Heterogenous Firms and the Dynamics of Investment

Matias Bayas-Erazo Fergal Hanks

Abstract

We show that standard heterogeneous firm models of investment feature investment responses that decrease steadily after a shock as positive relationship between the magnitude of the initial response to a shock and the speed of convergence of capital. This is inconsistent with the hump shaped responses found in the data. To resolve this tension tension we add stochastic time to build to the standard model. This feature generates hump shaped responses and enables the model to match micro moments of investment as well as the response dynamics.

1 INTRODUCTION

Investment models struggle to replicate both the lumpiness of investment at the microeconomic level and the sluggishness at the macroeconomic level. The classical adjustment cost formulation (Hayashi, 1982) and its variations, typically favored by macroeconomists to match aggregate data, is starkly at odds with the infrequent and large investments observed at the microeconomic level. Models of lumpy investment were designed to address these limitations and can explain key moments of the investment distribution (Caballero and Engel, 1999; Khan and Thomas, 2008). The extreme responsiveness of investment in these models, however, implies that the heterogeneity and state-dependence that distinguishes them is irrelevant for general equilibrium dynamics (House, 2014). Since an economy with lumpy investment behaves like a neoclassical economy with frictionless capital adjustment, the class of investment models consistent with the micro data appears inconsistent with the macro data. In an attempt to match micro lumpiness without sacrificing macro sluggishness, state-of-the art investment models rely on a combination of fixed and convex capital adjustment costs. Convex capital adjustment costs dampen the responsiveness mentioned earlier, helping bring the models in line with existing evidence and breaking the general equilibrium irrelevance result (Winberry, 2021; Koby and Wolf, 2020).

This paper brings renewed attention to the dynamics of aggregate capital in heterogenous firm models which is both important for the dynamics in general equilibrium but also useful for discriminating between different models of firms. It starts by showing that standard models of lumpy adjustment feature investment responses that peak on impact and then steadily decrease. Additionally these responses feature a strong positive relationship between the size of the initial response and the speed of convergence back to the steady state. These two features are inconsistent with the empirical evidence on investment dynamics such as Curtis et al., 2021. In order to resolve this inconsistency we add stochastic time to build to a model of lumpy capital adjustment. This leads to inertial dynamics in investment consistent with the empirical evidence on investment dynamics while still matching microeconomic moments of investment.

To analyse the dynamics of investment we use a general feature of firm investment models that the homogeneous of degree one in aggregate productivity and idiosyncratic capital. This assumption ensures that the importance of adjustment costs will not change with growth in aggregate productivity. In the absence of other idiosyncratic state variables this implies an equivalence between permanent shocks to aggregate productivity and capital destruction shocks[1]. Furthermore the response to a capital destruction shock captures the convergence dynamics of the model as it moves the distribution away from steady state but does not effect the policy functions. Thus the investment response to an permanent aggregate productivity shock will decrease over time as the distribution of capital converges to the new steady state distribution. Additionally these responses feature a strong positive relationship between the size of the initial response and the speed of convergence back to the steady state. As the size of the initial response

[1]This result has been used in papers such as Baley and Blanco (2021)

decreases, it takes longer for capital to converge back to the steady state level as the investment level at each distance from the new steady state level of capital will be lower. Thus, in models that produce large responses to shocks[2], such as models with only fixed costs that match micro moments, capital converges back to the steady state level quickly. On the other hand, models with smaller responses to shocks, such as models with convex costs, capital takes a long time to converge back to the steady state level.

These features are inconsistent with the empirical evidence on investment dynamics. Estimates from Curtis et al. (2021) show that response of investment does not steadily decrease in response to a permanent shock. Thus when calibrating to the initial response as estimated by Zwick and Mahon (2017) this implies a slow convergence of capital back to the steady state level. Additionally to explain dynamics of investment at the macro level Christiano et al. (2005) argued that investment needed to respond in a hump shaped manner to shocks.

We then resolve this inconsistency by adding stochastic time to build to a model of lumpy capital adjustment. When firms choose to start an investment project they choose the size of the project which is then fixed. In the next period the investment project may complete with an exogenous probability, otherwise the project is ongoing. This feature leads to inertia in investment in response to a shock as when the shock hits many firms will have ongoing investment projects. Since the size of these projects was fixed prior to the shock their realisation will not increase the level of investment. Instead investment increases as both the number of firms starting new projects and the share of projects started post the shock increases. Thus leading to a hump shaped response of investment to a shock.

