Macroeconomía y Costos de Ajuste
Cátedras 1 y 2
4 de agosto de 2008

Eduardo Engel

Yale University

CEA - DII - U. de Chile
Agosto, 2008.

I. Basics

1. Course Overview

- Standard in macro: representative agent model
- Advantages:
  - simplicity
  - some major success stories (e.g., growth theory)
- Limitation 1: ignores heterogeneity (in endowments, histories...)
  - could miss important aspects of aggregate dynamics
- Limitation 2: unrealistic micro foundations
  - macro, limited data and strong assumptions
- Limitation 3: heterogeneity is essential for some macro topics:
  - income distribution
  - sticky-price and sticky-wage models

An Old Example: Modigliani's Life-Cycle Theory

Deaton (2005):

Once upon a time, before we had quite the high status that we enjoy today, it was common for economists to be harassed by scientists (high status of yesterday) who wanted to know whether we had ever come up with anything that was neither trivial nor obvious. Such questions were asked in the clear expectation that the answer would be no, or would be so unsatisfactory as to lead quickly to that conclusion. When faced with such a challenge, I would always talk about Franco Modigliani and his life-cycle theory of saving.

An Old Example: Modigliani's Life-Cycle Theory

- Important motive for saving: retirement
- Life-cycle story: wealth of the nation gets passed around
  - very young: little wealth
  - middle aged people have more
  - wealth peaks at retirement
- Assets shed by the old are taken up by the young who are still in the accumulating phase of their life-cycle
- In an economy with no growth, assets just get passed around and there is no aggregate saving

An Old Example: Modigliani's Life-Cycle Theory

- Next consider an economy that is growing
- With population growth: more young people than old people, thus more saving than dissaving
- If incomes are growing: the young are saving on a larger scale than the old are dissaving
- In both cases: positive aggregate saving
- Saving determined by growth rate of aggregate income, not by income level: poor countries can save more than rich countries
- The theory does well: it is consistently found that saving rates are higher where growth is higher, from the first time that Modigliani looked at the evidence until today when we have more and better data

An Old Example: Modigliani's Life-Cycle Theory

- In simple cases, wealth as a fraction of income equals half the average length of retirement
- A prediction remarkable for its precision, simplicity, and lack of unspecified parameters
- Total wealth of an economy depends on
  - population growth rate
  - individual income growth rate
  - length of retirement

Derivations in the Simplest Case

Notation and Assumptions:

- Timeline: born, work L years, retire $T - L$ years, die.
- Zero interest rate
- $Y_{t,t+k}, C_{t,t+k}, S_{t,t+k}, A_{t,t+k}$: income/consumption/saving/assets in year $t+k$ of cohort born in t
- Y_t, C_t, S_t, A_t: aggregate income/consumption/saving/assets in year t

We assume:

$$Y_{t,t+k} = \begin{cases} 1; & k = 0, \cdots, L-1, \\ 0; & k = L, \cdots, T-1. \end{cases}$$

Derivations in the Simplest Case

Consumption smoothing implies that:

$$C_{t,t+k} = \frac{L}{T}, \qquad k = 0, \cdot, T-1.$$

And since, by definition, $S_{t,t+k} = Y_{t,t+k} - C_{t,t+k}$, it follows that:

$$S_{t,t+k} = \begin{cases} \frac{T-L}{T}; & k = 0, \cdots, L-1, \\ \\ -\frac{L}{T}; & k = L, \cdots, T-1. \end{cases}$$

and therefore

$$S_t = \sum_{k=0}^{T-1} S_{t-k,t} = L\frac{T-L}{T} + (T-L)\left[-\frac{L}{T}\right] = 0.$$

Derivations in the Simplest Case

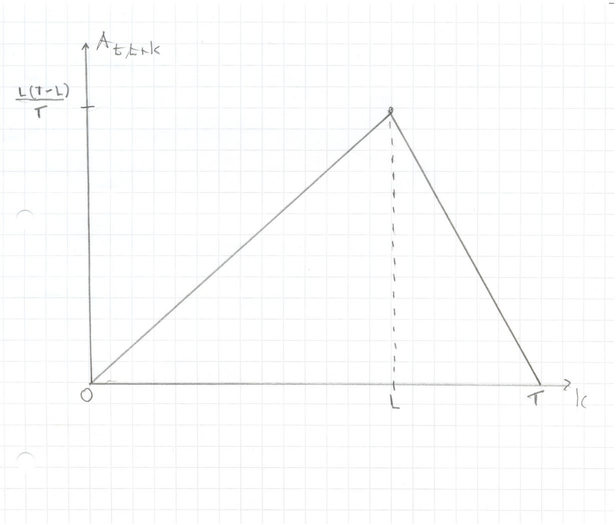
Also, adopting the convention that $A_{t,t+k}$ denotes assets at the end of period k, we have that

$$A_{t,t+k} = \sum_{j=0}^{k} S_{,t,t+j},$$

and therefore

$$A_{t,t+k} = \begin{cases} \frac{T-L}{T}(k+1); & k = 0, \cdots, L-1, \\ \frac{L}{T}(T-k); & k = L, \cdots, T-1. \end{cases}$$

Derivations in the Simplest Case



Derivations in the Simplest Case

We have:

$$Y_t = \sum_{k=0}^{T-1} Y_{t-k,t} = L.$$

From the figure it follows that:

$$A_t \cong \frac{1}{2} \times T \times \frac{T-L}{T} L$$

Combining both expressions above we have:

$$\frac{A_t}{Y_t} \cong \frac{1}{2}(T-L).$$

An Old Example: Modigliani's Life-Cycle Theory

Deaton (2005) concludes saying:

Of course, I was careful never to tell the scientists just how unusual was Modigliani's work. Or to admit that testing theoretical predictions is far from automatic in economics. Or to explain that many economists tend to think of the aggregate economy as if it were a single individual writ large, a "representative agent," instead of following Modigliani, and deriving a theory in which the distinction between individuals and aggregates is not only taken seriously, but is used positively, to derive predictions for the economy that are quite different from the predictions for an individual, or a family. That growth should increase saving rates is a prediction for the aggregate economy that has no counterpart for individuals or families, even though it follows from their behavior. That conversation is another one, in which Modigliani's admirers explain to other economists how economics ought to be done.

This Course

- Serious micro foundations to better understand macro variables
- Ultimate objectives:
  - better models: better informed policy decisions
  - better out-of-sample forecasts using only/mainly macro data better informed policy decisions
- Relevant for: investment, employment, prices, consumer durables, inventories, wealth distribution, ...

Why only during the last 15 years?

