Online Appendix for:

"Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model"

Rüdiger Bachmann, Ricardo J. Caballero, and Eduardo M.R.A. Engel

C Robustness

This section addresses various issues related to the robustness of our calibration. First we show that using alternative moments to summarize the range of heteroscedasticity values does not affect our findings. Second, we show that had we used model 1 instead of model 2 would have biased our results slightly against finding large conditional heteroscedasticity in the data. Third, we show that our sectoral calibration choices do not drive the results. Specifically, we study robustness to using 2-digit and 4-digit data to compute the statistics for the sectoral Solow residuals and the investment rates we use. Fourth, we also discuss robustness to how we detrend sectoral data. We show that using a weighted median as opposed to a weighted average to aggregate sectoral moments leaves the results unaltered. And finally, we experiment with the choice for the total annual standard deviation production units face, studying cases with 0.075 and 0.15, compared to our baseline choice of 0.10.

Table 9 shows the implied values for χ for alternative definitions for the heteroscedasticity range (log-ratio for 90th and 10th percentile, and log-ratio for largest and smallest values over the 184 quarters considered). It shows that $\chi = 0.50$ slightly overestimates the volatility of the responsiveness index, if we use the ratio of the maximum to minimum index – the 50% mentioned in the introduction and slightly underestimates it, had we use the ratio of the 90 percentile to the 10 percentile. But clearly, all calibration targets are close, whereas the frictionless and Khan-Thomas models cannot match conditional heteroscedasticity independently of how it is defined.

Table 9: HETEROSCEDASTICITY RANGE

Model	$\log(\sigma_{95}/\sigma_5)$	$\log(\sigma_{90}/\sigma_{10})$	$\log(\sigma_{\max}/\sigma_{\min})$
Data	*0.3021*	*0.2558*	*0.3841*
This paper:	0.2901	0.2183	0.4063
Frictionless:	0.0539	0.0405	0.0844
Khan-Thomas (2008):	0.0468	0.0391	0.0675

Table 10: HETEROSCEDASTICITY RANGE - TIME SERIES MODEL 1

Model	$\log(\sigma_{95}/\sigma_5)$	$\log(\sigma_{90}/\sigma_{10})$	$\log(\sigma_{max}/\sigma_{min})$
Data	*0.3175*	*0.2689*	*0.3964*
This paper:	0.2880	0.2184	0.4008
Frictionless:	0.0649	0.0493	0.0909
Implied χ	0.60	>0.60	0.50

Table 10 shows that had we used Model 1 in the conditional heteroscedasticity regression, we would have found a slightly stronger variation of the responsiveness index in the data, by 50% ($e^{0.3964} \simeq 1.49$), and, consequently, the calibrated maintenance parameter would have been higher.

We now go back to the calibration based on Model 2 and check robustness to the choices we made in obtaining the sectoral statistics we use in the calibration. Table 11 shows that our baseline calibrated value for the maintenance parameter, $\chi = 0.5$, is robust to these choices. Also, in any of these cases the frictionless model features a heteroscedasticity range around 0.05, well below the data's 0.3021.

Table 11: ROBUSTNESS - CALIBRATION

Specification	Cond. Adj. Costs/ Unit's Output	Calibrated χ	σ_S	ρ_S	Target Sectoral I/K Volatility	$\log(\sigma_{95}/\sigma_5)$ FL
Baseline	3.6%	0.50	0.0273	0.8612	0.0163	0.0539
Lin. detr. Solow res.	20.2%	0.40	0.0276	0.8979	0.0163	0.0537
Quadr. detr. I/K	4.4%	0.50	0.0273	0.8612	0.0152	0.0539
2-digit Data	7.4%	0.50	0.0166	0.8764	0.0098	0.0536
4-digit Data	1.4%	>0.60	0.0369	0.8492	0.0226	0.0539
Weighted Median	5.7%	0.50	0.0244	0.8727	0.0145	0.0537
Total $\sigma = 0.075$	3.3%	0.50	0.0273	0.8612	0.0163	0.0538
Total $\sigma = 0.15$	2.9%	>0.60	0.0273	0.8612	0.0163	0.0547

D Numerical Appendix

In this appendix, we describe in detail the numerical implementation of the model computation. Unless otherwise stated, the numerical specifications refer to the baseline calibration in the main text, although most of them are common across all models.