We then numerically confirm our earlier analysis by looking at the investment response as a function of the time to build parameter as well as the standard fixed and convex costs assumed in the literature. As the fixed and convex cost parameters increase[3] the initial response of investment decreases and the time taken for capital to converge back to the steady state level increases. When decreasing the probability of an investment project away from 1, the initial response of investment response decreases. In contrast to fixed and convex costs that lost initial response is concentrated in the periods after the initial period and not spread out over all periods. Thus for lower values of the time to build parameter the investment response is hump shaped consistent with empirical evidence.

Finally we calibrate three versions of the model. The first is a standard fixed cost model which cannot match both micro moments and even just the initial response of investment. The second adds convex costs to the model which allows it to match both micro moments and the initial response of investment. However the model misses the dynamics of investment as they are not hump shaped and the speed of convergence is too slow. Finally we allow for stochastic time to build in the model and find that it can match both the initial response of investment and the dynamics of investment.

[2]Koby and Wolf (2020) show that there are strong linkages between the responses to different shocks, justifying the use of the productivity response as a general measure of the model's responsiveness

[3]While in previous work convex costs are usually used to match empirical impact responses Fang (2023) shows that fixed costs models can match these moments at the cost of other moments of the distribution

2 DYNAMICS OF INVESTMENT

In models of heterogenous households, recent work such as Auclert et al. (2024) has emphasised the importance of understanding the dynamic response of heterogenous agent blocks for their impact on aggregate dynamics. In this section, we develop a framework in which to compare the dynamic response of investment in different models. A key feature of this framework is that the cumulative net investment response will be fixed so that the relative dynamics can be compared similar to how the iMPCs in Auclert et al. (2024) sum to one in net present value terms. We will then show that in standard models of capital investment these investment dynamics are inconsistent with empirical evidence.

Start by considering a model of a firm investing in capital subject to adjustment costs in capital facing idiosyncratic and aggregate productivity shocks. The value function of a firm is then given by

$$V(Z, e_{i,t}, k_{i,t}) = \max_{k'} F(Z_t, e_{i,t}, k_{i,t}) + AC(Z_t, k_{i,t}, k') \tag{1}$$

$$+ \beta \mathbb{E}_{e'|e_{i,t}} \left[V(Z_t, e', k') \right] \tag{2}$$

Where $F(Z_t, e_{i,t}, k_{i,t})$ is the production function of the firm, $AC(Z_t, k_{i,t}, k')$ is the adjustment cost function and β is the discount factor.

We then need to make two key assumptions. The first is that the economy is in the aggregate non stochastic steady state. This implies the firm does not have to form expectations over aggregate productivity. The second key assumption is that $F(Z_t, e_{i,t}, k_{i,t}) + AC(Z_t, k_{i,t}, k')$ is homogeneous of degree one in in k and Z. This ensures that the importance of the adjustment costs does not depend on the level of productivity and is satisfied by standard adjustment costs used in the literature. With these assumptions the firm's value function can be written as

$$V(Z, e_{i,t}, k_{i,t}) = ZV(e_{i,t}, \frac{k_{i,t}}{Z}) \tag{3}$$

This implies that the policy functions of the firm in $(e, \frac{k}{Z})$ space are invariant to changes in Z. So in response to a permanent unanticipated MIT shock to Z of 1% the long run average change in capital will be 1%.

Now we can define the dynamic investment response as the response of investment at date t to a permanent aggregate productivity shock at date s which generates a change in the capital stock of 1[4].