- New micro data sets (LRD, JOLS, BLS prices, BED, ...):
    - can test theories, simplifying assumptions, ...
- Computing equilibria and estimating models was impossible with computing power available in the early 90s:
    - equilibrium in DSGE model: infinite-dimensional object
    - estimation of DSGE model: even harder

Two Related Literatures

1. Emphasis on lumpy adjustment and inaction: this course
2. Emphasis on DSGE models with incomplete markets:
    - See Ch. 17 in Ljungqvist and Sargent's book

Recent work on investment (which we will cover in this course) reflects the convergence of both literatures

Some Philosophical Issues

- At what level should we impose rationality?
  - A country
  - Reasonably homogeneous groups of households, firms...
  - Mr. Brown, Mr. Jones, relatives and neighbors
- Blundell and Stoker (JEL, 2005) provide a good survey, focusing on three problems:
  - consumer demand analysis
  - consumption and saving analysis
  - analysis of wages and labor market participation

  and three sources of heterogeneity

  - individual tastes and income
  - wealth and income risks
  - market participation

2. Evidence on Lumpy Behavior and Non-Convex Adjustment Costs

- A. Lumpy micro:
  - macro variables vary “smoothly”
  - micro behavior: inaction most of the time, with infrequent bursts of lumpy behavior
- B. Non-convex adjustment costs:
  - big difference between acting and not acting (even if you act “a little”)
  - rationalizes lumpy behavior

A. Evidence of Lumpy Micro: Investment

- Doms y Dunne (1998)
- 12,000 continuous plants from US manufacturing (LRD)
- 1972–1989, annual frequency.
- Micro findings:
    - For more than half of the plants:

$$
\begin{aligned}
\max_t I_{it} &> 0.3 \sum_t I_{it}, \\
\max_t \Delta K_{it}/K_{it} &> 0.5.
\end{aligned}
$$

A. Evidence of Lumpy Micro: Investment

- Macro finding:
  - The number of spikes (defined as the year with largest investment) explains aggregate investment much better than the average magnitude of the spikes.
  - That is, the extensive margin matters more than the intensive margin

A. Evidence of Lumpy Micro: Investment

- Gourio-Kashyap (JME, 2007)
- Data from the U.S. and Chile
- Define spike: $I/K > 20\%$

A. Evidence of Lumpy Micro: Investment

F. Gourio, A. K Kashyap / Journal of Monetary Economics 54 (2007) 1–22 5

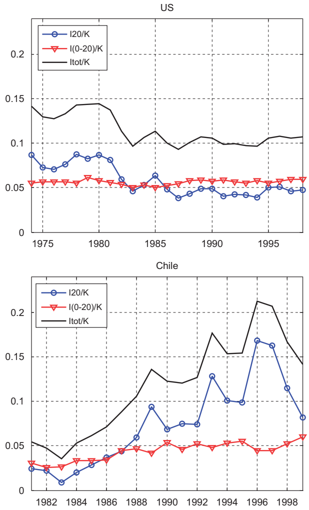


Fig. 1. Decomposition of aggregate investment for U.S. and Chilean manufacturing plant into investment spikes and remaining investment. *Note:* See Table 1 for definitions of the series.

A. Evidence of Lumpy Micro: Employment

- Davis and Haltiwanger (1999): flow approach, job creation, job destruction
- In representative agent models: either creation or destruction equals zero
- In reality:

$$
\begin{aligned}
\mathrm{CR}_t &\equiv \sum_i \mathrm{creation}_{it} \gg 0; \\
\mathrm{DE}_t &\equiv \sum_i \mathrm{destruction}_{it} \gg 0.
\end{aligned}
$$

A. Evidence of Lumpy Micro: Employment

- U.S.-manufacturing job creation explained by plants that increase L by more than 20%: 69.3%
- U.S.-manufacturing job destruction explained by plants that reduce L by more than 20%: 76.9%
- Canada: 75.2, 77.7; Denmark: 76.5, 79.6; Israel: 78.2, 84.7.

A. Evidence of Lumpy Micro: Durables

- Evident: most of the time consumers are inactive.
- For concrete data: Bar-Ilan and Blinder (1992), Eberly (1994), ...

A. Evidence of Lumpy Micro: Prices

- 1990s: many papers looking at frequency of price adjustments for very specific goods (e..g, newspapers)
- Only recently: broad database covering most of CPI
- U.S.: Bils and Klenow (JPE, 2004), Nakamura and Steinsson (2006)
- Europe: Dhyne et al. (2006) and Fabiani et al. (2006): Inflation Persistence Network.

Bils and Klenow

Data:

- Monthly prices used to build the CPI
- BLS, 1995–2001
- 70,000–80,000 prices per month from around 22,000 outlets in 88 geographic areas
- Covers 68.9% of consumer spending
- 350+ categories of consumer goods and services: Entry Level Items (ELI)

Main Findings

- Median frequency of price changes: 4.3 months
- Median frequency of price changes after adjusting for sales ('regular' prices): 5.5 months
- Frequency of price adjustments differs dramatically across goods

Nakamura and Steinsson

Use longer and more detailed version of data set used by BK, also consider producer prices (we'll focus on consumer prices only)

Obtain the following ffacts:

1. median duration of regular prices during 1998–2005 lies between 8 and 11 months, depending on how substitutions, sales and missing observations (stockouts) are treated; they provide 8 estimates
2. one-third of regular price changes are price decreases
3. the frequency of price increases covaries strongly with inflation while the frequency of price decreases and the size of price increases and price decreases do not
4. the frequency of price change is highly seasonal: highest in 1st quarter, lowest in 4th quarter

Explaining the Difference between BK and NS

- Median frequency of regular price adjustments: Bils-Klenow obtain 5.5 months, Nakamura-Steinsson obtain between 8.0 and 11.0 months, depending on how stockouts, sales and substitutions are dealt with
- An important part of the difference: BK only had data for sales on some sectors and extrapolated to the remaining sectors (the 'uniformity assumption')
- Excluding sales more than doubles the median duration of consumer prices even though only 20% of price changes are due to sales. This happens because sales are concentrated in a few sectors (food, apparel) and these sectors have frequencies of price changes close to the median.

