591	3.58%	DJ
35	Machinery	162,407	19.01%	DK, DL, DM, DN
36	Wood, Paper, Printing	61,672	7.22%	DD, DE
37	Textiles, Leather	46,173	5.41%	DB, DC
38	Food, Tobacco	37,708	4.41%	DA
40	Construction	54,569	6.39%	F
61	Wholesale Trade	213,071	24.95%	G51
62	Retail Trade & Cars	142,137	16.64%	G50, G51
70	Transportation & Communication	22,085	2.59%	I
	Total	854,105	100%	

[29] WZ 2003 is the industry classification from 2003 that the German national accounting system (*Volkswirtschaftliche Gesamtrechnung, VGR*) uses.

Table 16: Legal Form Distribution

Legal Form	Observations	Frequency
Publicly Traded (AG, KGaA, etc.)	18,582	2.18%
Limited Liability Companies (GmbH, GmbH&Co., etc.)	506,184	59.26%
Fully Liable Partnerships (OHG, KG, etc.)	327,526	38.35%
Other: unincorporated associations (e.V.) municipal agencies (Körperschaften öR) etc.	1,813	0.21%
Total	854,105	100%

Table 17: Size Distributions of Firms

Number of									
Employees	1-4	5-9	10-14	15-19	20-49	50-99	100-249	250-499	500+
Fraction	6.14%	9.46%	8.24%	7.30%	26.28%	17.04%	14.37%	5.68%	5.49%
Capital Stock									
(in 1000 1991-Euro)	0-299	300-599	600-999	1,000-1,499	1,500-2,499	2,500-4,999	5,000-9,999	10,000-24,999	25,000+
Fraction	8.23%	9.01%	9.67%	9.36%	13.08%	17.71%	13.87%	11.08%	7.99%
Real Value Added									
(in 1000 1991-Euro)	0-299	300-499	500-999	1,000-1,499	1,500-2,499	2,500-4,999	5,000-9,999	10,000-24,999	25,000+
Fraction	8.17%	7.93%	16.38%	11.56%	14.45%	16.28%	11.20%	8.25%	5.79%

A.4 Perpetual Inventory Method

In order to obtain economically meaningful stocks of capital series for each firm, we have to re-calculate capital stocks in a Perpetual Inventory Method (PIM); see Bayer (2006), for instance. The first step is to compute firm-level investment series, $I_{j,t}$, from the corporate balance sheets, which contain data only on accounting capital stocks, $k^a_{j,t}$, and accounting total depreciation, $d^a_{j,t}$. The following accumulation identity for the book value of capital allows us to back out nominal firm-level investment, $p^I_t I_{j,t}$:[30]

$$k^a_{j,t+1} = k^a_{j,t} - d^a_{j,t} + p^I_t I_{j,t}. \tag{1}$$

The next step is to recognize that capital stocks from corporate balance sheets are not directly usable for economic analysis for two reasons: 1) accounting depreciation, $d^a_{j,t}$, in corporate balance sheets is often motivated by tax reasons and is typically higher than economic depreciation, $\delta^e_{j,t}$, expressed as a rate; 2) accounting capital stocks are reported at historical prices. Both effects would lead to an underestimation of the real firm-level capital stock, if one were to simply deflate the current accounting capital stock, $k^a_{j,t}$, with a current investment price deflator, p^I_t (assuming that p^I_t increases over time). We therefore apply a Perpetual Inventory Method (PIM) to compute economic real capital stocks:

$$k^{(1)}_{j,1} = k^a_{j,1}. \tag{2}$$

$$k^{(1)}_{j,t+1} = (1-\delta^e_t)\, k^{(1)}_{j,t} + \frac{p^I_t}{p^I_{1991,t}} I_{j,t}. \tag{3}$$

$k^a_{j,1}$ is the accounting capital stock in 1991 prices at the beginning of an uninterrupted sequence of firm observations – if for a firm-year we have a missing investment observation, the PIM is started anew when the firm appears again in the data set. The investment-good-price deflator is $p^I_{1991,t}$, with 1991 as the base year. We estimate the economic depreciation rate δ^e_t for each year from national accounting data, *VGR*, separately for equipment and nonresidential structures (Table 3.1.3, *VGR*, *Nettoanlagevermögen nach Vermögensarten in jeweiligen Preisen, Ausrüstungen und Nichtwohnbauten*; Table 3.1.4, *VGR*, *Abschreibungen nach Vermögensarten in jeweiligen Preisen, Ausrüstungen und Nichtwohnbauten*). *VGR* contains sectoral and capital-good-specific depreciation data only after 1991, which is why

[30] Specifically, $k^a_{j,t}$ is the sum of balance-sheet items ap65, *Technische Anlagen und Maschinen*, and ap66, *Andere Anlagen, Betriebs-und Geschäftsausstattung*, for equipment; and balance-sheet item ap64, *Grundstücke, Bauten*, for structures. Since balance-sheet data are typically end-of-year stock data, notice that $k^a_{j,t}$ is the end-of-period capital stock in year $t-1$. $d^a_{j,t}$ is profit and loss account item ap156, *Abschreibungen auf Sachanlagen und immaterielle Vermögensgegenstände des Anlagevermögens*. In contrast to $k^a_{j,t}$, $d^a_{j,t}$ is not given for each capital good separately. For the solution of this complication, see below.

we decided to use only capital-good-specific depreciation rates for the entire time horizon. For the data sources for investment price deflators, see footnote 33 below. The drawback to this procedure is that we do not directly observe capital-good specific $d^a_{j,t}$ in the balance sheets, so that (1) is not directly applicable to the two types of capital good separately. We therefore split up $d^a_{j,t}$ according to the fraction that each type of capital good accounts for in the book value of total capital, weighting each type of capital good by its VGR depreciation rate. We finally aggregate both types of capital into a single capital good at the firm level.

