Ricardo Caballero
Aggregate Investment Theory

Report by Thijs Knaap (RuG)

1 Introduction

Professor Caballero's discussion of aggregate investment theory featured both old stock and innovations in the field. This report traces the main line of the argument and presents some of the models that can be useful in explaining today's aggregate investment. As with the lectures, this main line is contained, to a large extent, in Caballero (1997). The next section briefly highlights the importance of the subject matter. Sections 3 and 4 then cover the two main ideas from modern investment theory: the fact that investments are major and infrequent events and the occurrence of information and contract problems.

2 Investment Theory: History, Goals

Investment is an important subject in its own right: it is a substantial component of aggregate output and it accounts for much of its fluctuations. On top of that, it turns out that the theory of investment can be applied to a large number of other economic areas such as the analysis of R&D, labor hoarding, human capital accumulation and fiscal reforms. These subjects are similar in that they describe a situation where there is a cost in the short term, whose payoff is collected in the long term.

Early investment theory used an aggregate production function which explained the level of GDP by the levels of labor and capital. After inverting the relation and taking differences it is possible to explain changes in the stock of capital by changes in GDP. If the production function allows for substitution between factors, the cost of capital should also enter the equation. The bottom line of this approach is, however, that investment takes place in response to changes in GDP, rather than in response to market opportunities.

The q-theory of investment does not suffer from this deficiency. The average q, the value of the firm divided by the total value of its equipment and structures, or the marginal q, the ratio of extra value and added equipment, is used as an indicator of investment opportunities.

All the above models fail when confronted in a straightforward way with the data. To the dismay of many economists, investment correlates with changes in corporate cash flow and changes in GDP, but not with any measure of q or the cost of capital. Only when refined econometric measures are used can some of the latter variables be shown to exert influence, but only in the long run or in shorter periods which encompass a large change in their value. More importantly, they leave most of the variance in investment unexplained.

The theory that will be introduced in the next section aims at a way of explaining investment that is consistent 'from the ground up.' At the firm-level, we really see at least three sorts of investment take place:

- Ongoing investment, like machine maintenance, that takes place almost every day in a large plant. Costs are small relative to the total stock of the company's capital.

- Gradual investment, like small refinements of an existing machine, or a software upgrade.

- Major and infrequent investments, in which a large amount of new capital is purchased and costs are significant relative to the firm's total stock of capital.

Doms and Dunne (1993) use microeconomic data to show that the typical firm has a 'lumpy' investment pattern, *e.g.* spending 25-40% of seventeen years of investment in one year. This suggests that the latter type of investment accounts for an important share of the total expenditure. The next section will therefore characterize this type of investment at the micro level, and present a consistent macro model.

3 Major and Infrequent Investments

3.1 Theoretical Framework

The key to making investment take place in chunks rather than a continuous flow, is introducing a cost of adjustment over and above the rental cost of capital. Especially a cost that is significant even at infinitesimal adjustments will cause the firm to invest infrequently, and with large fractions of total

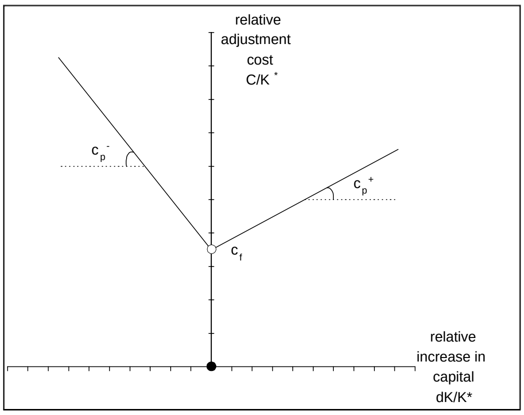

Figure 1: The adjustment costs relative to K^* as a function of the relative change in capital.

investment at a time. Caballero assumes an adjustment cost structure as in figure 1. The adjustment in capital is on the horizontal axis, and is taken as a fraction of the ideal stock of capital K^*. If the fraction is η costs are

$$C(\eta, K^*) = K^* \cdot \begin{cases} c_f + c_p^+ \cdot \eta & \text{if } \eta > 0 \\ 0 & \text{if } \eta = 0 \\ c_f - c_p^- \cdot \eta & \text{if } \eta < 0 \end{cases}$$