Explaining the Difference between BK and NS

- Klenow-Kryvtsov (2007) acknowledge that BK underestimated price durations because of the above, but argue that this brings their median duration of regular prices up to 7.2 months for 1988–2005, not to the 8-11 month range.
- The remaining difference is explained by the fact that there are many ways of accounting for stockouts, sales and substitutions. Also, KK consider a subsample of three major cities in their analysis, which is only 20% of total data set.
- My personal reading: which of the 8 estimates in NS you use, taking values between 8.0 and 11.0 months, will depend on the model you have in mind

A. Evidence of Lumpy Micro: Prices

- Average frequency of price adjustments:
  - U.S., 1995-2000: between 8 and 11 months
  - Euro Area, 1990-2003: approximately 12 months

B. Evidence of non-convex adjustment costs: Investment

Ramey and Shapiro (JPE, 2001):

- Consider aerospace plants that closed during the early 90's.
- Closed because of end of cold war (avoids selection bias).
- Non-trivial part: calculate replacement price of installed capital (equipment).
- Conclusion: average sales price 72% below the replacement price.
- Strong evidence in favor of investment irreversibilities.

B. Evidence of non-convex adjustment costs: Durable Goods

- Buying/selling a house: total (buyer and seller) commission between 1 and 5%.
- Let α denote total commission in any transaction, then cost of adjusting stock from K to $K + \Delta K$, $\Delta K \neq 0$ is:

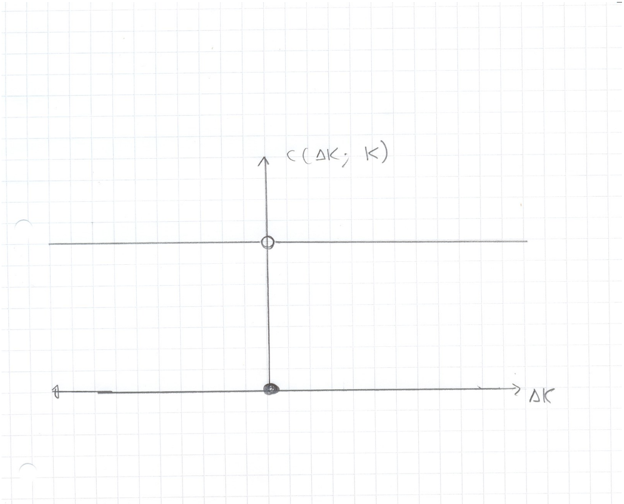
$$C(\Delta K; K) = 2\alpha K + \alpha \Delta K \cong 2\alpha K.$$

- Hence we have a discontinuity at $\Delta K = 0$:

$$C(\Delta K = 0; K) = 0.$$

and adjustment costs are non-convex.

B. Evidence of non-convex adjustment costs: Durable Goods



B. Evidence of non-convex adjustment costs: Prices

- Literal non-convex adjustment costs:
  - menu costs: small costs with large macro effects
  - cost of reviewing current price structure
- More general:
  - psychological costs of doing something, procrastination
  - goodwill costs and customer markets

Menu costs?

- Initially (1985): small ('menu') costs of adjusting prices.
- Why do small adjustment costs have large effects ('money matters')?
- One possibility: small differences at the micro level lead to large differences at the aggregate level.

Or significant adjustment costs?

- Levy et al (1997), Zbaracki et al (2004):
  - Quantify various components of cost of adjusting prices.
  - Consider supermarkets, pharmacies, ...
  - Components considered:
    - Labor cost of changing (re-labeling) price tags on shelves
    - Cost of printing and distributing new 'tags'
    - Cost of collecting information
    - Cost of making a decision
    - Cost of communicating the decision
  - Results:
    - 1st and 2nd comp.: 0.5% of sales; 10–15% of profits.
    - Costs of 3rd, 4th and 5th comp.: 0.7% of sales
    - Total: 20–30% of profits

B. Evidence of non-convex adjustment costs: Prices

Overcoming Fear of Terrorist Attacks

- Becker and Rubinstein (2004)
- *The Economist*, Economic Focus, 7/21/2005
- Response to 9/11: airline miles in the US
- Response to suicide bombers in Tel Aviv buses: bus rides
- Some people stop riding buses/airplanes altogether, others take as many trips as before
- Overcoming fear of terrorism as an “investment” that may (or may not) be worth making
- Data suggest investment in courage is a fixed cost, not a variable cost

3. Classic Aggregation Problem

- Classic problem in microeconomics (aggregate demand, see chap. 4 in Mas-Collel et al.)
- First of 3 'issues' considered by Mas-Collel et al:

> *"Individual demand can be expressed as a function of prices and the individual's wealth level. When can aggregate demand be expressed as a function of prices and aggregate wealth?"*

Some Simple Results

- From Stoker (1986)
- Considers a problem similar to the one stated by Mas Collel et al, emphasizing topics we'll cover in this course.
- Assumes that individual optimization leads to:

$$y_{it} = \alpha + \beta_1 x_{it} + \beta_2 x_{it}^2 \tag{1}$$

where:

  - α, β_1, β_2: common across individuals, constant over time.
  - y_{it}: consumption by individual i at time t.
  - x_{it}: wealth (or income) of i in t.
  - $i : 1, \ldots, n$, $t : 1, \ldots, T$.

Some Simple Results

- We define per capita consumption and income via:

$$\mu_t(y) = \frac{1}{n}\sum_i y_{it}, \qquad \mu_t(x) = \frac{1}{n}\sum_i x_{it}.$$

- Representative agent paradigm: the relation at the individual level also holds for the "representative agent", i.e., with per capita consumption $\mu_t(y)$ and per capita wealth $\mu_t(x)$:

$$\mu_t(y) = a + b_1\mu_t(x) + b_2\left[\mu_t(x)\right]^2 \tag{2}$$

- From (1), the correct relation is:

$$\mu_t(y) = \alpha + \beta_1\mu_t(x) + \beta_2\mu_t(x^2) \tag{3}$$

Some Simple Results

- The "aggregation problem" consists in relating the parameters a, b_1 and b_2 from the relation between aggregates derived from the representative agent paradigm with those that characterize micro behavior: α, β_1 and β_2.
- Comparing (2) and (3): square of average vs. average of squares.
- In general:

$$\mu_t(x^2) \geq [\mu_t(x)]^2 \tag{4}$$

with equality only if the x_{it}'s are identical across individuals.

Some Simple Results

- Hence, ignoring the correlation between x_{it} and x_{it}^2, (4) suggests that:

$$|\widehat{b}_2| > |\beta_2|$$

- For example, if x_{it} comes from an exponential distribution with mean θ_t, then the representative agent paradigm posits:

$$\mu_t(y) = a + b_1\theta_t + b_2\theta_t^2$$

while the true aggregate relation is:

$$\mu_t(y) = a + b_1\theta_t + b_2\left(2\theta_t^2\right)$$

so we'll have $\widehat{b}_2 \cong 2\beta_2$.