There is a final complication, which arises through relying on $k^a_{j,1}$ as the starting value of the PIM. The recalculation of capital stocks is motivated by the bias that historical cost accounting and tax depreciation induce, i.e., that the book value of capital is typically not a good estimate of the productive real capital stock of the firm at that time. To take this issue into account also for the first observation of a firm, we calculate the time-average factor ϕ (for each sector), by which $k^{(1)}_{j,t}$ is larger than $k^a_{j,t}$, and replace $k^a_{j,1}$ by $\phi k^a_{j,1}$ in the perpetual inventory method. We do this iteratively until ϕ converges, i.e., we calculate (using $k^{(0)}_{j,t} = k^a_{j,t}$ and $\phi^{(0)} = 1$):

$$k^{(n)}_{j,t+1} = (1 - \delta^e_t)\, k^{(n)}_{j,t} + \frac{p^I_t}{p^I_{1991,t}} I_{j,t} \tag{4}$$

$$k^{(n)}_{j,1} = \phi^{(n-1)} k^{(n-1)}_{j,1} \tag{5}$$

$$\phi^{(n)} = (NT)^{-1} \sum_{j,t} \frac{k^{(n)}_{j,t}}{k^{(n-1)}_{j,t}} \tag{6}$$

We stop when for each sector and each capital good category $\phi < 1.1$.[31]

Since we want to compute economic, i.e. productive, capital stocks, we then – as a final step – add to the capital stock series from this iterative PIM the net present value of the real expenditures for renting and leasing equipment and structures.[32]

[31] Extreme ϕ's indicate that constant depreciation is not a good approximation for this particular firm. Such a firm will have had an episode of extraordinary depreciation (e.g., fire, accident, etc.) and the capital stocks by PIM will be a bad measure of the actual capital stock after the accident. That is why we drop a small number of observations from the top and bottom of the ϕ-distribution (see Table 14).

[32] We take item ap161, *Miet- und Pachtaufwendungen*, from the profit and loss accounts, deflate it by the implicit investment good price deflator, which we compute from Tables 3.2.8.1 and 3.2.9.1 from *VGR*, and then divide it by a measure of the user cost of capital. The latter is simply the sum of real interest rates for a given year, which we compute from nominal interest rates on corporate bonds and ex-post CPI inflation data, and the time-average, accounting capital-good-weighted depreciation rate per firm.

A.5 Representativeness

How well does the USTAN aggregate represent the nonfinancial private business sector (NFPBS) in Germany? Table 18 shows that USTAN represents on average 70% of the value-added of the NFPBS, 44% of its investment, etc.[33]

Table 18: USTAN and the NFPBS

	USTAN/NFPBS
Value-Added	70%
Investment	44%
Capital	71%
Employment	49%
Payroll	54%

Table 19 shows that the cross-sectional *averages* of investment as well as output, employment and productivity growth, computed from USTAN, are strongly positively correlated with the cyclical component of the real gross value-added of the nonfinancial private business sector. This means that USTAN represents well the cyclical behavior of the sectoral aggregate it is meant to represent.

[33] NFPBS value-added is taken from *Bruttowertschöpfung in jeweiligen Preisen*, Table 3.2.1 of *VGR*, deflated by the implicit deflator for aggregate value-added, Table 3.1.1 of *VGR* (we apply the same deflator to the USTAN data). The base year is always 1991. NFPBS investment is *Bruttoanlageinvestitionen in jeweiligen Preisen* from Table 3.2.8.1, deflated with the implicit sector-specific investment price deflators given by *Bruttoanlageinvestitionen - preisbereinigt*, a chain index, from Table 3.2.9.1, *VGR*. NFPBS capital is *Nettoanlagevermögen in Preisen von 2000* from Table 3.2.19.1, *VGR*, re-chained to 1991 prices. In computing both the investment and the capital data for USTAN in the PIM, we use the implicit sector and capital-good-specific (equipment and nonresidential structures) deflators for investment: Tables 3.2.8.2, 3.2.9.2., 3.2.8.3 and 3.2.9.3., *VGR*. We also experiment with deflating USTAN data with a uniform investment price deflator, the *Preisindex der Investitionsgüterproduzenten*, source: GP-X002, *Statistisches Bundesamt*. NFPBS employment is the number of employed, *Arbeitnehmer*, from Table 3.2.13, *VGR*. Finally, payroll is taken from *Arbeitnehmerentgelt*, Table 3.2.10., *VGR*, deflated by the same general implicit deflator for aggregate value-added that we use to deflate value-added numbers.

Table 19: Cyclicality of Cross-Sectional Averages

Cross-sectional Moment	Correlation with Cycle
$mean(i_{j,t})$	0.756***
$mean(\Delta \log y_{j,t})$	0.663***
$mean(\frac{\Delta n_{j,t}}{0.5*(n_{j,t-1}+n_{j,t})})$	0.602***
$mean(\Delta \log \epsilon_{j,t})$	0.592***

Notes: The table shows the correlation with the cycle of the cross-sectional averages, linearly detrended, of, respectively, the investment rate, the log-change of real gross value-added (we deflate the profit and loss account item ap153, Rohergebnis, with the aggregate value-added deflator from VGR data), the net employment change rate, and the log-change of Solow residuals, all at the firm level. We have removed firm fixed and 2-digit industry-year effects from each variable. As a cyclical indicator we use the HP(100)-filtered aggregate real gross value-added in the German nonfinancial private business sector, computed from German VGR (*Volkswirtschaftliche Gesamtrechnungen*) data. *** indicates significance at the 1%level, resulting from an overlapping block bootstrap of four-year windows with 10,000 replications.