D.1 Decision Problem

Given the assumptions we made in the main paper: 1) $\rho_S = \rho_I = \rho$, and 2) approximating the distribution μ by the aggregate capital stock, $\bar{k}$, the dynamic programming problem has a 4-dimensional state space: $(k, \bar{k}, z, \epsilon)$. Since the employment problem has an analytical solution, there is essentially just one continuous control, k'.

We note that for all partial equilibrium computations the dimension of the state space collapses to three, $\bar{k}$ is no longer needed to compute prices and aggregate movements. Instead, we follow KT in fixing the intertemporal price and the real wage at their average levels from the general equilibrium simulations.

Since we allow for a continuous control, k, and $\bar{k}$ and z 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 large number of degrees of freedom problem, when using splines (see Judd, 1998). We compute the solution by value function iteration, using 20 steps of policy improvement after each actual optimization procedure. 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.

D.2 Equilibrium Simulation

For the calibration of the general equilibrium models we draw one random series for the aggregate technology level and fix it across models. We use $T = 600$ and discard the first 100 observations. For computing the conditional heteroscedasticity in the model simulations we use a much longer simulation horizon of $T = 10000$. We find that, generally, the statistics are robust to T. We start from an arbitrary individual capital distribution and the stationary distribution for the combined productivity level. The model economies typically settle fast into their stochastic steady state after roughly 50 observations. Since with idiosyncratic shocks, adjustment costs and necessary maintenance some production unit may not adjust for a very long time, we take out any individual capital stock in the distribution that has a marginal weight below 10^{-10}, in order to save on memory. We re-scale the remaining distribution proportionally.

As in the production unit's decision problem, we use a golden section search to find the optimal target capital level, given p. We find the market clearing intertemporal price, using a combination of bisection, secant and inverse quadratic interpolation methods. Precision of the market-clearing outcome is better than 10^{-7}.

To further assess the quality of the assumed log-linear equilibrium rules, we perform the following simulation: for each point in the $T = 500$ (we discard the first 100 observations) time

Table 12: Assessing agents' forecasting rules for capital

	FL	Baseline	Baseline-SKEW
$a_{\bar{k}}$	0.0065	0.0021	0.0112
$b_{\bar{k}}$	0.9061	0.9473	0.9388
$c_{\bar{k}}$	0.2199	0.1184	0.1106
$d_{\bar{k}}$	NaN	NaN	0.0075
$e_{\bar{k}}$	NaN	NaN	-0.0008
R^2	1.0000	0.9999	1.0000
SE	0.0000	0.003	0.0001
MAD(%)	0.09	0.61	0.34
MSE(%)	0.04	0.30	0.14
Correl.	1.0000	0.9956	0.9992

Table 13: Assessing agents' forecasting rules for p

	FL	Baseline	Baseline-SKEW
a_p	1.8438	1.8489	1.8748
b_p	-0.3357	-0.2442	-0.2701
c_p	-0.5836	-0.8020	-0.8215
d_p	NaN	NaN	0.0213
e_p	NaN	NaN	-0.0031
R^2	1.000	0.9992	0.9999
SE	0.0000	0.0006	0.0002
MAD(%)	0.02	0.19	0.11
MSE(%)	0.01	0.06	0.03
Correl.	1.000	0.9997	0.9999

series, we iterate for a time series of $\tilde{T} = 100$ aggregate capital and the intertemporal price forward, using only the equilibrium rules and assuming the actual time path for aggregate technology. We then compare the aggregate capital and p after $\tilde{T}$ steps with the actually simulated ones, when the equilibrium price was updated at each step. We then compute maximum absolute percentage deviations, mean squared percentage deviations, and the correlation between the simulated values and the out-of-sample forecasts. Tables 12 and 13 summarize the numerical results for each model. The rows contain: the coefficients of the log-linear regression, its R^2 and standard error and the three above measures that assess the out-of-sample quality of the equilibrium rules. They assess the log-linear approximation for future capital and current p, respectively. Baseline-SKEW refers to our baseline calibration, where agents use additionally the log standard deviation and skewness of the capital distribution for forecasting.