To compute the ideal stock of capital we construct the following model:

A firm would like to, at any point in time, maximize a flow of profit,

$$\begin{aligned} \Pi \cdot \Delta t &= \max_{K,L} \left(P(Y) \cdot Y - w \cdot L - r \cdot K \right) \cdot \Delta t. \\ Y(K, L, A) &= A \cdot K^{\alpha} \cdot L^{1-\alpha} \end{aligned}$$

Time is discrete and moves with increments Δt. We assume the firm actually performs this continuous maximization with respect to L. Use θ to denote a sufficient (scalar) statistic for conditions faced by the firm (including the shape of the function P, technology A and wages w). It is assumed that $\Delta\theta/\theta$ is IID so that $\ln\theta$ is a random walk, possibly with a drift. The maximization problem now becomes

$$\Pi \cdot \Delta t = \max_{K} \left(K^{\gamma} \cdot \theta - r \cdot K \right) \cdot \Delta t.$$

with $\gamma < 1$. This gives us K^*, the static optimum, the stock of capital that the firm would like to entertain:

$$K^* = \arg\max_{K} \Pi = \left(\frac{\gamma \cdot \theta}{r} \right)^{\frac{1}{1-\gamma}}$$

Being an affine transformation of $\ln\theta$, $\ln K^*$ follows a random walk.

At this point, it is important to remember the costs of capital adjustment that were introduced above. It will probably not be profitable to set $K = K^*$ at every point in time because unlike labor, every (small) adjustment in capital causes the firm to incur at least the fixed cost $c_f \cdot K^*$. So, instead of following the static optimum, the firm will choose maximization of its present value V as a guideline for investment. This new guideline will cause the firm's most preferred capital stock K^{**} to be different from K^* (due to, for instance, a drift in θ) and can make inaction the most profitable strategy.

From this setup, a few things can be seen immediately.

- If K is close to K^{**}, the increase in present value from adjusting K is probably less than the cost of adjustment. Thus, there will be a region around K^{**} in which the firm does not repair imbalances between K and K^{**}.
- If it is profitable to change the capital stock when it is at K_0, it must also be profitable to do so when it is at K_1 with $|K^{**} - K_0| < |K^{**} - K_1|$.[1]
- If the variable relative adjustment costs c_p^+ and c_p^- are positive and significant, adjustments will not be complete. It will be optimal to adjust K to a value closer to K^{**}, but not to K^{**} itself (the slope of the V-function, or marginal benefit, is zero there).

Summarizing, a firm now behaves as follows. Every period, it adjusts labor L to its optimal value. The optimal static stock of capital K^* is computed from conditions θ, as well as the imbalance $Z = K/K^*$. Then,

$$\begin{cases} \text{if } Z < L, & Z \text{ is adjusted to } Z = l \\ \text{if } L \leq Z \leq U, & Z \text{ remains unchanged} \\ \text{if } U < Z, & Z \text{ is adjusted to } Z = u \end{cases} \tag{1}$$

with $L \leq l \leq K^{**} \leq u \leq U$. Caballero argues that these variables may be found using the 'smooth pasting conditions' $V_Z(l) = V_Z(L) = c_p^+$ and $V_Z(u) = V_Z(U) = -c_p^-$.

[1] This because like the Π-function, the second derivative of the net present value function V w.r.t. K is always negative. The first derivative of V is zero at K^{**}. We know that K_0 is far enough away from K^{**} to incur both fixed and variable adjustment costs. Thus, $(V^{**} - V_0) / |K^* - K_0|$ is surely larger than the variable costs of adjusting. But from the second derivative V_{KK} we know that the slope the V-function along the stretch $K_0 \rightarrow K_1$ can only be steeper than the above sufficient measure. Thus, if it is profitable to adjust from K_0, it must be profitable to do it from K_1.