A.6 The Cross-sectional Investment Rate Distribution

As Table 20 shows, the distribution of firm-level investment rates from our USTAN sample is comparable to the one calculated for the U.S. from the LRD, reported in Cooper and Haltiwanger (2006). For comparability with Cooper and Haltiwanger (2006), we also show the distribution of investment rates for $\frac{I_{j,t}}{k_{j,t}}$. The USTAN sample exhibits less disinvestment activity, slightly more inactivity, and a little less frequent spikes. There are various reasons for these differences. First, LRD is plant-level data, whereas our data is firm level, so some difference may come from unit aggregation, in particular, observing fewer spikes in USTAN. Second, the U.S. and German manufacturing sector were facing fairly different compositional trends over the respective sample periods, which may explain why more negative investment due to stronger reallocation is observed in the LRD.

Table 20: Investment Rate Distributions in USTAN

	Negative		Inactivity	Positive	
	spike	intermediate		intermediate	spike
$i_{j,t}$					
All firms (USTAN)	0.4%	2.6%	14.8%	68.9%	13.4%
Manufacturing (USTAN)	0.4%	2.0%	11.1%	74.7%	11.9%
$\frac{I_{j,t}}{k_{j,t}}$					
All firms (USTAN)	0.3%	2.6%	15.1%	67.7%	14.2%
Manufacturing (USTAN)	0.3%	2.0%	11.4%	73.6%	12.7%
LRD	1.8%	8.6%	8.1%	62.9%	18.6 %

Notes: $i_{j,t} = \frac{I_{j,t}}{0.5(k_{j,t}+k_{j,t+1})}$ denotes our baseline investment rate definition. $\frac{I_{j,t}}{k_{j,t}}$ denotes the definition of the investment rate used by Cooper and Haltiwanger (2006). An investment spike is defined, for either investment rate, as being larger than 20% in absolute value. Inactivity is defined as an investment rate that is smaller than 1% in absolute value. The LRD data are taken from Cooper and Haltiwanger (2006).

A.7 Estimating the Production Function

We estimate the coefficients θ, ν of the production function by the median of the firm average share of factor expenditure in total value-added, as defined by:[34]

$$\hat{\nu}_j = T_j^{-1} \sum_t \frac{w_{j,t} n_{j,t}}{y_{j,t}}$$

$$\hat{\theta}_j = T_j^{-1} \sum_t \frac{(r_t + \delta_j) k_{j,t}}{y_{j,t}}$$

In a frictionless setup, this is estimating the production function coefficients from the first-order conditions. Importantly, this estimator is robust to classical measurement error in capital and labor. We take as the real interest rate, r_t, the average return on corporate bonds minus the ex post inflation rate and calculate firm-specific depreciation rates, δ_j, from capital-good-specific *VGR* depreciation rates, weighted by the firm-specific capital good portfolio.

[34] We use profit and loss account item ap153, Rohergebnis, for firm-level value-added and profit and loss account item ap154, Personalaufwand, for the firm-level wage bill.

Under the null hypothesis of our model, i.e., nonconvex capital adjustment frictions, these first-order conditions do not hold exactly for capital. However, when we estimate the production function coefficients from data simulated from our model in data sets of comparable size to USTAN, we find that this simple estimation procedure from factor expenditure shares works remarkably well. In fact, the labor share is estimated exactly, as labor can be adjusted without frictions in the model. The capital share estimated from the model-simulated data is $\hat{\theta} = 0.2238$, with the true θ being 0.2075.[35] Alternative estimation approaches, such as those advocated by Olley and Pakes (1996) or Levinsohn and Petrin (2003), suffer from the collinearity issues discussed in Ackerberg, Caves and Frazer (2006) and Gandhi, Navarro and Rivers (2011). In fact, when we estimate the production function parameters from model-simulated data using Olley and Pakes' (1996) estimator, we obtain a greatly upwardly biased estimate for ν: $\hat{\nu}_{OP} = 0.9962$. The estimated coefficient for capital is virtually zero. Also for the USTAN data we obtain fairly high coefficients on labor and low coefficients on capital (see also Burnside, Eichenbaum and Rebelo (1995) for similar findings in U.S. data). For manufacturing, for instance, we obtained Olley and Pakes estimates of $\nu = .744$ and $\theta = .069$, which is another indication that our model is a good description of firm-level behavior in the USTAN data set. The Olley and Pakes estimates are essentially unaffected by the inclusion (or lack thereof) of a selection term (cf. Appendix B) or the order of approximation used in the third stage of the estimator. In the first stage, we use only observations where the absolute value of the investment rate exceeds 20%.[36]

In summary, under the null hypothesis of our model, the factor shares are good estimators of the production function parameters despite the capital adjustment frictions.

A.8 The Idiosyncratic Productivity Process

We estimate the log-productivity residuals implied by the estimates for the production function coefficients as:[37]

$$\hat{\epsilon}_{j,t} = \log(y_{j,t}) - \hat{z}_t - \hat{\nu} \log n_{j,t} - \hat{\theta} \log k_{j,t}.$$

[35] We thus find a small upward bias for the capital coefficient θ. Yet, an exact calibration alongside the adjustment cost parameter using our model would be prohibitively time-consuming and, given the negligible size of the bias, would not change our results substantively.

[36] Using a 10% threshold yields basically the same results.