Table 12 shows that there exists a good log-linear approximation for aggregate capital as a function of last period's capital and the current aggregate shock. This may seem surprising in light of the time-varying impulse response functions we described in the main text. However, the numbers also show that in particular out-of-sample forecasts improve, when higher moments of the capital distribution are introduced. Furthermore, as we argue next, the goodness-of-fit for an equation analogous to (16a) and (16b), but with the aggregate investment rate as the dependent variable, is worse, even though the poorer fit has no bearing on aggregate investment dynamics.

Table 14: Assessing agents' forecasting rules for I/K

Highest moment	R^2						Autocorrelation	
	all	1st quart.	2nd quart.	3rd quart	4th quart.	average	1st	2nd
Baseline								
Mean:	0.9896	0.9535	0.7859	0.7259	0.9501	0.8538	0.906	0.816
St. deviation:	0.9992	0.9947	0.9869	0.9822	0.9961	0.9900	0.922	0.846
Skewness:	0.9998	0.9986	0.9975	0.9978	0.9988	0.9982	0.919	0.841

We simulated a series of 500 observations for our baseline model, assuming that agents use the first, the first two and the first three moments of capital in their forecasting rules.[38] We divided the simulated series into quartiles based on the magnitude of the actual investment rate, and calculated, for each quartile, the R^2-goodness-of-fit statistic between the aggregate investment rate series implied by the forecasting rule and the "true"aggregate investment rate series, which we assume to be the one generated, when agents use three moments of the capital distribution for forecasting.

Table 14 shows our results. The average (across quartiles) R^2 between the log-linear approximation and the true investment rate is only 0.85 for the baseline model. This average increases to 0.99 (0.97) when the log-standard-deviation of capital is added as a regressor, and to well above 0.99 when the skewness statistic is included as well.[39] The last two columns of Table 14 show that the estimated first and second order autocorrelations of the investment rate also improve significantly when using higher moments in the forecasting rules: the corresponding values for the actual investment rate series are 0.919 and 0.842, respectively, for the baseline calibration.

However, we also recomputed the evolution of the aggregate investment rate, when agents

[38]More precisely, the first case has the log-mean of capital holdings as a regressor, the second case adds the log-standard deviation and the third case also incorporates the skewness of capital holdings. Of course, $\log z_t$ is a regressor in all cases.

[39]For the frictionless model the first part of the first row would read: 0.9981, 0.9887, 0.9774, 0.9796, 0.9841, 0.9825. And the autocorrelations for forecasted investment rates are almost identical to the ones for the actual series.

use the rules that include higher moments of capital, and found no discernible differences with what we obtained with the log-linear forecasting rules: the correlation coefficient between the sample paths of I/K generated with forecasting rules with and without higher moments is above 0.9999.

D.3 Sectoral Simulation

Underlying the sectoral simulation are four assumptions: First, we assume a large enough number of sectors. Second, we assume a large $\sigma_S/sigma_A$ ratio, so that we can compute the sectoral implications of our model independently of the aggregate general equilibrium calculations. This is also reflected in our treatment of the sectoral data as residual values, which are uncorrelated with aggregate components. Third, we make use of the assumption that a sector is large enough to use a law of large numbers for the true idiosyncratic productivity shocks. Fourth, we assume that $\rho_S = \rho_I$, and that sectoral and idiosyncratic productivity shocks are independent, so that we can treat sectoral and idiosyncratic uncertainty as one state variable in the general equilibrium problem.