2.1 DYNAMICS OF INVESTMENT IN STANDARD MODELS

Having defined a framework to study the dynamics of investment we can now apply it to standard models of investment. With only capital as an endogenous state variable the response of the firm can be thought

[4]This choice to define it in levels is to match the iMPC definition, however defining the iMPI as a proportion of a 1% change in capital induced by a 1% change in aggregate productivity might be more natural the shock will be the same for all models rather than different depending on the initial steady state level of capital

about in $\frac{k}{Z}$ space. In this space two shocks will be equivalent, a positive productivity shock and a uniform proportional capital destruction shock. These shocks reduce $\frac{k_{i,t}}{Z}$ by the same fraction for all firms shifting the distribution to the left in $\frac{k}{Z}$ space. Additionally neither shock will change the policy functions of firms in $\frac{k}{Z}$ space. Thus the response of investment divided by Z will be the same. Furthermore since a capital destruction shock just moves the distribution of capital away from the steady state distribution this will also approximate the convergence behaviour.

Thus investment will response to a permanent shock in a steadily decreasing manner. Since initially the distribution of capital is far from the steady state but over time it will converge to the distribution. As the investment is increasing in this distance as this happens the level of investment will decrease. Additionally, this implies a relationship between the impact response of investment and the speed of convergence. The faster the convergence the larger the initial investment response.

These features are inconsistent with the empirical evidence on investment dynamics. Curtis et al. (2021) estimate the response of investment to changes in depreciation timing in the US. We compare this empirically estimated response to the response of investment in a standard model of investment in Figure 1.

They do not find a steadily decreasing response of investment to a permanent shock. Instead investment is elevated for a number of years before quickly returning to zero. In comparison a standard model of investment featuring only fixed costs such as Khan and Thomas (2008) generates a much larger response on impact and the investment response quickly returns to 0. On the other hand a model that adds convex costs such as Winberry (2021) or Koby and Wolf (2020) is able to much more closely match the response on impact. However the investment response declines at a steady rate in contrast to the empirical evidence. This leads to investment not returning to zero even after 20 years, much more drawn out than the empirical evidence.

There is also related evidence at the macro level. Christiano et al. (2005) argued that to match the dynamics of investment estimated in VARs, investment needed respond in a hump shaped manner. It was for this reason they assumed adjustment costs on investment.

3 MODEL

Consider a discrete time economy with a continuum of firms who choose sequences of investment projects to maximise the expected net present value of profits. The firm has an idiosyncratic productivity level $e_{i,t}$ which follows a AR(1) process in logs with normal innovations. The firm produces using a decreasing returns to scale production function $F(Z_t, e_{i,t}, k_{i,t}) = Ze_{i,t}k_{i,t}^{a}lpha$ where Z_t is the aggregate productivity level and $0 < alpha < 1$ is the return to scale parameter. The firm's capital depreciates at rate δ every period.

In order to increase their capital stock firms must start an investment project. If a firm doesn't have an ongoing project they draw a fixed cost of starting a new project ξ from a distribution G. We will assume