Distributional effects

- Aggregate consumption fluctuations depend on fluctuations in the income distribution.
- Representative agent model: fluctuations only depend on first moment of the income distribution (the mean).
- When higher moments matter (in this case the second moment): distributional effects.
- Fundamental result from classical aggregation theory (Gorman, 1953): no distributional effects if and only if the relation between individual variables (over which you aggregate) is linear.
- In our example, this corresponds to the case $\beta_2 = 0$. This is the only case where parameters estimated from an aggregate relation correspond to micro parameters.

Example with Intuition

- 50% of households: marginal propensity to consume out of income of 0.1.
- 50% of households: marginal propensity to consume of 0.9.
- $100 are transferred from each low propensity households to each high propensity households.
- $\Delta\mu_t(x) = 0$.
- $\Delta\mu_t(y) = 80$.
- According to the representative agent paradigm [see (2)],

$$\Delta\mu_t(x) = 0 \Rightarrow \Delta\mu_t(y) = 0$$

Conclusion

- Distributional effects in aggregate equations coincide with non-linear micro relations.
- For this reason, relations between macro variables that are estimated in practice are usually assumed linear: it doesn't make sense to include non linear expressions because these cannot correspond to what is obtained by aggregating a micro relation.
- Note that log-linear relations do *not* avoid distributional issues (even though we typically ignore such issues in macro)

Houthakker’s Approach

- Based on Houthakker (1955-1956)
- Main idea:
    - Agents are heterogeneous in some dimension
    - Find conditions under which the aggregate behaves like a particular type of agent
- Difference with classical approach:
    - Behavior of aggregate need not correspond to that of any of micro agents, but may instead be that of a totally different micro unit.

Concrete Example

- Industry with a fixed number (continuum) of firms.
- Firms produce (at most) one unit of a homogeneous good, using a Leontief-type fixed-proportions technology.
- The fraction of firms that requires a units of capital and b units of labor to produce one unit of the good is proportional to $\phi(a,b) = a^{\alpha-1}b^{\beta-1}$; $\alpha > 1$, $\beta > 1$ (i.e., Pareto distribution).
- The price of the good, p, is exogenous.
- Given an opportunity cost r of capital and a wage of w, a firm with parameters (a,b) produces if and only if it has positive profits:

$$ra + wb < p$$

Concrete Example

- Hence, aggregate capital and labor will be:

$$\begin{aligned} K &= \int_0^{p/r} \int_0^{(p-ra)/w} a\phi(a,b)\, db\, da \\ L &= \int_0^{p/r} \int_0^{(p-ra)/w} b\phi(a,b)\, da\, db \end{aligned}$$

  and aggregate production will be:

$$Q = \int_0^{p/r} \int_0^{(p-ra)/w} \phi(a,b)\, da\, db$$

- A patient (but ultimately straightforward) calculation, based on the Beta function, shows that there exists a constant $C = C(\alpha, \beta)$ such that:

$$Q = CK^{\gamma}L^{\delta}$$

  with $\gamma = \frac{\alpha}{(\alpha+\beta+1)}$, $\delta = \frac{\beta}{(\alpha+\beta+1)}$, $\gamma + \delta < 1$

Concrete Example

- Thus:
  - Micro: Leontieff.
  - Macro: Cobb-Douglas.
- A literature exploiting the above line of reasoning developed during the 60's:
  - Can be useful to obtain qualitative results.
  - Not useful for quantitative issues.
  - In most macro problems, the aggregate does not behave as a micro unit.
- Modeling adequately the aggregation problem posits computational problems that could be dealt with only beginning in the early 90's.

Recent Applications

- Ricardo Lagos (REStud, 2006): “A Model of TFP”.
- Explicit model of wage and employment determination
- Standard search and matching model based on Mortensen and Pissarides (1994)
- Shocks: productivity
- Crucial parameter: productivity threshold R
  - Match between firm and worker disolves if productivity falls below R.
  - New matches realized only if initial productivity is above R.

Recent Applications

- R is a function of institutions underlying matching and search technologies.
- Firms have Leontieff technologies, as in Houthakker.
- Aggregate is Cobb-Douglas aggregate, as in Houthakker.
- Show that, in equilibrium, multiplicative parameter in production function (TFP) is proportional to R.
- For another application, see Jones (2004)

4. Quadratic Adjustment Model

A. Static Case

B. Dynamic Case

A. Static Case

- Variable of interest: y_t
  - E.g.: capital stock, employment, price, stock of durable, ...
- Shocks that determine the variable of interest: x_t
  - E.g.: user-cost of capital, wealth, productivity shocks, ...
- Micro relation between both:

$$y_t = \alpha + \beta x_t \tag{5}$$

- Usually: linear relation between the logarithms of variables, for the time being we ignore this issue.

A. Static Case

- Empirical work shows that:
  - x_t explains less than predicted by (5).
  - Adding $x_{t-1}, x_{t-2}, \ldots$ on the r.h.s. of (5) improves the goodness of fit significantly.
  - There's more inertia than suggested by micro models
- Main candidates to solve this problem:
  - Expectations
  - Adjustment costs
    - Objective: include lags without loosing ability of estimating parameters of economic interest (e.g., β).

Intuition: partial adjustment model (PAM)

- Central Idea: (5) represents an equation for the target of y_t, which we denote by y_t^*.

$$y_t^* = \alpha + \beta x_t$$

- Agents close a fraction λ of the gap between y_{t-1} and y_t^*:

$$\boxed{y_t - y_{t-1} = \lambda(y_t^* - y_{t-1})} \tag{6}$$

where $0 < \lambda \leq 1$ is the adjustment speed.

- Taking first differences in (6):

$$\boxed{\Delta y_t = (1-\lambda)\Delta y_{t-1} + \lambda \Delta y_t^*} \tag{7}$$

Intuition: partial adjustment model (PAM)

- Assuming Δy_t^*'s i.i.d., uncorrelated with $\Delta y_{t-1}, \Delta y_{t-2}, \ldots$ often is a good approximation:
  - Economic theory: consumption and Hall's random walk result.
  - Calculate proxy for $\widetilde{y}_t$ (prices: Bils and Klenow (2004), investment/employment: Caballero, Engel and Haltiwanger (1995/1997)) and resulting process is close to a random walk.
- Then can easily show that $\Delta y_t \sim \mathrm{AR}(1)$ and hence $\rho(\Delta y_t, \Delta y_{t-1}) = 1 - \lambda =$ persistence.