We start by fixing the aggregate technology level at its average level: $z^{SS} = 1$. The converged equilibrium law of motion for aggregate capital can then be used to compute the steady state aggregate capital level that belongs to this aggregate productivity. It is the fixed point of the aggregate law of motion, evaluated at z^{SS}:

$$\bar{k}^{SS} \equiv \exp \frac{a_{\bar{k}}}{1 - b_{\bar{k}}}.$$

This, in turn, leads to the steady state $p^{SS} \equiv \exp(a_p + b_p \log(\bar{k}^{SS}))$.

Then we specify a separate grid for idiosyncratic and sectoral productivity in such a way that all new grid points and any product of them will lie on the original 19-state grid for the combined productivity, used in the general equilibrium problem. Given the equi-spaced (in logs) nature of the combined grid this is obviously possible. Thus, the idiosyncratic grid comprises 11 grid points, and the sectoral grid 9 grid points, both equi-spaced and centered around unity.

Next, we recompute optimal target capital levels as well as gross values of investment (see equation 17d) at $z^{SS}, \bar{k}^{SS}$, at the 19 values for ϵ. By construction, these are then also the values for any (ϵ_S, ϵ_I)-combination. Note that we use the value functions computed from the general equilibrium case. We draw a random series of $T = 2600$ for ϵ_S, which remains fixed across all models, start from an arbitrary capital distribution and the stationary distribution for the idiosyncratic technology level, and follow the behavior of this representative sector, using the sectoral policy rules. The details are similar to those of the equilibrium simulation.

Finally, we test the two main assumptions on which we base our sectoral computations: a continuum of sectors and fixing the aggregate environment at its steady state level. To this end, we compute the equilibrium with a finite number of sectors, N_S. Also, we introduce an additional state-variable, given by: $\bar{\epsilon}_{S,t} \equiv \sum_{i=1,\dots,N_S} \log(\epsilon_{S,t}(i))$, which captures changes in the aggregate environment, beyond the common aggregate shock. Obviously, $\bar{\epsilon}_{S,t} = 0, \forall\ t$, as $N_S \rightarrow \infty$, by the law of large numbers and assuming sectoral independence. This additional aggregate state is then integrated over by Gauss-Hermitian integration, which is facilitated by the fact that the $\bar{\epsilon}_{S,t}$-process is independent of the aggregate technology process (by assumption). For computational reasons - following a large number of sectors with a large number of production units each is considerably more onerous in a quarterly calibration than in a yearly calibration -, we run these robustness checks for the annual equivalents of our baseline models.

We choose two different values for N_S. First, 400, which roughly equals the number of 3-digit SIC-87 sectors in the U.S. (395). Since, however, sectors are of very different size and overall importance, and also often correlated, we decrease, secondly, N_S to 100 for robustness reasons. The resulting residual $\sigma_{\bar{\epsilon}_S}$ is 0.0026 and 0.0052, respectively. Notice that in both cases $\sigma_{\bar{\epsilon}_S}$ is considerably smaller than $\sigma_A = 0.0120$, the σ_A for the annual frictionless calibration, so that we should not expect too large an effect from this additional source of aggregate uncertainty.

The following table shows the aggregate and sectoral standard deviations for annual investment rates for the frictionless model and our baseline lumpy model ($\chi = 0.5$). The raw sectoral standard deviations are shown as a capital-weighted average (the unweighted averages are only insignificantly different). The residual sectoral standard deviations are shown with the same filtering operations as discussed in Print Appendix A.3.