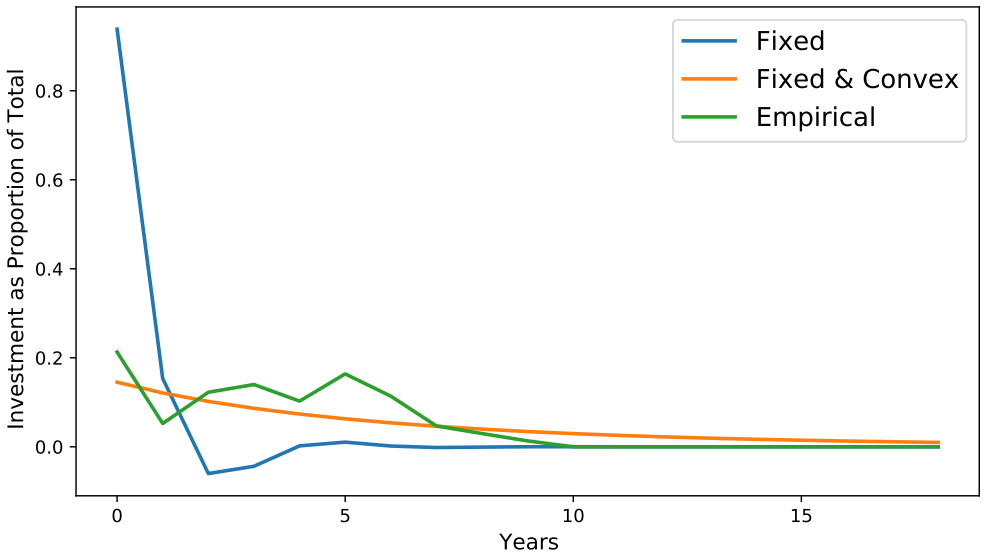


Figure 1: Comparison of Investment Dynamics in Standard Models and Empirical

this distribution is log normal with parameters μ and σ[5]. When the firm decides to pay the fixed cost to start a project they then choose the size of the project i which is fixed for the life of the project. The in the next period the project may complete with probability $kappa$. If the project completes the firms capital stock increases by the size of the project. The firm must also pay the costs of investment which we assume take the form $\gamma i + \phi \frac{i^2}{2}$. The allowance for a convex cost of investment is to nest recent models such as Winberry (2021) and Koby and Wolf (2020) which have shown that these costs can help match the response of investment on impact. We also allow for for a variable cost of investment to allow If the project does not complete then the firm continues with the project into the next period until it completes.

The timing of the model is that at the beginning of the period investment projects complete and then the firm produces. After production firms without an ongoing project decide whether to start a new project. The value function of a firm without an ongoing project can be written as

$$V(Z, e_{i,t}, k_{i,t}, 0, \xi_{i,t}) = Z e_{i,t} k_{i,t}^{\alpha} + \max\{\beta \mathbb{E}[V(Z, e_{i,t+1}, (1-\delta)k_{i,t}, 0, \xi_{i,t+1})], \max_{i}\{\beta \mathbb{E}[V(Z, e_{i,t+1}, (1-\delta)k_{i,t}, i, \xi_{i,t+1})] - \xi_{i,t}\}\} \quad (4)$$

Where the fourth argument of the value function is the size of the ongoing project with 0 indicating no ongoing project. The value function of a firm that enters the period with an ongoing project is then

[5]We depart from the usual assumption of a uniform distribution in order to make the value function of the firm continuously differentiable

$$V(Z, e_{i,t}, k_{i,t}, i, \xi_{i,t}) = \kappa(V(Z, e_{i,t}, i + k_{i,t}, 0, \xi_{i,t}) - \gamma i - \phi \frac{i^2}{2}) \quad (5)$$

$$+ (1 - \kappa)(Z e_{i,t} k_{i,t}^{a} lpha + \beta \mathbb{E}[V(Z, e_{i,t+1}, k_{i,t+1}, i, \xi_{i,t+1}]) \quad (6)$$

4 Numerical Illustration of the Dynamics of Investment

Having described the model we now illustrate the results relating to the shape of investment dynamics as well as the relationship between the initial response and the convergence speed numerically. To do this we vary one of the parameters of the model at a time and compare how the dynamics of investment change in response to each parameter. The three parameters we vary are μ of the fixed cost distribtuion [6], the convex investment cost parameter ϕ and the time to build parameter κ. Note that we keep $\kappa = 1$ when varying μ and ϕ and similarly keep $\phi = 0$ when varying μ and κ. This is to isolate the effect of each parameter on the dynamics of investment.

The results can be seen in 2. These three panels show the dynamics of net investment in response to a permanent as the three parameters are varied. The cumulative scale of the response is fixed as argued earlier and so the y-axis is the share of the cumulative net investment response that occurs in each period.

Staring with the first panel in which ϕ the convex cost parameter varies, we confirm the analytical results, that the investment response is decreasing after the shock. As well as there being a strong positive relationship between the size of the response and the speed of convergence back to zero net investment. Of particular note is that even moderate values of ϕ can lead to slow convergence speeds. This is unsurprising as convex adjustment costs strongly incentivise firms to spread the adjustment out over time. However this is a cautionary note for models that rely on convex adjustment costs to generate an empirically plausible response on impact to a shock.

In the second panel where μ one of the parameters of the fixed cost distribution is varied, we see similar results. The investment response is decreasing after the shock, though in this case there is a degree of overshooting. This overshooting may be related to the shape of the fixed cost distribution. Gourio and Kashyap (2007) found similar overshooting when they used a fixed cost distribution that features bunching around a single fixed cost. This can happen with a lognormal distribution but not a uniform distribution unless the cost is small enough so as to be almost irrelevant. This case also exhibits the relationship between the size of the initial response and the speed of convergence.