Intuition: partial adjustment model (PAM)

- That is:

$$\boxed{\text{Speed of adjustment} = 1 - \text{First-order correlation}}$$

- Take Var on both sides of (7), use that Δy_t is stationary and $\Delta\widetilde{y}_t$ is uncorrelated with $\Delta y_{t-1}, \Delta y_{t-2}, \ldots$, to obtain:

$$\sigma^2_{\Delta y} = \frac{\lambda}{2-\lambda}\sigma^2_{\Delta y^*}. \tag{8}$$

- Hence:
  - $\lambda < 1 \;\Rightarrow\; \sigma^2_{\Delta y} < \sigma^2_{\Delta y^*} \;\Rightarrow$ PAM smooths out shocks.
  - $\lambda \downarrow \Rightarrow$ slower adjustment $\Rightarrow \sigma^2_{\Delta y} \downarrow$.
- Will generalize below to arbitrary Δy^* process

B. Dynamic Case

- Given: $y_{i,t-1}$, period $t-1$ choice.
- Observe: $\widehat{y}_{i,t}$, the frictionless, static optimum, could be any stochastic process (finite conditional expectations).
- Assume $\widehat{y}$ is exogenous (PE)
- Choose $y_{i,t}$.
- Face following tradeoff:
  - Adjustment costs:
    - $C_1(y_{i,t}, y_{i,t-1}) = c(y_{i,t} - y_{i,t-1})^2$.
    - Minimum if don't adjust ($y_{i,t} = y_{i,t-1}$).
  - Deviating from frictionless optimum:
    - $C_2(y_{i,t}, \widehat{y}_{i,t}) = (y_{i,t} - \widehat{y}_{i,t})^2$.
    - Minimum at $y_{i,t} = \widehat{y}_{i,t}$.
  - c captures the relative weight of both costs.

Quadratic Costs Assumption

- If cost functions are *smooth* (continuous, differentiable, etc.), then a quadratic approximation is good (Taylor expansion).
- Yet if there are fixed costs of adjusting, C_1 is discontinuous at $y_{i,t} = y_{i,t-1}$.
- In some applications you can argue that C_2 is discontinuous at $y_{i,t} = \widehat{y}_{i,t}$ (e.g., if $y_{i,t}$ is the price charged by a firm in a competitive industry, then charging an ε above $\widehat{p}_t$, interpreted as the equilibrium price, means selling nothing).
- A solution to the criticism of assuming C_1 quadratic is to work with more realistic adjustment costs (with a discontinuity at $y_{i,t} = y_{i,t-1}$), that is, non-convex adjustment costs. If this is the case, what follows doesn't apply. (More on this later in the course).
- A solution to the criticism of assuming C_2 quadratic is to abandon perfect competition and work with monopolistic competition. Then what follows applies. (More later.)

Main Tradeoff

- Dynamic considerations: choice of y_t affects:
  - present and future adjustment costs
  - present and future costs from deviating from their frictionless optimum
- Sargent (1978): first application to macroeconomics.

Agent's Optimization Problem

The objective function:

$$\min_{y_{i,t}} \mathrm{E}_{i,t} \sum_{j \geq 0} \rho^j \left[c(y_{i,t+j} - y_{i,t+j-1})^2 + (y_{i,t+j} - \widehat{y}_{i,t+j})^2 \right] \tag{9}$$

with:

- $\mathrm{E}_{i,t}$: expected value for agent i conditional on information available at the beginning of period t.
- Information available at t:

$$y_{i,t-1}, y_{i,t-2}, y_{i,t-3}, \ldots ; \widehat{y}_{i,t}, \widehat{y}_{i,t-1}, \widehat{y}_{i,t-2}, \widehat{y}_{i,t-3}, \ldots$$

- $\rho \in (0,1)$ discount factor.

Optimal Policy

$$\boxed{y_{i,t} - y_{i,t-1} = \alpha \left(y^*_{i,t} - y_{i,t-1} \right)} \qquad (10)$$

with:

$$\boxed{y^*_{i,t} = (1-\delta) \sum_{k \geq 0} \delta^k \mathrm{E}_t \left[\widehat{y}_{i,t+k} \right]} \qquad (11)$$

where:

$$\delta = \frac{1 + c(1+\rho) - \sqrt{[1 + c(1+\rho)]^2 - 4c^2\rho}}{2c} \in (0, 1),$$

$$\frac{1}{1-\alpha} = \frac{1 + c(1+\rho) + \sqrt{[1 + c(1+\rho)]^2 - 4c^2\rho}}{2c} \in (1, \infty).$$

Optimal Policy

- Dynamic target: y^*
- Static target: $\widehat{y}$
- Conclusion: Quadratic adjustment costs $\Rightarrow$ PAM with dynamic target equal to a weighted average of expected present and future static targets.

An Important Particular Case

- Assume $\widehat{y}$ follows a random walk:

$$\widehat{y}_t = g + \widehat{y}_{t-1} + \nu_t,$$

  with ν_t i.i.d. $(0, \sigma_\nu^2)$.
- Then:

$$y_t^* = (1-\delta)[\widehat{y}_t + \sum_{k\geq 1} \delta^k(\widehat{y}_t + kg)] = \widehat{y}_t + \frac{\delta}{1-\delta}g.$$

- Hence:

$$\begin{aligned} y_t &= (1-\alpha)y_{t-1} + \alpha\widehat{y}_t + \frac{\alpha\delta}{1-\delta}g \\ \Rightarrow \Delta y_t &= (1-\alpha)\Delta y_{t-1} + \alpha g + \alpha v_t. \end{aligned}$$

- And we conclude that $\Delta y \sim \mathrm{AR}(1)$.
- Except for an irrelevant constant, in this case there is no distinction between the static ($\rho = 0$) and dynamic cases.

5. Calvo Model

A. Static Case

B. Dynamics and Aggregation

A. Static Case

In each period:

- y and $\widehat{y}$: as in Quadratic Adjustment Model.
- Each agent has a probability π of adjusting in a given period.
- Adjustment shock generated by $\xi_{i,t}$, which is equal to 1 with probability π and equal to 0 with probability $1 - \pi$: the 'Calvo fairy' or 'Calvo's green light'
- Adjustment shocks are i.i.d. across agents and for a given agent over time.
- Agent i's adjustment shocks at time t do not depend on previous values of y or current or previous values of $\widehat{y}$.
- When an agent adjusts, she faces no adjustment cost.