Table 15: ROBUSTNESS OF THE SECTORAL COMPUTATION

Model:	FL	FL	Lumpy	Lumpy
Number of sectors:	100	400	100	400
Aggr. St.dev.	0.0113	0.0102	0.0103	0.0099
Sect. St.dev. - raw	0.1824	0.1838	0.0190	0.0188
Sect. St.dev. - res.	0.1819	0.1834	0.0159	0.0160

The first important observation is that the numbers obtained here are not much different from what we have obtained in the simplified computation, which is in particular true for the lumpy model. Specifically, the frictionless model continues to fail to match observed sectoral volatility by an order of magnitude. And, secondly, the numbers deviate in the expected direction: the aggregate standard deviation increases (from 0.0098), because there is an additional aggregate shock, but only slightly so; the sectoral standard deviations decrease a little bit (from 0.0163), because now general equilibrium forces contribute also to sectoral smoothing. Overall, our simplified sectoral simulations seem justified.

E Matching Establishment Statistics

One argument we gave for using sectoral rather than plant level data to calibrate micro frictions is that matching micro moments may not be a robust way of pinning down microeconomic parameters when the goal is to use these parameters to identify aggregate effects of the mechanism. In this appendix we provide support to this claim by showing that a straightforward modification of the micro underpinnings of our baseline model leads to a satisfactory match of establishment level moments. More important, the match of sectoral and aggregate moments we obtained in the main text is unaffected by this extension. Our objective here is not to add realism to our original model, but to illustrate the potential lack of power of using (only) plant level data for our purpose.

E.1 A Simple Extension

A first choice we need to make when matching the model to micro data is how many micro units in the model correspond to one establishment. Choices by other authors have covered a wide range, going from one to a a continuum (see footnote 18 in the main text).

Two additional issues arise if we choose to model an establishment as the aggregation of many micro units. First, we must address the extent to which shocks—both to productivity and to adjustment costs—are correlated across units within an establishment.[40] Second, we must take a stance on the fact that establishments sell off and buy what in our model corresponds to one or more micro units.

Next we present a simple model that incorporates both elements mentioned above. The economy is composed of sectors (indexed by s), which are composed of establishments (indexed by e), which are composed of units (indexed by u). Data are available at the establishment level but not at the unit level. The log-productivity shock faced by unit u in establishment e in sector s at time t is decomposed into aggregate, sectoral, establishment and unit level shocks as follows:

$$\log z_{uest} = \log \varepsilon_t^A + \log \varepsilon_{st}^S + \log \varepsilon_{est}^E + \log \varepsilon_{uest}^U,$$

where $\log \varepsilon_t^A \sim \text{AR}(1;\rho_A,\sigma_A)$, $\log \varepsilon_{st}^S \sim \text{AR}(1;\rho_S,\sigma_S)$, $\log \varepsilon_{est}^E \sim \text{AR}(1;\rho_E,\sigma_E)$ and $\log \varepsilon_{uest}^U \sim$

[40] For tractability, we assume that decisions are made at the micro-unit level, not the establishment level.

AR(1; ρ_U, σ_U), and the usual orthogonality assumptions hold.[41] Consistent with the assumptions we made in the paper, we set $\rho_S = \rho_E = \rho_U$ and denote the common value by ρ.

An establishment is composed of a large number (continuum) of units. The extent to which the behavior of units within an establishment is correlated varies with the relative importance of σ_U and σ_E. The larger σ_E, the higher the correlation of productivity shocks across units and the more coordinated their investment decisions will be. The sectoral and aggregate investment series generated by the extended model are the same as those generated by the model developed in the main text as long as $\sigma_E^2 + \sigma_U^2 = \sigma_I^2$, since all we do in this extension is group micro units into "establishments" in a way that has no implication for sectoral aggregates. We consider the polar cases with uncorrelated productivity shocks ($\sigma_U^2 = \sigma_I^2$, $\sigma_E^2 = 0$) and perfectly correlated shocks ($\sigma_U^2 = 0$, $\sigma_E^2 = \sigma_I^2$). The degree of coordination also depends on how correlated adjustment costs are across units within an establishment; again we consider the polar cases where adjustment costs are perfectly correlated and independent.[42]

Regarding the sale and purchase of micro units, we assume that in every period the capital stock that is recorded at the establishment level is related to the capital stock for that unit determined by our baseline model via

$$K^r_{est} \equiv (1+\tau_{est})K^m_{est}, \tag{E.1}$$

where the superscripts r and m stand for "recorded" and "model". The τ's are i.i.d. draws from a normal distribution with zero mean and standard deviation σ_τ.[43]