[37] We also check whether unobserved utilization biases the properties of the idiosyncratic productivity risk process after controlling for measurement error. Specifically, we regress the growth rates of the firm-level Solow-residual on, respectively, the change in the material intensity / on the change in the log material expenditures (at 1991 prices), and take, respectively, the residuals. By construction, the residuals are orthogonal to the material intensity growth / to the growth in material expenditures, hence orthogonal to utilization if utilization impacts material intensity / material expenditures. Both the average level of idiosyncratic risk as well as its cyclicality and time series volatility are essentially unchanged compared to the baseline case.

We then specify for $\hat{\epsilon}_{j,t}$

$$\hat{\epsilon}_{j,t} = \epsilon_{j,t} + x_{j,t} + \mu_j$$

$$\epsilon_{j,t} = \rho\epsilon_{j,t-1} + u_{j,t},$$

where $x_{j,t}$ denotes measurement error. In order to estimate ρ we allow the measurement error to be a second- or third-order moving average process. This yields an estimation equation

$$\hat{\epsilon}_{j,t} = \rho\hat{\epsilon}_{j,t-1} + \sum_{s=1\ldots J} \xi_s \Delta\hat{\epsilon}_{j,t-s} + (1-\rho)\mu_j + \zeta_{j,t},$$

where we experiment with $J = 3$ and $J = 4$. An unbiased estimate of ρ can be obtained by an instrumental variable regression, where one uses lagged differences of $\Delta\hat{\epsilon}_{j,t-J-1}$ as instruments for $\epsilon_{j,t-1}$. The estimated ρ is 0.960 and 0.974 for $J = 3$ and $J = 4$, respectively. The coefficient used in our calibration is 0.9675 - the midpoint of these two estimates.

When measuring the size and the cyclicality of the dispersions of productivity growth, we apply a slightly simplified approach, given the relatively high values of ρ (and low values of ξ_s) that we estimate. Since a specification with a long moving average term for the measurement error would force us to discard many aggregate data points for generating the necessary lags, we specify $\epsilon_{j,t}$ as a random walk cum classical measurement error, for estimating the variance of the shocks as[38]

$$E\Delta\hat{\epsilon}^2_{j,t} - \sigma^2_{me},$$

where the variance of the measurement error, $\Delta x_{j,t}$, is estimated by the sample analogue to

$$\begin{aligned}
\sigma^2_{me} &= -E(\hat{\epsilon}_{j,t} - \hat{\epsilon}_{j,t-2})^2 + 2E\Delta\hat{\epsilon}^2_{j,t} \\
&= -E(\epsilon_{j,t} - \epsilon_{j,t-2} + x_{j,t} - x_{j,t-2})^2 + 2E(\Delta\epsilon_{j,t} + \Delta x_{j,t})^2 \\
&= -E(\Delta\epsilon_{j,t} + \Delta\epsilon_{j,t-1} + x_{j,t} - x_{j,t-2})^2 + 2E(\Delta\epsilon_{j,t} + \Delta x_{j,t})^2 \\
&= -E(-\sigma_{\Delta\epsilon,t} - \sigma_{\Delta\epsilon,t-1} - 2\sigma_x + 2\sigma_{\Delta\epsilon,t} + 4\sigma_x) = 2\sigma_x.
\end{aligned}$$

[38] We have conducted a Monte Carlo analysis that shows that the mistake one makes when our measurement error estimation is applied to not-quite-unit-root data is small. We therefore prefer the simple estimation procedure for the idiosyncratic shock variance.

B Robustness of the Empirical Findings

This appendix provides the details for the robustness checks discussed in Section I.C. Table 21 shows robustness with respect to the cyclical indicator. We experiment with different HP-filter smoothing parameters, with dynamic correlations, with excluding the post-reunification period and with excluding the years with the most extreme investment dispersion observations.[39]

Table 22 provides robustness checks with respect to sample composition. The first two rows (dropping the first three observations per firm or looking only at firms that are virtually always in the sample) show that our results are neither driven by firm entry (into the sample) nor by firm exit, nor do our results depend on how we remove outliers from the sample.

Further robustness checks for the procyclicality of the investment rate dispersion are available in Table 23. We check for the robustness with respect to replacing the standard deviation with the interquartile range as the measure of dispersion, and to alternative ways of detrending the dispersion series. Perhaps most important, the last row of Table 23 shows that the procyclicality of the investment rate dispersion is not due to cyclical variations in the sample composition. In the scenario 'Selection correction' we control for sample selection in the following way: we estimate a simple selection model, where lagged firm-level Solow residuals determine selection and the firm-level investment rate is modeled as a mean regression. We use the maximum likelihood estimator by Heckman (1976) to infer the selection-corrected variance of the residual in the firm-level investment rate equation. The latter is very close to the sample variance of firm-level investment rates, indicating that our results are not influenced by systematic sample drop outs. While the first stage of the regression shows that there is a positive selection in terms of levels (more productive firms being more likely to be in the sample), there is no strong selection with respect to changes or investment rates and their dispersions.

Finally, Table 24 shows that the ownership structure matters for cross-sectional results (focusing on publicly traded firms in Germany would eliminate the procyclicality of investment dispersion), making it important to use broader data sets for the study of cross-sectional facts (see Davis et al. (2006), for a similar point). Figure 1 plots the time-series of the cross-sectional standard deviation of firm-level investment rates, linearly detrended and detrended with an HP(100)-filter, the time-series of the cross-sectional interquartile range of firm-level investment rates, linearly detrended, and the fraction of spike adjusters, linearly detrended, against the cyclical component of aggregate real gross value-added.