A. Static Case

- Equivalent to having each agent draw an adjustment cost in each period from a distribution that takes two values: 0 with probability π and ∞ with probability $1-\pi$.
- There are a very large number of agents (we want to apply the Law of Large Numbers).
- Then, in a static context:

$$y_{i,t} = \begin{cases} \widehat{y}_{i,t} & \text{with probability } \pi \qquad (\text{when } \xi_{i,t}=1), \\ y_{i,t-1} & \text{with probability } (1-\pi) \qquad (\text{when } \xi_{i,t}=0). \end{cases}$$

- Hence:

$$y_{i,t} = \xi_{i,t}\widehat{y}_{i,t} + (1-\xi_{i,t})y_{i,t-1}.$$

A. Static Case

- Aggregating:

$$
\begin{aligned}
y_t &= \pi \times (\text{average } \widehat{y}_{i,t} \text{ of those who adjust}) \\
&\quad + (1-\pi) \times (\text{average } y_{i,t-1} \text{ of those who do not adjust}) \\
&\cong \pi \widehat{y}_t + (1-\pi) y_{t-1}, \qquad (12)
\end{aligned}
$$

  In the last step we used the indep. assumptions for adjustment shocks (and the Law of Large Numbers).
- We conclude that $y_t \sim$ PAM with $\lambda = \pi$.
- Intuition: Looking at aggregates you cannot distinguish between an economy where:
  - All agents adjust a fraction λ of their gap $(\widehat{y}_{i,t} - y_{i,t-1})$.
  - A fraction λ of agents adjust fully (close their gap to zero) and the remainder doesn't adjust at all.

B. Dynamics and Aggregation

- If the agent adjusts at t, she chooses $y_{i,t}$ keeping in mind that it may be a long time before she adjusts again.
- Hence, if the agent adjusts in t she solves:

$$\min_{y_{i,t}} \mathrm{E}_t \left[\sum_{k \geq 0} \{\rho(1-\pi)\}^k \left(y_{i,t} - \widehat{y}_{i,t+k}\right)^2 \right] \tag{13}$$

where the discount factor is $\rho(1-\pi)$ because the probability that $y_{i,t}$ is still relevant at time $(t+j)$ is $(1-\pi)^j$.

B. Dynamics and Aggregation

- Denoting by $y^*_{i,t}$ the value of $y_{i,t}$ that solves (13), the FOC is:

$$2\mathrm{E}_t \sum_{k\geq 0} \{\rho(1-\pi)\}^k \left(y^*_{i,t} - \widehat{y}_{i,t+k}\right) = 0$$

- Hence:

$$\boxed{y^*_{i,t} = [1-\rho(1-\pi)] \sum_{k\geq 0} \{\rho(1-\pi)\}^k \,\mathrm{E}_t\left[\widehat{y}_{i,t+k}\right].} \qquad (14)$$

- A derivation similar to that of (12) leads to:

$$y_t = (1-\pi)y_{t-1} + \pi y^*_t \qquad (15)$$

where, once again, we used the Law of Large Numbers.

B. Dynamics and Aggregation

- Comparing (14) and (15) with the results in Section 2, we conclude that, by assigning

$$\begin{array}{ccc} \alpha & \longleftrightarrow & \pi \\ \delta & \longleftrightarrow & \rho(1-\pi) \end{array}$$

the aggregate dynamics of both models are indistinguishable.
- An econometrician that only observes aggregate data *cannot* distinguish between a quadratic adjustment and a Calvo model.

Comments

- The equivalence result derived above (Quadratic Adjustment and Calvo) is due to Julio Rotemberg (NBER Macro Annuals, 1987).
- We will also refer to the probability π of adjusting as speed of adjustment and will denote it by λ.
- In the case where $\widehat{y}$ follows a random walk we have that $y^* = \widehat{y} + c$, where c is a constant that depends on the drift of the random walk (equal to zero if the drift is zero).

Comments

- Calvo's (1983) paper became popular more than a decade later, when his approach was used to model price adjustments in general equilibrium models (both keynesian and neoclassical), because it strikes a good balance between tractability (prices are only part of the story in these models) and price inertia (more later).
- Despite the previous point, Calvo's assumptions lack realism (we'll return to this later).

Prices and Calvo's Model

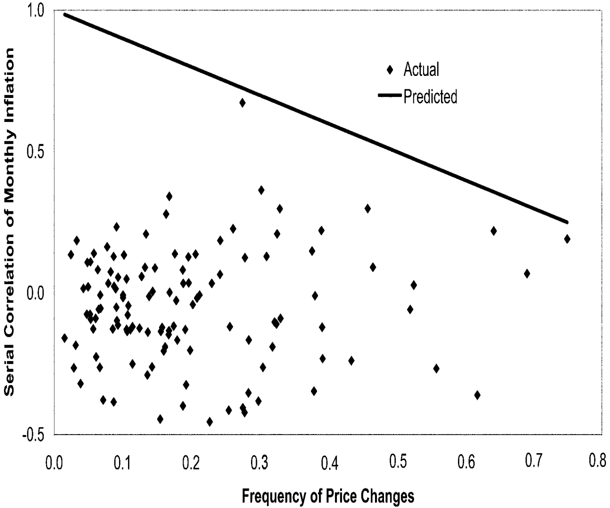
- Bils and Klenow (2004)
- Calculate micro-frequency of price adjustments, U.S., 1995–2000, 123 consumption categories
- Calculate first-order autocorrelations of inflation series
- Under Calvo (also under quadratic adjustment), assuming the log of nominal marginal costs follow a random walk:

$$\hat{\rho}_1 = 1 - \hat{\lambda}$$

  with:

  - $\hat{\rho}_1$: first-order autocorrelation from π_{it} series
  - $\hat{\lambda}$: micro-speed of adjustment

Testing the Calvo Model



Testing the Calvo Model

- In all (123) cases: $\hat{\rho}_1 < 1 - \hat{\lambda}$
- Bils-Klenow calculate first-order correlations of $\Delta\widehat{y}$: average close to zero
- Conclusion:
  - adjustment speed inferred from first-order autocorrelation much larger than true adjustment speed
  - product-level inflation series exhibit considerably less persistence than predicted by the Calvo model

6. Application: Labor Regulation and Adjustment

1. Motivation
2. Methodology and Data
3. Results
4. Robustness
5. Gauging the cost
6. Conclusion
7. Plant-level evidence

Sources:

- 6.1–6.6 based on Caballero, Cowan, Engel and Micco (2004)
- 6.7 based on Caballero, Engel and Micco (2005)

6.1. Motivation

- Creative-destruction at the core of economic growth:
  - Factors move from less to more productive firms
- More than half of productivity growth in U.S. Manufacturing
- Creative-destruction is hampered by adjustment costs:
  - Labor regulation: job security provisions

Impact of labor regulation on creative-destruction

- Difficult to find in the data:
    - Measuring restructuring?
    - Measuring labor regulation?
    - Data?
- Yet likely to be large:
    - Evidence of large idiosyncratic shocks

This Paper

Develops a methodology to bring together:

- New data on job security
- Cross-country sectoral data on employment and output
- Distinction: de jure and de facto regulation

Main Results

- Effective labor regulation matters a lot
- From 20th to 80th percentile in job security in countries with strong rule of law:
  - Speed of adjustment: down by 37%
  - Annual productivity growth: down by 0.85%

6.2. Methodology and Data

- A. Firm's Problem
- B. Employment Target
- C. Regressions
- D. Data

A. Firm's Problem

Country-sector representative firm solves:

$$\min_{e_t} \mathrm{E}_t \left[\sum_{j\geq 0} \rho^j \left\{ (e_{t+j} - \widehat{e}_{t+j})^2 + k_{t+j}(e_{t+j} - e_{t+j-1})^2 \right\} \right]$$

with

$$k_t : \quad \text{i.i.d.} = \begin{cases} 0 & \text{with prob. } \pi_0 \\ K & \text{with prob. } \pi_k \\ \infty & \text{with prob. } \pi_\infty \end{cases}$$

$\widehat{e}_t$: static employment target

Firm's Problem (cont.)