Finally we come to the third panel where κ the time to build parameter is varied. Most strikingly for the values of κ from 0.6 downwards the initial investment response is not the largest. Instead for two cases the second period is the peak of the response and for the lowest value of κ the third period is the peak. This humped shaped behaviour is a result of inertia in the distribution of ongoing investment projects. When the shock hits some firms will already have ongoing projects whose size was determined before the

[6] We choose μ over σ as this has a larger effect on the dynamics of investment and is thus more illustrative.

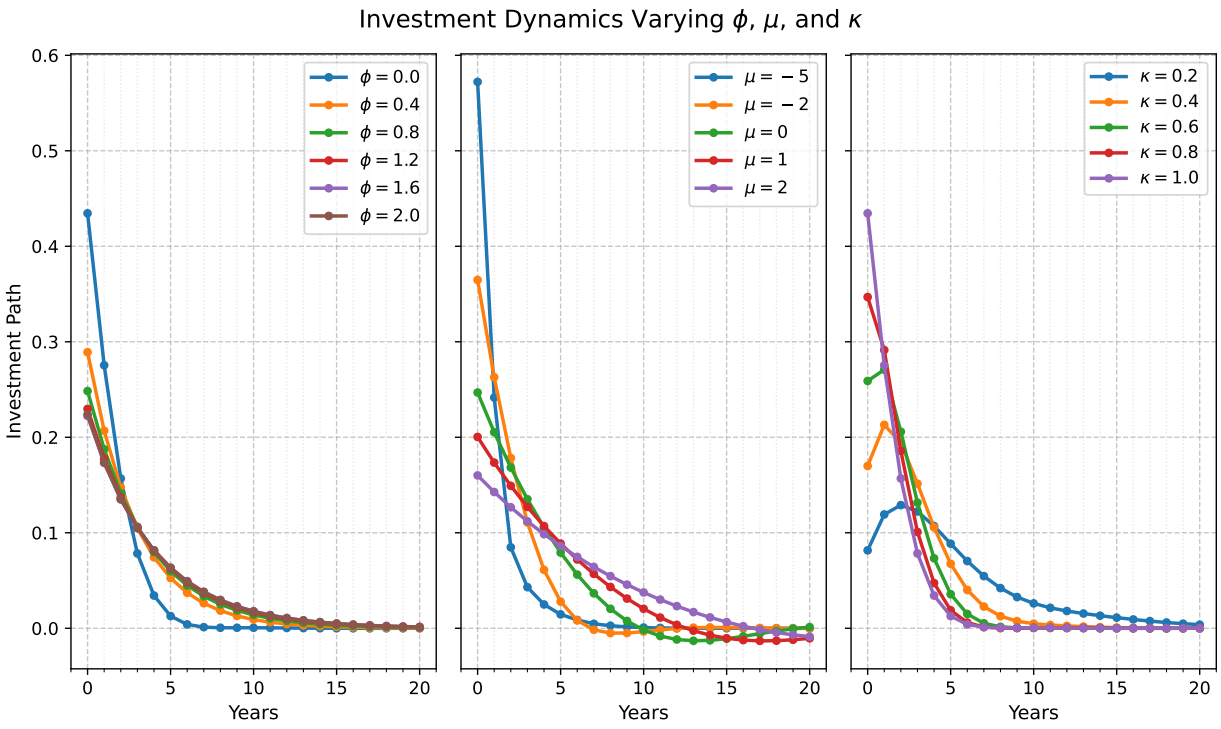


Figure 2: The effect of varying μ, ϕ and κ on the dynamics of investment in response to a permanent positive productivity shock.

realisation of the shock. This mutes the initial response as the realisation of these projects does not increase investment relative to the pre-shock level. Instead as the composition of ongoing investment projects shifts towards projects started after the shock the investment response increases. Thus generating a hump shaped response to the shock.

The second feature of a strong link between the size of the initial response and the speed of convergence is also affected. To see this compare the case of $\kappa = 0.4$ and the cases of $\phi > 0$. The initial response is smaller for $\kappa = 0.4$ is below 0.2 while all the initial responses in the first panel are above 0.2. However $\kappa = 0.4$ has a faster convergence speed than any of the cases in the first panel. The investment response reaches zero around year 10 while for the cases of $\phi > 0$ this takes around 14 years. The intuition for this result is simple, because some of the initial response shows up delayed due to time to build. Thus the investment level can be higher after the first period. Thus the level of the initial response is no longer as good a predictor of the speed of convergence as it is in the case of varying the other parameters.