[39] For the sake of brevity, we do not show results for firm-level Solow residual growth and fractions of adjusters. Results for both are robust: for the former, similar to the values for output growth, for the latter, similar to the values for the fraction of spike adjusters.

Table 21: Robustness Checks I – Cyclical Indicator

Cyclical Indicator / Sample	Correlation of ... with					
	$std(i_{j,t})$	$std(\Delta y_{j,t})$	$std(\frac{\Delta n_{j,t}}{0.5*(n_{j,t-1}+n_{j,t})})$	$std(\Delta \epsilon_{j,t})$	Fraction of Spike Adjusters	Std of Invest. rates cond. on spike adj.
Baseline: HP(100)-filtered Real Gross Value-Added, Y	0.45**	−0.45*	−0.50**	−0.47**	0.61***	−0.55***
HP(6.25)-filtered Y	0.37**	−0.47**	−0.53**	−0.47**	0.44***	−0.42***
HP(100)-filtered I	0.72***	−0.30	−0.31	−0.25*	0.78***	−0.33*
2 Lags of HP(100)-filtered Y	−0.01	0.41	0.35	0.51*	−0.01	0.01
1 Lag of HP(100)-filtered Y	0.26	0.12	0.04	0.18	0.36*	−0.25
1 Lead of HP(100)-filtered Y	0.35	−0.56**	−0.54**	−0.68***	0.48**	−0.39**
2 Leads of HP(100)-filtered Y	0.01	−0.35	−0.20	−0.47	0.11	0.08
Pre-reunification: 1973-1990	0.30	−0.65***	−0.63***	−0.68***	0.51**	−0.56***
Exclude min and max $std(i_{j,t})$ years	0.54***	−0.49**	−0.53**	−0.52**	0.66***	−0.60***
Exclude two max $\lvert std(i_{j,t})\rvert$ years	0.50***	−0.58***	−0.57**	−0.56***	0.63***	−0.55***

Notes: This table shows the correlation with the cycle of the cross-sectional standard deviations, linearly detrended, of, respectively, the investment rate, the log-change of real gross value-added, the net employment change rate and the log-change of Solow resiudals, all at the firm level. It also shows, in the last two columns, respectively, the correlation with the cycle of the fraction, linearly detrended, of firms exhibiting an investment spike, and the correlation with the cycle of the standard deviation of investment, linearly detrended, of firms exhibiting an investment spike. All data are from the Bundesbank's USTAN database. We have removed firm fixed and 2-digit industry-year effects from each variable. As a cyclical indicator we use aggregate real gross value-added and aggregate investment in the German nonfinancial private business sector, computed from German VGR (*Volkswirtschaftliche Gesamtrechnungen*) data. ***,**,* indicate significance at the 1%, 5%, and the 10% level, respectively, resulting from an overlapping block bootstrap of four-year windows with 10,000 replications.

Table 22: Robustness Checks II – Sample Composition

	Correlation of ... with HP(100)-filtered Y				
Sample	$std(i_{j,t})$	$std(\Delta y_{j,t})$	$std(\frac{\Delta n_{j,t}}{0.5*(n_{j,t-1}+n_{j,t})})$	$std(\Delta \epsilon_{j,t})$	Fraction of Spike Adjusters
Without entry entry: drop first 3 observations per firm	0.39^{*}	-0.48^{*}	-0.52^{*}	-0.42^{**}	0.61^{***}
Only firms with 20 and more observations.	0.39^{***}	-0.38^{**}	-0.37^{**}	-0.34^{*}	0.77^{***}
Stricter outlier removal (2.5 std.)	0.45^{**}	-0.45^{*}	-0.55^{**}	-0.49^{**}	0.62^{***}
Looser outlier removal (5 std.)	0.42^{**}	-0.44^{*}	-0.22	-0.47^{**}	0.62^{***}
Percentile outlier removal (5%)	0.59^{***}	-0.45^{**}	-0.60^{***}	-0.44^{***}	0.64^{***}
Percentile outlier removal (1%)	0.48^{**}	-0.47^{**}	-0.42^{*}	-0.50^{**}	0.62^{***}

Notes: See notes to Table 21.

Table 23: Robustness Checks III – Other

Correlation of investment dispersion with HP(100)-filtered Y	
Baseline	0.45 **
$iqr(i_{j,t})$	0.57 ***
Raw data - no fixed effects	0.45 ***
Uniform price index for investment	0.43 **
$std(\frac{I_{j,t}}{k_{j,t}})$ (dropping 5% outliers)	0.60 ***
$std(\frac{I_{j,t}}{k_{j,t}})$ (dropping 1% outliers)	0.48 **
$std(i_{j,t})_{\text{quadratic detrending}}$	0.56 ***
$std(i_{j,t})_{\text{cubic detrending}}$	0.60 ***
$std(i_{j,t})_{\text{HP(100)-detrending}}$	0.62 ***
Outlier $\geq$ 3 std means merger	0.42 **
Shorter in sample (2 obs.)	0.44 **
Selection correction	0.38 **