Generalizes quadratic adjustment and Calvo models:

- $\pi_k = 1$: quadratic adjustment costs
- $\pi_k = 0$: Calvo (lumpy) adjustment
- $0 < \pi_k < 1$: both smooth and lumpy adjustments

Solution

$$\boxed{\Delta e_t = \psi_t(e_t^* - e_{t-1})}$$

with

$$\begin{aligned} e_t^* &= (1-\tau)\sum_{j\geq 0} \tau^j \mathrm{E}_t[\widehat{e}_{t+j}], \\ \psi_t &\equiv \psi(k_t) = \begin{cases} 0 & \text{if } k_t = \infty \\ \nu & \text{if } k_t = K \\ 1 & \text{if } k_t = 0 \end{cases} \end{aligned}$$

Estimating Equation

- Estimate speed of adjustment from:

$$\boxed{\Delta e_{jct} = \psi_{jct}(e^*_{jct} - e_{jc,t-1}) = \psi_{jct}\mathrm{Gap}_{jct} = \lambda_c \mathrm{Gap}_{jct} + \text{error}}$$

  with ψ_{jct}'s:

  - i.i.d., indep. of $e_{jc,t-1}, e_{jc,t-2}, \ldots$
  - $\mathrm{E}[\psi_{jct}] = \lambda_c$;
  - $\mathrm{Var}[\psi_{jct}] = \lambda_c(1 - \lambda_c) - \pi_k \nu_c(1 - \nu_c)$

- Need:

  - Measure of the employment gap
  - Methodology to estimate the λ_c

B. Employment Target

Model

- Output, demand, wage schedule:

$$
\begin{aligned}
y &= a + \alpha e + \beta h \\
p &= d - \frac{1}{\eta} y \\
w &\cong w^o + \mu(h - \bar{h})
\end{aligned}
$$

- Key assumption:
  - adjustment costs when changing employment
  - no adjustment costs when changing hours
- $y = \log(Y), a = \log(A), e = \log(E), h = \log(H), \ldots$

Implication

- $\widehat{e}$ and $\widehat{h}$ determined from FOC of

$$\max_{E,H} = PY - W(H)EH = D(AE^{\alpha}H^{\beta})^{\gamma} - W(H)E,$$

  with $W(H) = H^{\mu} + \Omega, \gamma = (\eta - 1)/\eta$.
- Obtain: $\widehat{h}$ constant (indep. of shocks)
- h satisfies FOC for given value of e
- From 3 FOC (two involving $\widehat{e}$ and $\widehat{h}$, one involving e and h):

$$\boxed{\widehat{e} - e \;=\; \frac{\mu - \beta\gamma}{1 - \alpha\gamma}(h - \widehat{h})}$$

- Expression used in Caballero and Engel (1993)
- Problem: don't have data on hours

Static Target

Analogous derivation leads to:

$$\boxed{\widehat{e} - e \;=\; \frac{\phi}{1 - \alpha\gamma}(v - w^o)}$$

- v: marginal productivity
- $\phi = \mu/(\mu - \beta\gamma)$ decreasing in elasticity of marginal wage schedule to hours
- w^o: frictionless average productivity

From Static to Dynamic Target

- Problem
  - We derived an expression for the static target: $\widehat{e}$
  - Δe depends on the dynamic target:

$$e_t^* = (1-\tau)\sum_{j\geq 0} \tau^j \mathrm{E}_t[\widehat{e}_{t+j}].$$

- Solution:
  - Assume:
    - $\widehat{e}$ follows a random walk
    - Allow time-country specific drift
  - Data consistent with this assumption

Where We Stand

- We have derived:

$$\boxed{e^*_{jct} - e_{jc,t-1} = \frac{\phi}{1-\alpha\gamma_j}\left(v_{jct} - w^o_{jct}\right) + \Delta e_{jct} + \delta_{ct}}$$

- We need:
  - Proxy for w^o
  - Estimate for ϕ

Where We Stand

- Two candidates (both average across sectors) to proxy w^o:
  - Wage: competitive labor market
  - Productivity: more robust
- Used both, no discernible differences
- Statistical power comes from large idiosyncratic shocks
- Work with more robust option

Differences in Labor Quality

- **Problem**
  - Unobservable differences in labor quality (across sectors)
  - Would lead to downward bias in speed of adjustment
- **Solution**

$$(v_{jct} - v_{\cdot ct}) \longleftarrow (v_{jct} - v_{\cdot ct}) - \theta_{jct}$$

$$\theta_{jct} \equiv \frac{1}{2}[(v_{jct-1} - v_{\cdot ct-1}) + (v_{jct-2} - v_{\cdot ct-2})].$$

Estimating ϕ

- Taking first differences in

$$e^*_{jct} - e_{jct-1} = \frac{\phi}{1-\alpha\gamma_j}(v_{jct} - \theta_{jct} - v_{\cdot ct}) + \Delta e_{jct} + \delta_{ct}$$

- leads to

$$\begin{aligned}\Delta e_{jct} &= -\frac{\phi}{1-\alpha\gamma_j}(\Delta v_{jct} - \Delta v_{\cdot ct} - \Delta\theta_{jct}) + \kappa_{ct} + \upsilon_{it} + \Delta e^*_{jct} \\ &\equiv -\phi z_{jct} + \kappa_{ct} + \varepsilon_{jct}\end{aligned}$$