3.2 Empirical Verification

3.2.1 The hazard function

A reality check of the above model may be conducted with both micro- and macro-economic data. A first observation from micro data is that firms do not always react the same to an imbalance in capital. Thus, it is necessary to incorporate a random element in the (deterministic) adjustment scheme (1). Caballero (1997, p. 21) does this by letting c_f be a random variable of which a new value is drawn for each firm for every period. This means that trigger values L and U also become random, and we can talk about a probability of adjustment $\Lambda(x)$ with $x \equiv \ln(K/K^{**})$. Λ is the *hazard function*, a term borrowed from transition econometrics. For convenience, variable adjustment cost is set to zero, so that when a firm adjusts, its new stock of capital will be K^{**}. In terms of x, the net investment will be $e^{-x} - 1 \approx -x$. This allows us to write expected investment as $-x \cdot \Lambda(x)$.

3.2.2 Quadratic costs of adjustment

It is useful to see how the hazard function Λ would look in a rival model, the quadratic adjustment cost model. In this setup, adjustment costs are $C = c \cdot (K_t - K_{t-1})^2$. There are no fixed costs of adjustment, and c is positive. Rotemberg (1987, p. 92) shows that the optimal path of capital now is

$$K_t = \alpha \cdot K_{t-1} + [(1-\alpha)(\delta - 1)/\delta] \sum_{j \geq 0} (1/\delta)^j E_t K^*_{t+j}$$

Assuming $E_t K^*_{t+j} = K^*$ for all j and noting that $K^{**} = K^*$ we can write this, after log-linearization, as[2]

$$E\left(\frac{I}{K}\right) = -\beta \cdot \ln(K/K^*)$$

with $\beta > 0$. But here we can use the identities from above to get

$$-x \cdot \Lambda(x) = -\beta \cdot x$$

Dividing through by x we see that the quadratic adjustment cost model predicts a constant hazard rate.

[2] This trick is often applied to approximate the neoclassical model of growth. See, for instance, Barro and Sala-i-Martin (1995, p. 53).

3.2.3 Estimates

If we call the function that describes the distribution of capital imbalances $f(x,t)$, we can now write total expected investment in period t as

$$\frac{I_t}{K_t} = -\int_{-\infty}^{\infty} x \cdot \Lambda(x) \cdot f(x,t)\,\mathrm{d}x$$

Caballero, Engel and Haltiwanger (1995) estimate the hazard function $\Lambda(x)$ from panel data of approximately 7,000 firms during 17 years. The function is estimated both non-parametrically from micro data, and by fitting a 4^{th} degree polynomial on $\Lambda(x)$ using macro-data and several distributional assumptions. An artist's impression of the results[3] is in figure 2. From both

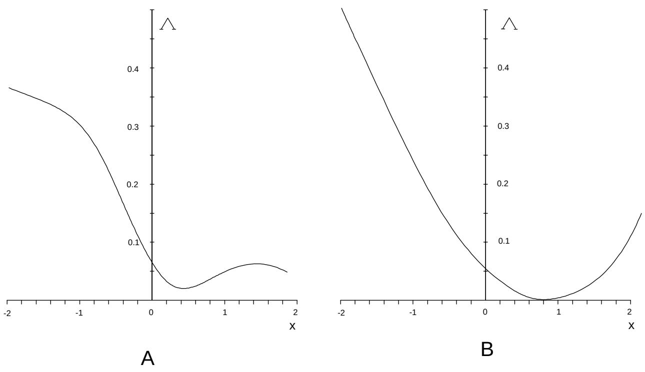

Figure 2: Hazard rate Λ versus imbalance x. A: estimated from micro data. B: estimated from aggregate data using distributional assumptions.

panels of this figure, it is clear that the hazard rate is not constant across x, thereby rejecting quadratic adjustment costs. Rather, a small value of x leads to a higher probability of investment. A large x does not often lead to disinvestment. This seems to indicate a strong irreversibility of investment.

4 Informational and Contractual Problems

Whereas the last section was much concerned with the technical side of investment, this section covers the consequences of strategic actions of investing agents. Strategies can involve waiting for others to gain information on the investment climate and reneging on imperfect contracts.

[3] I'm the artist.

4.1 Informational Problems

Gale (1995) discusses social learning, or situations in which private information is partly revealed by the actions of agents. In the investment context, it could be that all agents receive a private signal on the return of a certain kind of investment, and that the true return is the average of all these signals. These agents could all be better off by sharing the information, but assuming they do not it may be revealing to wait and see how many colleagues actually invest. Their action in this case carries extra information about the true return, information that can be added to one's private signal.