Notes: See notes to Table 21. $iqr(i_{j,t})$ refers to the cross-sectional interquartile range of firm-level investment rates. 'Raw data - no fixed effects' uses the standard deviation of the raw firm-level investment rates, no fixed effects removed. '$std(\frac{I_{j,t}}{k_{j,t}})$ (dropping 5% outliers)' uses $\frac{I_{j,t}}{K_{j,t}}$ the definition of the investment rate used by Cooper and Haltiwanger (2006). To take care of the higher sensitivity to outliers, we use a 5% outlier criterion here. '$std(\frac{I_{j,t}}{k_{j,t}})$ (dropping 1% outliers)' does the same with a 1% outlier criterion. 'Uniform price index for investment' refers to a scenario, in which we deflate firm-level investment and capital with an aggregate price deflator for investment goods, not with one-digit industry- and capital-good-specific ones. The next three rows show the results, when we detrend $std(i_{j,t})$ not with a linear trend, but, respectively, with a quadratic, cubic trend and an HP(100)-filter. 'Outlier $\geq$ 3 std means merger' refers to a scenario, in which we treat an observation of 3 standard deviations above or below the year-specific mean as indicating a merger and mark the firm henceforth as a new one. 'Shorter in sample (2 obs.)' refers to a scenario, in which we require firms to have two observations in first differences (instead of five) to be in the sample. 'Selection correction' refers to a scenario where we estimate a simple selection model, where lagged firm-level Solow residuals determine selection and the firm-level investment rate is modeled as a mean regression. We use the maximum likelihood estimator by Heckman (1976) to infer the selection-corrected variance of the residual in the firm-level investment rate equation.

Table 24: Cyclicality of Cross-Sectional Investment Dispersion – Legal Form

Aggregate	Publicly Traded	Limited Liability Companies	Fully Liable Partnerships
0.45**	0.10	0.32*	0.64***

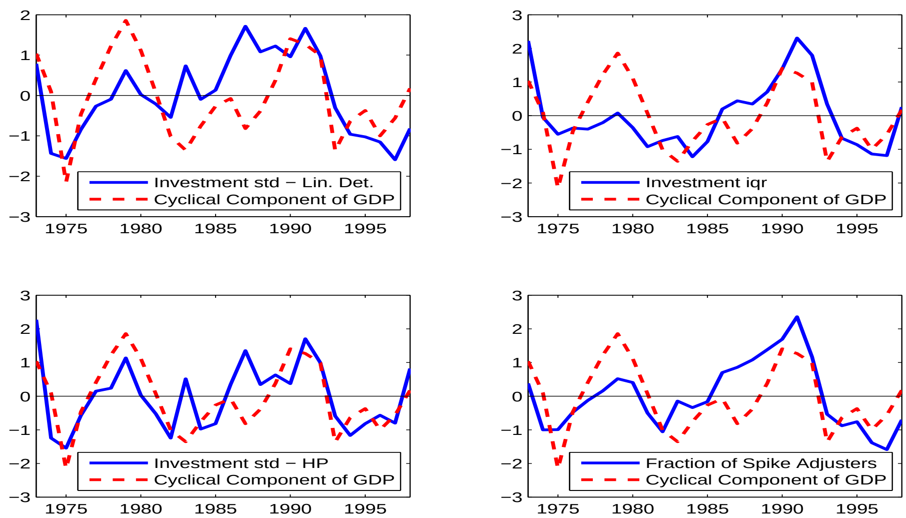
Notes: See notes to Table 21. 'Publicly Traded' means the German legal forms of *AG* and *KGaA*. 'Limited Liability Companies' means the German legal forms of *GmbH* and *GmbH & Co KG*. 'Fully Liable Partnerships' means the German legal forms of *GBR*, *OHG* and *KG*.

Table 25: Evidence from Disaggregation by Two-Digit Industry and Year – Robustness

Regression of $std(i_{j,t})$ on ...			
Fraction of Spike Adjusters (20%)	$.28^{***}$	$std(\Delta\epsilon_{j,t})$	$.20^{***}$
Fraction of Spike Adjusters (10%)	$.12^{***}$	$std(\Delta\epsilon_{j,t})$	$.25^{***}$
Fraction of Adjusters (1%)	$.23^{***}$	$std(\Delta\epsilon_{j,t})$	$.37^{***}$

Notes: The table displays the coefficients of an OLS regression of the cross-sectional investment rate dispersion for each two-digit industry and year on the fraction of (spike, defined, alternatively with a 20 percent and a 10 percent investment rate threshold in absolute value) adjusters in that industry-year and the dispersion of idiosyncratic productivity shocks. All data have been linearly detrended and demeaned at the industry level.

Figure 1: Cyclicality of Cross-sectional Moments - Data



Notes: See notes to Table 21. For better readability, in each panel we normalize the cyclical components of the cross-sectional moments and of aggregate real gross value-added by their respective standard deviations.

C Numerical Implementation

C.1 Details

C.1.1 Decision Problem

The computable dynamic programming problem of a firm in our model has a 6-dimensional state space: $\left(k, \bar{k}\right)$ are the endogenous states, while $\left(\epsilon, z, s, std(\epsilon)\right)$ are the exogenous states. Since the employment problem has an analytical solution, there is essentially just one continuous control, k'. We discretize the state space as follows:

1. k: $n_k = 41$ grid points from $[0, 40]$, with a smaller grid width at low capital levels, where the curvature of the value function is higher.

2. $\bar{k}$: $n_{\bar{k}} = 4$ grid points: $[0.8, 1.0, 1.2, 1.4]$.

3. ϵ: $n_\epsilon = 19$. The grid points are equi-spaced (in logs) and the width between the midpoint, which is normalized to unity, and the extreme points is given by $3 \times \sqrt{\frac{\bar{\sigma}(\epsilon)^2}{1-\rho_\epsilon^2}}$, i.e., three times the unconditional variance of idiosyncratic productivity.

4. z : $n_z = 5$ grid points: $[0.9561, 0.9778, 1.0000, 1.0227, 1.0459]$.

5. s : $n_s = 5$ grid points: $[0.0764, 0.0834, 0.0905, 0.0976, 0.1047]$.