Results

- Use $(\Delta w_{jct-1} - \Delta w_{\cdot ct-1})$ as instrument for $(\Delta v_{jct} - \Delta v_{\cdot ct})$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
z_{jct}	-0.280	-0.394	-0.558	-0.355	-0.387	-0.363	-1.168	-0.352
	(0.044)	(0.068)	(0.135)	(0.119)	(0.116)	(0.091)	(.357)	(0.103)
Obs.	22,810	22,008	8,311	6,378	7,319	7,730	6,883	7,036
Inc.	All	All	1	2	3	All	All	All
JS	All	All	All	All	All	1	2	3
Extr.	Yes	No	No	No	No	No	No	No

- In the remainder of the paper use $\phi = 0.4$
- Run robustness checks

Advantages of our Gap Measure

$$\boxed{\mathrm{Gap}_{jct} \;=\; \frac{\phi}{1-\alpha\gamma_j}(v_{jct}-\theta_{jct}-v_{\cdot ct})+\Delta e_{jct}}$$

- Summarizes all shocks in a single variable
- Only requires data on employment and nominal output

C. Regressions

$$\Delta e_{jct} = \lambda_{ct}(\mathrm{Gap}_{jct} + \delta_{ct}) + \text{error}$$

$$\lambda_{ct} = \lambda_1 + \lambda_2 \times \mathrm{JS}_{ct} + \lambda_3 \times (\mathrm{JS}_{ct}\mathrm{RL}_{ct})$$

Hence estimate:

$$\boxed{\Delta e_{jct} = \lambda_1\, \mathrm{Gap}_{jct} + \lambda_2 \mathrm{Gap}_{jct} \times \mathrm{JS}_{ct} + \lambda_3 \mathrm{Gap}_{jct} \times (\mathrm{JS}_{ct}\mathrm{RL}_{ct}) + \widetilde{\delta}_{ct} + \varepsilon_{jct}}$$

Parameters of interest: λ_2, λ_3.

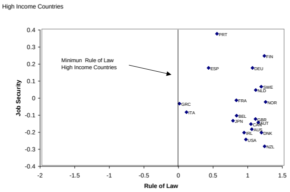
D. Data: Job Security

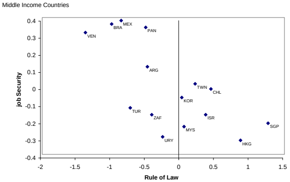
- Botero et al.: 60 ctries., no time variation (1997), broader:
  - grounds for dismissal
  - dismissal procedures
  - notice and severance payments
  - employment protection in the Constitution
- Heckman and Pages: 24 ctries., time variation, narrower:
  - expected cost (at hiring) of future dismissal

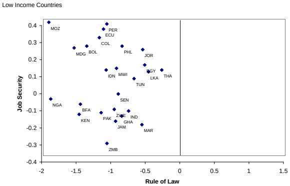
D. Data: Enforcement

- Kaufmann et al., two measures:
  - Rule of Law
  - Government Efficiency

Results







- Rule of Law and Job Security are far from perfectly correlated with income
- Rule of Law negatively correlated with Job Security in Middle Income Countries

Industrial Statistics

- Output, employment and wages
- UNIDO
- 2002 2-digit Industrial Statistics database
- Considered 1980–2000

6.3. Results

- Include country-year fixed effects
- Enforcement measures (RL/GE): use dummies determined by weakest OECD ctry. (Greece)
- Include $\mathrm{Gap} \times \mathrm{EM}$ control when $\mathrm{Gap} \times \mathrm{JS} \times \mathrm{EM}$ is included.

Main Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Change in Log-Employment				
λ_1:	0.600	0.603	0.607	0.611				
	(0.009)***	(0.008)***	(0.012)***	(0.012)***				
λ_2:		-0.080	-0.015	-0.025	-0.126	-0.027	-0.038	
		(0.037)**	(0.051)	(0.051)	(0.041)***	(0.052)	(0.051)	
λ_3-RL:			-0.514			-0.314		
			(0.068)***			(0.070)***		
λ_3-GE:				-0.515			-0.326	
				(0.068)***			(0.071)***	
λ_2-HP:								-0.022
								(0.007)***
Obs.	21,733	21,733	21,733	21,733	21,733	21,733	21,733	12,012
R2:	0.60	0.60	0.60	0.60	0.61	0.61	0.61	0.62
G-I Int.	No	No	No	No	Yes	Yes	Yes	Yes
G-S Int.	No	No	No	No	Yes	Yes	Yes	Yes

Comments on Main Results

- Col. 1: average speed of adjustment 0.60
- Col. 2 and 5: adjustment speed decreasing in JS
- Cols. 3–4 (and 6–7):
  - enforcement matters
  - de jure job security now insignificant
- Col. 8: use HP-JS measure
- Economic significance: 20th to 80th percentile in JS (Col. 3)
  - Strong RL: $\lambda \downarrow$ 0.22
  - Weak RL: $\lambda \downarrow$ 0.006

6.4. Robustness

See the paper.

6.5. Gauging the Costs

- A different metric
- Slower adjustment $\Rightarrow$ lower productivity
- Use AK-model to go from levels to growth rates
- Standard caveats

Explicit Expressions

$$
\begin{aligned}
\frac{\Delta Y}{Y} &\simeq \left[\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right]\xi \\
\xi &= \frac{\alpha\gamma(2-\alpha\gamma)}{2(1-\alpha\gamma)^2}\sigma^2 \\
\Delta g_Y &\simeq (g_Y + \delta)\frac{\Delta A}{A}.
\end{aligned}
$$

Gauging the Costs

Change in Job Security Index	Cost in Annual Growth Rate	
	Weak Rule of Law	Strong Rule of Law
20th to 40th percentile	0.002%	0.083%
40th to 60th percentile	0.007%	0.292%
60th to 80th percentile	0.008%	0.478%
20th to 80th percentile	0.023%	0.853%

6.6. Conclusion

New methodology to combine:

- New data on labor regulation (job protection)
- Cross-country sectoral data on employment and output

Main Findings:

- Effective labor regulation matters a lot
- 20th→80th percentile in JS in ctries. with strong RL:
  - Speed of adjustment: down by 37%
  - Annual productivity growth: down by 0.85%

6.7. Plant-level Evidence

- Establishment level data.
- Brazil, Colombia, Chile, Mexico, Venezuela: 140,000 obs.

Results: Flexibility in Latin America

- Brazil, Chile, Colombia more flexible than Mexico, Venezuela.
- Small firms substantially less flexible than large firms (except Venezuela).
- Behavior of large establishments explains x-ctry. differences.
- Increasing hazards in all countries, more pronounced in large establishments in more flexible economies.

Results: Evolution of Flexibility in Chile

- Significant decline in flexibility toward end of sample (1997–99).
- Bigger decline in larger establishments
- Main decline in adjustment to labor shortages
- Productivity cost: 2.7% upon impact, 0.44% every year thereafter.