5 Calibration

Having shown in the previous section that time to build can generate hump shaped responses in investment to shocks. We now formally calibrate two versions of the model to show that only the model with time to

build can match both the micro moments of investment and the dynamics of investment at the level of the firm block. We set the model period to be a year and set the discount factor to be $\beta = 0.977$. We set the depreciation rate $\delta = 0.067$. For the decreasing returns to scale parameter we set it equal to 0.67. For the parameters of the idiosyncratic productivity process we set the persistence $\rho_e = 0.890$ and the standard deviation $\sigma_e = 0.1$.

In the calibration we will have two sets of moments, micro and response moments. The first micro moment that we target is the average investment rate, defined as the mean of the ratios of investment to capital. In cases where an investment project is not completed the level of investment is considered to be 0. We also target the proportion of firms that feature a spike in investment defined as having an investment rate of more than 0.2. We take these targets from Zwick and Mahon (2017) who estimate the average investment rate to be 0.104 and the proportion of firms with a spike in investment to be 0.144.

The response moments that we target are the investment semi elasticity to a permanent shock to the cost of capital as estimated by Zwick and Mahon (2017) to be 7.2. We then use the 10 years of estimates of the investment from Curtis et al. (2021) normalised by the total net investment response similar to the figures in the previous section.

In the first calibration we set the time to build parameter $\kappa = 1$ and use the parameters of the fixed cost distribution μ and σ as well as the convex cost parameter ϕ as the parameters to be estimated.

As can be seen in **??** the model with only fixed and convex costs is unable to jointly match the micro and the response moments. In order to try and match the path of investment estimated by Curtis et al. (2021) the model uses high values of both μ and σ. This leads to a small initial response to the shock followed by slow convergence. This makes the dynamics of investment relatively flat. However this comes at the cost of a small initial response to the shock as in moving investment from the initial period to the next few periods, the model must also move investment to periods further in the future. This also leads to the investment in the model converging very slowly relative to the data. In the 10th year after the shock the model has only achieves 76% of the total net investment response.

By comparison the model with time to build is able to match the data much more closely as can be seen in 2. Because when $kappa < 1$ the model is able to generate a hump shaped response to the shock there is no longer a large trade off between the initial response and the convergence. Not only does the model better match the dynamics of investment but it also more closely matches the micro moments. This is because the model no longer has to rely on high values of μ and σ.

While the fit it much better than the model without time to build there are still some discrepancies. In particular the hump shape in the model is loaded more towards the first few years after the shock than in the data. This can be seen in that there is some overshooting of the data in the 2nd and 3rd years. Additionally the model doesn't match the very quick change in the level of investment after 7 years. However the esimated investment path from Curtis et al. (2021) is fairly noisy and so this feature of the data may not be robust. Additional empirical work more precisely estimating the investment path is needed to determine if

Table 1: Model Results and Moment Comparison

Parameters			
μ	0.493		
σ	9.999		
ϕ	10.000		
Moment	**Description**	**Model**	**Data**
1	Year 1 Convergence	0.132	0.215
2	Year 2 Convergence	0.247	0.268
3	Year 3 Convergence	0.347	0.390
4	Year 4 Convergence	0.433	0.530
5	Year 5 Convergence	0.509	0.633
6	Year 6 Convergence	0.574	0.796
7	Year 7 Convergence	0.631	0.910
8	Year 8 Convergence	0.681	0.957
9	Year 9 Convergence	0.724	0.987
10	Year 10 Convergence	0.761	1.000
11	Avg I/K	0.081	0.100
12	Avg Spike	0.242	0.144
13	Inv Semi-Elasticity	3.278	7.200

these discrepancies some further features of investment dynamics that are not captured by the model or if they are simply due to noise in the data.