6. $std(\epsilon)$: $n_{std(\epsilon)} = 3$ grid points: $[0.30, 0.35, 0.40]$.

The various transition matrices for the stochastic processes are calculated as follows: first, using a bivariate version of the procedure in Tauchen (1986), we compute a discrete bivariate Markov chain on $z \times s$, using the results in equation (9) in the main text. Second, we then compute (in the baseline case) for each s a transition matrix on the fixed (across s) ϵ-grid. The transition matrix takes two features into account: time-varying $\sigma(\epsilon)$ and the (small) excess kurtosis of the idiosyncratic productivity process in the data: 4.4480 on average. Since in our calibration strategy the fixed adjustment costs parameter is identified by the kurtosis of the firm-level investment rate (together with its skewness), we want to avoid attributing excess kurtosis in the firm-level investment rate to lumpy investment, when the idiosyncratic driving force itself has excess kurtosis. We incorporate the measured excess kurtosis into the discretization process for the idiosyncratic productivity state by using a mixture of two Gaussian distributions: $N(0, 0.0586)$ and $N(0, 0.1224)$ - the standard deviations are 0.0905 ± 0.0319, with a weight of 0.4118 on the first distribution.

Since we allow for a continuous control, k, and $\bar{k}$ can take on any value continuously, we can only compute the value function exactly at the grid points above and interpolate for in-between values. This is done by using a multidimensional cubic splines procedure, with a so-called "not-a-knot" condition, to address the degrees-of-freedom problem that arises when splines are used. We compute the solution by value function iteration, using 10 steps of policy improvement after each actual optimization step. The optimum is found by using a golden section search. Upon convergence, we check single-peakedness of the objective function, to guarantee that the golden section search is reasonable. We have also experimented with finer grids for the baseline case and found our results to be robust.

C.1.2 Equilibrium Simulation

For the equilibrium simulations we draw one random series for the aggregate states and fix it across models. We use $T = 1600$ and discard the first 500 observations, when we compute statistics from these simulations.

As in the firm's decision problem, we use a golden section search to find the optimal target capital level, given p, at every point in time during the simulation. We find the market-clearing intertemporal price, using a combination of bisection, secant and inverse quadratic interpolation methods. The precision of the market-clearing algorithm is better than 10^{-7} at every point in time during the simulation.

There is a final complication due to the nature of the bivariate aggregate shock process: given the correlated processes for aggregate productivity and idiosyncratic firm-level risk, not all of the 25 (5×5) distinct aggregate state combinations are reached with sufficient frequency during our $T = 1600$-simulations to compute the regressions (8a) and (8b) in the main text state by state. Since going much beyond $T = 1600$ would be prohibitively burdensome in terms of computational resources and time, we proceed as follows: if we have at least five observations on an aggregate state combination, we run the state-by-state regressions. Otherwise, we use a version of (8a) and (8b) where we treat z and $\sigma(\epsilon)$ as if they were continuous variables, include several, but not all possible interaction terms, and run OLS regressions on these modified rules to extrapolate the coefficients for the remaining aggregate state combinations. As we will show below, these somewhat restricted regressions nevertheless provide a good fit for the time-series of aggregate capital and the marginal utility of consumption.

C.2 Quality of Numerical Approximations

What role does $\log std(\epsilon)$ play in the Krusell and Smith (1998) or KS rules, (8a) and (8b) in the main text? Table 26 shows that both our cross-sectional results and, to some extent, our aggregate results depend crucially on the addition of $\log std(\epsilon)$ in the KS rules. For instance, the procyclicality of the investment rate dispersion is substantially lower without it.

Tables 27, 28 and 29 illustrate why. First, the $R2$ is substantially higher and the standard error of estimation is substantially lower when $\log std(\epsilon)$ is added in the KS rules, especially the KS rules for the aggregate capital stock. Second, as Den Haan (2010) argues, it is important to check the quality not only of the one-step-ahead forecasts of the KS rules but also of forecasts at longer horizons. The performance of the KS rules – measured by the mean squared percentage deviation of applying the forecasting rules t times (using the actual realizations of the driving processes) from the actual value on the equilibrium simulation path (assuming households use the converged one-step-ahead forecasting rules), measured by the maximum absolute percentage deviation and measured by the correlation coefficient between the t-forecast and the actual value on the equilibrium simulation path – deteriorates substantially, especially at longer horizons, when $\log std(\epsilon)$ is not used.

Table 26: Economic Implications of $\log std(\epsilon)$ Being in the Krusell-Smith Rules

correlation between Y and $std(i_{j,t})$	BL-Model w $\log std(\epsilon)$	0.53
	BL-Model w/o $\log std(\epsilon)$	0.28
	Data	0.45

		Y	C	I	N
volatility	BL-Model w $\log std(\epsilon)$	2.19	0.82	10.22	1.52
	BL-Model w/o $\log std(\epsilon)$	2.16	1.03	10.23	1.56
	Data	2.30	1.79	4.37	1.80
persistence	BL-Model w $\log std(\epsilon)$	0.29	0.55	0.22	0.20
	BL-Model w/o $\log std(\epsilon)$	0.22	0.65	0.08	0.06
	Data	0.48	0.67	0.42	0.61
correlation with Y	BL-Model w $\log std(\epsilon)$	1.00	0.87	0.98	0.96
	BL-Model w/o $\log std(\epsilon)$	1.00	0.74	0.93	0.89
	Data	1.00	0.66	0.83	0.68

Notes: The table displays – in the upper panel – the correlation of the dispersion of the investment rates with the cycle for the model under the baseline calibration ('BL-Model'), where $\log std(\epsilon)$ is included in the Krusell-Smith rules, a version of the baseline model, where $\log std(\epsilon)$ is not included in the Krusell-Smith rules, as well as from the USTAN data. In the lower panel the table does the same comparison for the percent standard deviations (volatility), autocorrelation (persistence), and correlation with aggregate output of HP(100)-filtered log aggregate output (Y), consumption (C), investment (I), and employment (N).