The capital accumulation identity for recorded investment and (E.1) lead to:

$$I^r_{est} = K^r_{es,t+1} - (1-\delta)K^r_{est} = (1+\tau_{es,t+1})K^m_{es,t+1} - (1+\tau_{est})(1-\delta)K^m_{est}. \tag{E.2}$$

It follows that if the establishment sells off a fraction of the units it holds in $t+1$, these sales will show up as lower (or even negative) investment in the recorded investment data.

Dividing both sides of (E.2) by K^r_{est}, using (E.1) and denoting investment rates I_t/K_t by i_t yields

$$i^r_{est} = \frac{(1+\tau_{es,t+1})}{(1+\tau_{est})}\frac{K^m_{es,t+1}}{K^m_{est}} - (1-\delta).$$

Using the capital accumulation identity to express $K^m_{es,t+1}/K^m_{est}$ in terms of the model's invest-

[41] $x_t \sim$ AR(1; ρ, σ) means that the process x_t follows an AR(1) with first order autocorrelation ρ and standard deviation of innovations equal to σ.

[42] There is a one-to-one match between micro units in the model and establishment level data when adjustment costs and productivity shocks are perfectly correlated. Otherwise a continuum of model micro units correspond to one establishment in the data.

[43] We choose a symmetric distribution so that asymmetries in the histogram of investment rates cannot be attributed to this choice.

ment rate and ignoring second order terms in τ then leads to

$$i^r_{est} \approx (1-\Delta\tau_{es,t+1})\, i^m_{est} + \Delta\tau_{es,t+1}(1-\delta), \tag{E.3}$$

with $\Delta\tau_{es,t+1} \equiv \tau_{es,t+1} - \tau_{est}$.

Summing up, our (admittedly simple) extension introduces three parameters —the degree of correlation of productivity and adjustment costs across units within establishments, and the volatility parameter for unit purchases and sales— that can be used to fit establishment level moments without affecting the match of sectoral and aggregate statistics.

E.2 Matching Establishment Level Statistics

For the four combinations of correlation across both sources of shocks we generate a histogram with 2,500 realizations of establishment level I/K using our model.[44]

Table 16: MATCHING LRD MOMENTS

Model		f_1	f_2	f_3	f_4	f_5	σ_τ	avge. abs. deviation
correl. adj. costs	correl. prod. shocks							
0	0	.022	.040	.015	.833	.090	.074	.380
0	1	.000	.090	.046	.729	.135	.056	.377
1	0	.031	.081	.043	.696	.149	.037	.331
1	1	.000	.090	.079	.726	.105	.035	.386
Data		.019	.090	.082	.622	.187	—	—

Denote by f_i, $i = 1, \ldots, 5$ the fraction of LRD establishments that adjusted less than −20%, between −20 and −1%, between −1% and 1%, between 1 and 20% and above 20%, respectively. And denote by $\pi_i(\sigma_\tau)$ the fraction of units with adjustment in the previous bins after applying the transformation described in (E.3). We choose the value of σ_τ that minimizes $\sum_i |f_i - \pi_i(\sigma_\tau)|/f_i$, that is, that minimizes the average absolute deviation.

Table 16 presents our results. We consider four combinations of correlation among adjustment and productivity shocks. Comparing the first four rows with the last row shows that the match we obtain for statistics of the plant level distribution is reasonable.[45] More important, this exercise illustrates that establishment level moments may not be useful to calibrate model parameters that play an important role determining aggregate dynamics.

[44] We compute these investment rates using the approximation described in Appendix D.3 with $\sigma^2_S + \sigma^2_E$ in the role of σ^2_S, and $\sigma^2_I - \sigma^2_E$ in the role of σ^2_I.

[45] The goodness of fit is similar to that obtained by KT, which is 0.303.