Gale argues that it may well be possible in a variety of models that imitation dominates private information. That is, even though every agent receives a positive signal, nobody invests because they all wait for someone else to make the first move. This situation is more likely if the space of actions is limited, as was the case with lumpy investments above. Any agent's information is then filtered and only observed as a binary variable: the investment go or no-go. The possibility that investments are irreversible also makes them prone to informational problems, as the potential losses of a bad investment project are obviously large compared to the costs of waiting one period to get information.

4.2 Contracts, Specificity, Opportunism

If an investment problem involves more than one party, contracts have to be made about each party's behavior after the costs have been sunk. Because investments are often irreversible and tailored to the other parties, it can be profitable to renege on a contract once the other party has committed itself. Professor Roberts gave an example of this 'hold-up problem' during his part of the workshop: after investing in a machine that makes seats that can only be used in a specific type of Toyota car, you have little outside opportunities if your sole possible customer, Toyota, unilaterally decides to lower the price it is willing to pay for the seats. Professor Caballero even admitted that M.I.T. itself had tried to renegotiate a contract with the Cambridge electricity company recently, after that firm had just invested to cope with the institution's demand for electricity.

4.2.1 Theory

Caballero and Hammour (1996) develop a model that shows the macroeconomic consequences of the hold-up problem. Production takes place with two factors, 1 and 2, both measured so that their supply equals one. An amount U_i is used of each factor in a factor-specific (autarkic) decreasing returns sector, the remaining units E_i are inputs in a constant returns joint production sector ($i = 1, 2$). The output in factor i's own sector is $F_i(U_i)$, and the joint production output is $y^n \cdot (1 - U_1)/x_1 = y^n \cdot (1 - U_2)/x_2$ with technical coefficients x_i given. Define $E_i = (1 - U_i)/x_i$.

Joint production requires some sunk costs: of E_i units sunk in this sector, only $\phi_i \cdot E_i$ can be retrieved to work in the factor's 'autarkic' sector ($0 < \phi_i < 1$). If contracts are incomplete, each party to a joint production project expects to be held up until the return to the project is no more than the return to his or her outside opportunity. Call factor compensation in the autarkic sector p_i, then, expecting a hold-up, the anticipated wage in the joint sector can never exceed $(1 - \phi_i) \cdot p_i$. This will cause a rent in that sector,

$$s^n = y^n - (1 - \phi_1) \cdot p_1 \cdot x_1 - (1 - \phi_2) \cdot p_2 \cdot x_2$$

which is assumed to be divided equally. Factor compensation in the joint sector now is $w_i^n = (1 - \phi_i) \cdot p_i \cdot x_i + \frac{1}{2}s^n$.

We compare situations with complete and incomplete contracts. With complete contracts ($\phi_i = 0$, $i = 1, 2$), we have that the returns in both sectors must equalize: $p_i = w_i$. With incomplete contracts, both parties require $p_i \geq w_i^n$ with the latter defined above. This is equivalent to

$$y^n \geq p_i \cdot x_i + p_{\neg i} \cdot x_{\neg i} + (\phi_i \cdot p_i \cdot x_i - \phi_{\neg i} \cdot p_{\neg i} \cdot x_{\neg i}) \tag{2}$$

with $\neg i = 3 - i$. In the perfect contract equilibrium the term in brackets does not appear.

From (2) we can see that there may exist projects that will be executed in a perfect contract equilibrium, and will not be executed in an imperfect contract equilibrium, if the term in brackets is different from zero. The term in brackets represents the difference in sunk costs between the two parties. If there exists such a difference, the party with the largest sunk costs (or, the largest appropriable sum) will demand the highest return. In figure 3, the general equilibrium versions of (2) are drawn out. The horizontal difference between lines (1) and (2) is the above 'term in brackets.' All else equal, factor 1 stands to lose the most from investing in the joint sector (because of, say, a high value of ϕ_1). For each y^n (vertical axis) this factor puts more into the autarkic sector than factor 2.[4] Thus, in equilibrium, there is an excess supply of factor 2 for the joint sector.