6 CONCLUSION

In this paper we show that the standard models of lumpy capital adjustment are not able to match the full dynamics of investment observed in the data. We resolved this issue by introducing time to build into the standard model. This feature generates hump shaped responses in investment to shocks and thus enables the model to match both the micro and macro moments of investment. Further research both more precisely estimating investment dynamics as well as understanding how these dynamics vary across depending on firm characteristics is needed to more fully understand the both macroeconomic dynamics as well as the impact of policies aimed at stimulating investment.

Table 2: Model Results and Moment Comparison

Parameters			
κ	0.777		
μ	-0.133		
σ	2.942		
ϕ	0.188		
Moment	**Description**	**Model**	**Data**
1	Year 1 Convergence	0.163	0.215
2	Year 2 Convergence	0.332	0.268
3	Year 3 Convergence	0.475	0.390
4	Year 4 Convergence	0.591	0.530
5	Year 5 Convergence	0.683	0.633
6	Year 6 Convergence	0.755	0.796
7	Year 7 Convergence	0.811	0.910
8	Year 8 Convergence	0.855	0.957
9	Year 9 Convergence	0.889	0.987
10	Year 10 Convergence	0.915	1.000
11	Avg I/K	0.089	0.100
12	Avg Spike	0.170	0.144
13	Inv Semi-Elasticity	6.270	7.200

REFERENCES

Auclert, A., Rognlie, M., and Straub, L. (2024). The intertemporal keynesian cross. *Journal of Political Economy*, 132(12):4068–4121.

Baley, I. and Blanco, A. (2021). Aggregate Dynamics in Lumpy Economies. *Econometrica*.

Caballero, R. J. and Engel, E. M. R. A. (1999). Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S, s) Approach. *Econometrica*, 67(4):783–826. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00053.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45.

Curtis, M. E., Garrett, D. G., Ohrn, E., Roberts, K. A., and Suárez Serrato, J. C. (2021). Capital investment and labor demand: Evidence from 21st century tax policy.

Fang, M. (2023). A note on nonconvex adjustment costs in lumpy investment models: Mean versus variance. *Macroeconomic Dynamics*, 27(4):1166–1177.

Gourio, F. and Kashyap, A. K. (2007). Investment spikes: New facts and a general equilibrium exploration. *Journal of Monetary Economics*, 54:1–22.

Hayashi, F. (1982). Tobin's Marginal q and Average q: A Neoclassical Interpretation. *Econometrica*, 50(1):213–224. Publisher: [Wiley, Econometric Society].

House, C. L. (2014). Fixed costs and long-lived investments. *Journal of Monetary Economics*, 68:86–100.

Khan, A. and Thomas, J. K. (2008). Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics. *Econometrica*.

Koby, Y. and Wolf, C. K. (2020). Aggregation in Heterogeneous-Firm Models: Theory and Measurement. *Working Paper*.

Winberry, T. (2021). Lumpy investment, business cycles, and stimulus policy. *American Economic Review*.

Zwick, E. and Mahon, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, 107(1):217–248.

A CONTINUOUS TIME MODEL

In the continuous time version of the model, the firm solves a combined stochastic control and impulse control problem. When there is no adjustment, capital evolves as

$$dk_t = -\delta k_t dt + \sigma dB_t.$$

With impulse Δk, the capital stock jumpts from k to $k' = k + \Delta k$ but the firm has to pay a cost $\mathcal{G}$ that can be given by

$$\mathcal{G}(\Delta k) = F + p \max\{\Delta k, 0\} + p(1-\gamma) \min\{\Delta k, 0\}$$

As in the discrete time model, we assume that the firm faces convex and symmetric adjustment costs in labour. It can hire workers at rate n_t but must pay a cost $\mathcal{C}$ that is allowed to depend on current employment at the firm. A fraction ρ of workers leave the firm each period, so the evolution of labour is governed by

$$dl_t = [n_t - \rho l_t]\, dt$$

The program of the firm can be written as

$$v(k_0, l_0) = \sup_{\{\Delta k_t, n_t\}_{t \geq 0}} \mathbb{E}_0 \int_0^\infty e^{-rt} \left[eF(k_t, l_t) - w l_t - AC(\Delta k_t, n_t, k_t, l_t) \right] dt$$

subject to the evolution of capital and labour above and with

$$AC(\Delta k_t, n_t) := \mathcal{G}(\Delta k_t, k_t) + \mathcal{C}(n_t, l_t)$$