Table 27: Quality of KS Rules - R2 and the Standard Error

	Rule for $\bar{k}$					Rule for p				
	KS rules include $\log std(\epsilon)$									
	s_1	s_2	s_3	s_4	s_5	s_1	s_2	s_3	s_4	s_5
z_1	-	-	-	0.9999	1.0000	-	-	-	0.9996	1.0000
z_2	-	0.9997	0.9996	0.9997	1.0000	-	0.9999	0.9991	0.9994	0.9998
z_3	0.9997	0.9993	0.9997	0.9997	1.0000	0.9997	0.9989	0.9994	0.9995	1.0000
z_4	0.9995	0.9996	0.9997	0.9997	-	0.9994	0.9994	0.9996	0.9999	-
z_5	0.9997	0.9997	-	-	-	0.9995	0.9997	-	-	-
	Results for the regression w/o all interaction effects									
	$R2 = 0.9998$			$S.E. = 3.14 * 10^{-4}$		$R2 = 0.9997$				$S.E. = 2.20 * 10^{-4}$
	KS rules do not include $\log std(\epsilon)$									
	s_1	s_2	s_3	s_4	s_5	s_1	s_2	s_3	s_4	s_5
z_1	-	-	-	0.9586	0.9769	-	-	-	0.9984	0.9989
z_2	-	0.9420	0.8786	0.8782	0.8890	-	0.9992	0.9959	0.9960	0.9958
z_3	0.8256	0.8241	0.8729	0.8763	0.9029	0.9952	0.9935	0.9956	0.9956	0.9970
z_4	0.8398	0.8745	0.8878	0.9592	-	0.9940	0.9952	0.9959	0.9985	-
z_5	0.9376	0.7452	-	-	-	0.9972	0.9886	-	-	-
	Results for the regression w/o all interaction effects									
	$R2 = 0.9062$			$S.E. = 8.00 * 10^{-3}$		$R2 = 0.9976$				$S.E. = 8.00 * 10^{-4}$

Notes: All simulation results reported in the paper refer to the setup in the upper panel.

Table 28: Quality of Forecasting Rules at Various Horizons

		Forecast Horizon t			
Rule for ...	Evaluation Criterion	$t = 1$	$t = 5$	$t = 10$	$t = 100$
$\bar{k}$	Mean Squared % Dev. forecast - actual	0.0003	0.0007	0.008	0.0008
$\bar{k}$	Max. Abs. % Dev. forecast - actual	0.0014	0.0034	0.0036	0.0037
$\bar{k}$	Correlation forecast - actual	0.9999	0.9995	0.9994	0.9994
p	Mean Squared % Dev. forecast - actual	0.0002	0.0003	0.0003	0.0003
p	Max. Abs. % Dev. forecast - actual	0.0007	0.00010	0.0011	0.0011
p	Correlation forecast - actual	0.9999	0.9998	0.9997	0.9997

Table 29: Quality of Forecasting Rules at Various Horizons - KS rules do not include $\log std(\epsilon)$

		Forecast Horizon t			
Rule for ...	Evaluation Criterion	$t = 1$	$t = 5$	$t = 10$	$t = 100$
$\bar{k}$	Mean Squared % Dev. forecast - actual	0.0080	0.0220	0.0270	0.0296
$\bar{k}$	Max. Abs. % Dev. forecast - actual	0.0273	0.0832	0.0966	0.0943
$\bar{k}$	Correlation forecast - actual	0.9525	0.5995	0.3689	0.2404
p	Mean Squared % Dev. forecast - actual	0.0008	0.0087	0.0116	0.0130
p	Max. Abs. % Dev. forecast - actual	0.0027	0.0309	0.0402	0.0394
p	Correlation forecast - actual	0.9989	0.8468	0.7201	0.6471

Notes: All simulation results reported in the paper refer to the setup in Table 28.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2006. "Structural identification of production functions." University Library of Munich, Germany MPRA Paper 38349.

Bayer, Christian. 2006. "Investment dynamics with fixed capital adjustment cost and capital market imperfections." *Journal of Monetary Economics*, 53(8): 1909 – 1947.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 1995. "Capital Utilization and Returns to Scale." *NBER Macro Annual*, 10: 67–124.

Cooper, Russell, and John Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies*, 73(3): 611–633.

Davis, Steven, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2006. "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms." National Bureau of Economic Research Working Paper 12354.

Den Haan, Wouter. 2010. "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents." *Journal of Economic Dynamics and Control*, 34(1): 79–99.

Gandhi, Amit, Salvador Navarro, and David Rivers. 2011. "On the Identification of Production Functions: How Heterogeneous is Productivity?" University of Western Ontario, CIBC Centre for Human Capital and Productivity University of Western Ontario, CIBC Centre for Human Capital and Productivity Working Papers 20119.

Heckman, James. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." In *Annals of Economic and Social Measurement, Volume 5, number 4. NBER Chapters*, 475–492. National Bureau of Economic Research, Inc.

Krusell, Per, and Anthony Smith. 1998. "Income and wealth heterogeneity in the macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

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

Olley, Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–1297.

Tauchen, George. 1986. "Finite state Markov-chain approximations to univariate and vector autoregressions." *Economics Letters*, 20(2): 177–181.