Caballero and Hammour (1996) derive a number of properties of this model. These properties include a segmented market for the factor in excess

[4] This means that the return in the autarkic sector for factor 1 will be lower, indicating a lower expected payof from the joint sector.

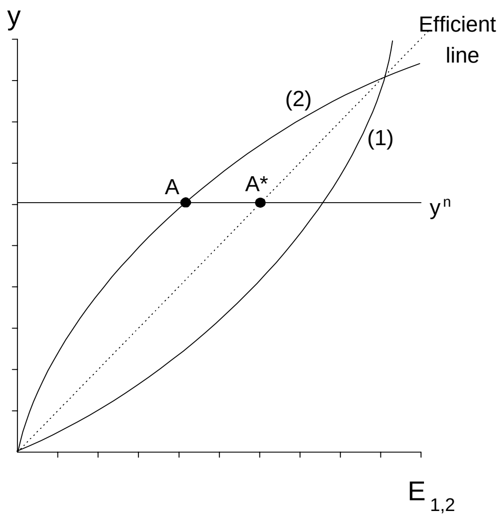

Figure 3: Factor 1 and 2's required y^n for different values of commited capital $E_{1,2}$. The efficient line is from the perfect-contract world, in which A^* is put into the joint sector. Imperfect contracts cause this amount to fall to A.

supply[5] and *sclerosis* or too little scrapping of old joint projects. When the factors are actually called capital and labor a number of properties of business cycles may be derived.

It is also shown that, when factors get to control their own level of appropriability ϕ_i, they will generally not aim at $\phi_i = 0$. This because a positive level of appropriability leads to positive rents earned. However, it is plausible that movements in ϕ_i are sluggish, and may not react quick enough to macroeconomic shocks (shocks in y^n). This causes inefficiencies in a volatile environment.

4.2.2 An empirical example

Professor Caballero argues that this model goes a long way in explaining some of the phenomena in Europe in the 1970s and 1980s. Looking specifically at France, he points out that institutions in the early seventies changed in favor of the factor labor. The changes took place in the Grenelle accords, and with the institution of labor inspectors. Indeed it seems like labor was the scarce factor in those years, and welfare-institutions were its way of raising ϕ_L. With

[5] In figure 3, at the going prices, factor 2 earns more in the joint sector than in the autarkic sector. There is no arbitrage because there are no more units 1 willing to cooperate. Thus the market for factor 2 is segmented.

the turning of the tide in favor of the factor capital, France and the continent found themselves suffering from 'Eurosclerosis'. Reducing appropriability again was now resisted by the fortunate 'insider' workers of the segmented labor market.

Some of these assertions were tested during the workshop with an amended version of the above model. Using 'suitable' parameters, the main features of French wages, profits, and interest rates could thus indeed be mimiced.

5 Concluding Remarks

This reprise of professor Caballero's lectures touched upon the history of investment theory and showed two new ways of looking at this age-old economic field. Having incorporated microeconomic observations and information- and appropriability-aspects, investment theory seems capable of generating more insight into a variety of economic problems.

References

Barro, R. J. and Sala-i-Martin, X.: 1995, *Economic Growth*, McGraw-Hill, New York.

Caballero, R. J.: 1997, Aggregate Investment: A 90's View. Forthcoming in the *Handbook of Macroeconomics*, J. Taylor and M. Woodford, eds.

Caballero, R. J., Engel, E. M. R. A. and Haltiwanger, J. C.: 1995, Plant-Level Adjustment and Aggregate Investment Dynamics, *Brookings Papers on Economic Activity* **0**(2), 1–54.

Caballero, R. J. and Hammour, M. L.: 1996, The Macroeconomics of Specificity. NBER Working Paper # 5757, Cambridge, MA.

Doms, M. and Dunne, T.: 1993, An Investigation into Capital and Labor Adjustment at the Plant Level. mimeo, Center for Economic Studies, Census Bureau.

Gale, D.: 1995, What Have We Learned from Social Learning? Mimeo, Boston University, Boston, MA.

Rotemberg, J.: 1987, The New Keynesian Microfoundations, *in* O. J. Blanchard and S. Fischer (eds), *NBER Macroeconomics Annual*, MIT Press, Cambridge, MA, pp. 69–104.